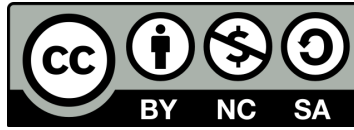


A History of the Middle East

Original Author: Robert J. Flynn, Ph.D.

Open Educational Resource released under the Creative Commons License
(Attribution-NonCommercial-ShareAlike 4.0 International)



Licensed by Portland Community College

Contents

Preface	iv
Chapter One: Introduction	1
Chapter Two: Muhammad and Islam	8
Chapter Three: The Arab Conquest and the Establishment of the Caliphate	37
Chapter Four: Religious, Social, and Intellectual Changes During the Caliphate, 632-1258	72
Chapter Five: The Decline of the Arab Empire, 809-1300	110
Chapter Six: The Gunpowder Empires.....	147
Chapter Seven: Ottoman Crisis and Adaptation, 1566-1768	185
Chapter Eight: Western Intrusions, Eastern Responses, 1768-1878.....	215
Chapter Nine: The Twilight of the Ottoman Empire, 1878-1916	256
Chapter Ten: “A Peace to End All Peace”.....	284
Chapter Eleven: The Independent States, 1918-1948	313
Chapter Twelve: The Middle East Under European Control, 1922-1948	353
Chapter Thirteen, The Establishment of Israel, 1918-1948.....	388
Chapter Fourteen: The Age of Nasser, 1949-1967	431
Chapter Fifteen: Arabia, Turkey, and Iran, 1949-1979.....	477
Chapter Sixteen: The Middle East After Nasser, 1967-1979	523
Chapter Seventeen: Transition and Stalemate, 1980-1990	568
Chapter Eighteen: The Gulf War and Its Legacy, 1990-2001	597

Preface

The History of the Middle East is a single volume account of the region's development from the time of Muhammad to the eve of the 2001 al-Qaeda attacks. It is intended to serve as the main textbook for a single semester or term class and is designed to be accessible to students who do not possess prior knowledge of the region or its history. Perhaps unusually for a textbook, this work includes footnotes. I made this choice to ensure that students would be able to understand and access the body of scholarship on which I based my interpretations. Sadly, owing to the fact that I drafted most the book during the Covid-19 lockdowns of 2020 and early 2021, I was frequently compelled to use older sources than I would have preferred.

This work was completed as an Open Educational Resource (OER) with the assistance of a grant from Open Oregon. As such, people are welcome to engage in the 5Rs of OER. They may retain, reuse, revise, remix, and redistribute this textbook so long as they adhere to the following conditions: that they do so for non-commercial purposes, that they assign the same licensing conditions to any revised versions as I have applied to the original work, and that they acknowledge Robert Flynn as the original author.

I am committed to keeping this work up to date. Accordingly, I would welcome people to contact me via email at robert.flynn@pcc.edu with suggestions, comments, and, especially, details of any errata or factual inaccuracies. I will address typographical and factual errors immediately and will take suggestions and comments into account when I make periodic larger revisions of the textbook.

I would like to acknowledge the assistance of a number of friends and colleagues in the production of this work. I am indebted to Rosa Bettencourt, Phil Seder, Chris Shelley, Terri Barnes, David Armontrout, Ben Weber, and, especially, Chris Brooks for reviewing draft chapters and providing feedback. Amy Hofer of Open Oregon provided assistance navigating the OER process. My dean, Dana Fuller, was kind enough to arrange for a single course release in the fall of 2020 that proved invaluable in permitting me to complete the first draft. Finally, I would like to thank my beautiful wife, Betsy Boyd-Flynn, and our children, Molly and Tommy for allowing me to spend countless hours in our basement engaged in drafting and revising this work.

Chapter One: Introduction

History of the Middle East is a survey of Middle Eastern History from the time of Muhammad in the seventh century until the eve of the War on Terror in 2001. It is designed to meet the requirements of a single-term introductory course and is aimed at students who have no more than a casual understanding of the history, cultures, and societies of the region. It offers a relatively concise narrative of the region's historical development that seeks to avoid oversimplification on the one hand and overcomplexity on the other.

Why Study History

This work rests on the assumption that a knowledge of history is an essential prerequisite for the proper functioning of a democratic polity. Self-governing political systems depend on the ability of their citizenries to make wise and productive choices about how to approach complex problems. Effective decision making, in turn, requires a strong comprehension of the context of the questions under debate and an appreciation of how people addressed similar situations in the past. In other words, citizens must understand the history of the issues that they confront if they are to make informed choices.

Importantly, however, they need to comprehend the subject not as students and laypeople typically perceive it, but instead in the way that professional historians do. These understandings are very different. The popular definition of history describes it as the accumulated events, actions, decisions, cultural constructs, relationships, and institutions that stretch back at least as far as the development of civilization some six-thousand years ago. According to this view, everything that people have ever done is history. Professional historians hold a different conception of the subject—one that is both quite distinct from the popular understanding and far more useful for those seeking to be engaged participants in a democratic system. According to this definition, history is the study of the past for the purpose of interpreting it—that is, of discerning meaning in it—in order to provide what historians call a “usable past” that can help people in the present-day contextualize and better comprehend the issues that they face. Historians are particularly interested in how current political structures, cultural practices, power relations, and social institutions developed. For example, scholars today study events such as the Paris Peace Conference of 1919 to better comprehend why many Middle Eastern nations have weak civil societies and, why, as a result, many of them lack genuinely representative governments. In other words, professional historians research the past not because they are history “buffs” or antiquarians who find the subject intrinsically interesting—though they do—but instead because the subject offers vital insights into the issues and problems that citizens in democratic societies confront today.

An understanding of history along these lines can certainly help Americans acquire a better grasp of the development of a region that is, today, of vital importance

to their country. Such has not always been the case. Indeed, prior to the First World War, US engagement with the region was limited largely to Protestant missionary activities and to related educational efforts such as the establishment of the American University of Beirut. With the discovery of petroleum in Saudi Arabia just before World War II and with the onset of the Cold War in the late 1940s, however, the US began to assert its interests in the Middle East both formally and informally. American involvement in the region steadily grew over the ensuing decades as the region's emergence as a major player in the global economy, and it rose sharply beginning in the early 1980s to the point that the Middle East became, after 2001, the primary focus of American foreign policy and military engagement. As the historian of international relations Andrew Bacevich perceptively notes, “[f]rom the end of World War II to 1980, virtually no American soldiers were killed in action while serving in that region. . . . Since 1990, virtually no American soldiers have been killed in action anywhere *except* in the Greater Middle East.”¹

Despite Washington's increasingly intensive involvement in the Middle East over the past few decades, however, most Americans still lack a basic understanding of the region's history. Few are aware of the ways in which Islam shaped the Middle East or understand the role that European imperialists played in forming its political and diplomatic contours. Even government officials and legislators who oversee American policy toward the Middle East are frequently ignorant of rudimentary aspects of the region's history and culture and fail to understand how those factors continue to influence contemporary identities and fault lines. As the journalist Jeff Stein noted in 2006, many congresspeople and policymakers—including the FBI's chief of counterterrorism—could not summarize something as basic as the difference between Sunnism and Shi'ism at a time when American troops were desperately trying to impose order on a country, Iraq, that was riven by communal violence between those confessional groups. In light of the importance that Washington continues to assign to the region, Americans and other Westerners—be they policymakers, legislators, or citizens—clearly need a substantially better understanding of the Middle East if they are to avoid the policy missteps of the recent past.²

Approach

This textbook is designed to satisfy that need and to offer students and faculty an affordable yet comprehensive text. I have written it explicitly with introductory courses in mind and have organized and presented the material accordingly. To meet the needs of all students, it requires no background knowledge of Middle Eastern history. To facilitate readability and to emphasize the role of causality as a force in history, it largely adheres to a traditional, chronological approach to the subject. To align with most Middle Eastern

¹ Andrew J. Bacevich, *America's War for the Greater Middle East: A Military History* (New York: Random House, 2016), 11.

² Jeff Stein, “Opinion | Can You Tell a Sunni From a Shiite?,” *The New York Times*, October 17, 2006, <https://www.nytimes.com/2006/10/17/opinion/17stein.html>.

history classes, it begins with the emergence of Islam in the seventh century CE and omits coverage of the region's long, pre-Islamic history.

The book focuses predominantly on the political history of the Middle East. Its basic narrative consequently emphasizes the region's political development, international relations, and the formulation and execution of state policies. It follows this approach for two reasons. First, for the entire period covered in this text, key political questions—Who should rule? How should the state be structured? What powers should the government have?—occupied the attention of the region's peoples. In other words, the Middle East's political history is significant because those who lived in the region and made its history believed it was important. Second, far from being antithetical to other ways of exploring the topic such as those that examine it through a social or cultural lens, a focus on political history can offer a solid starting point for inquiries based on those approaches. Specifically, this text provides students in courses organized along those lines with the foundational knowledge needed for them to contextualize—and thus better understand—the Middle East's cultural and social evolution. This is not to suggest that this work exclusively centers on politics. While social, cultural, intellectual, and religious history may not organize the broader structure of the work, they nonetheless receive significant attention and are woven into the broader narrative in ways that relate them to the textbook's focus on political developments.

This work also does not advance a strong unifying argument. At first blush, this choice may seem to contradict my earlier claim that interpretation is the main focus of professional historians; in fact, it is consistent with that view. I have opted to produce a work that is—such as can be the case—neutral in tone because I believe strongly that the interpretive direction of a given class is, in the end, the purview not of the author of the adopted textbook but instead of the person teaching that course. By deemphasizing my interpretation of events, this textbook frees professors and teachers who use it to articulate their own themes. Please note, however, that this approach does not mean that this book eschews historical argumentation. Instead, it seeks to expose students to some of the more important historiographical debates of recent decades by addressing them in the overall narrative. For example, Chapter Two largely follows the traditional story of Muhammad, but also explores the very different interpretation of Islam's origins that revisionist scholars such as Patricia Crone have developed.

While the text does not seek to advance a particular argument, it does develop two broad themes. First, it explores the part that faith—particularly the religion of Islam—played in shaping the political, social, cultural, and intellectual development of the region. Understanding religion's influence can help us to address several key questions. What impact has Islam had on the nature of modern Middle Eastern states? To what degree are present-day gender relations in the region the product of Islam? Did religion advance or impede scientific and philosophic knowledge in the Muslim world? To what extent did it shape political ideology and the structure of government in the region? Second, this book examines the nature of interactions between the Muslim Middle East and the West with the aim of helping students to develop a more

sophisticated understanding of that relationship. It is a commonplace in many quarters today that the West and the Muslim world have been in a constant state of conflict going back to the Arab conquest of Syria and Egypt in the 630s and 640s and that the two cultures are today locked in an ongoing “Clash of Civilizations.” This work argues implicitly that such an interpretation is, at best, a gross oversimplification that ignores the many instances in which the two regions have enjoyed cooperative rather than conflicting relations.³

Structure

This book is organized chronologically and divides the history of the Middle East into four periods: Islam and the Caliphate, The Ottoman Empire, the Western Intrusion, and Decolonization and America’s Bid for Mastery.

The first section focuses on the emergence of Islam and on the growth and transformation of the Arab-Muslim Caliphate. Chapter Two explores Muhammad (c. 570-632), the revolutionary political, social, and cultural message that he brought to Arabia, and the origins of the Muslim community. It also provides a very brief overview of Islam’s main tenets and characteristics. Chapter Three surveys the Arab Conquest and the formation of the Arab Empire from the death of the prophet through the high point of the Abbasid Caliphate. Chapter Four assesses intellectual trends during the Arab Empire and the ways in which Islam evolved, diversified, and took on many of its defining characteristics in the centuries following Muhammad. Chapter Five examines the gradual weakening of the Abbasid Caliphate, the growing political dominance in the Middle East of recently arrived Turkic people, the Crusades, and, finally, the Mongol destruction of the Abbasid Caliphate.

The second section assesses the Middle East under Ottoman rule. Chapter Six explores the growth of the empire, tracing the emergence of the Ottoman Dynasty, its rise to dominance, and the characteristics that made it successful. It also examines the similar Safavid Dynasty in Iran. Chapter Seven looks at how the Ottoman Empire evolved from the mid-sixteenth to the mid-eighteenth centuries. It also addresses the historical debate over whether changes in the empire are better understood as a case of decline or transformation. Chapter Eight focuses on two topics. First, it looks at the growing Western diplomatic and economic dominance of the Middle East from the 1760s through the 1870s. Second, it examines the bifurcated regional response to this challenge wherein some states, political players, and thinkers reacted by promoting policies designed to imitate the West while others instead argued that the region could best counter the European states through a revival of Islamic practices and values.

The third section reviews the period of European imperial domination from World War I through the late 1940s. Chapter Nine explores political changes in the Ottoman Empire in the late-nineteenth and early-twentieth centuries, the emergence of Zionism,

³ See Samuel Huntington, “The Clash of Civilizations?,” *Foreign Affairs* 72, no. 3 (1996): 22–49.

and the outbreak of the First World War. Chapter Ten examines the conclusion of the Great War and the subsequent diplomatic settlement that resulted in the European powers establishing control over nearly the entire region and imposing on it new frontiers and a Western-style state system. Chapter Eleven studies the efforts of Saudi Arabia, Iran, and Turkey to maintain their independence during the Interwar Period and World War II, while Chapter Twelve surveys the experience of Egypt, Lebanon, Iraq, Transjordan, and Syria under Western control between the Paris Peace Conference and 1948. Chapter Thirteen traces the struggle for control of Palestine during the 1920s and 1930s, and the 1948 War that accompanied the establishment of Israel.

The final section of the book centers on decolonization and America's effort to impose order on the region. Chapter Fourteen investigates the ways in which decolonization, the Cold War, nationalism, hostility among the Arab states, and the Arab-Israeli conflict shaped the development of the region from 1949 to 1967. Chapter Fifteen assesses the history of Saudi Arabia, Turkey, and Iran between World War II and 1979, paying special attention to the coup against Prime Minister Mohammad Mossadegh (1882-1967) in 1953 and to the Iranian Revolution in the late 1970s. Chapter Sixteen explores the transformative events of the 1970s—including the emergence of religious fundamentalism in the region—that culminated in the watershed year of 1979. Chapter Seventeen focuses on the dominant events and themes of the period from 1979 to 1990. It examines the Israeli invasion of Lebanon, the growing popularity of Islamism, the *Intifada*, and the Iran-Iraq War. Finally, Chapter Eighteen focuses on the period from 1990 to 2001. It reviews the Gulf War, the failed effort to negotiate a permanent settlement of the conflict between Palestinians and Israelis, the challenge that domestic jihadi movements posed to the region's established regimes, and the growth of the transnational al-Qa'ida terrorist organization.

Geographic Focus

Before we get into the history of the Middle East, we must define its territorial extent and understand how and why people came to view it as a distinct area. Of obvious Eurocentric origins, the term 'Middle East' is a comparatively recent neologism. Prior to the nineteenth century, Westerners had used the older terms 'Near East' or 'Nearer East' to refer to the European, African, and Asian territories then under the control of the Ottoman Empire. By the 1840s, however, some British colonial officials had become increasingly dissatisfied with those terms because they left out Iran and the Arabian side of the Persian Gulf. In response, policymakers in Britain's India Office coined the term 'Middle East' in the 1850s to describe an enlarged understanding of the region that included those areas. Though the expression gained some traction in certain parts of the British government during the latter part of the century, it failed to catch on at that time. Ultimately, it was the American naval strategist and historian Alfred Thayer Mahan's use of the term in his influential article "The Persian Gulf and International Relations" in 1902 that popularized it. Usage of the expression grew rapidly thereafter, culminating in the British government's decision to designate the body it established to coordinate the defense of the region during World War II as the Middle East Command—a choice that cemented the term's growing popular acceptance. By the

1950s, as a result, it had fully replaced the term 'Near East' as a descriptor for the region.⁴

This textbook delineates the region along lines very similar to those used by the Middle East Command during World War II. According to this definition, the region includes the territory encompassed by modern-day Iran, Iraq, Kuwait, Bahrain, Qatar, the United Arab Emirates, Oman, Yemen, Saudi Arabia, Egypt, Jordan, Israel, the Occupied Territories, Lebanon, Syria, and Turkey. Note, however, that this text will not confine itself exclusively to the area encompassed by that definition. Not only will it discuss events in non-Middle Eastern places such as Western Europe and the United States, but it will also, as necessary, stretch the definition of the region to include territories in North Africa, Central Asia, and the Balkans that were, at times, culturally or politically integrated into the Middle East.

The region's physical geography abounds in ironies. On the one hand, the Middle East is tremendously arid; Egypt and nearly all of the Arabian peninsula receive at best negligible rainfall, and much of the area cannot sustain agriculture. On the other, the great river valleys of the Tigris, Euphrates, and Nile have sustained large, complex, and highly productive agricultural societies since the dawn of civilization. Indeed, Mesopotamia—the land between the Tigris and Euphrates—was the place where people first practiced intensive agriculture, and Egypt and Mesopotamia remained the greatest wealth producing regions of western Eurasia until the late Middle Ages. Likewise, while the Middle East is famously devoid of basic natural resources such as timber and many ores, it today possesses oil deposits of enormous strategic, economic, and environmental significance.⁵

Less ironic has been the region's role in trade and cultural exchange. For millennia, the Middle East has been a literal crossroads. The Sinai Peninsula land bridge long served as a conduit for commerce between Southwest Asia and Egypt, while the easily crossed Bosphorus and Dardanelles Straits firmly connected the region to Europe. The Middle East also brought disparate regions together via shipborne trade. The Suez Canal and earlier waterways linking the Red Sea to the Nile River connected the Mediterranean to East Africa and South Asia by ship, while the Euphrates River provided easy passage for goods moving between the ports of the Persian Gulf and Asia Minor and the Black Sea. As we shall see, this fortuitous location proved to be a mixed blessing for the region. Under the Arab Empire, the region became the central hub in a massive system of commerce that connected Africa, Europe, South Asia, Central Asia, and China. In more recent times, however, the Middle East's oil reserves

⁴ Peter Beaumont, Gerald H. Blake, and J. Malcom Wagstaff, *The Middle East: A Geographical Study* (New York: Wiley, 1988), 16; Clayton R. Koppes, "Captain Mahan, General Gordon, and the Origins of the Term 'Middle East,'" *Middle Eastern Studies* 12, no. 1 (1976): 95–98.

⁵ Hosny Hasanean and Mansour Almazroui, "Rainfall: Features and Variations over Saudi Arabia, A Review," *Climate* 3, no. 3 (September 2015): 578–626, <https://doi.org/10.3390/cli3030578>.

and strategic location drew the unwanted attention of European imperial powers such as Britain, Russia, and France.

Culturally, the Middle East has always been very diverse. Today, most of the region's people identify as Arab. Arabs predominate in Egypt, the Arabian Peninsula, Syria, Jordan, Lebanon, and Iraq. However, the Middle East also includes large numbers of people who identify as Turkish and Iranian, as well as smaller numbers of Kurds, Armenians, and Jewish people. The region's religious diversity matches its ethnic heterogeneity. While the people of the Middle East overwhelmingly practice Islam, its Muslim population is divided between the Sunni majority and a Shi'i minority who are concentrated in Iraq, Iran, and the Persian Gulf region. The Middle East also has smaller numbers of Jewish people, Druze, and Alawites, as well as Maronite, Assyrian, and Coptic Christians.

A Word About Religion

This book is neither a religious history nor a Muslim Studies textbook, but it does touch upon faith, especially Islam, to a significant degree. As religious beliefs can be sensitive for observant students, I want to explain up front how this text will address the subject. Designed for higher-education classes, it adheres to academic rules regarding the use of evidence and argumentation. It thus cannot, and will not, weigh in on the validity of religious beliefs that are, by definition, articles of faith. As such, while it seeks to be respectful of peoples' beliefs, this textbook leaves faith to the reader and treats all religions as human constructions.

Transliteration

This book transliterates non-English words according to the International Journal of Middle East Studies (IJMES) Translation and Transliteration Guide.

Chapter Two: Muhammad and Islam

In 610 CE, an event would occur that would alter the trajectory of world history. That year, a successful Arab merchant named Muhammad (c. 570-632) claimed to have been visited by the angel Gabriel while meditating in a cave. According to Muhammad, Gabriel commanded him to “recite!” Terrified, he pleaded that he could not. Again, the angel made his demand:

Recite in the name of your Lord who created—Created man from a clinging substance. Recite and your Lord is the most Generous—Who taught by the pen—Taught man that which he knew not (Qur’an 96: 1–5).

This time, as Muhammad later claimed, words poured from his mouth like a torrent. To Muslims, this was a pivotal moment—the point at which Muhammad became God’s final messenger, or the Seal of the Prophets.

Thus began both the religion of Islam and its holy book, the Qur’an, which means, literally, “recitation.” Over the next twenty-three years, Muhammad would reveal the rest of the Qur’an, which he claimed originated with God, or, in Arabic, Allah, and which he declared was God’s final revelation to humanity. He called the religion Islam, or submission to the will of God, and referred to his followers as Muslims, or those who surrender to God. It was a modest beginning for an ardently monotheistic faith—one that would soon reshape the political, social, cultural, and economic map not just of the Middle East but of the entire globe.

The Pre-Islamic Middle East

Before we get to Muhammad and the religion he established, we must first place explore the world in which Islam emerged. All social movements are products of the times and circumstances in which they arise. In innumerable ways, they reflect the values, social hierarchies, and beliefs of the cultures from which they emerge. They also respond to and echo the frictions and challenges of their era and succeed most often when they are able to offer answers to those problems. The Islamic movement was no different. The culture, mentalities, and beliefs of the Arab people shaped the new religion even as Islam borrowed heavily from other monotheistic faiths—particularly Judaism and Christianity. Above all, the new religion that Muhammad espoused reflected the evolving socioeconomic circumstances of late-sixth and early-seventh century Arabia, and it drew strength from its ability to offer solutions to the rise in inequality that those changes produced.

The Arab People

Nearly the entire population of the Arabian peninsula at the time of Muhammad’s birth was Arab. A Semitic people, the Arabs were well acclimated to the harsh climate of the region in part because they had, around 3,000 BCE, domesticated an animal ideally

suited to that arid landscape: the camel. Capable of surviving for long periods of time without water, camels permitted the Arabs to travel between oases and to cross otherwise-inhospitable regions of the peninsula. Later, Arab merchants would carry out extensive trade using vast fleets of camels that could efficiently transport large quantities of goods.¹

Even in the time of Muhammad most Arabs lived a nomadic existence. Called bedouin, the nomadic Arabs followed a pattern common to most pastoral cultures. That is, they did not have complex, impersonal political systems of the type common among sedentary societies; instead, they organized themselves along less-formal, tribal political and social arrangements in which familial ties defined identity and group membership. In this system, the authority of elders and the threat of ostracism were generally all that were needed to maintain order. Likewise, the bedouin had a comparatively simple economic system typical of tribal societies. Similar to that of many other nomadic cultures, it mixed hunting and gathering, herding animals, and slash-and-burn agriculture.²

As was also often the case in other tribal societies, bedouin engaged in near-constant intertribal warfare. This form of conflict was very different from the territorial, trade, and resource-oriented fighting characteristic of both the sedentary empires of Late Antiquity and the states of the modern era. Rather than trying to conquer land or wipe out rivals—though they occasionally fought such wars—the bedouin instead typically engaged in low-intensity intertribal raids, or *ghazu*. They stole herd animals, fought and killed men from rival tribes in small engagements, captured women and children, and extracted tribute from the comparatively wealthy town Arabs. These low-intensity raids served several important purposes. On a material level, they provided a way to increase both the size of the tribe and its wealth. More importantly, the *ghazu* constituted a regulated system of violence that channeled the energy and discord of the tribes' young men outward in a way that ensured domestic tranquility while also giving those warriors a sanctioned opportunity to gain the status and social recognition needed to become respected and influential members of their tribe.³

Living in a tribal system, the bedouin neither had nor needed a formal, written legal system. Instead, they abided by a social code called *murawah* that governed behavior and public morality. The *murawah* code emphasized those values that were essential to the survival and prosperity of the tribe: bravery, preservation of honor, loyalty, care for the poor, hospitality, and protection of the weak. Valor, honor, and fealty are obviously critical in an environment in which *ghazu* are the norm—after all, group survival depended on such beliefs—but why did the *murawah* code place so much

¹ See Richard W. Bulliet, *The Camel and the Wheel* (Cambridge, MA: Harvard University Press, 1975).

² John L. Esposito, *Islam: The Straight Path*, Fourth Edition (New York: Oxford University Press, 2011), 2–3.

³ Robert G. Hoyland, *Arabia and the Arabs: From the Bronze Age to the Coming of Islam* (New York: Routledge, 2001), 99–101, 225.

stress on values such as generosity and concern for the group's most vulnerable members? It did so because the survival of both the tribe and its elite depended on unity and on every member contributing to the good of the collective. Only by caring for the weakest members of the tribe could the powerful ensure that those on the lowest rungs would help them during times of crisis.⁴

Not all Arabs were bedouin. A significant minority lived in towns and cities, particularly in the Hijaz, the part of Arabia that lies along the Red Sea coast. Given the region's scant rainfall—it receives less than 200 millimeters of rain annually—such settlements could only exist at the numerous oases that dotted the area. Taking advantage of the regular supply of water that the oases provided, towns like Taif, Yathrib, Mecca, and Badr sustained themselves by growing dates and cereal crops. They did not focus exclusively on agriculture, however. Instead, over time, these settlements became centers of long-distance commerce. They transshipped goods from north to south and used the profits they earned to buy Damascus swords, wine, and, especially, Syrian wheat. Town Arabs had a symbiotic relationship with the region's bedouin. While the two groups often clashed, they nonetheless engaged in important trade relationships in which the settled Arabs exchanged agricultural goods for the bedouins' animal skins, meat, and livestock. Town Arabs also retained close ties with those tribes to whom they were related.⁵

Muslims have long regarded the period before Muhammad as *al-Jahiliyya*, or the Age of Ignorance, a time during which the Arab people had lost their faith in the one God and had become polytheists. As the scholar of religion Reza Aslan argues, however, this interpretation fails to capture adequately either the rich diversity of pre-Islamic Arab religious beliefs or their syncretic nature—meaning their inclusion of elements from a variety of faiths. Most Arabs of this period were polytheists. They acknowledged the existence of 360 tribal gods that they believed oversaw the day-to-day functioning of the material world. Beyond the tribal deities there existed a supreme God, Allah, who had created the universe. The Arabs conceived of Allah as a remote deity that did not answer to supplication; instead, people prayed to one of the lesser gods for intercession if they needed help.⁶

Pre-Islamic Arab religious beliefs went well beyond polytheism, however. Some, particularly in Yemen, had converted to Christianity over the prior few centuries, and Christian theology permeated the beliefs of nearly all Arabs. The Arabian peninsula likewise had a significant number of Jewish communities composed of both Jewish immigrants and Arab converts. Christian and Jewish theology profoundly influenced the religious views of the Hijaz. Regardless of affiliation, for example, nearly all Arabs

⁴ Karen Armstrong, *Muhammad: A Prophet for Our Time* (New York: HarperOne, 2006), 15–17.

⁵ Hoyland, *Arabia and the Arabs*, 4, 102–9; Juan Cole, *Muhammad: Prophet of Peace Amid the Clash of Empires* (New York: Nation Books, 2018), 11–13.

⁶ Reza Aslan, *No God but God: The Origins, Evolution, and Future of Islam*, Updated Edition (Random House, 2011), 3–10.

believed that they descended from Abraham—the patriarch of the three Abrahamic faiths, Judaism, Islam, and Christianity—through Ismael, the product of Abraham’s union with his maid, Hagar. Indeed, by the sixth century, most town Arabs had come to believe that Allah was the same as the God that Christians and Hebrews worshipped.⁷

Muhammad’s hometown of Mecca had long stood as the dominant religious center in Arabia. Located about halfway up the Hijaz along the trade routes stretching from Yemen in the south to the Roman Empire in the north, the city owed its rise to prominence to its possession of Arabia’s principle religious shrine: The Ka’ba. Believed to have been erected by Adam, rebuilt by Noah after the Flood, and later restored by Abraham and Ismael, the Ka’ba was a small, roofless masonry structure that housed a black, meteorite rock believed to represent the distant creator God.⁸

The Ka’ba was already an important shrine when, in the late-fourth century, the Quraysh tribe that dominated Mecca persuaded the bedouin to house the idols representing the region’s 360 tribal gods in it. The success of this initiative dramatically raised the stature and economic position of the town. Thereafter the tribes of Arabia visited the Ka’ba each winter during an annual pilgrimage known as the hajj. The pilgrims frequented inns, food stalls, gambling dens, and brothels and, in the process, made many Quraysh rich. The trade that accompanied the annual pilgrimage was even more lucrative. Thanks to the presence of the Ka’ba, the entire town had become sacred ground in which violence and intertribal conflict was *haram*, or forbidden. Trade could thus occur in Mecca even between members of hostile tribes, with the result that the city became a vast marketplace during the hajj. The Quraysh profited handsomely from this commerce. Not only did it generate substantial revenue through a modest tax that the Quraysh imposed on the sale of all goods during the pilgrimage, but, more importantly, it lent Mecca a newfound economic significance that, like gravity, gradually pulled the region’s trade routes to it.⁹

The Byzantine and Persian Empires

While the Arabs dominated the Arabian peninsula at the time of Muhammad’s birth, the rest of the Middle East fell under the rule of two of the greatest states of Late Antiquity: the Byzantine and Persian Empires. From their capital of Constantinople, Byzantine emperors ruled Italy, the Balkans, Asia Minor, Syria, Palestine, North Africa, and grain-rich Egypt. Based in the city of Ctesiphon near modern-day Baghdad, meanwhile, the Persian Sassanid Dynasty controlled the Iranian plateau and the rich, grain-producing region of Mesopotamia. Both of these states were venerable. Cyrus the Great (c. 600-530 BCE) had founded the Persian Empire in the sixth century BCE, and the Sassanid Dynasty had ruled it since 224 CE. Heir to the Roman Empire, the Byzantine Empire was a continuation of the state that the Emperor Caesar Augustus

⁷ Aslan, 3–10.

⁸ Aslan, 23–28.

⁹ Aslan, 26–28.

(r. 27 BCE-14 CE) had created in the first century BCE out of the already very-old Roman Republic.¹⁰

Religion played an important role in both states. Most people who lived in the Persian Empire were adherents of its state religion: Zoroastrianism. Dating to the sixth-century BCE, Zoroastrianism was a dualist faith similar in many ways to Christianity that held that the god of light, Ahura Mazda, was locked in an eternal struggle with an evil deity named Ahriman. The Sassanid Empire was not uniformly Zoroastrian, however; instead, it also possessed a number of sizeable religious minorities—particularly in Mesopotamia. That region had a large Jewish community that dated to the Babylonian Captivity as well as a significant population of Christians. Persian shahs were well aware of both their empire's religious diversity and the challenges that confessional unrest could pose to their rule. Accordingly, though the dynasty had officially adopted Zoroastrianism as its state religion in the early third century, it pursued a policy of toleration toward other faiths designed to avoid internal dissent.¹¹

Meanwhile, by the sixth century, Christianity had become the overwhelmingly dominant religion in the Byzantine Empire. Initially an underground movement, it had grown slowly but steadily in the Roman world between the death of Christ and the end of the third century. Following the Emperor Constantine's (r. 306-337) declaration of religious toleration in the early fourth century, however, Christianity exploded in popularity. Indeed, by the end of the fourth century, it had become the official religion of the empire.

Even as Christianity was rapidly gaining converts and winning the support of the state, it was also experiencing significant internal conflict. Theological in nature, the debates centered on Christ's essence and the concept of the Trinity. First articulated in the third century by the orthodox theologian Tertullian (c. 155-220), the Trinity holds that God, Jesus, and the Holy Spirit are simultaneously both distinct entities and a single, unified deity, and that Jesus contained both divine and human elements.¹²

Many Christians did not agree with Tertullian's orthodox position. Some, called Monophysites, argued that a being could not be composed of two essences, and concluded that Jesus was wholly divine in nature. Monophysitism won few supporters in the empire outside of the Middle East, but its adherents constituted a majority of the population in Egypt, where they were called Copts, and in Syria, where they were known as Jacobites. The Monophysites soon found themselves in a difficult position. Both the Church and the Roman state placed great value on conformity of belief and sought to suppress theological positions that challenged orthodoxy; accordingly, under the direction of the emperor Marcian (r. 450-458), the Church reaffirmed the orthodox

¹⁰ Hugh Kennedy, *The Prophet and the Age of the Caliphates*, Second Edition (New York: Routledge, 2004), 1–6.

¹¹ Ira M. Lapidus, *A History of Islamic Societies*, 2nd ed. (Cambridge: Cambridge University Press, 2002), 6–8.

¹² Aslan, *No God but God (Updated Edition)*, 11.

view that Christ had two natures and declared Monophysitism to be a heresy at the Council of Chalcedon in 451. As a result, from the mid-fifth to the early-seventh centuries, the Jacobites and Copts endured periodic persecution at the hands of the ecclesiastical and civil authorities.¹³

Neither empire exerted direct control over the nearby Arabian peninsula. Given its sheer size, fractious tribal population, and challenging geography, conquering the region generated little interest among either Byzantine emperors or Sassanid shahs. Instead, both states sought to ensure the security of their Arabian border by entering into mutually beneficial relationships with Arab client tribes. For example, Constantinople provided the leaders of the largely Christian—though Monophysite—Ghassanid tribe with titles and an annual subsidy of gold; in exchange, the Ghassanids prevented other tribes from raiding the empire from Arabia. Ctesiphon enjoyed a similar arrangement with the predominantly Zoroastrian Lakhmid tribe along its frontier with the peninsula.¹⁴

The client tribes played an important part in the frequent wars between the Persian and Byzantine Empires. Relations between the empires had been poor since they had acquired a common border in the first century BCE. Thereafter, the Roman Empire and its Byzantine successor had waged a series of on-again, off-again wars with successive Persian dynasties for control of northern Mesopotamia. During some periods such as the third century CE, the Persians held the upper hand; at other points, Roman or Byzantine armies dominated the frontier. At no time, however, was one side able to decisively defeat the other or to substantially alter the border between the two empires.¹⁵

Trade in Arabia

In fact, the biggest consequence of the fighting between Rome and Persia related not to territory but instead to commerce. As a result of the Roman upper class's near-insatiable demand for Chinese silk and Indian spices, merchants from the empire imported vast quantities of expensive luxury goods from the east. Caravan merchants brought products originating in China west via the overland Silk Road network of trade routes that crisscrossed central Asia, while maritime traders carried spices and other luxuries from India by ship across the Arabian Sea and up the Persian Gulf. In both cases, the goods then moved through Persian territory along the Euphrates River before crossing the frontier into Roman Syria. In times of peace, this arrangement functioned reasonably well for the Romans and their Byzantine successors. During periods of conflict, however, Sassanid shahs could impose high taxes that merchants would then pass on to consumers; as a result, the trade in luxury goods through Persia

¹³ Chris Wickham, *The Inheritance of Rome: Illuminating the Dark Ages 400-1000* (New York: Penguin Books, 2009), 62–63.

¹⁴ Hoyland, *Arabia and the Arabs*, 78–83.

¹⁵ Hoyland, 80–82.

meant that Rome would, in effect, be footing part of the cost of the Sassanids' wars against the empire.¹⁶

Roman and then Byzantine emperors found this state of affairs intolerable and actively sought during times of conflict to foster alternate trade routes that lay outside the purview of Persian tax collectors. That strategy involved two components. In the north, imperial officials worked to move Silk Road trade with China from the Sassanid Empire to a new terminus on the Roman-controlled Crimean Peninsula. In the south, they induced merchants coming by ship from the Subcontinent to abandon the traditional route through the Persian Gulf in favor of a new one in which they traveled up the Red Sea to Roman ports in Egypt or Palestine.¹⁷

Not every ship entered the Red Sea, however. On the contrary, many instead offloaded their cargos at the port of Aden in southern Yemen. From there, massive camel caravans carried goods northward to trading centers in the Hijaz and thence into the empire. Imperial encouragement of the southern route thus had enormous consequences for the Arabs of the region. Thanks to it, trade from the Subcontinent to the Roman Empire via the Arabian peninsula surged from the first century CE until the late-fourth century CE, enmeshing the Arabs in one of the globe's great long-distance trade networks and creating an economic boom in the Hijaz.¹⁸

Economic Contraction and Recovery

The good times came to an abrupt end in the late fourth century, however. In 384, the Roman Emperor Theodosius negotiated a peace agreement with the new Persian Shah, Shapur III (r. 383-388). While such treaties between Rome and Persia were typically fleeting, this one proved durable. Benefitting from the presence of a common, existential threat in the form of the powerful Hunnic Empire, the peace between the two rivals would, save for a brief border skirmish from 421-422, hold until the early-sixth century.¹⁹

The treaty may have focused on diplomatic goals, but it nonetheless had a far-reaching impact on long-distance trade patterns. Shorter and thus significantly cheaper, the older route via the Persian Gulf and the Euphrates River to Rome gradually reasserted itself as the peace took hold. By the mid-fifth century, as a consequence, merchants bringing goods between India and Rome had largely abandoned the

¹⁶ Bernard Lewis, *The Middle East: A Brief History of the Last 2,000 Years* (New York: Scribner, 1995), 37–38.

¹⁷ Kennedy, *The Prophet and the Age of the Caliphates*, 22–23.

¹⁸ Kennedy, 22–23.

¹⁹ Stephen Williams and Gerald Friell, *Theodosius: The Empire at Bay* (New Haven: Yale University Press, 1998), 41.

overland route up the west side of the Arabian peninsula in favor of the more direct one through the Persia Empire.²⁰

This change had a disastrous impact on the Hijaz. Bereft of much of the trade that had nurtured it, the region entered a long depression in the fifth and early-sixth centuries during which its economy degenerated from a long-distance, cash-based, market-based system into a much poorer subsistence- and barter-oriented one. Many who had previously been involved in the caravan trade were consequently compelled to abandon their occupations in favor of farming or nomadism, with the result that the region's towns shrank significantly. Mecca was a rare exception to this trend. It too suffered from the economic depression, but, thanks to the Ka'ba and the revenue the hajj produced, the city managed to weather the hard times in a way that its commercial rivals could not.²¹

The economy of the Hijaz languished until events to the north finally delivered the region from its malaise in the sixth century. Salvation came as a result of renewed conflict in the Middle East. Beginning in the 520s, the Persian and Byzantine Empires slipped back into their old pattern of near-constant hostility. Between 526 and 591, they fought a series of wars for control of territory in northern Mesopotamia and for dominance of border states like Armenia. The renewal of conflict soon prompted Constantinople to resume the old Roman practice of encouraging the diversion of trade from Persia so as to avoid the stiff, 25 percent tax that the shah's government had started levying on goods bound for the Byzantine Empire. This effort proved more complicated than it had been in the past, however. In the early 570s, Ctesiphon had seized Yemen and with it control of the entrance to the Red Sea. As a result, ships could no longer travel directly from India to the Byzantine Empire without paying Persian levies.²²

These conditions soon led to the recovery of the Hijazi economy. Ctesiphon's control of Yemen proved key here. With trade via the Red Sea now subject to Persian tariffs, merchants searched for an alternate, tax-free route along which they could take goods from the Arabian Sea to the Byzantine Empire. They found it in the Hijaz. Thanks to a loophole that permitted neutral cities to import cargoes from Persia and then reexport them to the Byzantine Empire without having to pay the tariff, merchants could avoid Persian taxes by sending their goods north via the old overland routes across western Arabia. Very soon, large quantities of spices and other goods from India were traveling once again by ship across the Arabian Sea to Aden and then by camel caravan northward through the Hijaz to Byzantine Syria. Unsurprisingly, this sudden revival of the long-distance trade in luxury goods produced a rapid economic recovery in

²⁰ Lewis, *The Middle East*, 42–43.

²¹ Lewis, 42.

²² Cole, *Muhammad*, 10–13.

the region. In a matter of just a few decades, cities were repopulated, markets restored, and old commercial relationships renewed.²³

No part of the Hijaz benefited from the return of trade more than Mecca. The revitalization of long-distance trade soon flooded it with money and opportunity while the recovery of the other cities and towns in the Hijaz produced an attendant revival of the lucrative hajj. As a result, the booming city rapidly reemerged as the dominant economic nexus in the region.²⁴

The prosperity that Mecca enjoyed in the sixth and early-seventh centuries was not shared equally by all its residents, however. While some leading members of the Quraysh tribe grew fabulously rich, many of the city's residents experienced no benefit whatsoever from the revival of trade. Paradoxically, in fact, those on the bottom rungs of the economic ladder appeared to be in a worse position than they had been before the revival of commerce. Indeed, many were so destitute that they had little choice but to borrow money from the rich at usurious interest rates; when they inevitably proved unable to meet their obligations, their creditors forced them into debt slavery.²⁵

The Beginnings of Islam

It was into this rapidly evolving world that Muhammad was born in the late-sixth century. While he personally enjoyed the fruits of the economic revival that the Hijaz was experiencing, he was appalled by the inequality that had accompanied the city's reemergence as an important commercial center and alarmed by its leading families' repudiation of traditional Arab values. Indeed, the apparent abandonment of the ethic of generosity in favor of greed in late sixth-century Mecca would powerfully shape both his thinking and his religious message.

Muhammad

Most scholars today agree that Muhammad's difficult early life also profoundly influenced his religious views. Born sometime around 570 CE, he never met his father, Abdallah, who died shortly before he was born, and he lost his mother when he was just six-years old. Fortunately for Muhammad, strong family connections spared him from the grinding poverty that was typically the fate of orphans in sixth-century Arabia. Instead, he enjoyed a secure life—first with his paternal grandfather, Abd al-Muttalib (c. 497-579), and then, when he too passed away, with his uncle Abu Talib (c 535-619), who was the head of the Hashimite clan. This latter connection would prove critical for Muhammad. Aware of his nephew's preternatural wisdom and keen intelligence, Abu Talib arranged for the young man to help operate the clan's long-distance caravans.

²³ Cole, 13.

²⁴ Lewis, *The Middle East*, 42–43.

²⁵ Karen Armstrong, *Islam: A Short History* (New York: Modern Library, 2002), 3.

Muhammad excelled at the task. Bright and charismatic, he proved from an early age to be a capable manager and a shrewd negotiator.²⁶

His success eventually brought him to the attention of a wealthy woman named Khadija bint Khuwaylid (c. 555-619). Khadija was a remarkable person who had overcome the patriarchal customs of pre-Islamic Arabia and the deaths of three husbands to become Mecca's most prosperous merchant. Aware of the youthful Muhammad's probity and business acumen, she hired him around 595 to lead a caravan from Mecca to Damascus. Returning with twice the profits she had anticipated, he more than met her expectations. Taken with the young man, Khadija followed by suggesting marriage. Though her proposal was unusual—she was fifteen years older than Muhammad—he accepted. The two thereafter enjoyed a strong, loving relationship, and prospered under Muhammad's deft handling of the family business.²⁷

Thus, by his late twenties, Muhammad had risen to a position of comfort, and respect. Indeed, he held such a high reputation as a capable, forthright, and honest man, that he was frequently called upon to mediate commercial and social disputes in Mecca. In short, he had arrived.²⁸

He may have found personal success, but he was increasingly troubled by the changes that Mecca was experiencing at that time. Recalling his experience as an orphan, he could not help but to contrast the more-and-more sumptuous lifestyle of Mecca's leading families with the increasingly hard circumstances of the city's many impoverished orphans and widows and to wonder how poverty could have increased during a period of economic growth. After pondering the issue for some time, he concluded that the rise in inequality and the rapid growth of the economy were related; that is, the wealth pouring into the city was leading its people to abandon the *muwāḥ* code's emphasis on generosity and protection of the weak in favor of selfishness and greed. How else could he explain how some lived in opulence while others suffered in poverty, or, like many orphans, endured malnourishment?²⁹

Muhammad was very perceptive. He had correctly discerned that the economic changes that accompanied the revival of commerce in the Hijaz had grievously weakened the bonds that had heretofore tied clans together. Gone were the older ethics of egalitarianism and group cohesion embodied in the *muwāḥ* code—values that had ensured not only that the tribe survived but also that the strong cared for the weak. In their place, the merchants of Mecca instead lived by the values of individualism and acquisitiveness. As a result, the city had become a profoundly unequal place, one

²⁶ Jonathan A. C. Brown, *Muhammad: A Very Short Introduction* (New York: Oxford University Press, 2011), 8–10.

²⁷ Armstrong, *Muhammad*, 25–26.

²⁸ Cole, *Muhammad*, 5–8, 28–29; Armstrong, *Muhammad*, 25–26.

²⁹ Karen Armstrong, *Islam: A Short History* (New York: Modern Library, 2002), 3.

starkly divided between a tiny number of rich families and a growing mass of impoverished ones.³⁰

Revelation

Deeply disturbed by these changes, he began visiting the nearby Cave of Hira during the month of Ramadan where he ruminated over the problems that beset his city and considered possible solutions. These meditative sessions had enormous consequences. As Muhammad later claimed, it was during one of them in 610 CE that the angel Gabriel first demanded that he recite God's word. Described at the start of this chapter, this experience deeply shook him. Despairing that he had lost his sanity or been possessed by a desert spirit known as a jinni, he rushed back to Mecca in distress seeking the counsel and comfort of his wife, Khadija. She consoled him and reassured him that he was not mad; why, she argued, would God inflict such a punishment on so honest and upstanding a person as him? Still, despite her reassuring words, he continued to feel troubled. Accordingly, she sent for her Christian cousin Waraqa (?-610). Impressed by Muhammad's description of his meeting with the Angel Gabriel, an elated Waraqa declared that Muhammad should rejoice for "“thou art the prophet of this people.”"³¹

The Early Community of Believers

Heartened by Waraqa's words, Muhammad began to spread his message. At first, he limited his preaching to a small group of close family members and friends. Following Khadija, who had accepted his message from the start, his young cousin Ali Ibn Talib (601-661), his merchant friend and first caliph, or successor of Muhammad, Abu Bakr (c. 573-634), an Ethiopian slave named Bilal (c. 580-640), the future third caliph, Uthman ibn Affan (c. 579-656), and a few others adopted Islam. By late 613, perhaps fifty people—a disproportionate number of whom were the sons and daughters of Mecca's economic elite—had joined his fledgling movement.³²

Why did they choose to follow Muhammad? What caused them to abandon the religious beliefs of their families? They did so because they shared his disdain for the growing inequality and exploitation that had become commonplace in Mecca during its economic recovery. In other words, they joined his fledgling movement because the core message that he articulated—that God was a good deity who loved humanity, who denounced inequitable business practices such as usury, and who demanded not just that the rich cease exploiting the poor but that they affirmatively help the unfortunate—powerfully resonated with them and their desire to reform Meccan society.³³

³⁰ Aslan, *No God but God (Updated Edition)*, 33–34, 40–41.

³¹ Quoted in Esposito, *Islam*, 8.

³² Esposito, 9; Aslan, *No God but God (Updated Edition)*, 42.

³³ Armstrong, *Muhammad*, 42–43, 53–54.

Even after he began to preach publicly, the elite clans of the Quraysh tribe largely ignored Muhammad. They did not care for his calls for egalitarianism or demands for social justice; after all, leading clans like Banu Makhzum and Banu Umayya had grown wealthy under the prevailing economic system and any effort to achieve social justice and egalitarianism would inevitably come at their expense. Given the tiny size of Muhammad's following and its marginal impact on Meccan society, however, they concluded that he could be safely ignored.³⁴

By the mid 610s, however, changes in Muhammad's message and his growing success at gaining followers led his opponents to take a dimmer view of the nascent Muslim movement. Two issues in particular raised their hackles. First, starting in 613, Muhammad began to argue more explicitly that there existed only one God and that the three-hundred-and-sixty idols that the Arabs worshipped in the Ka'ba were mere pieces of wood or stone representative of nothing. The leaders of the dominant clans in Mecca viewed this claim as a serious threat to their future prosperity. Were the Arab people to accept his message, they would no longer come to Mecca for the annual hajj or take part in the accompanying trade that was so essential to the city's ongoing economic well-being. Second, Muhammad was winning more-and-more converts to his faith. Charismatic and well respected, he had an appealing message and was steadily increasing the size of the Muslim community. His success in gaining followers would not in itself have posed a huge problem; however, the conversions were dividing families between ardent Muslims and bitter opponents of Islam and thus causing clan ties—which the scholar Jonathan Brown argues “were the one real institution in Arab society”—to fray dangerously. Thus, to the Quraysh elite, Muhammad was a serious threat to both their continued prosperity and the stability of the community—one, they increasingly concluded, that they would need to eliminate.³⁵

For the time being, however, critics of Islam could not move directly against Muhammad. The issue was his family connections. Though the head of the Hashimite clan, Abu Talib, never himself became a Muslim, he did extend formal clan protection to Muhammad. Accordingly, were his opponents to kill Muhammad, they would initiate a blood feud that would likely go on for years and that would, in the process, surely disrupt the trade on which their status and wealth rested. Muhammad's enemies consequently refrained from attacking him directly and instead responded to his challenge by imposing an economic boycott on the Muslims and by targeting those among his following who did not enjoy the backing of the Hashimite clan or who were of low social rank. Enslaved Muslims confronted the greatest pressure. Many, such as the Ethiopian slave Bilal (580-640), endured torture and at least one, a woman named Sumayya (c. 550-615), was killed.³⁶

Circumstances dramatically worsened for Muhammad and the Muslim community in 619—a year so grim that Muslims refer to it as the Year of Sadness. Two

³⁴ Aslan, *No God but God (Updated Edition)*, 42.

³⁵ Brown, *Muhammad*, 17.

³⁶ Cole, *Muhammad*, 59–63.

deaths were responsible for their change of fortune. First, the death of Abu Talib resulted in Muhammad's uncle Abu Lahab (c. 549-624) becoming the head of the Hashimite clan. A bitter opponent of Islam, Abu Lahab promptly removed the protection that his brother had heretofore granted to Muhammad and the Muslims—thus leaving them exposed and vulnerable. Next, just a few months later, Khadija passed away. Her death robbed Muhammad of his emotional lodestone and denied him the critical advice that she had long provided. Her passing and the death of Abu Talib depressed Muhammad, and, more importantly, put him and his followers in an increasingly tenuous position. Indeed, it soon became clear that he and his fellow Muslims faced not just steadily rising persecution but—far worse—the outright destruction of their movement.³⁷

The Hijra

After three hard years of growing persecution at the hands of Mecca's leading families, the Muslims finally found deliverance in 622. Their salvation came as a result of a growing conflict in the Hijazi town of Yathrib. A sprawling date-farming community located about 340 kilometers north of Mecca, Yathrib was home to five tribes: two pagan-Arab ones and three ethnically Arab but religiously Jewish tribes. For several years, the two pagan tribes, the Khazraj and the Aws, had been locked in a destructive feud that had gradually intensified to the point that, by 621, it had become a serious threat to the stability of the settlement. Desperate to keep the peace, the tribal leaders of the town met and agreed to seek outside mediation that could end the conflict before it destroyed their community. They quickly settled on Muhammad. With a reputation for wisdom and integrity, he seemed like the ideal person to resolve the intertribal vendetta. The assembled leaders consequently sent a delegation to Mecca to ask him to come to Yathrib and arbitrate the feud. In exchange, they offered to let the Muslims relocate to their town. It was doubtless a bittersweet proposal for Muhammad. Though he was surely loath to leave the city of his birth, he saw no other way to ensure the survival of his struggling community. Accordingly, he agreed to the terms of the delegation's proposal.³⁸

Relocating would prove to be challenging, however. As Muhammad well knew, the Arab commitment to familial and clan bonds meant that his opponents would bitterly resist any attempt by their Muslim relatives to leave Mecca. The Muslims would thus have to make the move in a way that would not arouse suspicion. To do so, Muhammad arranged for the Muslims to gradually relocate in small, discreet groups of three or four over several months rather than en masse. To further allay suspicion, he himself planned to remain in Mecca until nearly all of his followers had slipped out of the city. It was a wise and ultimately successful strategy. Over several months, nearly the entire

³⁷ Brown, *Muhammad*, 22.

³⁸ Fred Donner, *Muhammad and the Believers: At the Origins of Islam* (Cambridge, MA: The Belknap Press of Harvard University Press, 2012), 43–44.

Muslim community quietly moved to Yathrib. Eventually, the only believers who remained in Mecca were Ali, Abu Bakr, and Muhammad himself.³⁹

As Muhammad grasped, this approach was a risky one. Without the protection of his followers, he would be dangerously exposed in Mecca and thus vulnerable to his enemies. He was right to fear for his safety. The leading families had in fact devised a plan in which they would take advantage of his growing isolation and lack of protection to kill him in his sleep. Fortunately for Muhammad, he caught wind of the plot before it went into effect. To counter it, he arranged for Ali to distract his enemies by pretending to be him—Ali went so far as to sleep in his bed—while he and Abu Bakr sneaked out of Mecca. The scheme worked. Ali drew the attention of Muhammad’s pursuers long enough for Abu Bakr and Muhammad to slip out of the city to safety. Later, after Muhammad had reached Yathrib, Ali quietly left Mecca and rejoined him and the other Muslims there.⁴⁰

Called the hijra, the migration to Yathrib—which was thereafter known as Medina, or the “city of the prophet”—proved vital to the growth and success of Islam. Indeed, Muhammad and his companions saw it as so pivotal that they devised a new, Muslim calendar in which 622 CE became year one. This interpretation of the hijra’s significance was not mere hyperbole. Indeed, in two critical ways, the move to Medina constituted the central event in Islam’s development. First, it saved the fledgling religion from destruction at the hands of its enemies. Had the Muslims remained in Mecca, their movement would have almost certainly collapsed in the face of the leading clans’ increasingly staunch opposition. Second—and perhaps more important—the hijra marked the full repudiation of pre-Islamic Arabia’s clan- and tribe-based social system. In its place, the Muslims established a new and far-broader form of identity, the *umma*, or community of believers, defined not by ties of kinship but instead by a shared set of religious beliefs. The move to Yathrib thus constituted a sea change in how its members saw themselves—“a revolutionary step,” in the writer Karen Armstrong’s words, that made Muhammad the “head of a collection of tribal groups that were not bound together by blood but by a shared ideology, an astonishing innovation in Arabian society.”⁴¹

While the Muslims constituted a narrowly religious movement while they were in Mecca, they transformed into a broader, religious-and-political community in Medina. Muhammad confirmed this change when he issued the Constitution of Medina shortly after he arrived. At its most basic level, the pact established the political and social foundations of his new community. It defined the relationship between the *Muhajirun*, the Muslims who accompanied Muhammad from Mecca, and the Ansar or “helpers,” those Medinans who had converted to Islam, and it put an end once and for all to the feud between the Aws and the Khazraj. It also established Muhammad as the leader of Medina and outlined the basic laws that would govern the *umma*. On a more fundamental level, it also refined the nature of Islam and set expectations about how its

³⁹ Aslan, *No God but God (Updated Edition)*, 48–49.

⁴⁰ Brown, *Muhammad*, 26.

⁴¹ Armstrong, *Islam*, 14.

adherents were to live. In particular, it made clear that the religion's main focus was not on the salvation of the individual, though that was important, but instead on the construction of a community along the lines that God had revealed to Muhammad.⁴²

The religion evolved rapidly in Medina. Beyond the changes he effected through the Constitution of Medina, Muhammad also established the basic structure of Islamic worship and articulated a number of new Qur'anic verses. Stressing the core elements of his message—the unity of God, the value of community, and the importance of social justice—these revelations included a broad array of legal rulings aimed at creating a revolutionary new society based on equity and the worship of the one God. These new laws included improvements in the legal status of women, the institution of a tax on wealth—the *zakat*—designed to provide charity for the indigent, and the prohibition of usury, or the lending of money at interest. Muhammad also had the first mosque constructed so that the Muslims could pray as a single, unified community in imitation of the unity of God.⁴³

The community that Muhammad established in Medina would loom large in the imagination of subsequent generations of Muslims. Put simply, it would thereafter constitute in the eyes of the faithful the model of the ideal Islamic polity. Indeed, as we shall see, when Muslims later confronted serious problems or when they believed that Islam had lost its way, they frequently concluded that the community of believers could best respond by eschewing new social practices and innovations in favor of a return to the ethics, values, and laws of the early *umma* as it existed in Medina under Muhammad's leadership.

War With Mecca

That future generations of Muslims would idealize the early community of believers in Medina did nothing to help the *Muhajirun* address the significant challenges that they faced in the present. They found life in their new home challenging to say the least. Though welcomed by a large portion of the Medinans—most of the town's polytheistic Arabs converted to Islam shortly after Muhammad's arrival—the Meccans nonetheless confronted a great deal of both overt and latent hostility in their adopted home. Many of the Ansar, particularly the leaders who had been compelled to cede political power, nursed deep grudges against Muhammad and the Meccan immigrants. The Jewish tribes also resented his political dominance of Medina and mocked him for what they saw as his poor understanding of the Abrahamic faith. More immediately, the *Muhajirun* faced the fundamental problem of securing food and other necessities. They had moved from a city with a trade-based economy to a farming settlement that revolved around date-palm cultivation. Lacking agricultural skills, they had no means of sustaining themselves in Medina and could hardly expect to survive on the charity of the Ansar. Muhammad thus found himself caught in a predicament even at this moment of genuine success. That is, while he had succeeded in establishing himself as the head of

⁴² Aslan, *No God but God (Updated Edition)*, 55–59.

⁴³ Armstrong, *Islam*, 14–15; Brown, *Muhammad*, 29.

the *umma*, he simultaneously had to find some way to sustain the *Muhajirun*—and fast—or he would risk seeing his fledgling community quickly fall apart.⁴⁴

Desperation soon led him to a solution. Aware that his community of merchants and shopkeepers were in no position to provide for themselves through farming, Muhammad began leading *ghazu*, or raids, against Quraysh-owned caravans travelling between Mecca and Syria. These attacks proved to be enormously successful right from the start. The Muslim raiders were able to seize more than enough goods from the Meccans to sustain the *umma* and thus ensured that the community of believers would survive during what proved to be its most trying period.⁴⁵

At the same time, the raids created new problems for the *umma*. Heretofore, the Quraysh had left Muhammad's followers unmolested in Medina; infuriated by the raids, however, they were now determined to crush his fighters in battle once and for all. Accordingly, in 624, they sent a large punitive expedition north with plans to defeat the Muslims. Muhammad rose to the bait, and the two sides met in battle just outside the small town of Badr. Though the Meccan army decisively outnumbered Muhammad's force, it failed to defeat the Muslims. On the contrary, Muhammad's well-disciplined troops won a clear victory in which they compelled the Meccan soldiers to retreat with substantial losses. The battle marked an important turning point for Muhammad and the fledgling community of believers. The triumph over such long odds seemed to demonstrate that he was a capable military leader and that the *umma* was a rising force in the region; perhaps more importantly, it also suggested that Muhammad enjoyed divine support and that he truly was God's messenger. As a consequence, the victory at Badr allowed him to strengthen his hold over Medina and led a growing number of bedouin tribes and clans to throw in their lot with him over the following months.⁴⁶

In spite of this triumph, the next few years proved to be trying ones for the *umma*. In 625, an army of 3,000 Meccan soldiers avenged the defeat at Badr by routing Muhammad's 1,000-strong force at the Battle of Uhud. It was a significant setback. The Muslims suffered substantial casualties, and only the Meccan army's failure to exploit its advantage permitted the survivors to escape. Nonetheless, Muhammad remained undaunted. Despite the defeat at Uhud, he and his troops continued to harass Meccan soldiers and to raid the city's caravans as they headed to and from Syria.⁴⁷

Increasingly frustrated with Muhammad's depredation, the Quraysh decided that the time had come to destroy the *umma* once and for all. Accordingly, in 627, they raised an army of 10,000 soldiers composed of both Meccan troops and the city's bedouin allies and set out to crush the Muslims. Muhammad's force had grown in size over the preceding two years, but it had no hope of defeating the vast Meccan host bearing in open battle. Only by finding some way to offset the Meccans' huge numerical

⁴⁴ Armstrong, *Islam*, 17–19.

⁴⁵ Donner, *Muhammad and the Believers*, 45–46.

⁴⁶ Kennedy, *The Prophet and the Age of the Caliphates*, 35–36.

⁴⁷ Brown, *Muhammad*, 33–34.

advantage could Muhammad possibly spare the *umma* from certain destruction. Once more, he rose to the challenge. To even the odds, he ordered the Muslims to dig a deep trench around Medina that would prevent the Meccan cavalry from being able to attack the town. The tactic worked brilliantly. Though they enjoyed the support of the Qurayza, one of Medina's Jewish tribes, the Quraysh and their bedouin allies could not find a way to force their way across the trench and eventually had to abandon the siege in disarray.⁴⁸

In two ways, the victory in the Battle of the Trench marked a critical turning point for Muhammad and the *umma*. First, the triumph allowed him to solidify his control of Medina. On the advice of one of the Ansar leaders, he imposed the standard seventh-century Arab penalty for treason on the Qurayza tribe immediately following the battle; that is, he ordered his followers to execute the men of the tribe and to sell the women and children into slavery. Thereafter, he no longer faced any organized resistance to his rule in Medina. Second, his victory over the huge Meccan army dramatically elevated his standing as a successful military leader throughout the region and seemed to suggest to a growing number of Arabs that he enjoyed God's favor just as he claimed. In response, many bedouin tribes converted to Islam and joined Muhammad's cause, thereby swelling both the size of the *umma* and the number of fighters in his now-substantial military. By the end of 628, as a result, he had the most powerful military force in the Hijaz—one so strong that he was now in a position to seize control of Mecca.⁴⁹

Muhammad was not interested in taking Mecca by force. Instead, he hoped to bring it under his control peacefully. Consequently, in 628, he made a bold and seemingly rash declaration. He announced that he intended to lead 1,000 Muslims on a pilgrimage to the Ka'ba during that year's hajj. On the surface, it seemed like a foolhardy proposal. As pilgrims, he and his followers would be traveling to Mecca unarmed and would thus be vulnerable to attack. In fact, it was a shrewd move that caught his Quraysh enemies in a bind. As both Muhammad and Mecca's frustrated leaders grasped, attacking the Muslims during the hajj would not only imperil the Quraysh tribe's status as the guardians of the Ka'ba but would also cost it what little support it still retained in the Hijaz.⁵⁰

Muhammad's gambit worked. As he had foreseen, the Quraysh did not attack the Muslims. Instead, desperate to impose some conditions on Muhammad in order to avoid further embarrassment and the defection of yet more of the city's bedouin allies to his side, they sent a delegation to negotiate with him at a place just outside Mecca called Hudaibiyyah. There, the Quraysh representatives proposed a peace treaty that called for an end to the war between Mecca and Medina and for the Muslims to return

⁴⁸ Kennedy, *The Prophet and the Age of the Caliphates*, 39.

⁴⁹ Armstrong, *Islam*, 20–21.

⁵⁰ Cole, *Muhammad*, 152–57.

immediately to their homes; in exchange, Muhammad and his followers could return the following year to take part in the hajj.⁵¹

Many of his followers blanched at accepting what they viewed as a one-sided deal. After all, the *umma* was the rising power while the Quraysh were in eclipse. Intent on taking Mecca peacefully, Muhammad nonetheless agreed to the terms of the proposal. His decision proved to be a wise one. When he and his followers returned in 629 for the hajj, their faith and discipline greatly impressed the people of the city. As a result, opposition in Mecca began to melt away.⁵²

Muhammad Victorious

It ended the following year. Using a Quraysh attack on a client tribe as a pretext, Muhammad moved to take Mecca with an army of 10,000 soldiers. His troops faced no organized opposition. Instead, the people of the city watched passively as the Muslims occupied Mecca, cleansed the Ka'ba of the pagan idols, and added the city to Muhammad's rapidly growing empire.⁵³

Three factors explain the Quraysh tribes' failure to resist Muhammad. First, by 630, the balance of forces in the Hijaz had swung so heavily toward the Muslims that further opposition was futile. Second, Muhammad had made an important concession to the tribe's commercial interests. Aware of the economic significance of the traditional pagan pilgrimage, he replaced it with a new, Islamic hajj that required all Muslims to visit Mecca at least once in their lives; he thus ensured that the Quraysh would continue to profit from the arrival of pilgrims and from the trade in which they engaged. Finally, Muhammad made clear that he did not come to Mecca seeking vengeance against those who had persecuted the Muslims over the prior two decades. Instead, save for a few cases involving diehard enemies of Islam, he treated his old persecutors with startling mildness. This approach proved doubly effective. Not only did it discourage the Meccans from resisting his army when it marched on the city, but it also led many of his former tormentors—including Banu Umayya head Abu Sufyan, (c. 565-c. 653)—to convert to Islam.⁵⁴

Having defeated the most powerful tribe in the Hijaz, Muhammad was now well positioned to extend his writ to all of Arabia. Over the next two years, he directed a series of triumphal military campaigns that saw his forces subdue the town of Taif and quell the few other remaining centers of organized opposition to Islam. These victories further cemented his standing among the Arabs. Understood by many to be a sign of divine favor, his continued success spurred most of the bedouin tribes that remained outside his control to accept both his suzerainty and, with some important exceptions, his religion. By 632, as a result, Muhammad had secured control of the entire Arabian

⁵¹ Cole, 152–57.

⁵² Kennedy, *The Prophet and the Age of the Caliphates*, 41–42.

⁵³ Aslan, *No God but God (Updated Edition)*, 167.

⁵⁴ Armstrong, *Islam*, 22–23.

peninsula. The consequences of this achievement were enormous. He had not merely joined the Arabs together into a single, socio-religious community—a significant accomplishment in its own right—but had, in the process, also replaced the tribal conflict that had characterized the region for centuries with a new era of intertribal peace. This was a signal achievement and one, as we shall see in the next chapter, which would have enormous consequences.⁵⁵

Muhammad was unable to enjoy the fruits of his victory, however. Soon after returning to Medina from the hajj in 632, he fell ill and retired to the home of one of his wives, ‘A’isha (c. 603-678). Abu Bakr, the man who had helped Muhammad escape to Medina during the hijra in 622, replaced the prophet as the prayer leader in the mosque on what he hoped would be a temporary basis. Muhammad did not recover, however. To the despair of the *umma*, he instead grew steadily sicker for several weeks before finally passing away. The Muslims thus faced the challenge of determining how to proceed without their leader—a prospect rendered all the more daunting by virtue of the fact that Muhammad had failed to leave instructions indicating who should replace him or even how the community should choose a successor.⁵⁶

A Brief Overview of Islam

So what were the characteristics, tenets, and values of the religion that Muhammad bequeathed to the Arab people? In broadest terms, as we have seen, early Islam stressed three overarching themes: monotheism, social justice, and community. Coupled with the practical necessities related to the administration of a growing state—like Moses, after all, Muhammad was both a political leader and a religious figure—these three central beliefs would inform all aspects of the religion.

Muhammad the Reformer

To be clear, first, Muhammad never claimed to be creating a new faith. On the contrary, he claimed that he was only seeking to purify and renew an older faith that had lost its way. This was evident in his description of the relationship between Islam and the other two Abrahamic faiths: Judaism and Christianity. According to Muhammad, God had long ago revealed the Abrahamic religion to humanity through a succession of prophets including Moses and Jesus. For reasons of self-interest, however, spiritual and secular leaders had perverted God's revelation at different points in the past by incorporating into it innovations such as the Trinity. As such, he maintained, Christianity and Judaism were theologically corrupted offshoots of the older, monotheistic religion. Meanwhile, some groups that had once followed the Abrahamic faith—including the Arabs—had subsequently become estranged from it and had abandoned it in favor of a return to polytheism. God had not forsaken those whom priests had led astray or who had taken up the false idols of polytheism, however. Instead, to bring humanity back into the light of the true religion, He had chosen Muhammad to serve as His final

⁵⁵ Donner, *Muhammad and the Believers*, 49–50.

⁵⁶ Brown, *Muhammad*, 61–63.

messenger and had tasked him with bringing His divine plan to humanity for the last time—thereby restoring and purifying the one true monotheistic faith.⁵⁷

The Qur'an

That perfect plan was outlined in Islam's holy book: the Qur'an. According to Muhammad, God revealed the Qur'an to him through the Angel Gabriel over a period of twenty-three years. It was, he maintained, the literal Word of God—an earthly copy of the eternal and uncreated Heavenly Qur'an. It also marked the culmination of God's revelation. That is, God had previously shown humanity elements of His message through the Torah, the Psalms, and the Gospels; the Qur'an, however, was the final, perfect, and most important revelation. Slightly shorter than the New Testament, it consists of one-hundred-and-fourteen sura, or chapters, arranged in order from longest to shortest. Individual suras were copied down or memorized while Muhammad was alive. However, a canonical, written version of the Qur'an was only created during the reign of Uthman, the third caliph.⁵⁸

Monotheism, Community, and Social Justice

The Qur'an outlines a religion that is stridently monotheistic. Islam then and now emphasizes tawhid, a concept that denotes the oneness and unity of God and that proclaims forthrightly that He has no partners. Indeed, the Qur'an explicitly rejects both polytheism and the Christian concept of the Trinity: "Say, 'He is God, the One. God, the Absolute. He begets not, nor was He begotten. And there is none comparable to Him'" (Qur'an 112). The Qur'an also characterizes God as an eternal and all-powerful being who is actively managing the universe, and it describes Him as loving and as al-Rahman, or merciful. God's benevolence was not absolute, however. Like Christianity, Islam holds that there will be a final reckoning during which all humans will be judged for their actions on earth. Those who had submitted to the divine will would spend an eternity in the blissful Garden of Paradise; those who had refused to accept Him or who had failed to abide by His law would instead burn in a fiery hell.⁵⁹

The emphasis on monotheism and the unity of God permeated Muhammad's thinking and had enormous consequences for Islam. Perhaps most obviously, it shaped his conception of how the community of believers should be structured. Just as God was unified, so would be the umma. To ensure that it was, he preached that humans were to live in a certain type of society and that they were to adhere to a set of laws that explicitly called on them to engage in some actions and to refrain from others. God required not just that individual adherents abide by these standards, moreover, but that the umma act collectively to ensure that the entire community behaved according to His

⁵⁷ Armstrong, *Islam*, 8.

⁵⁸ Donner, *Muhammad and the Believers*, 59; Daniel G. Bates and Amal Rassam, *Peoples and Cultures of the Middle East*, 2nd edition (Upper Saddle River, NJ: Prentice Hall, 2000), 52–53.

⁵⁹ Armstrong, *Muhammad*, 49, 62.

moral code. As the Qur'an states: "You are the best community evolved for mankind, enjoining what is right and forbidding what is wrong" (Qur'an 3:110). Individual salvation was important, in other words, but, from the very start, the religion that Muhammad outlined was explicitly communal in nature.⁶⁰

Much of the communitarian aspects of Islam were aimed at achieving social justice. As earlier noted, Muhammad was deeply disturbed by the Meccans' rejection of the muruwah code and its requirement that the rich and powerful care for the poor and weak. As we have seen, he was particularly troubled by the growing insecurity and difficult circumstances that confronted widows and orphans in a society that was based on kin relations. Absent strong family connections, such people were all-but-inevitably destined to a life of grinding poverty and, in many cases, enslavement. Accordingly, the Qur'an imposed a variety of socioeconomic requirements on the umma designed to ensure that it redistributed wealth to impoverished people and that it protected the weak. It outright banned usury, false contracts, and bribery, it made plain that those who accumulated wealth had an obligation to share some of their good fortune with the poor, and it explicitly required the just treatment of widows and orphans.⁶¹

Women

Muhammad paid particular attention to improving the position of women in Arab society. A woman's life in pre-Islamic Arabia was a difficult one. Effectively chattel owned either by her husband or father, she enjoyed neither legal protections nor the right of divorce and was utterly dependent on her male relatives for her well-being. A married woman did not even control the money or goods that her parents provided as a dowry; on the contrary, that wealth became the property of her husband. She also had no legal claim to inherit either her husband's estate or the dowry she brought into the marriage. Instead, one of her husband's relations could marry her and thereby acquire the deceased's property including the dowry; worse, if the widow was deemed too old, she and her dowry would instead become the property of her husband's clan, and she would be dependent upon it for her survival. Thus, widows—of whom there were many in the harsh environment of pre-Islamic Arabia—were often left destitute and unprotected.⁶²

Muhammad dramatically changed this state of affairs. The Qur'an included specific language granting women the right to end marriages even as it placed new limits on divorce for men and restricted the number of wives that they could have. It also gave women the right to inherit property and, importantly, granted them control over their dowry—thus permitting them to take it with them in the event of divorce. These reforms markedly improved the legal position of women in Arabia. Thereafter, widowed women were far less likely to be without some means of material support and those who had abusive husbands could end their marriages without losing their dowries. We must

⁶⁰ Esposito, *Islam*, 31–32.

⁶¹ Esposito, 32–33.

⁶² Aslan, *No God but God (Updated Edition)*, 61–62.

keep the extent of Muhammad's reforms in this area in perspective, of course. He did not grant women equivalent legal status to men, and the changes he effected failed to emancipate them fully. They continued to inherit only a portion of their husbands' wealth, for example, and their divorce rights were significantly more restricted than those that men enjoyed. Still, in comparison to the legal standing of women in other contemporary societies, Muhammad's reforms were nothing short of revolutionary. Indeed, women in much of Christian Europe would not win equivalently robust divorce and dowry rights until the nineteenth century.⁶³

Two topics regarding women that receive a great deal of attention in the West today merit further discussion: veiling and polygamy. The Qur'an explicitly endorsed the practice of polygamy, stipulating that a man could have as many as four wives. "If ye fear that ye shall not be able to deal justly with the orphans, marry women of your choice, Two or three or four; but if ye fear that ye shall not be able to deal justly (with them), then only [marry] one" (Qur'an, 3:4). One should not interpret this validation of polygamy as a blanket sanction of the practice, however, but should instead understand it as a limited, pragmatic endorsement that evinced the culture and circumstances in which the Qur'an emerged.⁶⁴

It did so in two ways. First, Muhammad's endorsement of polygamy reflected his larger message of social justice and the peculiar needs of the early umma. While successful, the war with Mecca had inevitably produced substantial casualties among his male followers and had thus resulted in a surge in the number of widows and orphans among the believers. This situation posed a serious challenge for Muhammad. How could the still tiny and impoverished umma ensure the well-being of the many women and children who had lost their husbands or fathers in the conflict? He found the answer in the existing practice of polygamy. Muslim men who were already married would take additional wives from among the women who had lost their husbands in the fighting, thus ensuring that they and their children enjoyed material support. Indeed, it was only after the disastrous Battle of Uhud, an engagement in which many Muslim men had died, that he shared Verse 3:4 with the community of believers and thus encouraged polygamy. Muhammad himself set an example for this practice when he married the widow Sawdah—who was past the customary age for marriage and thus otherwise unlikely to get remarried—to ensure that she and her children were adequately cared for.⁶⁵

Second, Qur'anic approval of polygamy reflected the political realities of seventh-century Arabia. At that time, powerful men—particularly tribal leaders—demonstrated their authority and success by having multiple wives. It was a practice that frequently put women in a tenuous situation. If a man could not maintain all of his wives, he might sell one or more into slavery or prostitution. Focused as he was on social justice, Muhammad found the tribal chiefs' practice of demonstrating status through multiple

⁶³ Esposito, *Islam*, 119–21.

⁶⁴ Armstrong, *Islam*, 16.

⁶⁵ Aslan, *No God but God (Updated Edition)*, 64.

marriages off putting and viewed the sale of wives as outright repugnant. However, he also understood that he was not in a position to oppose polygamy lest he alienate the tribal chiefs on whose support the success of his war with Mecca depended. Accordingly, he bent to prevailing custom and formally sanctioned the practice of political leaders taking multiple wives. Indeed, he himself married a number of women as part of a successful effort to bind certain key tribal leaders to him through political marriages. At the same time, however, he discouraged polygamy by limiting men to no more than four wives and by permitting them to take multiple spouses only if they could afford to do so—thereby making clear that he was not offering a blanket endorsement of the practice.⁶⁶

While commonplace in much of the Muslim world today, meanwhile, the practice of veiling did not emerge as a result of Qur'anic decree. The Qur'an did mandate that women dress chastely, declaring: "that they should . . . guard their modesty; that they should not display their beauty and ornaments except what (must ordinarily) appear thereof," and it explicitly called for Muhammad's wives to seclude and veil themselves (Qur'an 24:31). At no point did it stipulate that other Muslim women must cover their heads, however, and it most certainly did not include any passages requiring that they be attired in clothing such as the burqa. Instead, as most scholars now agree, veiling was an upper-class Persian and Byzantine custom that Muslim women only adopted several generations after Muhammad had died. Indeed, the very variety of veiling in the Muslim world today—the niqab in Saudi Arabia, the hijab in Egypt, the burqa in Afghanistan—indicates that the practice reflects local cultural values rather than a common, Qur'anic obligation.⁶⁷

As the institutionalization of veiling as a Muslim custom in the century after Muhammad suggests, the revolutionary changes to gender relations that he effected did not endure. What happened? In a nutshell, the patriarchy reasserted itself. During the immediate generations following Muhammad, longstanding notions that God had predestined men to run society and had assigned women to a subordinate position led people to reinterpret Muhammad's views. As a result, as we shall see in chapter four, men were able to gradually chip away at the prophet's reforms and eventually succeeded in reinstating many patriarchal pre-Islamic customs in the decades after his death. As a result, within just a few generations, women would lose nearly all of the rights that they had gained in the first years of the umma.⁶⁸

jihad

While Islam may have effected idealistic and revolutionary changes for women, its teachings also reflected the practical exigencies of the early and embattled umma. As previously noted, Muhammad was both a religious prophet and a political leader who—unlike Jesus or John the Baptist but like Moses—oversaw a political community

⁶⁶ Esposito, *Islam*, 116–19.

⁶⁷ Armstrong, *Islam*, 16.

⁶⁸ Esposito, *Islam*, 121–24.

and was thus tasked with making the difficult and often morally ambiguous decisions that are incumbent on secular rulers. He did so under particularly trying circumstances, moreover. As we have seen, for a brief but critical period, the community he led confronted an existential challenge from the Quraysh tribe of Mecca. As a consequence, Muhammad frequently took actions and articulated values based on the umma's immediate needs; for example, his execution of the Jewish men of the Qurayza tribe after the Battle of the Trench reflected his belief that he needed to eliminate a dangerous fifth column in order to ensure his movement's survival. Unsurprisingly in light of Muhammad's dual role as religious and political leader, such decisions shaped important aspects of Islam and affected the ways in which future generations would interpret the religion.

Such was the case with regard to jihad, or, more properly as we shall see, lesser jihad. Jihad, or warfare waged in defense of the faith, was very much a product of the umma's perilous circumstances in the 620s, and it reflected both Muhammad's need to justify the conflict with Mecca and to establish a set of rules regulating how his soldiers would wage that war. In other words, jihad emerged as a Muslim version of the Christian doctrine of Just War that St. Augustine (354-430) had articulated in the fifth century. Indeed, the two were remarkably similar. Like Augustine's concept, Muhammad's notion of just war required armies to refrain from attacking noncombatants and demanded that any conflict be defensive in nature. Indeed, the Qur'an explicitly holds that "[p]ermission to fight is only given to those who have been oppressed" (Qur'an 22:39).⁶⁹

Just as later religious scholars would roll back the reforms Muhammad had obtained for women, so would they distort jihad to justify aggressive military action. This change began shortly after Mohammed died. In the early 630s, Muslim leaders and religious scholars reinterpreted the concept of jihad to legitimize the Arab conquest of the Persian Empire and the Byzantine provinces of Syria and Egypt. Later, they divided the world into Dar al-Islam, or the House of Peace, and Dar al-Harb, or the House of War, and called on Muslim leaders to fight for the purpose of extending the territory of Dar al-Islam. Jihad waned when the Arab Empire stopped acquiring new territory not long thereafter, but, as we shall see, would reemerge in the twelfth century in response to the Crusades. It thereafter again faded in popularity until the Ottoman government revived the concept by declaring jihad against the allied powers during the First World War in a bid to create unrest in the Muslim parts of the British, French, and Russian empires. Today, despite the condemnation of most Muslim religious scholars, jihadi groups adhere to a reinterpreted conception of jihad that makes it an individual rather than a collective action and that disregards the explicit prohibition that Muhammad had placed on attacking noncombatants.⁷⁰

⁶⁹ Michael Bonner, *Jihad in Islamic History: Doctrines and Practice* (Princeton: Princeton University Press, 2006), 16–17.

⁷⁰ Aslan, *No God but God (Updated Edition)*, 85–86; John L. Esposito, *Unholy War: Terror in the Name of Islam* (Oxford: Oxford University Press, 2002), 35.

The Five Pillars of Islam

Islam is a religion of belief, of course, but its main focus is on proper action rather than on observance of dogma. As such, Islam eschews Christianity's emphasis on adherence to a particular doctrine and rejects theological inquiry as *zannah*, or idle speculation; instead, like Judaism, it stresses correct practice through the acceptance of a set of straightforward duties, rituals, and laws that define and exemplify what it means to be a Muslim. By far the most important of these obligations are the Five Pillars of Islam. These expressions of faith date to Muhammad's time in Medina, and they constitute the core rituals of Islam. Unsurprisingly, these duties reflect the prophet's emphasis on monotheism, community, and social justice.⁷¹

The first pillar is *shahada*, or the profession of faith. It calls for a genuine expression of belief in the monotheistic God through the pronouncement of the statement: "There is only one God, God, and Muhammad is his prophet." The declaration of this seemingly simple statement achieves two critical aims. First, in expressing it, Muslims profess their adherence to the religion of Islam and their belief in God. Indeed, genuinely stating the *shahada* is all that one must do to become a Muslim. Second, it affirms one's belief in *tawhid*, which emphasizes not merely the monotheistic concept of God's oneness and unity, but also His singular and indescribable nature and His indivisibility.⁷²

The second pillar of Islam, *salat*, or worship, calls on Muslims to pray five times per day. They are required to do so at dawn, noon, mid-afternoon, sunset, and nighttime. They must do so while facing Mecca, and only after they have ritually cleansed themselves. *Salat* also stipulates that Muslims pray collectively at a mosque at midday on Fridays. This pillar reflects two of the main themes in Muhammad's teachings. First, the prayer at the heart of it stresses the idea of *tawhid*, and thus reinforces Muhammad's message of strict monotheism and God's unity. Second, the call for collective public prayers on Fridays makes clear Islam's focus on community rather than on individual salvation.⁷³

The third pillar of Islam, the *zakat*, is an annual alms tax assessed to support the indigent. It began as a voluntary contribution in Mecca and became mandatory after Muhammad and his followers relocated to Medina—likely, as we have seen, because the war with the Quraysh had produced a large number of widows and orphans and thus a need for some mechanism by which they could be supported. Traditionally, local

⁷¹ Armstrong, *Islam*, 6.

⁷² Albert Hourani, *A History of the Arab Peoples* (Cambridge, Mass: Belknap Press of Harvard University Press, 2002), 147–48.

⁷³ Bates and Rassam, *Peoples and Cultures of the Middle East*, 46 Originally, men and women worshipped together in the mosque in Medina, albeit in gender-segregated sections. Not long after Muhammad, however, political and religious leaders took a dim view of mixed-gender prayer out of fear that the presence of women would distract the men and they consequently discouraged women from taking part in public prayers.

rulers assessed the zakat on each Muslim's net worth at a rate of 2.5 percent. Expected of all believers and collected by the state, the tax helps reinforce the ideas of community and mutual obligation, and, importantly, embodies both Islam's communal nature and its emphasis on social justice.⁷⁴

Sawm, or fasting, is the fourth pillar of Islam. *Sawm* occurs during Ramadan—the month in which Muhammad claimed that the Qur'an was first revealed to him—and has clear roots in the Jewish practice of fasting. From dawn until dusk during Ramadan, Muslims refrain from activities like sexual relations, smoking, eating, and drinking; exceptions are made for children, pregnant women, nursing mothers, the sick, the elderly and others who are physically vulnerable. It is a festive time in Muslim-majority countries. Ramadan includes nightly celebrations and communal meals capped by the *eid al-fitr*: a three-day holiday at the end of the month featuring feasts and family gatherings. *Sawm* is not just a period for celebrating with family and friends, however. Instead, it is also a time during which Muslims affirm their belief in *tawhid*, build communal bonds, and renew their commitment to social justice by experiencing the privation of the poor.⁷⁵

The fifth and final pillar of Islam, *hajj*, is the requirement that Muslims make a pilgrimage to Mecca at least once in their lifetime if they are able. Though the *hajj* was comparatively modest as recently as the 1940s, it is today one of the largest gatherings of people on earth. It takes place during the month of *Dhu al-Hijjah*, and first occurred in the late 620s. During the *hajj*, pilgrims in Mecca take part in a series of rituals including the *tawaf*, in which they circumnavigate the Ka'ba in a counterclockwise direction, and the *sa'ay*, the ritual running between two hills in imitation of Hagar's desperate search for water for her infant son, Ismail. A collective activity, the *hajj* evinces Islam's emphasis on community. Less obviously, its rituals are also designed to focus the pilgrims on *tawhid* and thus serve to reinforce the strict monotheism that lies at the heart of the faith. Communalism and monotheism were surely in the front of Muhammad's mind when he made the *hajj* a religious obligation, but, as previously noted, practical considerations likely also played a role in his decision. Put simply, requiring his followers to undertake an Islamic version of the *hajj* would offset the revenue the Quraysh stood to lose as a result of the demise of the older, polytheistic pilgrimage, and was thus a small price to pay to persuade them to abandon their resistance to his movement.⁷⁶

Though not a formal requirement of the religion, another ideal, greater *jihad*, is so important that many people frequently—if inaccurately—refer to it as the sixth pillar of Islam. *Jihad* literally translates as “struggle,” and greater *jihad*—in contrast to lesser *jihad*'s articulation of the Muslim doctrine of just war—refers to the ongoing struggle that individual believers wage within themselves to remain focused on God in the face of the temptations of the material world. Islam holds that devotion to success, money, luxuries,

⁷⁴ Esposito, *Islam*, 109–10.

⁷⁵ Hourani, *A History of the Arab Peoples*, 149.

⁷⁶ Aslan, *No God but God (Updated Edition)*, 150–52.

ego, or power distracts believers from centering their lives on God and His desires, and thus separates them from the divine. Only by ceaselessly waging greater jihad within themselves could they remain attentive to the goodness of God and thus honor their commitment to Him.⁷⁷

As noted earlier, Muhammad also imposed a number of other restrictions on the *umma* related directly or indirectly to the themes of monotheism, community, and social justice. Holding that the one God is the God of Abraham, Muhammad required Muslims to observe the cultural practices by which the patriarch and his descendants had lived. Accordingly, he barred the consumption of pork and required the circumcision of Muslim men. Other restrictions were related to Islam's focus on social justice. Muhammad banned usury, false contracts, bribery, alcohol, gambling, and men wearing silk clothing or gold jewelry as part of his effort to create a more equitable society. Enforcement of all of these restrictions ultimately fell to the *umma* rather than to individuals or families, thus reflecting Islam's communal rather than individual focus. As we shall see in chapter four, Muslim jurists would later formalize and expand these restrictions when they codified shari'a law in the eighth and ninth centuries.⁷⁸

The Historical Debate: The Revisionist School

So far, this chapter has outlined the conventional understanding of Muhammad and the emergence of Islam. That is, the preceding account broadly reflects both the traditional Islamic narrative articulated by Muslim historians in the ninth century and the views of most modern-day scholars who focus on this period. This perspective is not universally held, however. Instead, in recent decades, a number of historians and religious scholars have developed alternative interpretations of the emergence of Islam that challenge key parts of the dominant narrative.

Some scholars focus on the question of Muhammad's intention. They dispute the idea that he cast himself as the final prophet whom God had chosen to strip away accretions to the pure faith of Abraham and thus to restore the true monotheistic religion. Instead, they contend that his early conception of his mission was far-more limited in scope. For example, Karen Armstrong argues that Muhammad initially viewed himself merely as the prophet that God had selected to bring His message to the Arab people just as He had tapped Moses and Jesus to spread the monotheistic faith to other cultures. She argues that it was only after members of the Jewish tribes in Medina mocked Muhammad's understanding of the Torah that he declared that God had sent him to bring his message to all humanity as the final prophet—a shift he made manifest in 624 by changing the *qiblah*, the direction that Muslims faced while praying, from Jerusalem to Mecca.⁷⁹

⁷⁷ Esposito, *Islam*, 113–14.

⁷⁸ Esposito, 31–33.

⁷⁹ Armstrong, *Islam*, 8, 17–18.

Historians in the revisionist school go substantially further. These scholars draw attention to the fact that the extant sources on the life of Muhammad and the early *umma* are deeply problematic. In particular, they note that the best evidence available on Muhammad is a biography written by Ibn Ishaq (c. 704-767) more than a century after the prophet had died, and they argue that the gap between the era in which writers like Ibn Ishaq lived and the time of the early community of believers is simply so vast that such sources cannot be treated as reliable accounts. Revisionist scholars also contend that authors like Ibn Ishaq very likely shaped their interpretations to serve contemporary political interests, and that their works thus reflect the values and beliefs of the eighth century rather than those of the early *umma*. Those historians consequently view the Arabic sources with a skeptical eye and generally favor seventh-century, non-Muslim accounts that touch upon events in the Arabian peninsula over later Islamic narratives.⁸⁰

Interpretations based on those accounts dramatically contradict the traditional narrative. For example, the historian Patricia Crone finds the idea that Mecca was a bustling commercial nexus implausible because non-Muslim seventh-century sources—otherwise well-versed on the affairs of the Hijaz—made no mention of it whatsoever. Indeed, she concludes that Muhammad lived not in Mecca, but instead in northern Arabia close to the Mediterranean Sea. Meanwhile, John Wansbrough agrees with the traditional narrative that the Qur'an dates to the early-seventh century, but also maintains that eighth- and ninth-century Muslims thoroughly altered it through accretion and revision. Revisionist scholars who examine archaeological findings have gone even farther and argue that the material evidence casts doubt about even the most fundamental elements of the standard narrative. For instance, the archaeologist Yehuda Nevo notes that the earliest confirmed archaeological reference to Muhammad—coins bearing his name—date only to the late-seventh century. As a result, he concludes that the traditional story of Islam's emergence was a construct of the eighth and ninth centuries and suggests that Muhammad was at best a minor figure whom Muslim writers only later recast as a prophet to serve sectarian ends.⁸¹

The revisionist school of Islamic studies has certainly altered understandings of Islam's emergence, but most scholars continue to accept at least the broad outlines of the traditional narrative. They typically acknowledge the limits inherent in later sources such as Ibn Ishaq's biography, but also argue that they can draw valid inferences from such accounts through the careful use of historical methods and criticism. They likewise maintain that non-Arab sources accord with the overall contours of the conventional

⁸⁰ Donner, *Muhammad and the Believers*, 50–56; Brown, *Muhammad*, 78–79, 93–97.

⁸¹ See Patricia Crone, *Meccan Trade and the Rise of Islam*, 1st Edition (Princeton: Princeton University Press, 1987); John E. Wansbrough, *Qur'anic Studies: Sources and Methods of Scriptural Interpretation* (Amherst, NY: Prometheus Books, 2004); Yehuda D. Nevo and Judith Koren, *Crossroads to Islam: The Origins of the Arab Religion and the Arab State* (Amherst, NY: Prometheus, 2003).

understanding of Muhammad and the birth of Islam—at least in the broadest sense—and thus confirm the traditional narrative.⁸²

Conclusion

Muhammad had enjoyed astonishing success in the last twenty-two years of his life. Since beginning his religious mission, he had won the support of a core group of followers, relocated his community to Medina, conquered Mecca, purged the Ka'ba of polytheism, and secured control of the entire Arabian peninsula. Given the long odds that he faced during much of that period, it is little wonder that many Arabs concluded from his success that he enjoyed divine favor.

The religious movement that he had started continued to grow rapidly after his death, moreover. Indeed, neither his passing nor his failure to name a successor slowed Islam's growth to any degree. Instead, as we shall see in the next chapter, the community he had led would follow his death by initiating one of the greatest campaigns of conquest in history—in the process establishing a huge and enormously successful empire and institutionalizing and spreading the religion of Islam.

⁸² Brown, *Muhammad*, 85–93.

Chapter Three: The Arab Conquest and the Establishment of the Caliphate

Just a few years after the death of Muhammad in 632, the Arab people, heretofore marginal players in a region dominated by powerful empires, exploded out of Arabia. Striking at a time when the established powers of antiquity were peculiarly weak, small but highly mobile Arab armies won a series of decisive victories that, by 651, resulted in the umma taking control of the Byzantine Middle East and the entirety of the Persian Empire. After a pause in the mid-to-late-seventh century, the wars of conquest continued. When those campaigns finally came to a close a century later, they had produced a new empire, the caliphate, of breathtaking scale. Stretching from Spain in the West to the Central Asian city of Samarkand in the east and from the Pyrenees in the north to Yemen in the south, the Arab Empire was considerably larger in extent than Rome had been at its peak and every bit as sophisticated. This achievement was remarkable—not least because the Arabs took control of this territory during a period in which they experienced significant internal dissension that included intrigue, assassination, and even civil war.

Remarkable as it was, the conquest would have been of limited consequence had the Arabs proven unable to institutionalize their control over the far-flung territories that they acquired between 632 and 751. History is replete with examples of powerful conquerors whose empires were as fleeting politically as they had been dominant militarily. The Arabs proved to be more than up to the challenge of administering the territory that they had acquired, however. Despite having no more experience ruling a large sedentary empire than had earlier nomadic groups such as the Huns, they managed to establish a durable system of government based on a combination of borrowed and original practices and institutions. Critically, they also ensured domestic tranquility by developing stable relations with the empire's millions of diverse subject peoples who remained, until at least the tenth century, predominately non-Muslim.

The Birth of the Caliphate

The pivotal, early stages of the conquests occurred during the reigns of the first four caliphs, or “successors,” of Muhammad. These men, Abu Bakr (r. 632-634), Umar ibn Al-Khattab (r. 634-644), Uthman ibn Affan (r. 644-656), and Ali Ibn Talib (r. 656-661) played an outsized role in determining not only the structure of the early caliphal government, but also the nature of Muslim society. As a result, Muslim religious scholars called ‘ulama’ later came to refer to these early leaders as Rashidun, or “rightly guided.”¹

That the Rashidun caliphs were obliged to define the framework of the government stemmed from the fact that Muhammad had not left clear instructions as to how the polity he created should be governed after his death. In fact, beyond making

¹ Vali Nasr, *The Shia Revival: How Conflicts within Islam Will Shape the Future* (New York: W. W. Norton, 2006), 35.

clear that his successors would not be prophets, he had never articulated the role of the caliph and had thus left the *umma* with several critical, unanswered questions. Was the ruler's authority to be strictly political in nature, or was it also to encompass religious issues? If his² ambit was to extend to spiritual matters, what religious powers would he wield? Finally, was the caliph to be "first amongst equals" administering a loosely organized state or was he to rule as the all-powerful head of a tightly centralized government? With Muhammad having demurred in answering these questions, it fell to his immediate successors to address them. Through their actions, pronouncements, and decisions, the first four caliphs did so—establishing, in the process, the broad structure that would define caliphal authority going forward.

The Succession Crisis

Muhammad's refusal either to explain the nature and extent of his successors' powers or to make clear who was next in line would not have proven problematic had the *umma* been united and thus able to devise a system for determining who was best suited to follow Muhammad. The community of believers was anything but unified in 632, however. Instead, it was divided into three loose factions—the Ansar, the *muhajirun*, and the Banu Hashim—that vied with each other for power and influence. As a result, even as Muhammad's followers grieved following his death, they also engaged in a brief succession struggle that would have momentous political, social, and even religious implications for the fledgling movement.³

Each faction advanced a claim in favor of one of its members. The Ansar, the Medinan natives who had backed Muhammad from the time of his arrival in Yathrib, believed that their critical support during a peculiarly difficult period entitled them to select a successor from among their ranks. The *muhajirun*, Muhammad's earliest followers, strongly disagreed. They instead maintained that the successor should be a Muslim from the Quraysh tribe who had faced persecution alongside the prophet during Islam's earliest days in Mecca. Finally, a third group made up largely of members of Muhammad's Banu Hashim clan claimed that the prophet had intended that his cousin and son-in-law, Ali, should succeed him. Known as Alids or, more formally as the *Shi'atu Ali*, or Party of Ali, this faction argued that the Prophet had made his intention explicit when he had announced at the oasis of Ghadir al-Khumm during his return from his final pilgrimage to Mecca in 632 that "any of you who consider me a patron should consider Ali your patron."⁴

The prophet's son-in-law did not get the opportunity to take part in the deliberations at which the leading figures in the *umma* determined the succession, however. Instead, while Ali was preoccupied helping to prepare Muhammad's body for

² The use of a gendered pronoun is intentional here. All of the caliphs were men.

³ Hugh Kennedy, *The Prophet and the Age of the Caliphates*, Second Edition (New York: Routledge, 2004), 50–51.

⁴ Quoted in Tamim Ansary, *Destiny Disrupted: A History of the World Through Islamic Eyes* (New York, NY: PublicAffairs, 2009), 55.

burial, the leading *muhajirun* and *ansar* held a traditional consultation of important men called a *shura* to resolve the dispute before it turned into a full-blown crisis. They made two enormously consequential decisions. First, looking to freeze Ali out, they rejected his candidacy on the grounds that he was too young to head the community of believers. Second, they determined that they could only overcome the divisions in the *umma* that the succession had revealed by selecting someone who was both a member of the Quraysh tribe and an early, trusted companion of Muhammad. They quickly settled on the respected *muhajirun* Abu Bakr, who was both one of the first men to accept Islam and the father of Muhammad's favorite wife, 'A'isha (c. 613-678). Ali was disappointed that the *shura* had passed him over and unhappy that it had met without him; despite the anger of his Alid supporters, however, he reluctantly consented to the decision out of a desire to ensure the unity of the *umma*.⁵

The Ridda Wars

Acceptance of Abu Bakr as caliph resolved the dispute over the succession, but it did not prevent the emergence of a serious new crisis. It centered on the relationship between the central government in Medina and the bedouin tribes that had supported Muhammad. Many of the tribes had responded to the news of the Prophet's death by claiming that they had pledged allegiance not to the government in Medina but instead to Muhammad personally, and that they were, as a consequence, no longer bound to the state that he had created or obligated to forward the zakat tax to its treasury. Compounding this insubordination, several figures had emerged among the bedouin who claimed, like Muhammad, to be prophets tasked by God with delivering His divine message to the Arabs and who had begun to establish rival political-religious communities similar in structure to the *umma*. These threats were serious ones for the movement Muhammad had created. Were Abu Bakr to have permitted the tribes to secede or allowed new religious communities to take root, the *umma* would likely have collapsed in short order. Likewise, one of Muhammad's signal achievements—the establishment of intertribal peace in Arabia—would have all-but certainly ended in a return of the intertribal conflict that had characterized the peninsula in the pre-Islamic era.⁶

The new caliph was determined to prevent such a train of events from occurring. Accordingly, he began his brief reign by launching the *Ridda Wars*, or Wars of Apostasy against the bedouin tribes that had rejected his authority. Given the large number of tribes that had broken with Medina, this task promised to be a difficult one. Fortunately for the Muslims, Abu Bakr was a skilled military commander and a deft negotiator. Using a combination of, intimidation, diplomacy, and, when necessary, force, he succeeded in bringing the breakaway tribes to heel and in crushing the rival religious movements. As

⁵ Reza Aslan, *No God but God: The Origins, Evolution, and Future of Islam*, Updated Edition (Random House, 2011), 114–19.

⁶ Hugh Kennedy, *Caliphate: The History of an Idea* (New York: Basic Books, 2016), 9.

a result, the *Ridda Wars* ended in the spring of 633 in a decisive Muslim triumph. The *umma* had been saved.⁷

Still, despite the quick victory, the leadership of the *umma* found the secession of so many tribes deeply troubling. It not only suggested that the bedouins' acceptance of Muhammad's message of peace and unity was superficial, but, more critically, it raised questions about Islam's ability to endure. Accordingly, even before the fighting had ended, they began to consider the root cause of the *Ridda Wars* and to implement changes that would ensure the tribes' ongoing loyalty.⁸

After reflecting on the issue for some time, they determined that the war had emerged out of the tension that existed between Muhammad's message of peace and the bedouins' deeply entrenched social and economic practices. As we saw in chapter two, the nomadic Arabs had traditionally engaged in low-intensity *ghazu*, or raids that provided both a way for the tribes to acquire the wealth upon which survival depended and a regulated system of warfare that permitted young men to acquire respect and social standing through exploits in battle. While welcome, Muhammad's ban on violence within the *umma* had thus unintentionally created serious problems for the bedouins. Not only were the tribes now enduring material deprivation, but they were also experiencing internal conflict among the young men who suddenly lacked a sanctioned means of gaining prestige. It was these problems that had led the tribes to break with Medina, and it was these issues that the leadership of the *umma* would have to resolve if the community of believers were to endure.⁹

This assessment was sobering. It suggested that there existed an intrinsic tension between the tribes' socioeconomic needs and Muhammad's call for peace within the Muslim community. Importantly, it also indicated that, absent significant changes, every caliph would have to wage a destabilizing civil war with separatist elements upon acceding to office—a prospect that boded ill for the success of both the community of believers and Muhammad's religious message.¹⁰

But how could the *umma* solve this problem? How could its leadership reconcile Muhammad's vision of internal peace with the tribes' need for raiding? It did so by devising a plan wherein the caliphate would direct the bedouins outward against external, non-Muslim enemies. This approach would provide the young men with the regulated warfare they needed to gain prestige just as it would permit the tribes to

⁷ Fred Donner, *Muhammad and the Believers: At the Origins of Islam* (Cambridge, MA: The Belknap Press of Harvard University Press, 2012), 100–102.

⁸ Karen Armstrong, *Islam: A Short History* (New York: Modern Library, 2002), 25–27.

⁹ Hugh Kennedy, *The Great Arab Conquests: How the Spread of Islam Changed the World We Live In* (New York: Da Capo Press, 2007), 56–57.

¹⁰ Kennedy, 56–57.

secure the wealth that they required to survive; now, however, they would achieve those ends in a way that was consistent with Muhammad's call for peace within the *umma*.¹¹

The Arab Conquests

Accordingly, after concluding the *Ridda Wars*, Abu Bakr and the other leaders of the *umma* prepared to launch a sustained campaign of raiding against the nearby Byzantine and Persian Empires. It was a bold decision to say the least. As the merchants of Mecca knew from firsthand experience, those empires were fabulously rich, and the Arabs could expect to win great exploits and to acquire great wealth by raiding them. At the same time, however, the targets of those *ghazu* were two of the mightiest powers that had ever existed. Attacking them was thus a highly dangerous move—one that could potentially imperil the fledgling community of believers.¹²

Though he had played an important part in shaping the *umma's* plans, it was not Abu Bakr who would shoulder that risk. Instead, following his death in 634, direction of the campaign fell to his successor, Umar. One of the *muhajirun*, Umar was a severe, disciplined, and pious man who famously demonstrated his dedication to Islam's revolutionary message by wearing homespun clothing and by taking naps in the corner of the mosque in Medina. Deeply committed to the success of the *umma*, he was as dedicated as his predecessor had been to ensuring that there would be no repetition of the *Ridda Wars*. Accordingly, after taking the title of Commander of the Faithful in addition to that of caliph, he proceeded with the plan to launch simultaneous attacks against the Persian and Byzantine Empires. It was a fateful decision—one that would mark the beginning of what turned out to be one of the most successful sustained military campaigns in history: the Arab Conquest.¹³

The Conquest of the Byzantine Middle East

While the campaign against the Byzantine Empire began with modest goals it quickly transformed into an outright invasion. The attacks started in early 634 when mobile Arab forces entered southern Palestine from the Arabian desert in search of plunder. Meeting with immediate success, they quickly amassed large quantities of loot. More importantly, they also handily defeated a small Byzantine detachment and seized control of several border towns and cities. These early victories were modest in scale but significant in terms of their impact. Making clear that Byzantine control of Palestine was more tenuous than the Arabs had earlier assumed, they suggested to Umar that his soldiers could do more than merely make off with some stolen goods. Accordingly, the caliph ordered one of his top commanders, Khalid ibn al-Walid (585-642), to lead his small force across the desert from Iraq, where it had been raiding the Persian Empire, and to take charge of the war in Palestine. Thus reinforced, the Arabs quickly scored a

¹¹ Kennedy, *The Prophet and the Age of the Caliphates*, 59.

¹² Kennedy, *Caliphate*, 10–11.

¹³ Aslan, *No God but God (Updated Edition)*, 70–77, 124–25; Kennedy, *The Prophet and the Age of the Caliphates*, 57–58.

stunning victory over imperial troops at the Battle of Ajnadayn in July 634—a triumph that resulted in them taking control of most of Palestine. Khalid did not rest on his laurels with this victory, moreover. Moving with characteristic aggressiveness, he took advantage of Byzantine disarray following the battle to lay siege to Damascus in August 634; Just a month later, he secured the city's surrender.¹⁴

The occupation of Damascus was a serious threat to the Byzantine position in the Levant. Emperor Heraclius (r. 610-641) could tolerate some raids and even the seizure of a few towns on the edge of the desert, but he could not stomach the loss of one of the empire's most important commercial and administrative centers. Accordingly, determined to restore Byzantine control of the region, he responded by bringing his main field army into the Levant in 635. Its arrival set the stage for a decisive battle—one that would determine whether Syria would remain a province of the venerable Byzantine Empire or whether it would become a part of the emerging Arab Empire.¹⁵

The clash took place in southern Syria. Khalid initially responded to the arrival of the Byzantine army by relinquishing Damascus and retreating into Palestine. His withdrawal should not be interpreted as a permanent abandonment of the territory that his troops had conquered, however. On the contrary, it was a shrewdly designed strategic redeployment aimed at gaining the time needed for reinforcements to arrive and at ensuring that the decisive engagement occurred at a place of his choosing. Khalid was successful on both counts. Not only was he able to greatly reinforce his army thanks to his decision to defer the decisive engagement, but he was also able to maneuver Heraclius's force into battle at a place conducive to the Arabs' more mobile soldiers. Fought in the summer of 636, the resulting contest, the Battle of Yarmuk, ended in a decisive victory for the Arabs that shattered Heraclius's army.¹⁶

The Byzantine defeat at the Battle of Yarmuk settled once and for all the question of who would control Syria. Lacking sufficient troops to continue to contest the Arab attack, Heraclius abandoned the province and withdrew into Asia Minor through the Cilician Gate, a strategic pass in the Taurus Mountains. Gazing down on Syria for the last time from the top of the pass, he reputedly acknowledged the province's permanent loss by declaring “Peace unto thee, O Syria, and what an excellent country for the enemy.” Whether he actually uttered those words or not,¹⁷ his withdrawal made clear that Byzantine rule in the Levant had come to an end—a reality that the remaining cities in Syria and Palestine quickly grasped. With the return of imperial authority now an impossibility, their leaders opted to come to terms with the Arabs rather than to fight what they concluded would be a doomed effort. Even Jerusalem, seat of one of the Church's five patriarchates and site of the Crucifixion, negotiated an agreement with

¹⁴ Fred Donner, *The Early Islamic Conquests* (Princeton: Princeton University Press, 1981), 111–32.

¹⁵ Kennedy, *The Prophet and the Age of the Caliphates*, 60–61.

¹⁶ Kennedy, *The Great Arab Conquests*, 82–85.

¹⁷ He didn't.

Umar's troops. Thus, with shocking suddenness, a millennium of Greco-Roman rule in Syria and Palestine came to a close.¹⁸

Byzantine Egypt came next. If Syria had been important to the empire, Egypt was its economic linchpin. Thanks to the steady supply of water that the Nile River afforded, the province produced far more wheat and other agricultural goods than its people consumed. That surplus was critical to the well-being of the empire in two ways. First, it supplied the food that sustained the hundreds of thousands of residents of the Byzantine capital, Constantinople. Second, it produced substantial annual tax receipts that provided the government with the revenue needed to pay for the army. Without Egypt and its surplus, in other words, the Byzantine government would have been unable either to fund its military or to keep its capital fed. Unsurprisingly, given its importance to the empire, the Byzantine government devoted a great deal of attention to the province's defense. It stationed a large force of troops there and maintained a series of strategically placed fortifications throughout Egypt. By far, the most important of these were the walled city of Alexandria, located on the coast, and the imposing fortress of Babylon situated along the Nile River just south of the Delta. These two well-garrisoned positions dominated Egypt. Indeed, so long as they remained in imperial hands, Egypt could not be captured.¹⁹

Undaunted, Arab forces under the direction of the veteran commander Amr Ibn al-As (c. 573-664) invaded the province in 640. The attack seemed a fool's errand. The army he led was tiny and appeared to be no match for Egypt's extensive fortifications and powerful garrison. Nonetheless, he and his troops thrust their way into Egypt where they quickly defeated a Byzantine field army and laid siege to the pivotal fort of Babylon. Physically, the fortress was beyond the capabilities of al-As's small army. With ample provisions and stout walls, it could have held out against the Arabs indefinitely. The morale of the defenders was a different matter, however. As the months passed without any sign of a relief army coming to their rescue, the garrison gradually began to lose hope. By April 641, they had had enough. That month, the defenders agreed to surrender the fortress to al-As in exchange for safe passage. As a result, central Egypt was now under Arab control.²⁰

Al-As was pleased with his victory but was unwilling to settle for anything less than full control of the entire province. He consequently followed his triumph by besieging Alexandria. With imposing walls, a large garrison, and a port able to receive supplies, the city appeared even more invulnerable than the recently captured fortress of Babylon. Despite the apparent solidity of its defenses, however, the Arab commander once again achieved a nearly cost-free victory. Taking advantage of dissension among

¹⁸ Ahmad Al-Baladhuri, "The Battle of The Yarmuk (636) and After," accessed October 23, 2019, <https://sourcebooks.fordham.edu/source/yarmuk.asp>; Robert G. Hoyland, *In God's Path: The Arab Conquests and the Creation of an Islamic Empire* (New York: Oxford University Press, 2015), 47–48.

¹⁹ Hoyland, *In God's Path*, 68–70.

²⁰ Kennedy, *The Prophet and the Age of the Caliphates*, 64–65.

the defenders, he was able to secure the city's surrender from its leader, the Patriarch Cyrus (?-642), in September 642. Thus, in just two short years, Egypt and its massive grain production had become part of the rapidly expanding caliphate.²¹

The Conquest of the Persian Empire

Astonishingly, at the same time that Arab armies were overrunning the Byzantine Middle East, they were simultaneously dismembering the venerable Persian Empire. Like the attack on the Byzantine Empire, the invasion of the Sassanid domain began with a series of raids. Disruptive, these incursions prompted the Persian Shah, Yazdgard III (r. 632-651), to deploy his army to the border to repel the raiders. Prudence dictated a quick retreat lest the shah's powerful formations destroy the Arab raiding forces; sensing opportunity, however, Umar instead decided to reinforce the Arab troops in Mesopotamia—thus setting the stage for a decisive clash. It occurred just outside the small settlement of Qadisiya on the border between Mesopotamia and Arabian Desert. Details of the battle are hazy due to an acute absence of sources—historians are not even sure if it occurred in 636, 637, or 638—but the outcome was clear: Umar's forces won a crushing victory that shattered the Persian army and left the empire defenseless. The invaders wasted no time in exploiting their triumph. Moving quickly, they followed the battle by seizing Mesopotamia, including, pivotally, the imperial capital of Ctesiphon.²²

Yazdgard III and the four-century old Sassanid Dynasty were in serious trouble. The Arab occupation of Mesopotamia meant that the young shah could no longer access the resources of what was by far the wealthiest part of his empire. Meanwhile, the loss of the capital not only massively complicated his efforts to provide central direction to the defense of his realm, but also scattered his tax officials—thereby leaving him bereft of the bureaucrats needed to raise the revenue on which the restoration of his shattered armies depended.²³

The Arabs did not stop with the conquest of Mesopotamia, moreover. Taking advantage of the disarray that the capture of the Persian capital had created, they surged out of Iraq and onto the Iranian plateau. While they met significant resistance in a few areas, Umar's forces pushed inexorably eastward over the course of the late 630s and 640s. Yazdgard III could do little to stem the tide. With Arab armies nipping at his heels, he had no choice but to flee eastward while attempting—unsuccessfully—to rally the Persian nobility and to win the support of the Turkish principalities of Central Asia. The end finally came in 651. With his last supporters having deserted him, the shah took shelter in a watermill just outside the caravan city of Merv in modern-day Turkmenistan. Accounts differ about what happened next, but all end with the last Persian *shahansha*, or king of kings, dying, abandoned and alone, at the hands of the

²¹ David Levering Lewis, *God's Crucible: Islam and the Making of Europe, 570-1215* (New York: W. W. Norton & Company, 2007), 81–84.

²² Donner, *The Early Islamic Conquests*, 173–76.

²³ Hoyland, *In God's Path*, 66–67.

millar. So ended the Persian Empire, and so, at least for a time, ended the Arabs' astounding surge out of Arabia.²⁴

Sources of Arab Success

The extent of the Arab Conquest at the time of Yazdgard III's death was staggering. Between 632 and 651, Arab armies had established an empire of almost incomprehensible size. They had conquered a territory that stretched from Yemen in the south to the Caucasus Mountains in the north, and from Egypt in the west to Central Asia in the east. Even more shocking, they had, in the process, reduced one of antiquity's greatest empires to a rump state and had utterly destroyed another. How did they do it? What accounted for the tremendous sequence of victories that Arab armies achieved during this period?

From the start, popular explanations for the Arabs' success have tended to stress religious factors. Unsurprisingly, Muslims have long maintained that the Arabs' string of victories was a product of divine will. That is, they believed that the caliphs' soldiers had defeated their enemies because they were followers of the true religion and thus had God on their side. Indeed, so far as the early Muslims were concerned, the sheer scale of their conquests demonstrated conclusively that Islam was the correct path. How else could they explain their armies' seemingly preternatural success?²⁵

Many people in the West have also long posited a religious explanation for the Arab Conquest, though they reject the idea that the Arabs had God on their side. Instead, those arguing from this perspective have asserted that Arab success was a product of Islam's newness and its militarist ideology of jihad, understood here to mean 'holy war.' According to this interpretation, Arab soldiers fighting against Byzantine and Persian forces prevailed for two reasons. First, they had the 'zeal of the convert,' meaning that they had the extra motivation and fervor typically ascribed to those new to a faith. Second, they were inspired by an Islamic tradition that holds that a Muslim fighter who died while on jihad would go immediately to heaven where he would enjoy the company of seventy-two virgins. This interpretation concludes that these factors ensured that Arab Muslim soldiers entered battle without fearing death—a conviction that dramatically boosted their morale and all-but assured them of victory.²⁶

Present-day scholars take issue with these explanations. They dismiss the traditional Muslim interpretation that divine favor explains Arab success on the grounds that such an argument is, *ipso facto*, unprovable, and thus inconsistent with modern rules of scholarship and argumentation. For different reasons, they also reject the popular Western argument that Arab success stemmed from a combination of religious zeal and the enticement of immediate admission into heaven. Noting that the promise of

²⁴ Kennedy, *The Great Arab Conquests*, 169–99.

²⁵ Ansary, *Destiny Disrupted*, 47.

²⁶ Donner, *Muhammad and the Believers*, 199; Ansary, *Destiny Disrupted*, 47; Armstrong, *Islam*, 29.

seventy-two virgins came from a *hadith*—a report of the words and actions of Muhammad—that was recorded only in the ninth century, they reject the idea that it could have been a motivating factor in the wars of expansion. Instead, while acknowledging that Islam played a major role in motivating and organizing the Arab Conquest, they generally agree that three key factors made possible the caliphate's remarkable military success in the decades immediately following Muhammad's death: the Arab's mobility and effective leadership; deep-seated religious discontent in Syria and Egypt; and the peculiar weakness of the Byzantine and Persian Empires in the 630s and 640s.²⁷

Arab Mobility and Leadership

First, historians today argue that the Arabs benefited from their superior mobility and unusually effective leadership. Possessing large herds of camels and operating without cumbersome and slow-moving supply trains, the Arabs easily outmaneuvered their Byzantine and Persian enemies and were able to mass troops quickly at the point of battle. They also enjoyed highly capable leadership from far-sighted political figures such as Umar and from skilled battlefield commanders such as Khalid and al-As.²⁸

An episode recounted by the Muslim historian Al Tabari (839-923) illustrates the critical role that mobility and strong leadership played in securing one of the Arabs' earliest victories against the Byzantine Empire. As he relates, Umar's decision in 634 to order Khalid to rapidly shift his force from Mesopotamia to Palestine so that they could help the Arab soldiers there take advantage of the province's weak defenses presented the latter with a dilemma. If he took the long-but-secure route along the Euphrates River, his army would safely make the journey but would arrive too late to help defeat the Byzantine forces massing in Palestine. If he instead marched directly across the waterless Syrian Desert, he risked seeing a large part of his force die of thirst. A savvy leader, Khalid quickly came up with a clever way of ensuring that his soldiers could safely take the direct route through the desert. Before setting out, he had his camels drink their fill; he then ordered his men to periodically kill some of the animals during their journey across the desert and to fill their canteens with drinking water from the deceased animals' stomachs. In this way, his army was able to make it to Palestine in time to help win the critical Battle of Ajnadayn.²⁹

Religious Dissension

Second, historians today also stress the role that religious discord within the Byzantine Middle East played in facilitating the Arab Conquest. As noted in chapter two, the majority of Christians in that part of the empire were Monophysites who believed that Jesus was wholly divine in nature. Known as Copts in Egypt and Jacobites in Syria,

²⁷ Aslan, *No God but God (Updated Edition)*, 165–66; Kennedy, *The Great Arab Conquests*, 62, 70.

²⁸ Kennedy, *The Great Arab Conquests*, 366–73.

²⁹ Donner, *The Early Islamic Conquests*, 121–22.

they had faced concerted persecution from both the imperial government and the Church hierarchy ever since the Council of Chalcedon had declared their beliefs to be heresies in 451. They did enjoy a brief respite under Persian occupation during the early-seventh century—a point we shall return to in a moment—but then experienced a new and particularly aggressive period of religious oppression when Byzantine officials and clerics returned in the late 620s and early 630s. Persecution was especially severe in Egypt. Installed as both Patriarch of Alexandria and provincial governor in 631, the Bishop Cyrus oversaw a brutal campaign that centered on torturing and even executing Coptic clerics who refused to recant their beliefs.³⁰

There was another religious minority that suffered even more than the Monophysites from the reinstatement of Byzantine control: Jewish people. Despised by Christians of all denominations, Syria's long-suffering Jewish population had enjoyed a brief respite under Sassanid rule before enduring renewed persecution following the Persian withdrawal. In many ways, in fact, their circumstances had gotten considerably worse following Syria's return to Byzantine rule than they had been before the Persian occupation. Most notoriously, in 632, Heraclius decreed that all Jewish people living within the empire had to undergo forcible conversion to Christianity.³¹

Unsurprisingly, given this situation, the people of the Middle East did little to oppose the Arabs. Experiencing severe persecution at the very moment that the invaders were surging out of Arabia, the Copts, Jacobites, and Jews who lived in Syria and Egypt were disinclined to risk their lives resisting the Arabs who, while hardly liberators, did free them from their tormentors. Indeed, recalling recent Byzantine persecution, many cities opted to come to terms with the Arabs rather than resist them. Some later sources even suggest that the conquered people may have preferred the Arabs to their previous imperial masters. For instance, the medieval Arab historian Al-Baladhuri (820-892) related that the Syrian city of Hims requested Arab help when threatened with recapture by Heraclius, declaring that “[w]e like your rule and justice far better than the state of oppression and tyranny in which we” lived under Byzantine authority. Meanwhile, fear that Alexandria's huge Coptic population would help the Arabs enter the city played an important part in leading Cyrus to surrender the city in 642. Concerns that religious minorities would help the Arabs were not overblown. Textual evidence shows that the Arabs captured the city of Caesarea only after a Jewish resident eager to end Christian harassment revealed to them a secret passage through its walls.³²

³⁰ Kennedy, *The Great Arab Conquests*, 8–9, 67, 142–46.

³¹ Tom Holland, *In the Shadow of the Sword: The Birth of Islam and the Rise of the Global Arab Empire* (New York: Doubleday, 2012), 296.

³² Quote from Al-Baladhuri, “The Battle of The Yarmuk (636) and After”; Lewis, *God's Crucible*, 81–82; Kennedy, *The Great Arab Conquests*, 89.

Exhausted Empires

Finally, and most importantly, modern-day historians maintain that the peculiarly feeble state of the Byzantine and Persian Empires in the early-seventh centuries proved pivotal in making the Arab Conquest possible. Their weakness had two sources. The first was a catastrophic outbreak of bubonic plague in the sixth century known in the Byzantine Empire as the Plague of Justinian. The pandemic was a product of global trade. Travelling along the Silk Road, rat-borne fleas brought it from the east to Persia and the Mediterranean world in the early 540s. There, it spread rapidly in a series of devastating waves that resulted in the deaths of millions. Ruinous for the Persians, it proved particularly damaging to the highly urban, trade-oriented Byzantine Empire. Moving from city to city, it brought death on an almost-unimaginable scale—so much death, in fact, that the empire’s population declined by a dispiriting 25 percent over the course of the sixth century alone. The plague not only demoralized the people of these empires, but also confronted both states with a brutally challenging fiscal situation. Each still had to guard the same borders and to maintain the same infrastructure that had existed before the outbreak, yet they now had to do so with substantially reduced populations that generated commensurately less tax revenue than they had in the past.³³

Second, just before the Arab invasions, the two empires had waged a debilitating, twenty-six-year-long conflict known as the Byzantine-Sassanid War. Enormously costly, it left both empires in political, economic, and, especially, military disarray. The conflict had its origins in a Byzantine coup. In 602, a low-ranking army officer named Phocas (r. 602-610) seized power by deposing and killing Emperor Maurice (r. 582-602). The Persian shah, Chosroes II (r. 590-628), sensed opportunity in the turmoil that subsequently engulfed Constantinople; accordingly, having been helped into power by Maurice, he used the coup as a pretext to declare war. Phocas reacted incompetently. Better at seizing power than at ruling, his main response—stripping the Danube River of troops for use against Persia—backfired badly. Not only did the powerful nomadic Avars take advantage by invading the Balkans, but the repositioned troops failed to prevent Chosroes II from making substantial incursions into Byzantine Mesopotamia. These setbacks did little to help the already-unpopular Phocas, and he soon confronted a series of plots and outright revolts that only ended when Heraclius, the powerful governor of North Africa, landed in Constantinople and seized the throne.³⁴

The situation worsened considerably during the early years of Heraclius’s reign. Taking advantage of the instability that followed Phocas’s downfall, Chosroes II’s armies broke through the Byzantine defenses in northern Mesopotamia in the early 610s. Now

³³ William Rosen, *Justinian’s Flea: The First Great Plague and the End of the Roman Empire* (New York: Penguin Books, 2007), 1–3, 315–19; See also Kyle Harper, *The Fate of Rome: Climate, Disease, and the End of an Empire* (Princeton: Princeton University Press, 2017).

³⁴ John Moorhead, *The Roman Empire Divided: 400-700 AD* (New York: Pearson, 2001), 172–73.

able to operate unhindered, they quickly occupied Syria, Egypt, and Palestine. It was a huge defeat for Constantinople. The loss of these provinces and the vital grain and tax revenues that they produced enormously complicated efforts to mount a counterattack. The situation was so bleak in the early 620s, in fact, that Heraclius gave serious consideration to relocating his government to the city of Carthage in North Africa. Ultimately, however, the emperor opted not to pursue that drastic course. Instead, he implemented a series of military and fiscal changes that he hoped would reinvigorate the army and thus lead to victory.³⁵

Heraclius's reforms proved successful. While the Persians and their new allies, the Avars, unsuccessfully besieged Constantinople in 626, the emperor launched a bold counterstroke through Armenia. After ravaging the Persian Empire's northern provinces, he marched into Iraq and decisively routed the Sassanid army at the Battle of Nineveh in 627. That defeat marked the end for Chosroes II. Weary of the war and the taxes it required, Persian aristocrats deposed and executed him and then elevated his son to the throne. The new shah promptly sued for peace. Relieved, Heraclius accepted, and the two empires returned to their prewar borders.³⁶

While the long war imposed huge costs on both empires, its impact was greatest on Persia. Politically, defeat thoroughly discredited its ruling house and tipped power toward the nobility; as a consequence, it experienced four years of court intrigue and civil war during which a succession of seven short-tenured shahs ruled. Only with the accession of Yazdgerd III to the throne in 632 did Persia once again experience a semblance of political stability. The conflict's fiscal and economic consequences were every bit as dire. The long war had been tremendously expensive and had resulted in the devastation of several economically critical provinces and in a concomitant steep decline in tax receipts; as a result, Persia lacked the revenue needed to rebuild its thoroughly depleted military. Thus, Persia was uncharacteristically feeble following the war with the Byzantine Empire. With its army weakened, its political system destabilized, and its economy gutted, the empire that Yazdgerd III inherited was a shadow of its former self—a state that was in no position to repel a serious attack.³⁷

While Heraclius avoided the kind of political challenges that Chosroes II's successors faced thanks to his victory, he too faced serious fiscal and military problems in the conflict's wake. Fiscally, the expense of the lengthy war had placed enormous strains on the treasury, while the temporary loss of the empire's two most economically vital provinces—Syria and Egypt—had produced a precipitous decline in tax revenue. As a result, the Byzantine treasury came out of the conflict in no position to support another war. Militarily, meanwhile, the empire's once firm defenses had suffered greatly as a result of neglect and wartime exigencies. Compelled to maintain the bulk of its

³⁵ Holland, *In the Shadow of the Sword*, 289–91.

³⁶ Edward N. Luttwak, *The Grand Strategy of the Byzantine Empire* (Cambridge: The Belknap Press of Harvard University Press, 2011), 399–408.

³⁷ Chris Wickham, *The Inheritance of Rome: Illuminating the Dark Ages 400-1000* (New York: Penguin Books, 2009), 258.

army on the Persian front, Constantinople could no longer keep the Avars and Slavs at bay and had to accept the loss of its traditional defensive line along the Danube River. The situation was even more dire in Syria and Egypt. There, the empire confronted the daunting task of completely rebuilding its shattered defensive positions immediately following the conflict—a task that promised to take many years. Thus, like Persia, the Byzantine Empire emerged from the war ill prepared to resist a concerted attack.³⁸

Scholars today largely agree that the feebleness of these venerable empires following their long war with each other was the most pivotal factor in setting the stage for the success of the Arabs in the 630s and 640s. Militarily weakened, economically strapped, and, in Persia's case, politically destabilized, they proved to be easy pickings for the Arabs' small, but well-led and highly mobile armies. Thus, contingency—luck—played an outsized part in the astonishing success of the Arab Conquest. As the historian Hugh Kennedy puts it, “[i]f Islam had been born fifty years earlier, and the early Muslims had attempted to raid Syria and Palestine in the 580s not the 630s, there can be little doubt that they would have been sent off very quickly, as the provinces were firmly controlled by the government and the defenses well organized.”³⁹

Early Islamic Government

Arab leaders were pleased with the extent and speed of the conquests, but they also understood that the rapid acquisition of a huge empire created significant practical challenges that they would need to quickly address if they were to retain possession of the valuable territory they had seized. Two issues were particularly pressing. First, they had to determine how they were going to administer the huge number of non-Muslims that they had conquered. Second, they had to develop political structures to ensure that power did not decentralize as the empire expanded but instead remained firmly in the hands of the central government in Medina. Resolving those challenges fell initially on the shoulders of Umar and his successor, Uthman—the two men who oversaw the most explosive phases of the early Arab conquest. Astonishingly, in light of his background as a relatively modest merchant, Umar proved remarkably adept at managing the growth of the empire and at addressing some of the problems that came with rapid expansion. Uthman, in contrast, was unable to deal effectively with issues arising from the caliphate's rapid growth; as a consequence, the empire would experience significant upheaval during the latter part of his reign.

Umar (r. 634-644)

As his armies began racking up victory after astonishing victory, Umar grappled with the question of how his government could best administer the newly acquired territories and the huge, non-Muslim populations that they contained. According to the traditional Muslim narrative of the conquests, he addressed these issues by hewing

³⁸ Moorhead, *The Roman Empire Divided*, 205; Kennedy, *The Great Arab Conquests*, 70–71.

³⁹ Kennedy, *The Great Arab Conquests*, 70.

closely to Qur'anic decree and to Muhammad's earlier practices. Thus, he did not order the forcible conversion of those now living under Arab rule. Instead, honoring the Qur'an's declaration that "there is no compulsion in religion," he gave defeated people—particularly the "People of the Book," meaning Christians, Jews, and, later, Zoroastrians—three choices: conversion to Islam, death, or submission to Muslim political authority as *dhimmi*, meaning, literally, protected people. (Qur'an, 2:256).⁴⁰

Umar's decision to rule the conquered territories indirectly and with a light hand made this last choice a hard one to refuse. Cities that surrendered could continue to administer themselves as they had in the past, and merely had to agree to collect and forward the *jizya*, a poll tax paid by the *dhimmi*, to the central government in Medina. These were remarkably easy conditions made even more palatable by the fact that the *jizya* may have been less financially burdensome than the levies that the Byzantines and Persians had previously imposed. Meanwhile, aware that his government lacked the institutional capacity to rule the new lands at the local level, he opted to retain the existing Byzantine and Persian administrative institutions—after all, the bureaucrats who staffed them excelled at the critical task of collecting taxes. Finally, he chose to continue to use Greek and Persian as the languages of administration in the conquered territories. Collectively, these were unusually lenient terms, which, by encouraging many cities to surrender rather than to resist, paid substantial dividends for Umar and the Arabs.⁴¹

Umar also took the first, critical steps toward ensuring that political power did not decentralize as the empire expanded. He refused to let Arab soldiers seize land—which would have provided them with enough wealth to be independent of Medina—and instead stationed them in newly founded garrison towns such as Kufa and Basra in Iraq and Fustat (Cairo) in Egypt. This decision had far-reaching consequences. It kept the soldiers concentrated and thus ready for further military operations, and it ensured that they did not assimilate into the local culture—thus preserving their Arab-Muslim identity. Umar next set up a sophisticated fiscal system wherein an office called the *diwan* used the proceeds of the *jizya* to pay salaries to Arab warriors. By keeping the soldiers financially dependent, this arrangement ensured that they remained firmly under the control of the central government. Finally, Umar appointed governors, or amirs, to oversee the conquered territories and to ensure that the Arabs carried out his will.⁴²

Uthman (r 644-656)

Though Umar had bequeathed him a solid foundation, his successor, Uthman, proved far less successful in administering the evolving empire and at balancing the different interests within it. Indeed, discontent—particularly among the Alids—beset his

⁴⁰ Kennedy, 373.

⁴¹ Ira M. Lapidus, *A History of Islamic Societies*, 2nd ed. (Cambridge: Cambridge University Press, 2002), 36.

⁴² Donner, *Muhammad and the Believers*, 136; Aslan, *No God but God (Updated Edition)*, 125.

rule from the very start. Still seething that the leading Muslims had passed Ali over in 632 and again in 634, they anticipated that he would finally become caliph following Umar's death at the hands of an insane Persian slave in 644. Unsurprisingly, they greeted the news that Ali had once again been sidelined with incredulity and fury. They were especially incensed by the fact that Uthman was from Banu Umayya—the clan that had most stridently opposed Muhammad in the 610s and 620s and whose members had resisted Islam until, conveniently it seemed, they had finally felt moved to convert just as the prophet was taking control of Mecca. Ali shared his followers' disappointment. Not wanting to stoke dissension, however, he agreed to acknowledge Uthman as the new leader of the *umma*. This move eased the tension that had followed the new caliph's succession and helped ensure that Uthman enjoyed broad support at the start of his reign. Still, below the surface, there remained a latent, if as-yet inchoate dissatisfaction with Uthman's selection among the Alids and some other important groups.⁴³

A series of controversial decisions on Uthman's part soon brought that discontent to the surface. First, early in his reign, he had authorized the production of an official, canonical version of the Qur'an. This action was uncontroversial in itself, but his related decree that all other variants had to be destroyed provoked heated opposition both from Muslims who preferred those versions and from those who believed that he was overstepping his authority. Second, lacking Umar's charisma and forceful nature, he had found maintaining control of the empire increasingly difficult. To compensate, he replaced the existing amirs with members of Banu Umayya. This decision was not an unreasonable one—after all, he could rely on his kinsmen to carry out his wishes—but it alienated many of the older Muslims who resented the fact that they were increasingly under the rule of the Umayyad clan that had so bitterly resisted Muhammad and his religious message. Finally, seeking to consolidate the empire before it expanded further, Uthman brought the first phase of the Arab Conquest to a close in 651. This change did not adversely affect those already on the *diwan* and thus engendered little opposition among the veterans of the earlier campaigns. However, it produced significant discontent among new recruits who thereafter did not have the opportunity to earn the military experience needed to start receiving a salary—a state of affairs that they found particularly galling in light of the fact that a growing share of those who received compensation did so not because they had fought in the wars of expansion but because their fathers had.⁴⁴

Uthman's position rapidly deteriorated after 650 as discontent with these decisions spread. By 655, the situation had gotten so bad that the garrison towns of Iraq were in outright revolt against his government and factions like the Alids and the Ansar were loudly demanding a series of sweeping changes to the government. The end finally came in 656. Seeking redress for their grievances, armed delegations from Egypt and Mesopotamia descended on Medina. There, they discovered that the rest of the Muslim elite—including, most notably, Ali—had forsaken Uthman and that the caliph

⁴³ Aslan, *No God but God (Updated Edition)*, 126–27.

⁴⁴ Kennedy, *Caliphate*, 17–19, 24.

was thus vulnerable and unprotected. Still, weak though his position may have been, Uthman refused to accede to their demands. Infuriated, the mob forced its way into his house and murdered him while he was reading the Qur'an.⁴⁵

The First *Fitna*

The assassination of Uthman came as a great shock to the *umma* and initiated a long period of instability marked by political conflict and outright *fitnas*, or civil wars, that lasted from 656 to 692. This era was pivotal in the political evolution of the *umma* and—as we shall see in chapter four—in the emergence of competing sects of Islam. It was a time in which divergent attitudes about who could lead the community, about how that person was to be selected, and about the nature of the caliph's responsibilities began to take clear form. It was also a period in which the government transformed from a personal and consensual one based on traditional Arab practices to an absolutist, monarchical one patterned after the government of the Byzantine state.

Ali (r. 656-661)

The first order of business following Uthman's death was the selection of a new caliph. Ali was the obvious choice, and he at last became the new leader of the *umma*. Delighted that he had finally assumed the office that they believe he had been unjustly denied, his Alid backers rejoiced. Their elation would prove premature, however. Dominated by the First *Fitna*, his tenure as caliph was marked by a series of challenges to his legitimacy that included the first large-scale bloodshed between rival Muslim forces. As a result, while of enormous political and religious significance, his reign would also prove to be a brief and unhappy one.⁴⁶

The first challenge to Ali came from Muhammad's widow, 'A'isha, and two of the prophet's former companions, Talhah (594-656) and Zubayr (594-656). Though they appeared nominally supportive of Ali immediately following Uthman's murder, they were in fact looking for an opportunity to rally people against him. The new caliph soon obliged. Eager to restore order in chaotic Medina, he felt compelled to pardon the men involved in Uthman's murder. This decision gave 'A'isha, Talhah, and Zubayr the excuse they were seeking. Using Ali's failure to prosecute Uthman's killers as a pretext, they declared themselves in revolt and went to the garrison town of Basra to raise an army. Having not yet fully consolidated power, Ali viewed this challenge as a serious one. Accordingly, desperate to nip it in the bud, he recruited troops in Kufa—including some who had participated in Uthman's murder—to bolster his forces before marching to Basra in late 656. There, in a contest known as the Battle of the Camel because the fiercest fighting took place around the mount 'A'isha was riding, Ali's army won a decisive victory. Talhah and Zubayr both died in the battle while 'A'isha, whose status

⁴⁵ Kennedy, *The Prophet and the Age of the Caliphates*, 73–75.

⁴⁶ Aslan, *No God but God (Updated Edition)*, 131–32, 134.

as the Prophet's wife insulated her from more serious consequences, was compelled to live out the rest of her life in quiet retirement in Medina.⁴⁷

The Battle of Siffin, 657

Almost immediately after defeating 'A'isha and her supporters, Ali confronted a considerably more serious challenge: a revolt by the powerful amir of Syria, Muawiyah ibn Abi Sufyan (r. 661-680) of Banu Umayya. A savvy leader, the amir had deftly caught the caliph in a bind. Pointing to Ali's reliance on the very soldiers involved in the death of his relative Uthman, Muawiyah declared that he would not acknowledge Ali as caliph until the latter had first punished the men who had killed Uthman—a demand that the amir knew the caliph could not meet in light of his dependence on his predecessor's murderers. Initially a war of words, the dispute devolved into civil war by 657. That summer, Muawiyah began marching his troops down the Euphrates River from Syria while Ali brought his up from Kufa. The two armies met at the Battle of Siffin near modern-day Raqqa in late July. The fighting did not go well for Muawiyah, and his army soon stood on the brink of defeat. In response, the amir pursued a desperate gambit. He ordered his soldiers to stick pages of the Qur'an on their spears and to call for arbitration in a bid to sow dissension in Ali's army. The ploy worked. Realizing that many of his soldiers would not continue the fight if he refused the offer of arbitration, a frustrated Ali felt that he had no choice but to accept Muawiyah's proposal.⁴⁸

New Challenges

In two ways, this decision proved to be a serious misstep on Ali's part. First, Muawiyah's arbitrator, the conqueror of Egypt, al-As, ran rings around Ali's representative, Abu Musa Ashari (?-c. 662 or 672). As skilled at negotiating as he was at leading troops into battle, al-As managed to trick Ashari into accepting an agreement that called for both Muawiyah and Ali to renounce their claims to the caliphate, after which a *shura* would choose a new leader. Ali was furious. Though presented as a compromise, the deal was, as he quickly intuited, a bad one that caught him on the horns of a dilemma. If he accepted it, he would lose the caliphate; if he instead refused to abide by a process that he had freely entered into, he stood to lose credibility within the *umma*. His assessment was correct. Disenchanted with him when he made clear that he refused to accept the terms of the deal, many of his Iraqi supporters either abandoned him or—worse—switched sides.⁴⁹

Second, and more troubling, was the reaction of a small, radical group of Ali's soldiers who soon came to be known as Kharijites, or "seceders." Their objections centered on the caliph's willingness to enter into arbitration with his rebellious subordinate. In their view, Ali had been obligated as caliph to enforce divine justice by attacking the rebellious Muawiyah. If the amir was guilty, God would ensure that he was

⁴⁷ Kennedy, *The Prophet and the Age of the Caliphates*, 75–77.

⁴⁸ Lewis, *God's Crucible*, 90–92.

⁴⁹ Lewis, 92.

defeated; if Muawiyah was instead in the right, God would see his army to victory. Arbitration was thus tantamount to a usurpation of God's power to judge, a sin so grave in the Kharijites' eyes that it rendered Ali's rule illegitimate and entitled—indeed *required*—the community of believers to depose him. Accordingly, chanting “only God can decide,” the Kharijite extremists abandoned Ali in disgust once he agreed to arbitration and began plotting his overthrow.⁵⁰

To do so, the Kharijites followed a course patterned after Muhammad's actions four decades earlier. Just as Muhammad had retreated during the hijra and reestablished his movement in Medina when faced with a greater power, so, too, did the Kharijites withdraw—initially to central Iraq and later to Arabia and Persia. Likewise, just as the Prophet had established a new, ideal community of believers in his new home, so, too, did they act to make their state the perfect Muslim society. Finally, just as Muhammad had waged war against his Meccan enemies from his new base, so, too, did the Kharijites mount a campaign against Ali from their new home.⁵¹

Thus, Ali was in an increasingly weak position following the Battle of Siffin. Not only did both Muawiyah and the Kharijites continue to deny that he was the legitimate caliph, but, more broadly, his refusal to abide by the arbitrators' decision had seriously damaged his legitimacy within the *umma*. Even his successes came with caveats. Most notably, while his effort to bring the Kharijites to heel met with substantial success in 658 and thus solidified his hold on Iraq, it also drew his attention away from the more serious challenge that Muawiyah posed. As a result, his Umayyad rival had time not only to rebuild his army but, worse, to gain control of the wealthy province of Egypt. By 661, as a result, Muawiyah had once again emerged as a serious menace to Ali's position.⁵²

Now focused on the threat that Muawiyah posed, the caliph massed his army in Kufa in 661 in preparation for a campaign against Syria designed to finally bring the insubordinate amir to heel. He never got the chance to lead it, however. Instead, the day before the campaign was to begin, a Kharijite assassin struck Ali with a sword as he entered the mosque in Kufa to pray. It was a glancing blow that did little physical harm; however, the weapon had been treated with a powerful poison that slowly killed the caliph over the next two days. The political consequences of his death were enormous. With Ali gone and his supporters hopelessly divided, there was no one left to stand up to Muawiyah; as a result, the way was now paved for the amir to become caliph.⁵³

⁵⁰ Gerald R. Hawting, *The First Dynasty of Islam: The Umayyad Caliphate AD 661-750*, 2nd Edition (New York: Routledge, 2000), 28–30.

⁵¹ John L. Esposito, *Islam: The Straight Path*, Fourth Edition (New York: Oxford University Press, 2011), 46–47.

⁵² Aslan, *No God but God (Updated Edition)*, 137.

⁵³ Aslan, 137–38.

Political Debates

The First *Fitna* proved enormously significant for the rapidly evolving Islamic polity. It not only settled the short-term question of who would control the empire, but, more importantly, also gave form to three competing schools of thought regarding both who was eligible to serve as Commander of the Faithful and the extent of that person's authority. The first, held by Ali's supporters, maintained that the rightful successors descended from Muhammad through his daughter Fatimah (?-632) and her husband, Ali, and that the caliph served as both a secular ruler and as a divinely-endowed religious leader, or Imam.⁵⁴ The second position, held by those who supported Muawiyah, asserted instead that the leader of the *umma* had to come from the Quraysh tribe, and that while the caliph was obligated to uphold the Qur'an and the sunna—the laws and customs that Muhammad had established—the position was a fundamentally secular one with no special religious role or powers. The third view, advocated by the Kharijites, likewise held that the caliph's role was secular, but took a very different view regarding the question of who could serve. For them, the caliph's tribal affiliation and ancestry were irrelevant. Instead, the sole qualification that mattered to the Kharijites was piety; in their eyes, only the most devout were eligible to serve as caliph and those that did had to remain free of sin for the duration of their reigns.⁵⁵

Inchoate as they were in the 650s, these schools of thought regarding who should rule proved to be enormously influential. As we shall see in chapter four, the fact that religion and politics were so tightly bound together at that time meant that the opposing views held by the supporters of Muawiyah and Ali would produce parallel religious understandings that would gradually evolve into competing Islamic sects. For its part, the Kharijit conception of interpretation of the caliph's authority never gained widespread support and thus did not produce a distinct religious perspective. Still, even it remained powerfully influential. Indeed, as we shall see, the Kharijites' puritanical beliefs, endorsement of terrorism, and strategy of retreating before counterattacking would remain influential and would continue to profoundly shape extremist movements in the Muslim world through the present day.⁵⁶

The Establishment of the Umayyad Dynasty

While Kharijism would exert a strong influence on extremist groups in the future, it was the new caliph, Muawiyah, who would define the political structure of the Arab Empire as it emerged from the First *Fitna*. He did so in the process of responding to two related issues that the civil war with Ali had revealed. First, the conflict had demonstrated that the *umma* would likely once more devolve into internal strife following his death if he did not first find some means of stabilizing and institutionalizing the succession. Second, the factionalism and anarchy of the First *Fitna* had made clear that

⁵⁴ Not to be confused with ordinary *imams*, or prayer leaders.

⁵⁵ Kennedy, *Caliphate*, 5, 26–29.

⁵⁶ John L. Esposito, *Unholy War: Terror in the Name of Islam* (Oxford: Oxford University Press, 2002), 41–42.

the traditional, tribal system of personal loyalty that lay at the heart of the caliphate's existing administrative arrangement was not up to the task of overseeing a massive, sedentary empire.⁵⁷

Ruling from Damascus, he based his reforms on the institutions of the caliphate's main rival: the Byzantine Empire. Drawing on three of its defining political practices, he transformed the government he inherited into a highly stratified and centralized system that lodged enormous power in the caliph's hands. First, he adopted a new, administrative apparatus modelled on the Byzantine bureaucracy that replaced the caliphate's informal system of governance with a stratified and highly efficient administration. Central to that effort was the creation of a chancery to oversee government finances and a postal service to improve communications. Second, he remade the position of caliph from one that ruled on the basis of persuasion and consensus into one patterned after the Byzantine Empire's absolutist emperors. Thereafter, he and his successors would no longer present themselves as first-among-equals; instead, they would rule as powerful, unchecked monarchs. Finally, determined to institutionalize a system of succession that would avoid the conflict that had ensued following Uthman's death, he adopted the Byzantine system of hereditary succession and named his son, Yazid (r. 680-683), to succeed him—thereby establishing the Umayyad Dynasty.⁵⁸

The majority of Muslims came to accept these changes, if perhaps grudgingly. Many were unenthusiastic about Muawiyah, and many grumbled that his assumption of imperial powers was alien to Muhammad's vision of a just and equitable society. Despite these reservations, however, opposition was muted. Wishing to avoid a repetition of the anarchy that had accompanied the First *Fitna*, most welcomed the return to order that Muawiyah had brought to the *umma*.⁵⁹

The Battle of Karbala, 680

Still, there remained a minority who balked at the changes that he had effected. The most vociferous among them were the *Shi'atu Ali*. Unlike many other Muslims, they did not object to having an hereditary system of government; however, they stridently opposed the idea of having the succession run through Banu Umayya and argued with growing fervor that it should instead follow the descendants of Muhammad through the line of Fatimah and Ali. The leader of the Alids at that time, Ali's son Husayn (626-680), reluctantly came to share this view. Though he had been content to focus quietly on his religious studies in Medina, Yazid's accession to the caliphate repelled him and led him to declare that he was the rightful caliph.⁶⁰

⁵⁷ Lapidus, *A History of Islamic Societies*, 47–49.

⁵⁸ Lapidus, 47–51.

⁵⁹ Nasr, *The Shia Revival*, 36.

⁶⁰ Ansary, *Destiny Disrupted*, 67–69.

Husayn moved quickly to organize the Alids against the Umayyad government. Lacking significant military support in the Hijaz, he left Medina in the late summer of 680 with a small contingent of soldiers and marched northward hoping to rally the strongly pro-Alid garrison town of Kufa to his cause. He never made it. Instead, the Umayyad governor of Iraq foiled his plan by first intimidating the soldiers in Kufa into inaction and then intercepting Husayn's force shortly after it had entered Iraq. In the ensuing Battle of Karbala, Yazid's troops handily crushed Husayn's tiny force. The Umayyad victory was so total, in fact, that it appeared to have ended once and for all the Alid movement. Events soon made clear that this impression was inaccurate, however. As we shall see in chapter four, Husayn's defeat did not mark the end of the Alids; instead, Karbala became a pivotal moral victory in their eyes—one that marked the moment that the the *Shi'atu Ali* transformed from a political movement into a full-blown religious sect.⁶¹

The Second *Fitna*

Yazid likely would have institutionalized the stable system of succession that his father had sought to put in place had he enjoyed a long reign, but his death in 683 and the passing of his son and successor only two months later created a power vacuum that left Umayyad rule in complete disarray. As a result, the fighting between Yazid's troops and the Alids that had begun in 680 morphed into a broader series of challenges to Umayyad legitimacy known as the Second *Fitna* (680-692). It was a chaotic time for the caliphate. While Umayyad forces retained control of Syria and Egypt, they faced challenges in every other part of the empire. In Mecca, the prominent Qurayshi Abd Allah ibn al-Zubayr (624-692) declared himself caliph shortly after Yazid's death. Meanwhile, a pro-Alid revolutionary named Mukhtar ibn Abi Ubayd (622-687) seized control of Iraq, and the revived Kharijites reasserted themselves in central Arabia. Hopes that the accession of the new Umayyad caliph, Marwan (r. 684-685) would arrest the decline of the dynasty's fortunes died with him the following year. Indeed, by the time his son Abd al-Malik (r. 685-705) took the throne in 685, the Umayyad Dynasty appeared to be doomed to imminent collapse.⁶²

To the surprise of many observers, it did not. Instead, Abd al-Malik's accession to the caliphate proved to be a turning point for the dynasty. A dynamic and charismatic leader, he had the determination, intelligence, and strength of will to restore Umayyad control of the empire. He began by quickly consolidating his position in Damascus. With Syria firmly under his control, he then ordered his trusted lieutenant, al-Hajjaj ibn Yusuf (661-714), to reestablish Umayyad authority over Iraq and to bring al-Zubayr and the Kharijites to heel. As decisive as Abd al-Malik, al-Hajjaj easily intimidated the Iraqis into capitulating in 691 before taking his army into the Hijaz. There, he demonstrated his characteristic ruthlessness by bombarding Mecca with catapults—some projectiles even

⁶¹ Nasr, *The Shia Revival*, 40–43.

⁶² Kennedy, *Caliphate*, 42–44.

struck the Ka'ba itself—before defeating al-Zubayr in 692. Finally, with Mecca secured, he brought the Second *Fitna* to a close by crushing the Kharijites in central Arabia.⁶³

The Great Leap Forward

With Abd al-Malik having solidified the dynasty's position, he and his Umayyad successors were finally in a position to continue Muawiyah's effort to remake the caliphate. Much of what they did built directly on his earlier efforts. For example, they further institutionalized the system of hereditary succession and absolutist rule that he had pioneered, and they expanded the bureaucratic system that he institutionalized. At the same time, however, these caliphs also broke new ground by implementing a series of enormously consequential new policies. They restarted the wars of conquest, undertook a number of important symbolic actions aimed at demonstrating the extent of the caliph's authority, and, most importantly, adopted new laws that began to erode the sharp line that had heretofore existed between the Muslim Arab conquerors and the vastly more numerous subject peoples. This last action proved particularly influential. With it, the Umayyads set in motion the forces that would result in the development of a new, Islamic civilization—a change so pivotal that the historian Robert Hoyland refers to it as “The Great Leap Forward.”⁶⁴

Asserting Caliphal Authority

Abd al-Malik initiated this period of rapid change with a series of high-profile symbolic initiatives designed to enhance the authority of the caliph. First, he decreed that the issuance of currency was a prerogative reserved solely to the caliph and ordered the minting of coins with his name on them. Thereafter, the gold *dinar* replaced the Byzantine *solidus* and the silver *dirham* supplanted the Sassanid *drachm*. Of great symbolic significance, this change not only demonstrated the central government's authority following the instability of the Second *Fitna* but, in banning the amirs from minting currency, also made clear that they were subordinate to the caliph. Second, he and his successor, al-Walid I (r. 705-715), commissioned monumental works of architecture—most notably the Dome of the Rock Mosque (692) on the Temple Mount and the Damascus Mosque (715). Designed to exalt Islam and to make clear its permanence, these monuments served an important secondary purpose as tangible signs of the caliphs' authority.⁶⁵

Renewed Wars of Conquest

Abd al-Malik also moved to restore the power of the caliph by restarting the wars of expansion. Since the first phase of the Arab Conquest had ended in 651, caliphal armies had won almost no new territory. In part, this lengthy pause in the empire's expansion was a function of the civil wars that had dominated the 650s and 680s.

⁶³ Armstrong, *Islam*, 44; Kennedy, *Caliphate*, 45.

⁶⁴ Hoyland, *In God's Path*, 138.

⁶⁵ Wickham, *The Inheritance of Rome*, 292–93.

During both *fitnas*, the Arabs had been too preoccupied with internal conflict to engage in efforts aimed at securing new territory. Even when they had been able to wage offensive wars during this period, moreover, they had been unsuccessful. Indeed, the most notable effort to expand the empire at that time—Muawiyah’s lengthy siege of Constantinople in the 670s—had ended in abject failure. Still, lingering memories of this setback aside, al-Malik saw great value in restarting the wars of conquest. Reasoning as Abu Bakr had five decades earlier that he could best maintain internal peace and assert caliphal authority by directing the energies of the Arabs outward, he followed his victory in the Second *Fitna* by ordering Muslim armies to return to the offensive.⁶⁶

They started in North Africa. There, Arab soldiers made rapid gains against the Byzantine Empire. Moving with characteristic speed and boldness, they secured possession of the city of Carthage and the rich grainlands of the Maghreb in 698. They followed by crossing the Strait of Gibraltar in 711 to invade the Visigothic Kingdom that ruled Spain. According to Arab sources, they were assisted in this bold move by Julian, the rogue Byzantine governor of the North African city of Ceuta. Furious that Roderic (?-711), king of the Visigoths, had impregnated his daughter while she was serving as a lady-in-waiting in his court, Julian helped the Umayyad army cross to Spain as an act of revenge, declaring that “I do not see how I can punish him and pay him back except by sending the Arabs against him.” While this story is almost surely apocryphal, the speed of the Arab conquest of Spain—thereafter known as al-Andalus—was anything but. By 716, Umayyad forces had completely destroyed the Visigothic Kingdom and occupied the entire peninsula save for the upland region in the far northeast. Even then, the caliph’s army did not stop pressing forward. On the contrary, Arab forces pushed on into Western Europe—stopping only when the Frankish ruler Charles Martel (r. 718-741) managed to check their advance at the Battle of Poitiers in south-central France in 732.⁶⁷

In the meantime, al-Walid I had restarted the wars of conquest in the east. These new attacks focused on conquering the regions of Sind along the lower-Indus River in present-day Pakistan and Transoxiana north of modern-day Afghanistan. Both operations were successful. After a difficult march through southern Iran, an Arab army took possession of the valuable region of Sind during the early years of al-Walid I’s reign in a campaign that marked the first major Muslim incursion into the Indian Subcontinent. To the north, meanwhile, a separate army concurrently surged across the Oxus River and into Transoxiana. In the process, they took control of a number of key commercial centers including the entrepôts of Bukhara and Samarkand and thus secured control over the western half of the Silk Road.⁶⁸

The only notable setback for Arab armies during this period was the failure in 718 of their second effort to conquer Constantinople. Massive in scale, the campaign came at the behest of the new caliph, Sulayman (r. 715-717), who sought a great victory to

⁶⁶ Lewis, *God’s Crucible*, 94, 98; Kennedy, *The Great Arab Conquests*, 216.

⁶⁷ Lewis, *God’s Crucible*, 120.

⁶⁸ Hoyland, *In God’s Path*, 148–51, 190–95.

enhance his power. The results were disappointing. Though the army he dispatched was substantial and though it enjoyed the support of a huge fleet, it proved unable to break through Constantinople's stout defenses. Instead, benefiting from the recent development of Greek Fire—a secret incendiary weapon that proved especially effective against ships—the city's defenders succeeded in turning the tables on the Arabs and in compelling them to retreat in disorder in 718. Thereafter, the caliphate made no further effort to attack Constantinople and instead consigned itself to frequent raids into Byzantine territory in eastern Anatolia.⁶⁹

Islamic Civilization

Meanwhile in parallel with the renewed campaigns of conquest, the Umayyad caliphs of the late-seventh and early-eighth centuries also effected a series of legal and social changes that would have far-reaching social and political ramifications. As with the renewed wars of conquest and the symbolic elevation of the caliph, it was Abd al-Malik who initiated this process. Seeking to elevate the culture of the conquerors, he ordered that Arabic would henceforth replace Greek and Persian as the language of administration throughout the empire. It was a sweeping ruling that had far-reaching consequences. It required those who wished to work in the administration or who hoped to do business with it to learn the language of the conquerors; over time, as a result, Arabic gradually displaced Coptic, Greek, and Aramaic as the dominant tongue in most of the empire. Only Persia resisted. Thanks to a proto-nationalist cultural rebirth it underwent in the ninth century, that part of the empire proved impervious to spoken Arabic. Even there, however, written Arabic made sufficient inroads that literate Persians gradually came to accept it as the basis for their system of writing. Thus, by the year 1000, the process that Abd al-Malik had set in motion was nearly complete. Arabic had become not only the dominant language of the Middle East, but also—thanks to the caliphate's central position in long-distance trade networks—the *lingua franca* of commerce in a zone stretching from Central Asia in the north to East Africa in the south, and from the Maghreb in the west to the ports of India in the east.⁷⁰

Even more significant for the people of the Arab Empire was a push to attract non-Muslims to Islam. This new effort marked a significant change for the *umma*. Through the early-eighth century, Muslims had made little effort to convert non-Arabs to Islam. On the contrary, jealous of the prerogatives that accompanied membership in the ruling religion, most sought to prevent *dhimmi* from adopting their faith. Even the state itself took steps to limit conversion. Concerned about the impact that the mass adoption of Islam would have on tax receipts, the government actively discouraged conversion and required those who became Muslims—known as *mawali* or clients—to continue to pay the *jizya* tax assessed on the *dhimmi*. This approach was effective. Through the

⁶⁹ Lewis, *God's Crucible*, 136–37.

⁷⁰ Albert Hourani, *A History of the Arab Peoples* (Cambridge, Mass: Belknap Press of Harvard University Press, 2002), 27.

early eighth century, conversion rates remained low, and the overwhelming majority of the subject people retained the religion into which they had been born.⁷¹

That approach began to change under Umar II. Noted for his piety, the caliph hoped to encourage conversion to Islam and wished to end the *mawalis*' second-class status. Accordingly, he decreed that Muslims would henceforth pay the same taxes and be accorded the same status whether they had been born into the religion or had converted.⁷²

Meanwhile, though its provenance is disputed—tradition holds that Umar I issued it—Umar II was also likely responsible for the Covenant of Umar. Designed to formally establish the status of non-Muslims in the empire, it enumerated the rights and restrictions that would apply to the *dhimmi*. It barred them from riding horses or marrying Muslim women, imposed on them a series of sartorial laws that required them to identify themselves by wearing certain clothing, prohibited them from building prominent houses of worship, and required them to show deference to Muslims. In exchange, the caliph promised to protect them and guaranteed that they would enjoy the right to worship freely. The rights that the Covenant of Umar accorded to non-Muslims were substantial—particularly during an era hardly noted for its religious tolerance. Still, the restrictions that it imposed were also significant and made life considerably less pleasant for the caliphate's non-Muslims.⁷³

Together, the religious and fiscal changes that Umar II effected marked a critical turning point for Islam's growth in the Middle East. Over time, the carrot of lower taxes and the stick of greater restrictions on non-Muslims would help to spur a steady increase in conversion. It is important to keep in mind that this change was very gradual; indeed, Egypt would remain predominantly Christian through the fourteenth century. Nonetheless, the demographic evidence makes clear that it was during Umar II's reign that conversion—heretofore a trickle—rose to a rate sufficient to ensure that Islam would eventually become the dominant religion in the Middle East.⁷⁴

More broadly, coupled with Abd al-Malik's reforms, Umar II's policies substantially weakened the barriers between conquered and conqueror and led to the gradual emergence of a new, hybrid, Islamic civilization. Importantly, though the Arabs were clearly the dominant party, its creation was not simply a one-way street in which the conquerors imposed their culture and religion on the *dhimmi*. Instead, it was a complex, two-way exchange in which, in the words of the historian Robert Hoyland, the conquered people "absorbed the Arabs and reshaped their values" even as the Arabs adopted much of the existing culture of the region. Thus, Islamic art and literature derived more from pre-Islamic Persian than Arab traditions, while the legal system incorporated vast numbers of Roman rulings that belatedly gained approval as Muslim

⁷¹ Hawting, *The First Dynasty of Islam: The Umayyad Caliphate AD 661-750*, 76–78.

⁷² Hoyland, *In God's Path*, 161; Kennedy, *Caliphate*, 54.

⁷³ Hourani, *A History of the Arab Peoples*, 47.

⁷⁴ Hoyland, *In God's Path*, 161; Kennedy, *Caliphate*, 54.

law when scholars interpreted Muhammad's words in ways that made him appear to be their source. The result was something new: an amalgamated civilization that was Arab in language and Islamic in religion, but with strong antecedents in the pre-Islamic cultures of both the conquerors and the subjugated.⁷⁵

Umayyad Decline

While the Umayyads had enjoyed substantial success during the late-seventh and early-eighth centuries, they began to confront growing problems in the 730s and 740s. Three issues in particular hamstrung the Umayyads. First, the dynasty gained a reputation as degenerate pleasure seekers whose lives of opulence and conspicuous consumption were at sharp variance with the simplicity and morality of Muhammad and the early community of believers. This perception was more caricature than reality, but, at great cost to the Umayyad Dynasty's legitimacy, it nonetheless gained widespread acceptance. Second, thanks to distance and the resilience of the remaining states that bordered the Arab Empire, the wars of conquest that Abd al-Malik had restarted gradually petered out by the end of the 720s. As a result, the Umayyads no longer enjoyed the fiscal ease that the steady influx of loot had afforded earlier caliphs. Finally, the dynasty confronted a peculiarly challenging political dilemma. On the one hand, the *umma* could not thrive without the order, efficient tax collection, and firm enforcement of the laws that only a highly centralized government could provide; on the other, Islam was a revolutionary and egalitarian religion that explicitly challenged the political and economic hierarchies that were intrinsic to such a government. As a result of this tension, efforts to consolidate and enhance caliphal power—however reasonable or necessary—produced bitter opposition among those who saw such initiatives as a return to the arbitrary and capricious practices of the empires that the Islamic revolution had usurped.⁷⁶

Thus, the Umayyads found themselves in an increasingly tenuous position by the middle of the eighth century. Facing a steadily worsening fiscal situation, the ruling dynasty now confronted opposition from many Muslims who seethed at a political order that they believed ran counter to Islam's ethic of social justice and who resented a ruling class that seemed to have abandoned Islam for hedonist excess. Indeed, though the Umayyads did not realize it, by the 740s, they were on very thin ice. At that point, all it would take to overthrow the dynasty was the emergence of an organized, ideologically appealing, and militarily powerful opposition movement within the empire.⁷⁷

⁷⁵ Hoyland, *In God's Path*, 157–58.

⁷⁶ Wickham, *The Inheritance of Rome*, 292–95.

⁷⁷ Esposito, *Islam*, 54–55.

The Abbasid Dynasty

The Abbasid Revolution

Such a movement emerged with great suddenness in the mid-740s in the province of Khurasan. Encompassing the far-northeast part of the empire, the province contained a large population of Arab transplants and Iranian *mawali*, who, as residents of a frontier province, were quite proficient at warfare. By the 740s, both groups fumed with resentment at the Umayyads. The Arab residents of Khurasan objected to the Syrian Arabs' dominance of the government and to the high levies that the dynasty assessed on them while the *mawali* chafed at the lingering discrimination they experienced and at the government's recent decision to reimpose the poll tax on them.⁷⁸

Seeking to tap into that resentment, a mysterious Persian provocateur named Abu Muslim (?-755) arrived in Khurasan in the mid 740s. He bore a straightforward and apocalyptic message for the disgruntled people of the province: the Umayyad caliphs were corrupt pleasure seekers who had perverted Muhammad's vision of the ideal, egalitarian society. As such, devout Muslims were obligated to band together to overthrow the reigning Umayyad caliph and to replace him with a new ruler from Muhammad's clan, Banu Hashim. Abu Muslim's message fell on receptive ears throughout the empire. Its revolutionary implications drew the disgruntled Arabs and *mawali* of Khurasan to his cause, while his call for a new government headed by a "chosen one from the Family of the Prophet" was—though vague—sufficiently pro-Alid to win over many of the *Shi'atu Ali*.⁷⁹

Bolstered by this support, Abu Muslim felt strong enough to shift from propaganda and underground organization to open revolt in the late 740s. Under the black banners that would become the hallmark of the dynasty that he helped put in power, his army seized control of the strategic city of Merv in 747 and then drove the Umayyad governor out of Khurasan in 748. Greatly strengthened by this success, Abu Muslim and his soldiers followed by marching into the heart of the empire. Gaining strength as they moved westward, they swept inexorably through Iran and into southern Iraq. Finally, along the bank of a tributary of the Tigris River, Abu Muslim's warriors decisively defeated the Umayyad army in 750—thus bringing the dynasty to an ignominious end.⁸⁰

Only then did Abu Muslim reveal that the movement he led had a narrower aim than he had previously suggested. To the disappointment of the Alids, he made clear after his victory that his goal was not to install as caliph someone from the line of Ali and Fatimah but instead to put on the throne a member of the Abbas family, which descended from Muhammad's uncle, Abbas ibn Abd al-Mutatalib (c. 565-c. 653) and which had secretly masterminded the revolt that he led. Accordingly, shortly after taking

⁷⁸ Hoyland, *In God's Path*, 201–6.

⁷⁹ Kennedy, *Caliphate*, 66.

⁸⁰ Hawting, *The First Dynasty of Islam: The Umayyad Caliphate AD 661-750*, 115–18.

control of Iraq, Abu Muslim arranged to have one of Abbas's great grandchildren, Abu al-Abbas (r. 750-754), proclaimed caliph in the mosque in Kufa—establishing, in the process, the Abbasid Dynasty that would rule the caliphate for the next five centuries.⁸¹

Abu al-Abbas spent the next few years seeking to ensure that the Umayyads could not mount a comeback. Accounts of this effort differ. One holds that he had all of the remaining members of the dynasty hunted down and killed; a competing story instead maintains that he invited them to his palace for a dinner party and had his guards murder them after they had taken their seats. Regardless of which account is accurate, the only member of the Umayyad family who managed to escape the dragnet was Abd al-Rahman I (r. 756-788), who, after an arduous trek across North Africa, established a competing caliphate in Spain.⁸²

Abu al-Abbas's successor, al-Mansur (r. 554-575), completed the consolidation of Abbasid power. He began by moving to ensure that the dynasty faced no danger from within the movement that had put it in power. Most notably, he had the popular—and, thus, potentially dangerous—Abu Muslim executed while the latter was passing through Iraq in 755. Having eliminated the most serious internal threat to his authority, he next dealt with the Alids. Bitter that Abu Muslim's call for a new caliph from Muhammad's family had failed to bring a descendant of Ali and Fatimah to power, they launched revolts against the Abbasids in 762 in the Hijaz and in Iraq. Once again, Al-Mansur acted decisively. Taking advantage of dissension among the rebels, he first crushed the uprising in the Hijaz and then, in 763, defeated the Alid forces in Iraq. He followed with an extended campaign of persecution against the remaining members of *Shi'atu Ali* that ended only when his successor, al-Mahdi (r. 775-785), took the throne.⁸³

The Abbasid Golden Age

Despite the ruthlessness with which the regime consolidated power, the Abbasid dynasty has long enjoyed a glowing reputation among non-Shi'i Muslims. In fact, while the conventional Muslim narrative of world history is highly critical of the Umayyads, it lauds the early centuries of Abbasid rule as Islam's golden age—an idealized time of prosperity, stability, and order during which the *umma* thrived under the dynasty's enlightened, Islamic government. According to this interpretation, the good times peaked under the reign of the fourth Abbasid caliph, Harun al-Rashid (r. 786-809). Cultured, intelligent, sophisticated, and, above all, empathetic to the needs and wants of his people, al-Rashid became the quintessence of the wise and caring Islamic ruler.⁸⁴

Ironically, a close examination of the dynasty reveals that its system of rule bore far more similarities than different from the rule of the despised Umayyads. Like their

⁸¹ Kennedy, *Caliphate*, 67–70.

⁸² Kennedy, *The Prophet and the Age of the Caliphates*, 128.

⁸³ Hugh Kennedy, *When Baghdad Ruled the Muslim World: The Rise and Fall of Islam's Greatest Dynasty* (Cambridge, MA: Da Capo Press, 2004), 18–26.

⁸⁴ Ansary, *Destiny Disrupted*, 86–87.

predecessors, the Abbasids oversaw a highly centralized and unrepresentative government that was fundamentally at odds with the egalitarian message of early Islam. They also enjoyed lives of luxury that at a minimum equaled the widespread-if-exaggerated stories of Umayyad excess and pleasure-seeking. As we have seen, moreover, the Abbasids were every bit as capricious and tyrannical as their predecessors had been. Even that most celebrated of Abbasid caliphs, Harun al-Rashid, resorted at times to arbitrary and brutal methods. When he felt threatened by his old friend and vizier, Ja'far ibn Yahya Barmakid (767-803), for example, he had Ja'far and the entire Barmakid family brutally murdered—despite the fact that it had faithfully served the Abbasids since the beginning of the dynasty.⁸⁵

So what accounts for the Abbasid Dynasty's continued good reputation? After all, the fact that the royal executioner stood at the side of the caliph during public ceremonies hardly suggests that the dynasty embraced restrained rule or Islamic egalitarianism. To some degree, the Abbasids' enduring good press is the product of a sense of political cynicism that settled onto the Muslim world in the early years of the dynasty. When the social and political revolution that the Abbasids had promised failed to arrive, most Muslims grudgingly came to accept that all secular rulers were power hungry and thus bound to behave just as the Umayyads had. Accordingly, the majority of Muslims thereafter assumed a pragmatic attitude toward the ruling elite: so long as the Abbasid caliphs engaged in their worst and most excessive behavior behind closed doors and so long as they continued to uphold the stability upon which the success of Islam depended, they would look past the dynasty's moral failings and give it their support. As a result, the Abbasids never faced the criticism that had bedeviled their Umayyad predecessors. More importantly, as we shall see in chapter four, the Abbasids' willingness to leave key religious questions—including the articulation and interpretation of Islamic law—to religious scholars known as 'ulama' secured for them the support of the influential clerical establishment, which was thereafter only too happy to celebrate the Abbasids and to buttress their rule.⁸⁶

The caliphate's strong economy during the first few centuries of Abbasid rule also strengthened the dynasty's reputation. In this regard, the Islamic world really did enjoy a golden age under Abbasid rule. The rapid growth of the Arab Empire had created a huge, unified, and stable market that was conducive to internal trade. The resulting expansion reached its apogee under the Abbasids. Indeed, by the early years of the dynasty's rule, the economy of the empire had grown so large and sophisticated that it had begun to pull both nearby and distant regions of the globe into a growing, Middle East-centered commercial order. With Iraq acting as its center, a massive new network of trade routes rapidly emerged to connect the economies of the Middle East, North

⁸⁵ Kennedy, *When Baghdad Ruled the Muslim World*, 71–79.

⁸⁶ Armstrong, *Islam*, 53–55, 58–59.

Africa, Sub-Saharan Africa, India, Central Asia, and even China into a cohesive trading bloc for the first time.⁸⁷

The new commercial order that the Abbasids oversaw benefited both the rural and urban economies. In the agricultural sector, connection with distant regions brought valuable new crops such as rice, bananas, sugarcane, and cotton to the Middle East for the first time. Coupled with the attention that the dynasty lavished on the irrigation works of Egypt and Iraq and the development of new innovations such as windmills and watermills, the spread of these crops had enormous benefits for the empire's rural people. Indeed, these changes markedly increased agricultural productivity, diversified and improved diets, and raised living standards. Meanwhile, growing contact with other cultures spurred the development of new industries in the empire's rapidly growing cities. For example, sugarcane and cotton brought to the Middle East from India stimulated the development of sugar refining and textile manufacturing in the cities of Iraq. The rapidly expanding paper industry likewise had external origins. A technology of Chinese origin, paper had spread to Central Asia in the centuries prior to the arrival of the Arabs in that region. The Arabs were quick to see its practical benefits. Indeed, within just a few decades, a substantial number of papermills existed in the cities of Iraq, and paper had become cheap and plentiful throughout the empire.⁸⁸

Finally, the Abbasids' enduring reputation stemmed from the fact that their approach to governance, while similar in broad outlines to that of the Umayyads, also broke with their predecessors' policies in a number of important ways. For example, aware of simmering popular anger at Syrian dominance of the empire under the Umayyads, the Abbasids had begun the dynasty's reign by quickly relocating the center of power to Iraq. Likewise, they acted to placate the heretofore disgruntled *mawali* by permanently abolishing the remaining special rights that the Arabs had enjoyed and by replacing the Arab tribal forces that had constituted the core of the empire's army since the time of Umar with a new, professional military dominated by the Muslims from Khurasan who had put the dynasty in power.⁸⁹

Less popular but still tolerable was the gradual Abbasid embrace of older Persian political ideas and practices. The dynasty quickly made the Sassanid shahs' claim that they ruled by divine right their own; they took imperious new titles such as God's Deputy on Earth and revived older Persian designations like Shadow of God on Earth. By Harun al-Rashid's time, the Abbasids had even adopted complex Persian ceremonies and court dress designed to exalt the status of the monarch and to cast him as detached from the day-to-day world of his subjects. Likewise, the caliphs soon required all subjects—even high-ranking ones—to prostrate themselves and kiss the ground when they approached him. Clearly, the position of caliph had changed considerably from the days when Umar took naps in the corner of the mosque in his homespun

⁸⁷ Peter Frankopan, *The Silk Roads: A New History of the World* (New York: Knopf, 2015), 91–98.

⁸⁸ Hourani, *A History of the Arab Peoples*, 45–46; Kennedy, *Caliphate*, 106–8.

⁸⁹ Lapidus, *A History of Islamic Societies*, 58–59.

clothing! One might think that the gap between the severe-but-approachable Umar and the wealthy-and-aloof Harun al-Rashid would have produced resentment toward the caliph. Perversely, however, the sense of distance and remove that the Abbasids' elaborate, Persian-style rituals created may have helped to insulate the caliph from popular anger. Thereafter, rather than blaming the ruler for actions they did not like, people held his viziers—the officials who were in charge of implementing those policies—responsible.⁹⁰

Baghdad

All of these trends—the embrace of the Umayyads' centralized system of government, the expansion of the economy, the political shift to the east, and the adoption of Persian political practices—came together in the Abbasids' new capital city: Baghdad. Al-Mansur established it in 762. Seeking a strategic location for his seat of government, he ordered his chief vizier to build the city on the west bank of the Tigris River at the point where it was nearest to the Euphrates. His choice of location was excellent. The city not only enjoyed ample water, but also had easy access via river and canal to the abundant food of Iraq. Indeed, it was a natural spot for an important city—a fact attested to by its proximity to both the ruins of Babylon, the capital of Hammurabi's (c. 1810-1750 BCE) ancient empire, and to Ctesiphon, the seat of the Sassanid Persian Empire.⁹¹

The city itself was very different from the Roman and Sassanid ones that the Abbasids inherited. Laid out as a perfect circle, it was an orderly, planned city that contained the caliph's palace, a mosque, and buildings for the administration of the empire. As such, it was well geared both to the centralized system of government that the Abbasids borrowed from the Umayyads and to the ritualized court ceremonies that they adopted from the Persians.⁹²

Baghdad did not serve strictly administrative and ritualistic functions for very long, however. Instead, its central location and access to the court quickly drew large numbers of people. Likely the most populous city on Earth for much of the Middle Ages, its population soon topped 500,000. A disproportionate share of those residents were intellectuals. Enticed by the patronage that the palace provided, legal scholars, scientists, mathematicians, and philosophers flocked to the city. As we shall see in chapter four, those thinkers would make Baghdad the greatest intellectual center of the Middle Ages. The city also had a huge population of merchants. Attracted by Baghdad's central location, political importance, and easy access by river, they soon turned the city into a major global entrepot. Traders based in the city exchanged silk and porcelain from China; gold, enslaved people, and salt from Africa; metalworks from Syria; spices and textiles from India; and leather goods from Spain. Indeed, for several hundred

⁹⁰ Armstrong, *Islam*, 54–55.

⁹¹ Kennedy, *The Prophet and the Age of the Caliphates*, 134–35.

⁹² Kennedy, *Caliphate*, 72–73.

years, it was the emporium of the world—standing, without challenge, as the nexus of the globe’s most important trading system.⁹³

The Historical Debate: The Arab Conquest

The aforementioned account reflects the standard understanding of the Arab Conquest and the development of the caliphate through the reign of Harun al-Rashid. That is, it broadly summarizes both the conventional Muslim narrative and the dominant interpretation that Western scholars articulated in the twentieth century. This view has not gone uncontested, however. On the contrary, a number of historians have challenged the standard account in a variety of ways over the past few decades—particularly its treatment of the Arab Conquest. Some of these reinterpretations accept the general thrust of the traditional understanding of the Arab invasions of the Byzantine and Persian Empires while still making important corrections to it. Others go much further in reassessing the conquest. Building on the work of the Islamic Revisionist historians discussed in chapter two, these scholars have instead developed new interpretations that upend long-held understandings of who participated in the invasions, what their motives were, and how they dealt with the newly conquered subject peoples.

These reinterpretations of the Arab Conquest emerged as a result of the doubts that many scholars have about the reliability of the sources on which the standard account is based. Put simply, historians who study the first century of the Arab Empire face the unfortunate reality that there exists precious little firsthand evidence of the invasions. As the historian Hugh Kennedy notes, the extant primary sources on the conquest are fragmentary and “can be counted on the fingers of one hand.” What does exist, instead, are narrative histories authored by Muslim scholars like al-Baladhuri and al-Tabari. Writing centuries after the events they described, they based their accounts on long-lost works produced in the eighth century that were themselves derived from earlier oral accounts passed from generation to generation. Unsurprisingly, many modern historians doubt the reliability of these narratives. They question whether the original oral accounts were, in the scholar Gerald Hawting’s words, “mangled in transmission” and argue that writers like al-Tabari and al-Baladhuri were likely shaping their histories to support one side or another in the religious and political debates of the centuries in which they lived.⁹⁴

Some historians have applied methods of critical analysis to these sources to draw conclusions that make relatively modest changes to the traditional narrative. Pointing to al-Baladhuri and al-Tabari’s mention of Persian nobles joining forces with the conquerors without having to convert to Islam or pay the *jizya* tax, for example, Hugh Kennedy concludes that the Arabs treated the conquered people with substantially more flexibility than the tripartite, ‘convert, die, or submit’ ultimatum suggests. Meanwhile, Fred Donner’s critical reading of the ninth-and-tenth century sources leads him to

⁹³ Lapidus, *A History of Islamic Societies*, 56–58.

⁹⁴ Quote from Kennedy, *The Great Arab Conquests*, 13, 16–18; Hawting, *The First Dynasty of Islam: The Umayyad Caliphate AD 661-750*, 16–17.

conclude that the traditional narrative's emphasis on religion as a major force in organizing and shaping the conquest is correct. However, he disputes the standard view that the invasion of the Byzantine and Persian Empires was a narrowly Islamic initiative. Instead, asserting that "the Believers"—his term for the early *umma*—initially practiced a "non-confessional" monotheism that accorded equal status to Muslims, Christians, and Jews, he argues that the Arab Conquest was, at least at first, an ecumenical venture and claims that Islam only began to develop as a distinct religion in the late-seventh century.⁹⁵

Other scholars are even more dubious of the traditional account and the evidence on which it rests. Relying on non-Arabic, seventh-century sources, for instance, Robert Hoyland goes considerably further than Donner in downplaying Islam's role in the Arab Conquest. While acknowledging that the religion's similarity to the other Abrahamic faiths gave it an important "integrative capacity" that allowed it to secure the conquered people's acceptance of the Arabs as rulers, he argues that the Muslim community, which he calls "Muhammad's west Arabian coalition," was but one group among many seeking to exploit the weakness of the established states in the mid-seventh century. Indeed, he claims that incursions by other, often Christian Arab forces predated the *umma's* raids. However, he continues, ninth-century Muslim historians erased the significant role that non-Muslims played by revising the chronology of the conquest to give it a strictly Islamic character—a change in interpretation, Hoyland asserts, that gradually "airbrushed out of history" the contributions of Arabs who were not part of Muhammad's group. Hoyland also takes issue with the conventional narrative's implicit claim that the Arabs were outsiders seeking to conquer the established order; instead, he contends that they were insiders—clients or, in many cases, residents of the Byzantine and Persian Empires—who used the sophisticated weapons and tactics that they had acquired while serving in the armies of those states to effect what turned out to be a wildly successful effort to win a share of political and economic power from within.⁹⁶

Conclusion

Regardless of the reason for the Arab Conquest, the results were enormously consequential. Indeed, the political, economic, and cultural changes that it brought about were so significant to the course of global history that its emergence conventionally marks the dividing line between Late Antiquity and the Middle Ages. Politically, it established a new superpower of immense scale that ended the Persian Empire and reduced the Byzantine Empire to a rump state. Economically, it produced a new system of global trade centered on Baghdad that displaced long-established commercial relations and that massively enriched the Middle East. Most importantly, the Arab Conquest resulted in the emergence of a new, enduring, Islamic-based civilization.

⁹⁵ Kennedy, *The Great Arab Conquests*, 180, 192; Donner, *Muhammad and the Believers*, 109.

⁹⁶ Hoyland, *In God's Path*, 41,56-57, 227–30; See also Holland, *In the Shadow of the Sword*.

It is to the evolution of the faith and to the social and cultural changes that accompanied its transformation that we shall next turn.

Chapter Four: Religious, Social, and Intellectual Changes During the Caliphate, 632-1258

At the same time that Arab armies were winning an empire, Muslims were also grappling with a series of fundamental questions about their society and religion. How should they organize their community to best please God? How were people to determine correct behavior in situations for which the Prophet—having never experienced them—had not provided guidance? Would God judge them based on their faith alone, or a combination of faith and works? What role should human reason play in understanding God's plan?

Above all, they struggled with the momentous question of leadership. As we have seen, Muhammad had died without naming a successor or articulating what authority that person would possess; as a result, the leading Muslims who survived him had no clear political path to follow. How and by whom, they debated, was the *umma* to be governed? Was the caliph to hold only Muhammad's political power, or was that person also to inherit his religious authority? As recounted in chapter three, these issues would produce intense factional infighting and even civil war among the Muslim elite. The question of leadership did not remain confined to the ruling class, however. Instead, with enormous consequences for the *umma*, the debate spilled over into the broader Muslim community where it produced a conflict over religion of such ferocity that it split Islam into rival variants.

The Division of Islam

How was it that a debate within the *umma's* leadership had transformed into a broader, popular debate about religion? It did so as a function of the peculiar structure of the early community of believers. Because Islam had begun as a movement in which religion and politics were tightly intertwined, any dispute over politics inevitably acquired a religious dimension and vice versa. Thus, a conflict about an issue as pivotal as the leadership inexorably produced an equally fraught debate over religion. The results would be enormously consequential for the *umma*. As it intensified, the dispute over who should lead produced a corresponding debate over the nature of the religion—one that would soon divide the community of believers between Sunnis, Shi'i, and the small, but-influential movement known as Kharijism.¹

Sunnism

Sunnism has been the largest branch of Islam since the religion's gradual separation into different denominations in the seventh and eighth centuries. Today, Sunnism accounts for about 85 percent of the world's estimated 1.8 billion Muslims. The term itself is derived from the word *sunna*—the habits, practices, and customs of Muhammad—and reflects the view that Muslims should live in emulation of the Prophet

¹ Karen Armstrong, *Islam: A Short History* (New York: Modern Library, 2002), 45–46.

and the early *umma*. Sunnis hold that all Muslims can follow the religion without the need for intercession by a formal priesthood or religious hierarchy, and they rely on ‘ulama’—male religious scholars more akin to rabbis than to Christian priests or ministers—to interpret religious law and to resolve legal disputes. It is a belief system centered far more on behavior than theology. That is, in keeping with the prevailing Muslim rejection of theological debate as *zannnah*, or idle speculation, Sunnis focus their attention more on living according to the Straight Path that Muhammad had laid out than on adhering to correct doctrine.²

The Emergence of Sunnism

Sunnism began in the eighth and ninth centuries as a broad, populist movement that explicitly—almost cynically—rejected any special religious role for the caliph. It grew out of popular revulsion with the fighting and elite political intrigue that took place during the *Fitnas*—conflicts that in their eyes directly contradicted the idea that the *umma* should be unified in imitation of God’s *tawhid*, or divine oneness. The movement subsequently gained strength as a result of growing dissatisfaction with Umayyad rule. The dynasty’s embrace of pre-Islamic political practices, its pursuit of luxury, and its incorporation of Byzantine legal practices nurtured the sense that the caliphate was an intrinsically worldly institution that was ill suited to the articulation of the social laws and customs that would structure life under Muslim rule. For a brief period following the demise of the Umayyads, Sunnis hoped that the Abbasids might govern as devoutly as the first four caliphs and thus be able to provide the guidance they sought; they quickly abandoned that expectation, however, when the first few leaders of the new dynasty made clear that they were every bit as focused on earthly concerns as the Umayyads had been. By the latter half of the eighth century, as a result, the Sunnis had abandoned the expectation that the caliph would establish the type of community Muhammad had called for. Instead, they would create that society themselves based on the explicit commandments outlined in the Qur’an and on the sunna of Muhammad.³

Sunni Political Theory

What role, then, did the early Sunnis see the caliphate playing? In terms of defining the obligations and responsibilities of the community, very little. Indeed, the ‘ulama’ accorded the state almost no power in shaping the legal codes by which the *umma* lived, and limited its social and religious responsibilities to collecting the zakat tax and punishing those who broke religious laws. Sunnism’s opposition to the caliphs having a religious role is most obvious in the popularization of the term *Rashidun*, or rightly guided, to describe the first four caliphs. People used that word as a sign of

² Michael Lipka and Conrad Hackett, “Why Muslims Are the World’s Fastest-Growing Religious Group,” *Pew Research Center* (blog), accessed November 19, 2019, <https://www.pewresearch.org/fact-tank/2017/04/06/why-muslims-are-the-worlds-fastest-growing-religious-group/>; Armstrong, *Islam*, 6, 57–64.

³ Albert Hourani, *A History of the Arab Peoples* (Cambridge, Mass: Belknap Press of Harvard University Press, 2002), 60–61.

reverence and respect for those influential leaders, of course, but they also employed it as a potent rhetorical device aimed at denying later rulers the power to intrude in the social and religious sphere. That is, the first four caliphs were Rashidun because they had been companions of Muhammad and, as such, had the authority to establish the customs, legal practices, and obligations that were binding on the *umma*. Later caliphs, in contrast, were not Rashidun and thus enjoyed no corresponding influence over religion or daily life.⁴

The caliph did have one very important religious responsibility, however: the maintenance of order. To Sunnis, political stability was essential. Indeed, it was so paramount, that a popular Sunni proverb declared “better sixty years of tyranny than a single day of civil strife.” Why was order so important to them? They stressed its importance because they feared that, in its absence, people would violate the legal and social requirements that lay at the heart of Muhammad’s vision of the proper community of believers. As the scholar Vali Nasr explains, “[t]he Sunni conception of authority has centered on a preoccupation with order. Religion does not depend on the quality of political authority but only on its ability to help the faith survive and grow.” Accordingly, by the late-eighth century, Sunni ‘ulama’ were glad to uphold the legitimacy of the caliph so long as he remained—at least outwardly—a devout Muslim and, more importantly, so long as he preserved the order needed for the *umma* to function as Muhammed had intended.⁵

Perhaps surprisingly in light of the determination of the Sunni ‘ulama’ to prevent the caliph from having a meaningful voice in religious affairs, the Abbasid Dynasty would—after a time—throw its support behind them. It had three reasons for doing so. First, the nascent Sunni movement that the religious scholars led enjoyed popular backing; supporting the ‘ulama’ would thus pay dividends for the Abbasids by helping them to secure broad approval of the dynasty’s rule. Second, Sunni emphasis on order and stability appealed to the Abbasids because it suggested that they could rule autocratically without risking the emergence of a populist rebellion similar to the one that had toppled the Umayyads. Finally, the Sunni ‘ulama’ seemed well positioned to solve a vexing problem that had emerged in the first century of Islam. During the caliphate’s rapid expansion, the different regions, towns, and cities of the empire had haphazardly developed competing and even contradictory Islamic customs and laws drawn as much from existing Roman or Persian legal codes as from the Qur’an or the practices of Muhammad. Not only was the existence of multiple legal systems inconsistent with either *tawhid* or the religion’s universal nature, but, more practically, it also profoundly complicated the caliphate’s effort to administer justice. As a result, the

⁴ Armstrong, *Islam*, 61.

⁵ Vali Nasr, *The Shia Revival: How Conflicts within Islam Will Shape the Future* (New York: W. W. Norton, 2006), 36–39.

dynasty supported the ‘ulama’ and encouraged them in their efforts to devise a single legal and social code, shari‘a law, by which all Muslims could live.⁶

Shari‘a Law

With its emphasis on correct living, shari‘a law became the beating heart of Sunni Islam as it took shape over the course of the eighth century. Intended to establish the ethical and legal standards under which Sunni Muslims were to live, it mandated certain activities while regulating or outright prohibiting others. Thus, it required Muslims to pray daily and to pay the *zakat*, stipulated inheritance and divorce rights, and forbade sex outside of marriage and the consumption of pork or alcohol.⁷

Its formulation proved lengthy and contentious. The problem was one of sources. When Muhammad was alive, ascertaining how believers should act or crafting appropriate laws to handle new situations had been simple: one merely asked the Prophet for his opinion on the subject. After his death and the passing of his companions, however, the community needed to find a different way to determine the traditions and laws by which the *umma* should abide. The Qur’an provided the obvious starting point and would go on to serve as the primary source for the shari‘a code. It included some eighty explicitly legal verses that addressed issues such as marriage, inheritance, and female infanticide. It also established a series of moral precepts that could provide the religious scholars with critical guidance for the construction of Islamic law. The ‘ulama’ of the eighth century consequently set about systematically interpreting the Qur’an for the purpose of refining the shari‘a code. In the process, they won popular backing for a particular understanding of the Qur’an that held that it was not a living document or the product of a particular time and place—the Hijaz of the early seventh century—but that it was instead the timeless, literal word of God and thus beyond interpretation. Gradually, moreover, they also secured acceptance of the idea that the Qur’an was a divine attribute of God, and that it was, as such, eternal and uncreated; like God, in other words, it had forever existed and was unchanging in form, and its teachings—as interpreted by the religious scholars—could not be questioned.⁸

At the same time, the Qur’an quickly revealed itself to be of limited utility in the effort to craft an Islamic legal and social code. Focused largely on establishing the broad moral principles by which Muslims were to live rather than on providing explicit legal rulings or guidance, it offered only minimal assistance for resolving many pressing questions of behavior and law. As a result, the ‘ulama’ soon found themselves stymied in their effort to construct a comprehensive shari‘a code.⁹

⁶ John L. Esposito, *Islam: The Straight Path*, Fourth Edition (New York: Oxford University Press, 2011), 76; Armstrong, *Islam*, 58.

⁷ Esposito, *Islam*, 106.

⁸ Reza Aslan, *No God but God: The Origins, Evolution, and Future of Islam*, Updated Edition (Random House, 2011), 163–65.

⁹ Esposito, *Islam*, 77–78.

The legal scholar Malik ibn Anas (711-795) was the first to suggest a way out of this situation. He argued that when the Qur'an did not provide explicit direction on an issue, the 'ulama' should look to the practices and laws of the city in which the Prophet's community first existed: Medina. He subsequently founded a legal school in that city, the Maliki School, and published the *Muwatta*, in which he collected the customs, laws, and traditions of contemporary Medina that he believed dated to the time of Muhammad. Malik's approach would prove highly influential in the development of Islamic *fiqh*, or jurisprudence. At the same time, however, doubts soon emerged regarding his reliance on Medina for the construction of Muslim law and custom. As critics asked, was it not possible that the traditions and laws of Medina included accretions and innovations made in the century after Muhammad's death? If so, contemporary Medina could not serve as a model for Islamic behavior and law. These objections proved persuasive. In the end, they convinced most 'ulama' that the practices of contemporary Medina could not serve as the basis for shari'a law. Thus, despite Malik's important contributions, Islamic jurisprudence appeared once again to be stuck.¹⁰

One of Malik's students, Muhammad ibn Idris al-Shafii (767-820) provided a way out of the impasse. The founder of the Shafii School of jurisprudence, he argued that shari'a law had four legitimate sources. The first two were direct ones: the Qur'an and the hadith—the documented stories of the sayings, actions, and teachings, or sunna, of Muhammad. The other two, *qiyas*, or analogical reasoning, and *ijma*, or consensus of the Muslim community, were indirect interpretive methods derived from the Qur'an and the sunna that he argued could be used to determine the law in areas where the direct sources were silent.¹¹

Like Malik, Al-Shafii argued that the Qur'an was the revealed word of God and should thus be the first source to which the religious scholars should look in the creation of shari'a law. When it was silent on an issue or when its dictates needed clarification, however, he contended that they should seek guidance not in the customs of Medina but instead in the sunna of Muhammad as recorded in the hadith. He maintained that since the Prophet was the archetype of the perfect Muslim, the documented stories of his life constituted—in law professor David Forte's words—"indirect divine revelation" that outlined the kind of society in which God wanted humanity to live.¹²

Al-Shafii's argument was persuasive, and soon won broad support among the 'ulama'. The hadith provided answers to many questions on which the Qur'an was silent, and proved pivotal in breathing new life into the effort to develop a comprehensive religious and social code. At the same time, however, the hadith were surprisingly problematic as a source for Islamic law. The primary issue was the sheer quantity of them. By the time that scholars such as Muslim ibn al-Hajjaj (815-875) began

¹⁰ Esposito, 77–78.

¹¹ Hourani, *A History of the Arab Peoples*, 67–68.

¹² Forte quote from David F Forte, "Islamic Law: The Impact of Joseph Schacht," *Loyola of Los Angeles International and Comparative Law Review* 1 (1978): 4.

compiling hadith in the ninth century, their number had exploded into the hundreds of thousands. Making matters worse, the hadith frequently contradicted one another. While religious scholars agreed that many were valid, they also believed that large numbers were forgeries fabricated by rulers or other individuals to advance their own narrow interests. How, then, could the ‘ulama’ use them to devise the shari‘a code? That is, how could they separate the legitimate hadith from the many forgeries?¹³

Here, too, al-Shafii provided a path forward. He developed a rigorous approach called Hadith Scholarship by which the ‘ulama’ could assess the validity of the hadith using techniques similar to the methods that modern-day historians use to critically assess the credibility of sources. Al-Shafii’s approach called on the ‘ulama’ to examine carefully the *isnad*, or chain of people who transmitted the story from the time of Muhammad to the ninth century. Valid hadith needed to be traced through a succession of people of unimpeachable moral character back to a person who knew the prophet personally. If an individual in an *isnad* could not have met the person from whom they purportedly heard a hadith because they lived at different times or in different places, then that hadith could not be authenticated. Likewise, if a person in the *isnad* was of poor moral reputation, then the hadith could not be trusted. This process was reputed to be quite strict. One story holds that the religious scholar Ismail al-Bukhari (810-870) rejected a hadith because he discovered that one of the transmitters beat his horse; to al-Bukhari, a person who treated his mount in that way could not possibly possess the ethical and moral character of a reliable transmitter.¹⁴

When both the Qur’an and the hadith failed to clarify a legal or ethical question, al-Shafii devised two indirect interpretive methods that the ‘ulama’ could use. The first technique, *qiyas*, or analogical reasoning, allowed the ‘ulama’ to extend the precepts of the Qur’an and the sunna to subjects that those sources did not explicitly address. For example, religious scholars used *qiyas* to forbid the consumption of drugs such as marijuana that were unknown to the early community of believers and that Muhammad consequently never discussed. They argued that if alcohol was forbidden because it was an intoxicant, then other intoxicants—including marijuana—should likewise be prohibited. The second indirect method that al-Shafii articulated was *ijma*, or consensus. Based on Muhammad’s declaration that “my community will never agree on an error,” *ijma* held that any decision that a substantial majority of the religious scholars agreed to was binding as Islamic law. Ironically, *ijma* permitted the ‘ulama’ to validate legal decisions—such as one that called for adulterers and prostitutes to be stoned to death—even when those decisions appeared to contradict to the Qur’an.¹⁵

For al-Shafii, the only appropriate sources for shari‘a law were the Qur’an, the hadith, *ijma*, and *qiyas*. In contrast, other religious scholars in the ninth and tenth

¹³ Esposito, *Islam*, 95–96.

¹⁴ Tamim Ansary, *Destiny Disrupted: A History of the World Through Islamic Eyes* (New York, NY: PublicAffairs, 2009), 95–96; Tom Holland, *In the Shadow of the Sword: The Birth of Islam and the Rise of the Global Arab Empire* (New York: Doubleday, 2012), 35.

¹⁵ Aslan, *No God but God (Updated Edition)*, 166–67.

centuries argued that the ‘ulama’ could use a third, indirect, interpretive method when those sources failed to provide a clear answer: *ijtihad*, or independent reasoning. Supporters of *ijtihad* argued that a well-respected ‘ulama’ could render a judgement about an unprecedented situation based on his understanding of Islamic law; if, over time, most other religious scholars supported his reasoning via *ijma*, then that decision would acquire legal precedence and thus permanence. For a time, *ijtihad* enjoyed widespread support among the ‘ulama’ and constituted an important source of shari‘a law. As we shall see, however, enthusiasm for its use would soon fade.¹⁶

Using al-Shafii’s approach, four Sunni schools of jurisprudence developed in the ninth century to articulate shari‘a law and to erect the theological foundation on which it rests. Along with the aforementioned Maliki and Shafii Schools, there emerged the Hanifa School founded by Abu Hanifa (699-767) and the Hanbali School established by Ahmad ibn Hanbal (780-855). Each of these schools adopted approaches to the formulation of the shari‘a code that differed in terms of the sources that they emphasized. The Shafii School stressed the importance of the sunna, for example, while the Maliki School drew most heavily on the customs and practices of the people of Medina. Regardless of their differing views on the appropriate sources, however, the schools developed remarkably similar codes. Indeed, they differed more in terms of tone than in the substance of their legal rulings. The four schools thus enjoyed equal status and authority; that is, the sum of their rulings collectively constituted the core of shari‘a law, and all of them gained widespread support among Sunni Muslims.¹⁷

Importantly, the ‘ulama’ presented the code that they had developed as one that was not subject to debate or challenge. As they argued, it was founded not on fallible, mortal faculties nor was it developed incrementally in response to the specific circumstances of a given place or historical period. Instead, they maintained that the shari‘a law was a direct and indirect extension of two perfect sources, the Qur’an and the sunna of Muhammad—the Qur’an because it was the timeless, literal word of God, and the sunna because God Himself had chosen the Prophet to serve as his final messenger. Thus, the religious scholars declared, the shari‘a law that they had developed was “divine in origin” and, as such, beyond the capacity of humans to question or dispute.¹⁸

The Historical Debate: Hadith Scholarship

The preceding understanding of the code’s origins dominated the Muslim world without serious question for over a millennium. In recent decades, however, scholars have mounted a sustained challenge to it. Central to their critique have been doubts about the validity of the hadith from which the ‘ulama’ had constructed the sunna. Scholars have raised two main objections to their reliability. First, they have persuasively demonstrated that many of the laws ostensibly derived from the hadith in

¹⁶ Aslan, 167.

¹⁷ Armstrong, *Islam*, 65; Aslan, *No God but God (Updated Edition)*, 167–68.

¹⁸ Forte, “Islamic Law: The Impact of Joseph Schacht,” 4–8.

fact predate Islam and originated in the Talmudic, Roman, or Persian legal traditions—meaning that the code itself reveals that it had roots that extended well beyond the sunna. Second, scholars have uncovered evidence that indicates that many of the *isnad* that the ‘ulama’ had analyzed with such apparent rigor almost certainly did not extend back to the time of the Prophet. For example, Robert Hoyland cites the case of a legal scholar of the mid-eighth century who wrote that a prominent older ‘ulama’ he knew had never uttered a single hadith over the course of his life while “the young men round here are saying [them] twenty times an hour”—a statement that strongly suggests that the articulation of the hadith began well after the time of Muhammad.¹⁹

These findings offer insights into the goals that motivated the religious scholars to engage in the effort to link the shari‘a code to the practices of the early *umma* through Hadith Scholarship. As the religious scholar Reza Aslan has shown, while the sunna on which the shari‘a code is based may correctly reproduce some of the customs and practices of the early community of believers, any instances of accuracy on the part of the ‘ulama’ are more a function of accident than intent. This is so, he continues, because the real purpose of Hadith Scholarship was not to separate authentic hadith from the chaff but instead, through the manufacture of hadith and *isnads*, to connect ninth-century legal and religious views to the early *umma* and, in doing so, to give sanction to ‘ulama’-approved laws and customs that either pre- or post-dated Islam. In other words, Hadith Scholarship was a tool used to create a past consistent with the present rather than the reverse. Tellingly, the hadith that appeared to enjoy the strongest claims to legitimacy were the least reliable. As the twentieth-century German scholar Joseph Schacht succinctly put it, “the more perfect the *isnad*, the later the tradition.” Regardless, the religious scholars’ campaign worked. Their effort to give the existing laws and customs that they endorsed the imprimatur of Islamic precedence gained broad acceptance, and their control of the past through Hadith Scholarship came to form a central plank in their claim to hold sole authority over shari‘a law.²⁰

Aslan adds that their ability to win adherence to their understanding of the Qur’an constituted a second critical element of that claim. As previously noted, the ‘ulama’ had succeeded in gaining broad approval both for their interpretation of the Qur’an and for their assertion that it was the literal, uncreated, eternal word of God. That is, they had managed to simultaneously advance a particular understanding of the Qur’an and to assert that it was beyond interpretation—thereby establishing their view as the only valid one. In Aslan’s view, the religious scholars’ success in institutionalizing their interpretation of the Qur’an proved pivotal in their effort to cement their dominance of the religion of Islam and the shari‘a code. Coupled with their monopoly over Hadith

¹⁹ Robert G. Hoyland, *In God’s Path: The Arab Conquests and the Creation of an Islamic Empire* (New York: Oxford University Press, 2015), 137.

²⁰ Aslan quote from Aslan, *No God but God (Updated Edition)*, 166; Schacht quote from Joseph Schacht, “A Revaluation of Islamic Traditions,” *Journal of the Royal Asiatic Society of Great Britain and Ireland* 2 (1949): 149.

Scholarship, it ensured that they controlled the main sources of Islamic law and afforded them a degree of authority over the shari‘a code that even the caliph had to respect.²¹

As modern-day scholars point out, their control over the sources of the shari‘a code would give Sunnism a conservative, backward-looking character. That is, by limiting the law directly or indirectly to the Qur’an and the life of Muhammad, the religious scholars ensured that the main venue for legal and social debate would not be the present day or the use of reason, but instead the interpretation of the Salaf—the first three generations of Muslims—which the ‘ulama’ had successfully assumed control over through Hadith Scholarship. In other words, in the process of creating the shari‘a code, the legal scholars had also institutionalized the view that the Sunni community should address challenges not by developing new practices or institutions, but instead by doing the precise opposite: eliminating accretions and innovations—*bid‘ah*—and returning to the customs, however ahistorical, of the Salaf. This framework would prove enormously influential. Indeed, as we shall see later, it would shape many of the ways in which Sunnis would respond to difficult challenges such as the Mongol conquest in the thirteenth century and the growing power of the West in the nineteenth and twentieth centuries.²²

Shi‘ism

Not all Muslims have shared this emphasis on the first few generations of believers. Instead, in the late-seventh century, a minority of Muslims known as Shi‘a began to outline a very different understanding of the religion that deemphasized the Salaf in favor of a focus on the descendants of Muhammad through the line of his daughter Fatima (605-632) and his son-in-law Ali ibn-Talib (r. 656-661). In the process, they developed a very different set of rituals, articulated a distinct interpretation of Islam’s nature, and developed a radically contrasting set of beliefs regarding who should rule and the extent of their religious authority than did the Sunni majority. Though their differences with the Sunni majority meant that the sect faced frequent persecution, Shi‘ism has nonetheless persevered. Indeed, according to a 2009 Pew Forum study, the world’s 154 million to 200 million Shi‘a account for 10-13 percent of the global Muslim population.²³

Origins

Shi‘ism originated in a political dispute that went all the way back to Islam’s earliest days. As discussed in chapter three, a crisis had emerged following Muhammad’s death regarding who should succeed him. With the Prophet having died without first explicitly naming a successor, the *umma* quickly divided into factions over

²¹ Aslan, *No God but God (Updated Edition)*, 163–64, 169–72.

²² Ansary, *Destiny Disrupted*, 98–99.

²³ “Mapping the Global Muslim Population,” *Pew Research Center’s Religion & Public Life Project* (blog), 2009, <https://www.pewforum.org/2009/10/07/mapping-the-global-muslim-population/>.

the question of who should replace him. One of those groups, the Alids—also known as the *Shi'atu Ali*, or Party of Ali—argued that Muhammad had in fact made clear that he wanted his son-in-law, Ali, to succeed him. Other prominent Muslims balked at the idea, however. Believing that Ali was too young, they instead chose Muhammad's father-in-law, Abu Bakr (r. 632-634), as the Prophet's successor. Grudgingly Ali and his backers acceded to the decision. Later, and with greater reluctance, they also accepted the decision of the leading Muslims to pass over Ali in favor of Umar (r. 634-644) in 634 and Uthman (r. 644-656) a decade later.²⁴

Finally, following Uthman's murder in 656, Ali at last became caliph. As we have seen, his tenure as leader of the *umma* was neither happy nor successful. Ali faced significant rebellions, including one led by 'A'isha (c. 613-678) and two of the Prophet's former companions and an even more formidable one headed by the powerful governor of Syria, Muawiyah ibn Abi Sufyan (r. 661-680). He also had to deal with an uprising by militant former supporters known as Kharijites, or "seceders," who had broken with him over his willingness to enter into arbitration with Muawiyah. Ali's short and difficult reign ended in 661 when a Kharijite assassin murdered him in the Mosque in Kufa using a poisoned sword. With Ali dead, leadership of the *Shi'atu Ali* fell to his eldest son, Hasan ibn Ali (624-670). A deeply religious man, Hasan had been horrified by the fighting that had occurred between Muslims during the First *Fitnah*. Eager to avoid a continuation of the civil war, he proposed a settlement wherein he would accept Muawiyah as the new caliph on the condition that a *shura* would choose his successor. With Muawiyah agreeing to abide by this proposal, Hasan retired to Medina where he died in 670. Following his death, his brother, Husayn ibn Ali (626-680), succeeded him as the leader of the *Shia' tu Ali*.²⁵

Ultimately, Muawiyah did not stand by his agreement with Hasan. Instead of calling for a *shura* to determine who should succeed him, he arranged during his final months as caliph for his son, Yazid (r. 680-683), to take power—thereby establishing the Umayyad Dynasty. This action upset many Muslims, but it particularly infuriated the *Shia' tu Ali*. Their complaint about Yazid's accession was distinct from that of other believers. Unlike most Muslims, who objected on the grounds that dynastic rule was antithetical to Islam, the Alids had no issue with hereditary succession. Instead, their complaint was that the wrong family had assumed control of the caliphate. That is, rather than having the leadership of the *umma* descend through Banu Umayya, they argued that it should follow the descendants of Muhammad through the line of Fatima and Ali.²⁶

Furious at Muawiyah's betrayal, Husayn moved to challenge Yazid shortly after the latter acceded to power. Launching a rebellion was a tall order for the prophet's grandson, however. Living in distant Medina, he was isolated from the main Muslim population centers and was thus in a poor position to organize an effective revolt. Iraq,

²⁴ Ansary, *Destiny Disrupted*, 35–52.

²⁵ Aslan, *No God but God (Updated Edition)*, 176–77.

²⁶ Ansary, *Destiny Disrupted*, 67–68.

however, had large numbers of Alids whose backing could possibly give him the numbers needed to triumph over the Umayyads. Accordingly, after he received a message from troops in the pro-Alid garrison town of Kufa expressing their support and inviting him to come and take command, he and a small band of supporters departed for Iraq in the late summer of 680. It was a risky move. Yazid would surely attempt to intercept him before he reached Kufa. Accordingly, Husayn instructed his Kufan supporters to march out to meet him when he neared Iraq. They failed to do so, however. While Husayn was traveling north, the powerful Umayyad governor of Iraq, Ubayd Allah (?-686), brought his army to Kufa where he intimidated the Alids into inaction through a show of force. As a result, when Husayn finally arrive in Iraq, he met not the Kufan supporters he expected but instead Ubayd Allah's large and powerful army. Unsurprisingly, the fight that followed was one sided. On October 10, 680, the Iraqi governor's army slaughtered Husayn and his seventy-two companions at the Battle of Karbala—thereby defeating the last serious Alid effort to secure control of the caliphate.²⁷

By all reasonable expectations, Karbala should have been the end the Alids; instead, the battle marked the point at which the *Shia 'tu Ali* transformed from a political faction into a new religious sect: Shi'ism. This process started quietly. In 684, penitents from Kufa assembled in Karbala on the anniversary of the battle to mourn the death of Husayn and to seek atonement for their failure to help him. In doing so, they turned what had been a political and military defeat into a noble act of moral suffering and righteousness. According to this new interpretation of the battle, Husayn had understood clearly that he faced certain death if he continued into Iraq; nonetheless, he pushed on and sacrificed himself as an act of devotion to God. In this way, the early Shi'a remade the defeat at Karbala into what the scholar Vali Nasr characterizes as the central paradigm of Shi'ism: "the triumph of moral principles over brute force."²⁸

Shi'i Rituals and Theology

In keeping with this idea, the nascent sect rapidly developed a set of beliefs that centered on the themes of atonement, martyrdom, and sacrifice but also, for those who kept the faith through hard times, redemption. This emphasis on loss and penance reflects the Alids' bitter experiences in the first few decades of the caliphate. While the Sunnis could point to a string of successes—Muhammad's triumph over Mecca, the defeat of the Byzantines, the conquest of the Persian Empire, triumph over the Alids—the Shi'a could instead look back only to a series of defeats and setbacks. As a result, they placed less emphasis than the Sunnis on right behavior and strict observance of Islamic law and instead stressed the creation of emotional rituals through which they could simultaneously express their devotion to God and seek forgiveness for their failure to aid Husayn in his hour of need. In reenactments of the Battle of Karbala staged during the holiday of 'Ashura', for example, penitent Shi'a engage in acts of self-

²⁷ Fred Donner, *Muhammad and the Believers: At the Origins of Islam* (Cambridge, MA: The Belknap Press of Harvard University Press, 2012), 177–80.

²⁸ Nasr, *The Shia Revival*, 46–49.

flagellation such as whipping their backs with chains in order to demonstrate their faith and to atone for the community's failure in 680. Similar rituals occur at the shrine of Husayn in Karbala during Arbaeen, the holiday that marks the end of the traditional forty-day period of mourning following 'Ashura'.²⁹

The Imamate

The differences between Sunnism and Shi'ism went well beyond the latter's embrace of ritual, however. Indeed, the most obvious point of disagreement between the two sects was over the question of political leadership. As we have seen, Sunnis argued that any outwardly devout male member of the Quraysh tribe could serve as caliph so long as he maintained the peace. The Shi'a disagreed. After Karbala, they elaborated an increasingly complex political-religious doctrine that held that leadership of the Muslim community rightly belonged to the Imams. To Sunni Muslims, the term imam, or leader, refers to the people who direct the communal prayers in the mosques. To the Shi'a, in contrast, it is a proper noun—a term of great respect reserved only for the line of male descendants of Muhammad through Ali and Fatima, who, they believed, God had ordained to rule. Thus, to the Shi'a, Ali should have become caliph immediately following Muhammad's death, and, thereafter, the succession should have remained confined to his descendants with Fatima.³⁰

Central to Shi'ism then and now is the belief that the Imams are more than just political leaders. To the Shi'a, they are instead divinely appointed figures endowed with the same authority and power as Muhammad. Most notably, Shi'ism holds that Ali had inherited from the Prophet his *baraka*—his special power to see beyond the Qur'an's literal, surface meaning and to understand and interpret its more profound, hidden, esoteric message. This belief had important ramifications. In contrast to the Sunni view that the age of revelation had ended with the Prophet's death and that the Qur'an was to be understood literally, it led the Shi'a to maintain that the Muslim community continued to receive direct, divine guidance—and even new revelation—through the Imams' continued interpretation of God's message. It also led the Shi'a to claim that the Imams had a special role to play in helping others to achieve salvation. This position emerged in the ninth century when Shi'i religious scholars began to argue that ordinary believers lacked the ability to understand the hidden message of the Qur'an and were incapable of securing God's grace on their own. God had not forsaken them, however. Instead, He had given the Shi'i saints—Muhammad, Fatima, and the line of infallible Imams—the power to intercede in order to help people attain salvation.³¹

Unsurprisingly, the Shi'a did not approve of the government that ruled the community of believers. On the contrary, they were unsparing in their opposition to the caliphate. They did not characterize the first three caliphs as Rashidun, and they outright denied the legitimacy of both the Umayyad and Abbasid Dynasties. Save for Ali

²⁹ Nasr, 18–19, 43–49.

³⁰ Aslan, *No God but God (Updated Edition)*, 185–86.

³¹ Ansary, *Destiny Disrupted*, 71; Nasr, *The Shia Revival*, 38–39.

and Muhammad, in fact, they believed that *all* of the leaders of the *umma* had been usurpers. For a time, they translated this sentiment into action. During the eighth and ninth centuries, the Shi'a organized numerous opposition movements and rebellions in Iraq aimed at restoring the line of Imams to power. Predictably, all of those revolts failed in the face of the Abbasid Dynasty's enormous power. Indeed, the caliphate's response to their subversive activities was so overwhelming that it compelled the Shi'a to abandon militant opposition at the end of the ninth century. Thereafter, they responded to the caliphate's ongoing efforts to suppress them by adopting a combination of quietism, meaning the acceptance of the existing order, and *taqiyya*, or dissimulation, the practice of concealing their religious identity to avoid persecution.³²

Competing Sects: Zaidism, Ismailism, and Twelver Shi'ism

Despite their common veneration of the Imams and their shared experience of being regularly persecuted by the caliphate, the Shi'a split into competing variants over the course of the eighth century. They did so in part because Shi'ism experienced a regular event that provided the opportunity for its adherents to factionalize: the succession to the Imamate. The death of each Imam raised the question of which son or relative should follow him as leader of the Shi'i community. In most cases, a consensus existed regarding who should succeed to the Imamate. Twice during the eighth century, however, a significant minority broke with their fellow Shi'i regarding the succession. As a result, Shi'ism divided into three main variants: Zaidism, also known as Fiver Shi'ism because of its belief that the line of infallible leaders ended with the Fifth Imam, Ismailism, or Sevener Shi'ism, so named because of its view that Ismail ibn Jafar (c. 722-c. 762) was the rightful seventh Imam, and Ithna Asharism, or Twelver Shi'ism, so called because it holds that there have been twelve Imams.

The Zaydis were the first group to split with the majority of Shi'a. Following the death of the fourth Imam, Ali Zayn al-Abadin (c. 659-713), most Shi'a recognized his son Muhammad al-Baqir (677-733) as the Fifth Imam. However, a small minority instead claimed that Muhammad's brother, Zayd ash-Shahid (c. 698-740), was the rightful Fifth Imam, and established a separate Shi'i sect called Zaydism. Never accounting for more than a small fraction of the global Muslim population, Zaydis differed from other Shi'i not merely over the correct succession of infallible Imams but also over how they should react to a hostile political order that had, in their eyes, usurped the Imams' rightful position as the head of the *umma*. While most Shi'a responded to Shi'ism's precarious position in the Sunni-dominated caliphate by practicing political quietism and *taqiyya*, the Zaydis instead followed the Kharijites in arguing that the Qur'anic commandment to "enjoin what is good and forbid what is evil" required them to rebel against unjust rule. In keeping with this view, the Zaydis spearheaded several revolts against both the Umayyad and Abbasid Dynasties in the eighth and ninth centuries. Predictably, given the power imbalance between them and the caliphal government, none of these rebellions came anywhere close to succeeding; instead, they only served to leave the Zaydis weak and vulnerable. The Zaydis were

³² Esposito, *Islam*, 45–46.

resilient, however, and managed to stage a comeback in the late ninth century. Taking advantage of disarray in the Abbasid court, they succeeded in establishing a long-lasting state in distant Yemen in 897. It survived in various forms until a coup overthrew it in 1962.³³

Another small sect, the Ismailis, or Sevener Shi'a, also promoted a militant response to what they argued was an unjust political order. Like Zaydism, Ismailism emerged in the mid- to late-eighth century as a result of a dispute over the succession to the Imamate. In this case, the issue was who should follow the Sixth Imam, Jafar al-Sadiq (702-765) as the leader of the Shi'i community. Jafar had already answered this question by naming his elder son, Ismail ibn Jafar, to succeed him as Imam. However, Ismail had predeceased his father. In response, the majority of Shi'a acknowledged Ismail's younger brother, Musa al-Kazim (745-799), as the Seventh Imam. This decision did not enjoy unanimous support within the Shi'i community, however. On the contrary, a minority vigorously disagreed and argued that Ismail had to be the rightful Seventh Imam; after all, Jafar was infallible, and, as such, could not have made a mistake in choosing his successor. This reasoning was consistent with the belief that the Imams are divinely appointed figures endowed with unerring judgement, but it raised a difficult question: how could one of the infallible Imams have made a decision that seemed to have left the faithful without a leader?³⁴

For most Ismailis, the answer was simple. Arguing that the Imamate continued through Ismail's lineage, they acknowledge one of his sons, Muhammad ibn Ismail (740-813) as the new Imam. For a small minority of Ismailis later known as Qarmatians, however, the apparent paradox inherent in Jafar al-Sadiq's decision to choose a successor who predeceased him was deeply troubling. They soon arrived at a novel explanation. Ismail had not died, they maintained, but had instead gone into Occultation, or spiritual hiding, as the Hidden Imam, and would return at the end of days as the *Mahdi*, or messiah, to restore God's rule for all time.³⁵

Among the Shi'a, the Ismailis proved most effective at opposing the Abbasids. As we shall see in more detail in chapter five, an Ismaili group seized control of modern-day Tunisia in 909 and proclaimed the establishment of a dynastic rival to Abbasid rule known as the Fatimid Caliphate. At first, it posed little threat to the government in Baghdad. Far from the center of the empire, it was no more than a nuisance to the Abbasid Dynasty during its first six decades. However, the Fatimid Caliphate's seizure of Egypt and its agricultural bounty in 969 almost instantly transformed it from a minor threat into a serious rival that, for a time, constituted an existential challenge to the

³³ Esposito, 46.

³⁴ Aslan, *No God but God (Updated Edition)*, 188–89.

³⁵ Esposito, *Islam*, 48–51.

Abbasids. Indeed, for much of the tenth through twelfth centuries, the Fatimids enjoyed more power and influence in the Muslim world than did the Abbasids.³⁶

Even following the collapse of the Fatimid Caliphate in 1171, the Ismailis continued to constitute a serious challenge to Abbasid rule. The new Ismaili threat came in the form of the Nizari Ismailis, or, as their enemies called them, the Assassins, who had broken away from the Fatimid Caliphate in the late-eleventh century over another disputed succession. Establishing a chain of nearly impregnable fortresses in modern-day Iran, Iraq, and Syria—including the famous Alamut Castle—they waged highly effective, low-intensity, terror campaigns against the Abbasids, Seljuq Turks, and European Crusaders during the twelfth and thirteenth centuries. These operations were so successful, in fact, that the Nizaris managed to maintain an independent, non-contiguous state that, though surrounded by enemies, thrived for nearly two centuries. Indeed, only with the coming of the mighty Mongols to the Middle East in the 1250s were the powerful Assassins finally brought to heel.³⁷

The Ismailis later evolved into a very different sect. Having long since abandoned violence, Ismailism's 15 million adherents now live peacefully under the leadership of the 49th Imam, the wealthy and influential Karim Aga Khan (r. 1957-). As they have for centuries, most reside in Iran, Central Asia, and the Indian subcontinent. However, many Ismaili merchants took advantage of their membership in the British Empire to pursue opportunities in other colonies. As a result, Ismaili communities exist today in Africa, the Middle East, North America, Asia, and Europe.³⁸

As noted earlier, adherents of the largest variant of Shi'ism, Ithna Ishari, or Twelver Shi'ism, did not acknowledge Ismail as the seventh Imam. Instead, they accepted his brother Musa as Jafar's rightful successor to the imamate. However, they, too, eventually ran into a theological conundrum. In their case, the problem stemmed from the disappearance of the twelfth Imam, Muhammad al-Mahdi (870-?). His disappearance before he could designate an heir compelled them to confront the same essential dilemma that the Qarmatians had faced at the time of Jafar's death: how could they reconcile their belief that the Imams were infallible with the apparent end of the Imamate? The answer came through the appropriation and elaboration of the Qarmatian concept of Occultation. Twelver 'ulama' started claiming that Muhammad al-Mahdi was not dead but was instead in Occultation as the Hidden Imam. Eventually, they continued, the Mahdi, meaning "the expected one," would reappear as the Messiah and bring a period of peace and justice to the earth until Jesus returned and the world ended. In the meantime, the Shi'a 'ulama' would direct the spiritual and legal life of the Twelver community as the representatives of the Hidden Imam. Today, the vast majority of Shi'a, about 85 percent, are Twelvers. They constitute a minority of Arabs living

³⁶ Chris Wickham, *The Inheritance of Rome: Illuminating the Dark Ages 400-1000* (New York: Penguin Books, 2009), 336–38.

³⁷ Ansary, *Destiny Disrupted*, 129–32, 155.

³⁸ Nasr, *The Shia Revival*, 75–77.

along the Gulf Coast, a substantial plurality of the Arabs of Iraq, and an overwhelming majority of Iranians.³⁹

Similarities and Differences Between Shi'ism and Sunnism

In a number of ways, the beliefs of Shi'i sects such as the Twelvers stand in sharp contrast to the tenets of Sunnism. Four differences are particularly stark. First, Sunnism vigorously maintains that no intermediaries—human or divine—stand between Muslims and God. Shi'a, in sharp contrast, hold that believers can be redeemed only through the intercession of the Shi'a saints. Second, while Sunnism has no religious establishment and argues that all adherents are equally capable of understanding the beliefs of Islam, Shi'ism, especially its Twelver variant, instead has a formal priesthood that establishes and upholds religious doctrine and that claims to represent the Hidden Imam until his return. Third, Shi'a adhere to their own version of the shari'a law, the Jafari School, established in the mid-seventh century by the Sixth Imam, Jafar al-Sadiq. Moreover, in contrast to the Sunni legal tradition's rejection of independent reasoning in the thirteenth century—a topic we shall cover presently—Shi'a 'ulama' continue to view *ijtihad* as a legitimate tool for rendering legal decisions. Finally, and most importantly, the early Sunnis and Shi'a split dramatically over the question who should rule and what powers they should wield. As we have seen, Sunnis asserted that the empire should be headed by a fundamentally secular caliph whose primary responsibility was the maintenance of order; Shi'a, in contrast, claimed that the infallible Imams who descended from Muhammad through the line of Fatima and Ali were the rightful leaders of the *umma* and that they exercised both political and religious authority.⁴⁰

It is easy to overstate the differences between Sunnism and Shi'ism, however. In many ways, in fact, their contrasts are largely surface ones that mask the degree to which their social customs and religious beliefs are fundamentally similar. For example, the Shi'i legal tradition's origins in the distinct, Jafari School and its continued embrace of *ijtihad* suggest at first blush that it diverges dramatically from the Sunni tradition; in reality, those differences were more ones of process than of function, and the Shi'i and Sunni versions of shari'a law ended up looking remarkably similar. Likewise, Sunnis and Shi'a frequently hold similar beliefs, typically practice the same social customs, and largely adhere to common rituals such as the Five Pillars of Islam, albeit in slightly different forms. Indeed, much of the all-too-real contemporary strife between members of the two sects is not the product of genuine religious differences but is instead a function of what Sigmund Freud called the Narcissism of Small Differences: the phenomenon wherein similar groups magnify minor dissimilarities to accentuate their distinct identities.

³⁹ Nasr, 67–68.

⁴⁰ Nasr, 34–40, 68–70.

The Kharijites

The final significant movement to take shape during the seventh century, Kharijism, did not constitute a distinct sect in the way the Shi'a and Sunni did. Outwardly, they appeared to be conventional Sunnis who adhered to the same beliefs and followed the same customs and shari'a law as did the majority of Muslims. In certain critical ways, however, they differed dramatically from both Shi'a and Sunnis. Most notably, their unforgiving interpretation of how Muslims should live and their equally strict perspective regarding who should serve as caliph constituted a construction of Islam that not only differed sharply from the views of other Muslims but, importantly, also established a template for puritanical opposition that remains influential through the present day.

The Kharijites emerged in the 650s as proponents of a rigid, extremist, and strongly egalitarian understanding of Islam. They held that every Muslim was equal before God and that each one was obligated to scrupulously obey every injunction and prohibition outlined in the Qur'an and the sunna of Muhammad at all times. They also interpreted the Qur'anic commandment to "enjoin what is good and forbid what is evil" to mean that sinning was not merely a moral failure but also an act of apostasy (Qur'an, 3:110). As a result, they believed fervently that the community was strictly obliged to deal with sinners through *takfir* or excommunication: the formal declaration that a person was an unbeliever and thus no longer a Muslim. A declaration of *takfir* was a serious charge to the Kharijites. Indeed, in their eyes, one subject to *takfir* faced the penalty of death unless they could demonstrate that they had genuinely repented.⁴¹

Stressing Islam's egalitarian roots, they made no exceptions for the powerful. In contrast to both the Sunni position that any male member of the Quraysh tribe could rule so long as he maintained order and to the Shi'i belief that only the infallible Imams could legitimately head the umma, the Kharijites argued that leadership of the community should fall to the most pious Muslim regardless of lineage. Those who ruled had to remain scrupulously devout if they wished to remain in power, however, for the Kharijites believed that the caliphs were subject to the same laws and punishments as any other Muslim. Indeed, they contended that the community was *required* to depose or even kill them if they sinned—a position that they held with unwavering commitment. Thus, as we have seen, they renounced Ali following the Battle of Siffin because they saw his willingness to enter into arbitration with Muawiyah as the usurpation of God's will and, as such, as an act of apostasy. This was a sin so grievous that, to the Kharijites, it justified their declaration of *takfir* against Ali and his subsequent assassination.⁴²

⁴¹ John L. Esposito, *Unholy War: Terror in the Name of Islam* (Oxford: Oxford University Press, 2002), 41–42.

⁴² Hugh Kennedy, *Caliphate: The History of an Idea* (New York: Basic Books, 2016), 28–29.

Despite their success in deposing Ali, the Kharijites proved almost completely incapable of imposing their conception of Islam on their coreligionists. Their extremist views remained highly unpopular, and they never managed to win more than a tiny fraction of Muslims to their cause. As a result, Kharijism remained a marginal movement confined largely to isolated communities in North Africa and the fringes of Arabia. Nonetheless, it has had an influence in the Muslim world wildly out of scale with its modest numbers. As we shall see, Kharijite ideas would shape later puritanical movements like the Wahhabi of eighteenth-century Arabia, the views of twentieth-century jihadi thinkers such as Sayyid Qutb (1906-1966), and the actions of contemporary terrorist groups like al-Qaeda.⁴³

Sufism

While the early Kharijites, Shi'a, and Sunnis argued about Islam's outward attributes, other Muslims, the Sufis, instead explored the religion's inner, spiritual element. Named for the simple, unrefined woolen clothing called *tasawwuf* that they favored, the Sufis eschewed politics and even Islam's traditional emphasis on the community in favor of an emotional and mystical focus on the individual's relationship with the divine. In the process, they developed a powerful and popular new conception of Islam that proved peculiarly capable of attracting converts.⁴⁴

Sufism emerged in reaction to two powerful trends in the Muslim world of the eighth century: surging materialism and the development of the shari'a code. To many believers, it seemed that worldly affairs had come increasingly to dominate the *umma*. Indeed, the latter decades of the Umayyad Dynasty had coincided with a growing emphasis in the Arab Empire on material concerns such as wealth and power—a shift that the ruling dynasty's increasing use of extravagant court rituals and embrace of luxury exemplified. Pious Muslims recoiled at the growth of materialism. Many responded by embracing the movement to codify the outward manifestation of Islam through the development of shari'a law. Not all spiritually minded Muslims supported that effort, however. On the contrary, a minority—the early Sufis—found the religious scholars' emphasis on external behavior and legalism every bit as objectionable as the Umayyads' embrace of wealth and opulence. In their view, the shari'a code was stiflingly rigid and formalistic, and its narrow focus on correct behavior left no room for a personal relationship with God.⁴⁵

Early Sufi masters responded by developing *tariqa* or “the Way”: the different methods through which people could connect directly to the divine. They started with the premise that rationalism and legalism were of no use in helping believers attain personal contact with God; after all, how could one *experience* the divine through the

⁴³ Esposito, *Unholy War*, 42–43, 48. 50.

⁴⁴ Armstrong, *Islam*, 74.

⁴⁵ Marshall G. S. Hodgson, *The Venture of Islam, Volume 2: The Expansion of Islam in the Middle Periods* (Chicago: The University of Chicago Press, 1974), 202–4; Aslan, *No God but God (Updated Edition)*, 204–5.

study of law or knowledge of philosophy? Instead, they believed that the embrace of the divine was a fundamentally intimate act that could only be achieved through intense emotional expression and through scrupulous adherence to proper techniques. It was not a task for the uninitiated. On the contrary, navigating the *tariqa* was a challenging venture that required the guidance of a Sufi saint who had already attained union with God. Known as *shaykhs* or, in Persian, *pirs*, saints guided those seeking communion with God on an emotional, inner, spiritual journey that involved passing through a series of stages designed to strip away their sense of self. Eventually, with perseverance and discipline, those seeking God's embrace would achieve *fana*: the annihilation of the ego and, with it, direct union with the divine.⁴⁶

From the start, Sufism was divided into a multiplicity of schools each of which followed a distinct *tariqa* created by a different master. These orders shared a number of commonalities. All held that *fana* required the intercession of a Sufi master who had already attained union with God and all revered the saints as intercessors between the individual and the divine. All Sufi schools also began with the belief that material temptations prevented people from maintaining the discipline needed to attain union with God; they consequently pursued asceticism, or the rejection of worldly comforts and pleasures, as the essential first step toward *fana*. Simple clothing, the embrace of poverty, and the practice of celibacy were the rule for the Sufis. All Sufis likewise practiced *dhikr*, the remembrance of God through ritual physical action. Here, however, the various Sufi orders followed very different *tariqa*. Most engaged in some form of vocal *dhikr*, which involved rhythmically and repetitively chanting God's name while simultaneously undertaking careful breathing exercises designed to break down the individual's sense of self. Others instead relied on music or, in the case of the famous Whirling Dervishes, on a rapid, spinning dance through which they sought to strip away their egos and to induce, in the scholar John Esposito's words, "ecstatic states in which the devotee could experience the presence of God or union with God."⁴⁷

While the numerous Sufi orders were heterogeneous, they had, by the ninth century, divided into two broad approaches. Championed by al-Junayd (834-910), so-called sober or conservative Sufism embraced customary Sufi approaches such as asceticism and *dhikr*, but also adhered to traditional Muslim beliefs and practices such as observance of the shari'a law. Drunken Sufism, in contrast, rejected orthodox Islamic traditions as impediments to the attainment of *fana*. Indeed, they viewed Islam as little more than the starting point on their journey to the annihilation of the ego and readily incorporated popular local customs and religious beliefs into their canon. Some even went so far as to argue that the world's great religions were merely different manifestations of a common search for unity with the divine.⁴⁸

⁴⁶ Esposito, *Islam*, 128–32.

⁴⁷ Aslan, *No God but God (Updated Edition)*, 221–22; Quote from Esposito, *Islam*, 130–32.

⁴⁸ Ira M. Lapidus, *A History of Islamic Societies*, 2nd ed. (Cambridge: Cambridge University Press, 2002), 92–93.

Unsurprisingly, neither the caliphal government nor the ‘ulama’ approved of these views. Indeed, from their perspective, drunken Sufism was a serious menace. Its rejection of the laws not only imperiled the social order, but also threatened the religious scholars’ authority and prestige. As a result, the ‘ulama’ and the caliphs joined together at different times to persecute drunken Sufis for apostasy. In the early tenth century, for example, the Sufi mystic Mansur al-Hallaj (c. 858-922) ran afoul of the religious scholars by claiming that the devout could complete the hajj without actually traveling to Mecca and by elevating Jesus’s stature to such a degree that some accused him of being a Christian. The religious scholars publicly condemned him for these declarations but did not take action against him. However, when he declared in the markets of Baghdad that “I am God!”—by which he meant that he had attained union with the divine—he crossed the line. At the behest of the ‘ulama’, caliph al-Muqtadir (r. 908-932) responded by first imprisoning al-Hallaj and then, after he refused to disavow his statements, by declaring him guilty of heresy and ordering his execution. To make an example of him, the caliph had him tortured, crucified, beheaded, dismembered, and, finally, cremated.⁴⁹

Despite al-Hallaj’s grisly fate, Sufism steadily gained new adherents among those who found the legal scholars’ emphasis on formalism off-putting. Indeed, the movement not only continued to grow, but, thanks to its embrace of syncretism—meaning its willingness to adopt religious practices from other faiths—it also came to play a vital role in Islam’s expansion. It was particularly important in winning converts in India, Southeast Asia, Central Asia, and Africa. People in those regions saw Sufism as a way to continue to practice the religion into which they were born while simultaneously adopting elements of Islam that they found attractive. Later, they, or more typically, their descendants, gradually converted to orthodox Sunnism. Sufism thus constituted a religious way station for many people who were otherwise unlikely to embrace orthodox Islam.⁵⁰

Islamic Society

Meanwhile, while that Sunnis, Shi‘a, and Kharijites debated questions of politics and Sufi saints and initiates sought communion with God, a new, urban-centered, Islamic society took form in the Middle East. Though it emerged out of the movement that Muhammad had led in the early-seventh century, it was not patterned along the egalitarian lines that the Prophet had championed. Instead, like its contemporaries, this new civilization was organized along starkly hierarchical lines of class, religion, and gender.

The Elite

While the primary social and political fault line lay—at least initially—between the Arab Muslims and the people they had conquered, even the ruling elite was highly stratified. At the very pinnacle of the social hierarchy stood the Abbasid caliph and his

⁴⁹ Quoted in Aslan, *No God but God (Updated Edition)*, 209.

⁵⁰ Esposito, *Islam*, 72.

court. This tiny elite was unmatched in terms of political power, social status, cultural gravity, and economic pull. Just under them were the religious scholars. Through their control of the shari‘a law, they constituted a decentralized, alternate elite that sometimes rivaled and sometimes collaborated with the caliph’s government. Below them, in turn, were the rich merchants of the major commercial cities and a class of politically influential large landowners to whom the court had, in exchange for service, given the estates of the Byzantine nobles who had fled during the conquest. Rounding out the elite were those Arabs who received military pensions from the *diwan* either because they had fought in the conquest or, as was more frequently the case as the years passed, because one of their ancestors had. Having one’s name on the register was highly prized. Not only did it entitle one to receive an annual pension, but it also carried with it a great deal of prestige.⁵¹

Ultimately, however, this last group’s membership in the elite proved to be transitory. For a combination of reasons both fiscal and political, the Abbasids gradually struck different groups from the *diwan* register. Over the years, as a consequence, thousands lost their pensions, their prestige, and, along with them, their status as members of the elite. Even then, however, they still held a relatively privileged social position. Most notably, in contrast to the caliph’s non-Muslim subjects, they enjoyed both the perquisite of higher social status and the practical benefit of paying lower taxes.⁵²

Dhimmi

Before the tenth century, the overwhelming majority of people outside the elite were *dhimmis*: non-Muslims obligated to pay the *jizya* tax. For the *dhimmis*, life in the Arab Empire came with both costs and benefits. As we have seen, they held certain rights—particularly the freedom to practice their faith—that religious minorities in other Medieval societies typically did not enjoy. Educated Christians and Jewish people also had the opportunity to serve in the bureaucracy. Still, the *dhimmis* had to endure a number of restrictions designed to make clear that they were second-class members of the Arab Empire. State and religious authorities barred them from riding horses, denied them the right to give testimony against Muslims in court, prohibited them from organizing Easter parades, compelled them to show deference to Muslims, maintained a series of sartorial laws that required them to wear clothing that marked their confessional identity, and, of course, imposed the *jizya* tax on them.⁵³

Conversion to Islam increased over time thanks in part to these restrictions. As we saw in chapter three, the proportion of the population that consisted of *dhimmis* declined only slightly during the two centuries following Muhammad’s death. Initially, as we have seen, the conquerors had discouraged conversion in hopes that Muslims could remain a separate, ruling caste; as a result, only a relatively small number of elite

⁵¹ Kennedy, *Caliphate*, 82–83.

⁵² Hourani, *A History of the Arab Peoples*, 45; Kennedy, *Caliphate*, 82–83.

⁵³ Hourani, *A History of the Arab Peoples*, 47.

*dhimmi*s who perceived practical benefits in converting adopted the religion of the conquerors. Even after Umar II (r. 717-720) threw open the door to conversion, the vast majority of the population in the Middle East remained Christian or Zoroastrian. Around 900, however, a second, larger wave of conversion started. Thanks to the collapse of non-Islamic social structures in Iraq and to growing Muslim hostility to Christianity—particularly after the start of the Crusades—the rate of conversion increased dramatically. Within a few centuries, as a result, Islam had become the dominant religion in the Middle East—achieving majority status in Iraq by 1200 and in Syria and Egypt by 1400.⁵⁴

Enslaved People

As was the case elsewhere in the world, enslaved people occupied the lowest rung in the social hierarchy. Slaves were commonplace in the Arab Middle East—just as they had been in the Byzantine and Sassanian Empires. Like slave owners in those states, the Arabs at first used gangs of enslaved people to work large, cash-crop estates. A massive and difficult-to-subdue revolt in southern Iraq by East African slaves known as the Zanj spurred a change in that practice in the ninth century, however. Concluding from the rebellion that maintaining large concentrations of enslaved people was too risky, the Arabs thereafter relied on peasants for agricultural labor and shifted enslaved people to domestic service and manual labor in the cities. Despite this change, slavery itself remained a vast enterprise in the empire.⁵⁵

Frequent manumission and Islamic prohibitions on enslaving Muslims meant that the caliphate could not meet the demand for slaves domestically. As a result, an expansive, long-distance trade in enslaved people from Europe, Central Asia, and Africa took shape over the course of the seventh and eighth centuries. Increasingly, those purchased from Central Asia were not destined for domestic service or manual labor in the cities. Instead, as we shall see in chapter five, the caliphate began purchasing a steadily growing share of them for the purpose of serving as slave soldiers or, as they were later known, Mamluks.⁵⁶

Women

Like slaves, Muslim women also found themselves in a subordinate legal and social position during much of this period. That they were in such a situation marked a significant change from their circumstances during the early years of the *umma* and was contrary to Muhammad's wishes. As we saw in chapter two, the Prophet had secured a substantial degree of legal emancipation for women. Most notably, he granted them control over their dowries and accorded them inheritance and divorce rights. Their improved legal and social standing did not merely endure while he was alive, moreover,

⁵⁴ Lapidus, *A History of Islamic Societies*, 200–201.

⁵⁵ Hourani, *A History of the Arab Peoples*, 36, 116–17.

⁵⁶ Lapidus, *A History of Islamic Societies*, 103–4; Hourani, *A History of the Arab Peoples*, 116–17.

but instead persisted through the early decades of the Rashidun Era. Thus, during the *umma's* early years, Muslim women enjoyed a degree of autonomy and legal standing that had little parallel in other parts of the globe.⁵⁷

Those gains proved to be short lived, however. Over the ensuing decades, men gradually reasserted traditional patriarchal values and rolled back Muhammad's reforms. The 'ulama' led the way. While they had felt compelled to include in shari'a law some of the improvements he had instituted for women, they also reinterpreted Qur'anic decrees to water those reforms down and to return women to the social-and-legal control of their male guardians. In addition, they steadily incorporated into the shari'a code pre-Islamic customary laws that made divorce easier for men and more difficult for women—despite the fact that those changes clearly violated both Qur'anic injunctions and Muhammad's wishes. In similar fashion, the right that women had won to determine whom they married was soon ignored, and fathers returned to the pre-Islamic custom of arranging their daughters' marriages. Finally, the 'ulama' also reinterpreted the Qur'an to accord Islamic sanction to longstanding Byzantine and Persian customs such as veiling and female seclusion in the *harem*, the private part of the home into which only close male relatives could enter. Often rigidly enforced to ensure female honor—meaning sexual faithfulness—the practice of seclusion in the Muslim world differed on the basis of class status. That is, upper-class women were far more typically confined to the *harem* than their lower-class counterparts, who, by dint of economic circumstances, often had to work outside the home.⁵⁸

It is important to keep in mind that that the position of women varied dramatically and that not all were powerless. For example, a number of queen mothers parlayed their influence over their sons who served as caliph into significant political authority. Most notably, al-Khayzuran bint Atta (?-789) dominated politics during the reigns of her husband and sons and served as de facto co-caliph during her son al-Hadi's (r. 785-786) tenure and during the first three years of the caliphate of her younger son, Harun al-Rashid (r. 786-809). These cases, of course, were very much the exception rather than the rule. During the century-and-a-half that followed Muhammad's death, most women instead endured the gradual reassertion of patriarchal values and the concomitant reduction of the rights and liberties that the Prophet had granted them until, by the late-eighth century, they had lost nearly all of the legal and social gains that they had made during Islam's first few decades.⁵⁹

⁵⁷ Lapidus, *A History of Islamic Societies*, 25–26.

⁵⁸ Esposito, *Islam*, 119–24.

⁵⁹ Hugh Kennedy, *When Baghdad Ruled the Muslim World: The Rise and Fall of Islam's Greatest Dynasty* (Cambridge, MA: Da Capo Press, 2004), 181–99; See also Fatima Mernissi, *The Forgotten Queens of Islam* (Minneapolis: University of Minnesota Press, 2003).

Science

Beginning in the ninth century, the Arab Empire experienced an efflorescence of scientific achievement that stood in stark contrast to the intellectual poverty of contemporary Europe. It began when a multicultural school of loosely affiliated Muslim, Christian, and Jewish scholars called *faylasuf*—derived from the Ancient Greek word *philosophia*—launched a concerted effort to translate ancient works of science, math, philosophy, and geography into Arabic. They quickly moved beyond mere translation, however. Building on the works of classical antiquity, those thinkers soon began to contribute vast new knowledge and to make enormous and lasting breakthroughs in a wide variety of fields. Their efforts reached a peak from the tenth through the twelfth centuries, when the Arab Empire stood as the most intellectually active and productive place on Earth—a time when its scientists, geographers, doctors, mathematicians, and linguists would not merely surpass the achievements of the ancients but also lay the groundwork for the Scientific Revolution.

Syriac-speaking Christians began the process of translating the Classical heritage into Arabic in the ninth century. Led by the greatest of the translators, the Nestorian Christian Hunayn ibn Ishaq (809-873), these scholars developed highly refined methods of producing Arabic versions of existing Greek, and, in some cases, Persian works. They copied the writings of mathematicians such as Pythagoras (c. 570-495 BCE) and Euclid (c. 325-265 BCE), the books of doctors such as Galen (129-c. 210 CE), Hippocrates (c. 460-370 BCE), and Shushustra (c. 800-c. 700 BCE), the charts of astronomers like Ptolemy (c. 100-170 CE), the thinking of philosophers such as Aristotle (384-322 BCE) and Plato (c. 427-347 BCE), and the theories of philologists such as John Philiponus (c. 490-570 CE). In the process, the translators developed new idioms and vocabulary that rendered Arabic suitable for the discussion of the concepts and methods of the different scientific disciplines.⁶⁰

From nearly the beginning of this effort, moreover, they did not limit themselves to merely creating Arabic copies of existing works or repeating the ideas of the ancients. Instead, as their approach and knowledge became more sophisticated, they increasingly found errors and flawed reasoning in the sources that they translated. In response, they started to gloss their works, meaning that they began to make marginal comments that corrected and expanded on the manuscripts they translated.⁶¹

The Beyt al-Hikma

Perceiving the practical benefits of the translation movement—particularly in the realms of geographic and medical knowledge—the Abbasid Dynasty threw its support behind the scholars beginning in the early-ninth century. Its backing was critical. The dynasty not only lavished substantial financial patronage on the movement but also ensured its continuation by protecting it from its numerous religious critics. More

⁶⁰ Esposito, *Islam*, 56–57.

⁶¹ Hourani, *A History of the Arab Peoples*, 75–77.

importantly, under Caliph al-Mamun (r. 813-833), the Abbasid government constructed the *Beyt al-Hikma*, or House of Wisdom. Over the next two centuries, this combination of library and school would be the most important center of learning in the world—a place where scholars could translate ancient works, undertake research, conduct experiments, and synthesize knowledge. Thinkers working there not only benefited from the ability to interact with other leading scientists, but, thanks to the steady acquisition of new manuscripts from India, Persia, and the Byzantine Empire, also drew on what was, for several centuries, the most expansive collection of books on earth.⁶²

By that point, the scholars affiliated with the House of Wisdom had begun to move beyond the Classical Greek thinkers who had originally inspired them. Building on the practice of glossing, Arab scientists, doctors, geographers, and astronomers had begun producing original works of enormous creativity and lasting influence. Ibn al-Hataym (c. 969-1040) did pioneering work in optics, for example, while Abu Zayd al-Balhki (850-934) established a school of cartography and developed new methods that produced vastly more-accurate maps. Known as Avicenna in the West, Ibn Sina (c. 980-1037) wrote the *Canon of Medicine*—a work so pathbreaking that it remained the standard medical text in Europe and the Middle East until the nineteenth century. Muslim intellectuals also developed astronomical tables using an observatory attached to the House of Wisdom that the Caliph Mamun had ordered built in 828.⁶³

Scholars affiliated with the House of Wisdom had perhaps their greatest influence in the field of mathematics. Arab and Persian mathematicians such as Umar Khayyam (1048-1131), al-Karaji (953-1029), al-Haytham (965-1040), and others produced major advances in areas including linear and quadratic equations, geometry, and the use of fractional place notation. None was more significant than al-Khwarazmi (c. 800-847). The leading mathematician of his day, he made major contributions to the study of trigonometry and to the development of algorithms and was instrumental in the adoption of Indian numerals—later passed on to the West where they are known today as Arabic numerals. Most importantly, he thoroughly revolutionized the study of algebra. A thinker of the first rank, his contributions to the development of mathematics were pivotal and played an outsized part in laying the groundwork for the achievements of later Western mathematicians such as Isaac Newton (1642-1726). Indeed, his impact on Western thinking was so enormous that the word algorithm comes from the Latinized version of his name, *Algoritmi*, while the term algebra derives from the title of his most influential book, *Hisob al-Jabr wal' Mugabalah*.⁶⁴

⁶² Lapidus, *A History of Islamic Societies*, 78.

⁶³ Ahmad Dallal, *Islam, Science, and the Challenge of History* (New Haven: Yale University Press, 2010), 38–39; Hourani, *A History of the Arab Peoples*, 202–3; E. Edson and E. Savage-Smith, *Medieval Views of the Cosmos* (Oxford: Bodleian Library, University of Oxford, 2004), 61–63.

⁶⁴ Uta C. Merzbach and Carl B. Boyer, *A History of Mathematics*, Third Edition (Hoboken, New Jersey: John Wiley and Sons, 2011), 206.

While many of the *faylasufs* connected to the House of Wisdom were non-Muslims, the knowledge that they created and the larger project to which they contributed fit comfortably within the structure of the emerging Islamic society. The close relationship between the House of Wisdom and Islam is most evident in the fields of mathematics and astronomy. Astronomic research often centered on practical questions related to religion such as the maintenance of religious calendars or the establishment of the correct time for daily prayer. Math likewise often served religious purposes. For example, mathematicians affiliated with the *Beyt al-Hikma* such as Abu al-Wafa (940-998) learned to use all six trigonometric functions in part so that Muslims could accurately determine the *qibla*—the direction of the Ka‘ba and thus the point on the horizon to which they should pray—from any point on earth.⁶⁵

The knowledge that the *faylasuf* created spread quickly to thinkers in other parts of the Arab Empire and even to those living in nearby states. That it did so owed in no small part to the growing availability of paper, a technology of Chinese origin that had spread to Central Asia in the centuries prior to the arrival of the Arabs in that region. Significantly cheaper to produce than the parchment that it replaced, paper soon created an explosion in the quantity of titles produced in the caliphate and a concomitant growth in the number of bookstores in Arab cities. More pertinently, it permitted thinkers working at the *Beyt al-Hikma* to inexpensively share their ideas with scientists and other intellectuals throughout the caliphate, including those who lived in other important centers of learning such as Cairo, Bukhara, and Damascus.⁶⁶

The Historical Debate: the *Beyt al-Hikma*

Based on primary-source descriptions of the *Beyt al-Hikma* and on the achievements of scholars affiliated with it such as al-Khwarazmi, historians have long viewed the House of Wisdom as the central institution in the development of Muslim science and mathematics. Written evidence suggests that it played an outsized role both in funneling caliphal support to the scholars who worked there and, through its library, in providing them with access to both ancient and contemporary thinking. However, there are historians today who do not merely question the centrality of the *Beyt al-Hikma* to the development of Arab science and mathematics but, more fundamentally, also raise doubts whether the institution actually existed. The most prominent such critic is the scholar of Islamic Studies Dimitri Gutas. Pointing out that “we have *exceedingly little historical* information about the” House of Wisdom, he argues that descriptions of it as a center of learning similar to a modern university are “fanciful and sometimes wishful projections.” He maintains instead that the *Beyt al-Hikma* was at most a royal archival library established to translate and store Sassanid government documents and that, as such, it played no role in the translation of Greek

⁶⁵ Hourani, *A History of the Arab Peoples*, 77.

⁶⁶ Kennedy, *Caliphate*, 106–8; S. Frederick Starr, *Lost Enlightenment: Central Asia’s Golden Age from the Arab Conquest to Tamerlane* (Princeton, NJ: Princeton University Press, 2013), 47.

works or in the development of Arab thought. In support of this argument, he notes that *Beyt al-Hikma* is merely the Arabic translation of the Sassanian term for library.⁶⁷

Gutas's critique has thus raised real doubts about whether the House of Wisdom even existed. Still, in the broadest sense, the significance of his assessment is more limited than it may appear at first glance. Whether a physical, caliphally supported institution of learning existed or not, he agrees with other scholars that Baghdad was not merely the intellectual heart of the Arab Empire during the Abbasid era, but, more importantly, constituted for several centuries the globe's leading center of learning and research.⁶⁸

Al-Andalus

Baghdad's only intellectual peers at that time were the multicultural cities of Cordoba, Seville, and Toledo in Muslim-controlled Spain. Indeed, many of the most prominent thinkers of that era lived in al-Andalus. Hailing from Toledo, for example, the Muslim scholar al Zarqali (1029-1087) revolutionized the field of astronomy. Originally trained as a metalsmith, he developed an improved astrolabe that greatly aided navigation, and he showed that the planets followed elliptical rather than perfectly circular paths when they traveled through the night sky. Even more influential was the Muslim polymath Ibn Rushd (1126-1198), or, as he was known in Western Europe, Averroes. Born in Cordoba, he made pathbreaking contributions in fields including physics, philosophy, math, jurisprudence, and medicine. His contemporary, the Jewish intellectual and bureaucrat Maimonides (c. 1136-1204) was a similarly towering figure. Also born in Cordoba, he developed groundbreaking new ideas in a wide variety of fields including medicine, philosophy, logic, and astronomy.⁶⁹

While the Muslim, Christian, and Jewish intellectuals of Spain constituted an integral part of the broader scholarly community of the Islamic world, they also played a vital role in reintroducing to Western Europe critical works of classical antiquity that had been lost to the people of that region since the fall of the Western Roman Empire in the fifth century CE. This process began in the twelfth century when a multicultural group of scholars in Toledo began translating works like Ptolemy's *Almagest*, Aristotle's *Nicomachean Ethics*, and Euclid's *Elements of Geometry* from Arabic to Latin. Importantly, they did not merely provide verbatim copies of those works. On the contrary, their translations also included the annotations and commentaries of major Arab thinkers updating, correcting, and challenging those manuscripts. Likewise, they did not confine themselves to providing Latin copies of the lost works of classical antiquity. Instead, they also translated into Latin many books written by Medieval Arab

⁶⁷ Quotess from Dimitri Gutas, *Greek Thought, Arabic Culture: The Graeco-Arabic Translation Movement in Baghdad and Early 'Abbasaid Society* (New York: Routledge, 1998), 54; Kennedy, *When Baghdad Ruled the Muslim World*, 246.

⁶⁸ Kennedy, *When Baghdad Ruled the Muslim World*, 54.

⁶⁹ Dallal, *Islam, Science, and the Challenge of History*, 65, 88–89.

and Persian scholars including al-Khwarazmi's *Algebra*, al-Kindi's (801-873) *On Optics*, and Ibn Sina's *Canon of Medicine*.⁷⁰

These translated works struck Western Europe like a bombshell. Their powerful challenge to the prevailing Christian understanding of natural science produced a firestorm of controversy that sharply divided scholars in Europe's fledgling universities. By far the most divisive works were those produced by the Greek philosopher, Aristotle. Put simply, he posed a serious dilemma for European scholars at that time. On the one hand, his thinking appeared to challenge the prevailing Christian idea that the physical universe operated as a function of God's will; on the other, no educated person could ignore the brilliance of what Aristotle had written or the practical benefits that his method of study promised to bring. As a result, scholars at Europe's fledgling universities quickly divided between radicals who wished to use Aristotelian methods in their studies and conservatives who hoped to banish what they believed to be subversive and highly dangerous ideas. Intense as it was, the fight did not last long. Western European thinkers would resolve the seeming contradiction between Aristotle and received theological views by reframing his ideas in Christian terms through Scholasticism. Safely domesticated, his works—and the other books translated into Latin in Spain—would thereafter help to spark the Renaissance of the Twelfth Century and, more broadly, set the stage for the Scientific Revolution of the sixteenth and seventeenth centuries.⁷¹

Islamic Philosophy

The Problem of Aristotle

Aristotle had created a not-dissimilar controversy in the Middle East a few centuries earlier; there, however, the debate had a decidedly different outcome. One of the greatest thinkers of ancient Greece, Aristotle had articulated a comprehensive and persuasive system for investigating and explaining the physical world. According to him, the universe had begun when a distant, creator deity had first constructed it and set it in motion. Thereafter, it operated according to a series of internal laws that determined the behavior of all matter, all life, and all people. This understanding had far-reaching implications. If, as Aristotle believed, a set of rules governed the operation of the universe, then people could—through careful study, experimentation, and classification—understand how it functioned. In other words, properly structured and grounded in logic, human reason could be relied upon to fathom the workings of all creation.⁷²

⁷⁰ Richard E. Rubenstein, *Aristotle's Children: How Christians, Muslims, and Jews Rediscovered Ancient Wisdom and Illuminated the Dark Ages*, First Edition (Orlando: Harcourt, 2003), 15–19.

⁷¹ Rubenstein, 5–11, 79.

⁷² Rubenstein, 77–82.

The *faylasufs* found Aristotle's way of understanding the world electrifying. Almost from the start, they had made fruitful use of his thinking in their study of science, math, politics, society, and culture. In fact, his ideas were so powerful that, over time, the *faylasufs* even applied them to more explicitly religious topics. Here, however, differences between Aristotle and contemporary Islam would make the going far-more difficult. Those differences were indeed substantial. For instance, Aristotle's central scientific idea that a set of laws governed the operations of the universe seemed to leave no room for God to perform miracles or to otherwise affect the workings of creation as both Muhammad and the Qur'an maintained He did on a regular basis. Early *faylasufs* consequently took great care in applying Aristotelianism to the study of religion. For example, al-Kindi contended that scholars and theologians could make profitable use of it to explore religious questions that the Qur'an and the sunna did not explicitly address, but he also maintained that the Qur'an was the literal word of God, and that, as such, it lay beyond the realm of rational inquiry.⁷³

Mu'tazilism

While the 'ulama' viewed even this fairly conservative position with suspicion, they found the ideas of a new group of thinkers, the Mu'tazilites, to be significantly more alarming. Influenced by Aristotelian philosophy, the Mu'tazilite School first emerged as a challenge to the traditionalist religious scholars in the eighth century. Rooted in Greek rationalism, it rejected outright the orthodox assertion that religious study began and ended with the sunna of Muhammad and the revealed text of the Qur'an. Instead, arguing that God was rational and that humans could understand Him using the methods pioneered by classical Greek philosophy, the Mu'tazilites maintained that reason—and not revelation—was the correct basis for theological inquiry.⁷⁴

Starting from such dramatically different positions, the Mu'tazilites and the orthodox 'ulama' advanced sharply different understandings of Islam. Indeed, they disagreed about nearly everything. They differed about issues such as the origins of the universe, the nature of physical reality, and even human psychology. Above all, they clashed over three pivotal issues: whether humans enjoyed free will, whether the Qur'an was eternal or created, and whether its references to God's attributes—its descriptions of him that appear to render him in anthropomorphic terms—must be taken literally.⁷⁵

Based on their embrace of moral absolutes, the religious scholars maintained an uncompromising position on these issues. They argued that revelation fundamentally trumped reason, that the Qur'an was uncreated and had existed for all time, and that, as the literal word of God, it could not be subject to interpretation—though, more accurately, as we have seen, what they were really contending was that *their* understanding of the Qur'an could not be challenged. By extension, they also asserted

⁷³ Lapidus, *A History of Islamic Societies*, 79.

⁷⁴ Esposito, *Islam*, 87–89.

⁷⁵ Hourani, *A History of the Arab Peoples*, 63–64.

that while the Qur'an was emphatic in maintaining that God was in no way anthropomorphic or even graspable in human terms, its description of His attributes—references to His hands, speech, face, and so forth—nonetheless had to be understood in literal terms. Finally, taking at face value both the Qur'an's declaration that God is omnipotent and its apparent embrace of determinism—"God sends astray whom He wills and guides whom He wills"—they concluded that humans did not have free will and that God instead predestined their actions (Quran, 35:8).⁷⁶

On the basis of their understanding of Greek rationalism and their Islamic faith, the Mu'tazilites found these positions to be intellectually and theologically untenable. They raised three primary objections. First, their emphasis on *tawhid*, or the unity of God, led them to criticize the religious scholars' position that the Qur'an was uncreated and eternal. The Mu'tazilites maintained that this argument was *shirk*, or idolatrous, because it separated the Qur'an from God and thus treated it as a second deity much as Muslims believed that Christians had done in transforming Jesus into a divine being. Based on this reasoning, they concluded not only that the Qur'an had been created but that it was best understood in metaphorical rather than literal terms. Second, the Mu'tazilites rejected the religious scholars' argument that references to God's attributes had to be understood literally. God was not anthropomorphic and could not be interpreted in such terms, they countered; hence, discussions of His attributes in the Qur'an, such as references to His speech or face, were meant to be understood allegorically rather than literally. Finally, the Mu'tazilites disputed the religious scholars' view that God predestined all human actions. To the Mu'tazilites, that interpretation was inconsistent with Islam's emphasis on divine judgement. How, they asked, could a just God judge people if they are not responsible for their actions? More troublingly, such thinking appeared to make God not merely the basis for all that is good, but also the source of evil. Accordingly, noting that the Qur'an explicitly held that people have free will, they rejected the religious scholars' support for predestination and instead argued that God had given people the power to embrace or reject salvation through their actions.⁷⁷

For the 'ulama', the emergence of Mu'tazilism in the eighth century constituted a mortal peril. The rationalists' contention that reason was a better guide to religion than the Qur'an or knowledge of the sunna constituted a direct threat to their position. Indeed, it put at risk both their recent success in institutionalizing their interpretation as the only legitimate understanding of the Qur'an and their assumption of control over the sunna through Hadith Scholarship. In response, they launched an aggressive campaign against the Mu'tazilites beginning in the late-eighth century. Initially, their effort was enormously successful. Using their control of the courts and mosques and their dominant position in education, they waged a relentless and effective drive against rationalist ideas. Its impact was substantial and immediate. Indeed, having never

⁷⁶ Aslan, *No God but God (Updated Edition)*, 156–60.

⁷⁷ Hourani, *A History of the Arab Peoples*, 62–64.

commanded much in the way of popular support, the Mu‘tazilites withered in the face of this sustained assault.⁷⁸

The situation reversed dramatically with the accession of al-Mamun. As caliph, he lavished support on the Mu‘tazilites, accorded their movement official status, and, most importantly, made adherence to its doctrines a requirement for teachers and religious scholars. Why did he choose to throw his support behind the rationalists? After all, Mu‘tazilism enjoyed negligible public support and thus seemed to provide him with little advantage. He did so in part because he found its rationalist arguments about free will and the uncreated nature of the Qur’an to be persuasive; that is, he genuinely believed that the Mu‘tazilites were correct. He had another, more practical reason to support the movement, however. He believed that he could use it to assert caliphal authority over the increasingly autonomous ‘ulama’ and to challenge their control of shari’a law.⁷⁹

He stepped up his confrontation with the ‘ulama’ near the end of his reign. Seeking to assert his authority over the religious scholars, he initiated a campaign of persecution in 833 called the *mihna* designed to enforce acceptance of the rationalist position on the Qur’an. During the *mihna*, ‘ulama’ were summoned to the court and compelled to declare whether they believed that the Qur’an was created. Those who affirmed the Mu‘tazilite position that it was a divine creation were free to go; those who instead clung to the view that it was eternal and uncreated were arrested and subjected to torture. Faced with the prospect of spending time in the caliph’s dungeons, many ‘ulama’ declared that they had undergone a change of heart. A small number stood fast in their beliefs, however. Most prominent among them was the founder of the Hanbali School of jurisprudence, the influential religious scholar ibn Hanbal. Though he endured years of isolation and brutal torture in prison, he refused to crack and remained steadfast in his refusal to recant his belief that the Qur’an was uncreated.⁸⁰

Ultimately, the *mihna* proved unsuccessful either in bringing the ‘ulama’ to heel or in institutionalizing Mu‘tazilism. Ibn Hanbal’s imprisonment marked the turning point in the campaign. Not only did his time in jail fail to change his views, but it made him a popular hero and a powerful symbol around which those who backed the traditionalist position could rally. That support proved critical in securing his release. Fearing civil unrest were he to die in captivity, Mamun’s successor, al-Mutasim (r. 833-842), finally agreed to free ibn Hanbal near the end of his reign. A massive and enthusiastic crowd greeted him upon his release—a demonstration of support that further impressed upon the Abbasids the popularity of the traditionalist position. Indeed, the enthusiastic response to his relief proved pivotal. Well aware of the importance of public opinion to the ongoing stability of the empire, Caliph al-Mutawakkil (r. 847-861) eventually responded to the public’s clear preference for the religious scholars’ traditionalist views by bringing the *mihna* campaign to an end early in his reign. He soon went further. A

⁷⁸ Ansary, *Destiny Disrupted*, 105.

⁷⁹ Aslan, *No God but God (Updated Edition)*, 143–45.

⁸⁰ Wickham, *The Inheritance of Rome*, 329–30.

few years later, he formally ended the government's support for the Mu'tazilites and publicly announced his endorsement of the orthodox 'ulama' position that the Qur'an is uncreated.⁸¹

The consequences of al-Mutawakkil's actions were enormous for both the Mu'tazilites and the traditionalists. Having never enjoyed mass appeal, Mu'tazilism rapidly went into eclipse once the caliph ended state support for it. For the orthodox 'ulama', on the other hand, the dynasty's reversal amounted to a major victory. They had not only defeated the Mu'tazilites but, more importantly, had succeeded in asserting their autonomy in the face of a sustained challenge from the court.⁸²

The Faylasuf

The traditionalists may have beaten back the immediate threat of Mu'tazilism, but they soon confronted another, more enduring challenge in the form of the increasingly bold *faylasufs*. While early *faylasufs* like al-Kindi had been circumspect and careful in their application of Aristotelian rationalism to religious inquiry, later ones were willing to employ such thinking much more aggressively in their examination of Islam. For example, al-Farabi (872-951) explicitly argued that philosophy offered a superior path to salvation than did religion. Known as the "second teacher" of philosophy—Aristotle being the first—he maintained that God had given humans a tool, reason, that they could use to cleanse themselves of sin and thus ready themselves for union with the divine. Only a small minority had the education and intellect to understand God in this way, however. For all others, the Qur'an and the teachings of Islam provided a simplified, symbolic means of accessing the truth.⁸³

The brilliant ibn Sina built on and extended al-Farabi's ideas. Fitting Islam into Greek philosophy rather than the reverse, he contended that the prophets were people whose superior minds gave them the insight needed to fully understand God and His plan. They, in turn, created simplified, accessible rituals and taught them to the masses—not because those practices were derived from the literal word of God but instead because they presented, in the historian Ira Lapidus's words, "an imaginative presentation of truths" that the uneducated common people could grasp.⁸⁴

From the late-ninth through the late-eleventh centuries, the *faylasuf* and, to a lesser extent, the reinvigorated Mu'tazilites, posed a serious challenge to the orthodox 'ulama'. Their command of Greek-derived rationalism, their rejection of received-but-logically untenable religious views such as the idea that the world was created *ex nihilo*—from nothing—and their ability to offer a persuasive, natural-science based model of the physical operation of the universe lent them great credibility. As a result,

⁸¹ Hugh Kennedy, *The Prophet and the Age of the Caliphates*, Second Edition (New York: Routledge, 2004), 162–63, 167–68.

⁸² Ansary, *Destiny Disrupted*, 106.

⁸³ Quote from Esposito, *Islam*, 57; Hourani, *A History of the Arab Peoples*, 78.

⁸⁴ Lapidus, *A History of Islamic Societies*, 178.

they attracted the support of a small but growing number of educated and influential urban Muslims.⁸⁵

The rising appeal of the *faylasufs* troubled many ‘ulama’. One of them, the religious scholar Abu al-Ashari, (c. 873-935), responded by developing a new approach known as the Asharite School. Designed to establish a compromise position between the rationalists and the traditionalists—and, thus, to coopt the arguments of the *faylasufs* and Mu‘tazilites—this new approach articulated a limited acceptance of human reason as a means of understanding religion. In other words, al-Ashari and his followers had conceded that rationalism was a powerful tool that could be used to better understand Islam. Importantly, however, they also maintained that its application had significant limits. Most notably, it could only support and defend understandings of revealed truths, and could not supplant, contradict, or challenge the path to salvation outlined in the Qur’an and the sunna.⁸⁶

Asharism quickly gained a large following and soon became the dominant theological school within Sunni Islam. However, it failed to make inroads among the *faylasuf* and Mu‘tazilites. Instead, finding Asharism to be entirely unpersuasive, they criticized it and challenged its proponents in a series of public arguments during the late-tenth and early-eleventh centuries. These debates ended up being embarrassingly one sided. Using their command of Greek logic and rational inquiry, the *faylasufs* and Mu‘tazilites regularly ran rings around their flustered Asharite opponents. Their success against the traditionalists did little to help them attract mass support, but, to the consternation of the Asharites, it did win for them substantial support among the influential educated, urban classes.⁸⁷

Al-Ghazali

That the Mu‘tazilites failed to do so owed largely to the efforts of one adherent of the Asharite School: al-Ghazali (c. 1058-1111). Deeply troubled by the growing influence of the *faylasufs* and Mu‘tazilites, he moved in the late-eleventh century to rebut their views. Doing so promised to be a difficult task. As the Asharites had so painfully discovered, the *faylasufs* were highly skilled at using Greek logic and rationalism to undermine their opponents. Al-Ghazali was well prepared to parry their criticisms, however. Shrewdly concluding that he could best refute the *faylasufs*’ arguments by using against them the very rationalism on which their philosophical and theological views rested, he first produced a summation of Classical philosophy called *The Aims of the Philosophers*. It was a powerful distillation of Greek rationalism. Indeed, as the writer Tamim Ansary notes, *The Aims of the Philosophers* was so clear and

⁸⁵ Armstrong, *Islam*, 63–64.

Esposito, *Islam*, 90–91.⁸⁶ Armstrong, *Islam*, 63–64.

⁸⁷ Ansary, *Destiny Disrupted*, 110.

balanced that many contemporary Europeans who read translated copies of it assumed that the author was a supporter rather than a critic of Aristotelian inquiry.⁸⁸

Having mastered *falsafa*, al-Ghazali next set out to destroy it. His attack on Muslim philosophy came in his next book, *The Incoherence of the Philosophers*. In it, he identified the ideas that, collectively, constituted the foundation of the *faylasufs'* reason-based views. Using the logic and methods of the rationalists against them, he then offered a devastating critique of those premises.⁸⁹

The core of his argument was his attack on causality. At heart, Aristotelian natural science and philosophy were rooted in the belief that the world operated according to a series of immutable laws. That idea, in turn, rested on the concept of causality, which maintained that a given effect is the consequence of a particular cause, and that a certain cause will, all other things being equal, always produce the same effect. For example, since flame is always present when cotton burns, fire—the cause—will result in cotton bursting into flame—the effect—any time the two come into contact.⁹⁰

Al-Ghazali completely rejected cause and effect. Instead, using Greek rationalism, he articulated a refined version of an older idea called Occasionalism. In direct contrast to causality, Occasionalism holds that natural laws do not govern the physical universe; instead, the world exists as a series of discrete moments bound together only by God's will. Cotton burns when it comes into contact with fire not because of causality, but instead because God, for reasons that mere humans cannot fathom, wills that it would burn. Occasionalism thus maintains that the fire had nothing whatsoever to do with the cotton burning, and its proximity to the cotton at the point that it burst into flame was merely a matter of correlation rather than one of causation. To argue otherwise, al-Ghazali concluded, was to deprive God of agency and to render Him as something less than the all-powerful figure described in the Qur'an.⁹¹

Al-Ghazali's takedown of Aristotelian rationalism was persuasive, effective, and enormously far-reaching. Using the concept of Occasionalism with devastating effectiveness, he undermined the idea that the universe operated according to a set of immutable laws; by extension, he also discredited the foundation on which all rational, Aristotelian theology rested. As a result, educated Muslims not only began to dismiss causality as an illusion, but, more importantly, increasingly rejected the entire basis of the *faylasufs'* rationalist theology.⁹²

⁸⁸ Ansary, 111.

⁸⁹ Esposito, *Islam*, 91.

⁹⁰ Rubenstein, *Aristotle's Children: How Christians, Muslims, and Jews Rediscovered Ancient Wisdom and Illuminated the Dark Ages*, 85; Ansary, *Destiny Disrupted*, 111–12.

⁹¹ Ansary, *Destiny Disrupted*, 111–12.

⁹² Esposito, *Islam*, 89–92.

Al-Ghazali's critique also offered a successful rejoinder to the Mu'tazilite argument that people enjoyed free will. In direct contrast to the rationalist position, Occasionalism held that since God was all powerful, He determined all actions—including those that appeared to be the product of human agency; free will thus did not exist. Importantly, however, al-Ghazali rejected the idea that people were mere moral bystanders who were not accountable for what they did. As the religious scholar John Esposito writes, he instead adhered to the Asharite position that individuals became responsible for their choices through "a theory of 'acquisition,' which maintained that people acquire responsibility and thus accountability for their actions." Though God predestined all actions, in other words, individuals nonetheless assumed responsibility for them.⁹³

Ibn Rushd

The Incoherence of the Philosophers did not go unchallenged. Instead, the Spanish-Muslim *faylasuf* Ibn Rushd offered a powerful, point-by-point response in his twelfth-century rebuttal, *The Incoherence of the Incoherence*. Building on the work of earlier scholars such as Avicenna and al-Farabi, he argued in it that as "'truth does not oppose truth but accords with it,'" the revealed knowledge of the Qur'an and the reasoned findings of philosophical inquiry cannot be in opposition. Any apparent conflict between them does not reflect genuine contradiction, he continued, but instead indicates that the prevailing understanding of the Qur'an is inaccurate and should be changed. Like the Mu'tazilites, ibn Rushd thus believed that educated people needed to comprehend the Qur'an in allegoric rather than literal terms. Not everyone had the intellect to understand it metaphorically, however. Accordingly, he maintained that while the *khassa*, the educated elite, should interpret the Qur'an using philosophy, the *amm*, or masses, should instead continue to approach it on a simpler, literal level.⁹⁴

Ibn Rushd's riposte was brilliant, sharply reasoned, and convincing. It came, however, at an extremely inauspicious moment for arguments of such subtlety and sophistication. As we shall see in chapter five, the Arab Empire was no longer the power it had once been. Instead, weakened by the fact that it had fractured into warring principalities, it confronted a significant external challenge in the form of Western European crusaders, and an existential one in the shape of the awesomely powerful Mongols. Simply put, it was a time of catastrophe—one that left the *umma* unreceptive to the sophisticated positions that ibn Rushd had developed. As a result, however brilliantly argued it may have been, *The Incoherence of the Incoherence* proved unable either to gain purchase or to arrest philosophy's declining influence in the *umma*.⁹⁵

⁹³ Esposito, 89–92.

⁹⁴ Hourani, *A History of the Arab Peoples*, 174–75.

⁹⁵ Hodgson, *The Venture of Islam, Volume 2*, 318–23.

The Closing of the Gates of *Ijtihad*

The climate worsened considerably for the *faylasufs* and their allies in the decades following ibn Rushd's death. To a degree, this change was the product of a longstanding effort by traditionalist religious scholars to limit the use of reason in the analysis of questions related to theology and shari'a law. Beginning in late-tenth century, they had responded to the caliphate's growing problems by calling for *taqlid*, or the uncritical acceptance of received views. As part of that effort, some had started to argue that since the shari'a code had been perfected, the use of *ijtihad*, or independent reasoning, was no longer a necessary or even appropriate tool for the articulation of religious law or the study of Islam. Known as the "closing of the gates of *ijtihad*," this turn away from independent reasoning made comparatively limited headway at first. However, it began to gain wide acceptance in the Muslim world after the Mongols sacked Baghdad and ended the Abbasid Dynasty in 1258. Indeed, that event marked a critical watershed for independent thinking in the Muslim world. As the historian Toby Huff asserts, from that point forward, "[b]oth theology and law specifically rejected the idea of a rational agency attributable to all men in favor of the view that man should follow the path of traditional authority . . . and not attempt to fathom the mysteries of external nature or sacred writ."⁹⁶

The consequences of the closing of the gates of *ijtihad* were as enormous as they were adverse. While the Muslim Middle East had led the world in the study of science and philosophy during the Abbasid Dynasty, it would make comparatively negligible intellectual contributions thereafter. Instead, it would be the *umma's* former pupil, Christian Europe, that would use the Islamic scientific and philosophic texts it had acquired from the Muslims to vault ahead intellectually and materially—in the process setting the stage for the enormously consequential intellectual and technological breakthroughs of the modern era.

Ibn Taymiyyah

Still, though support for the use of *ijtihad* declined precipitously after 1258, a minority continued to insist that it remained a valid and necessary tool of jurisprudence. The most prominent such person was the religious scholar ibn Taymiyyah (1263-1328). Forced to flee his home while a child to escape the Mongols, he had come to conclude that the *umma's* failure to live as God had intended had led directly to the catastrophes that had befallen the Muslim world. Accordingly, like many others at that time, he argued that the community of believers needed to purify itself by returning to the practices and customs of the Salaf. In this regard, ibn Taymiyyah agreed with the traditionalists that innovations in religion and social practices ran counter to God's will. In other ways, however, he was anything but backward looking. Most notably, while nearly all of his fellow 'ulama' maintained that the shari'a was timeless and perfect, he

⁹⁶ First quote from Aslan, *No God but God (Updated Edition)*, 167; Huff quote from Toby E. Huff, *The Rise of Early Modern Science: Islam, China, and the West* (Cambridge: Cambridge University Press, 1993), 113.

argued instead that Islamic law needed to be continuously modified through independent reasoning so that it could better accord with the times. In other words, he maintained, the shari'a code should not adhere dogmatically to laws that had been drafted in a bygone era but should instead reflect the spirit of God's vision for the *umma* in the context of present-day circumstances.⁹⁷

At the same time, however, he strongly opposed innovations in the practice of Islam that he believed violated the religion's basic tenets. Here, ibn Taymiyyah advanced an argument that while framed in conservative terms was in fact a highly innovative one. He contended that certain Muslims—the Shi'a, those Mongols who had converted but still followed their traditional laws, and members of syncretic religious groups such as the Alawites and Druze—had strayed so far from the vision of Islam laid out in the Qur'an and the hadith that they had become *kafirs*, or apostates. What should the community of believers do with such Muslims? Contending that the *umma* needed to purify itself in order to please God, he asserted that Muslims not only should but were indeed obligated to *takfir*—or excommunicate—and kill apostates. Thus, for the first time, a mainstream religious scholar had joined the Kharijites in articulating a legal justification for using violence against people conventionally understood to be Muslims.⁹⁸

This idea would have only limited appeal during ibn Taymiyyah's lifetime. Indeed, his views were so frequently at odds with those of his fellow religious scholars and the civil authorities that they resulted in him spending substantial periods of time in jail. Beginning in the eighteenth century, however, many of his ideas would win a new following among important puritanical reformist groups. Most notably, the stern Wahhabi 'ulama' of central Arabia would revisit ibn Taymiyyah's conception of *takfir* to justify attacking Muslims deemed to have committed apostasy through their actions or beliefs. In more recent times, his thinking would also inform the ideology of Islamist organization like al-Qaeda and the Islamic State of Iraq and Syria (ISIS).⁹⁹

Conclusion

By the end of the thirteenth century, Muslims had answered the questions that they had confronted in the decades after the Prophet's death. They had come up with criteria for determining who could serve as Muhammad's successor and had outlined the extent of their powers. They had also established how Muslims were to live and how they were to practice the religion. They had even identified how human reason was to be used in exploring God's plan. What they had not done, however, was come to a consensus in addressing those questions; instead, they had arrived at a multiplicity of conflicting answers. In the process, they had divided along a variety of fault lines:

⁹⁷ Esposito, *Islam*, 100–102; Armstrong, *Islam*, 104–5.

⁹⁸ Fawaz A. Gerges, *Making the Arab World: Nasser, Qutb, and the Clash That Shaped the Middle East* (Princeton: Princeton University Press, 2018), 250–51; Ansary, *Destiny Disrupted*, 162–64; Nasr, *The Shia Revival*, 95–97.

⁹⁹ Esposito, *Unholy War*, 45–46.

between behaviorally minded Sunnis and devotional Shi'a, between Sufi mystics and legal traditionalists, and between 'ulama' literalists and intellectual *faylasufs*. At times, those divisions undoubtedly weakened the caliphate. On balance, however, their impact was anything but negative. On the contrary, the divisions that emerged among Muslims during this pivotal era played an enormous part in ensuring that the *umma* developed into a rich, complex, and heterogeneous civilization—one that continued to thrive even as the caliphate was crumbling around it. It is to the gradual weakening and eventual collapse of the empire that we shall next turn.

Chapter Five: The Decline of the Arab Empire, 809-1300

The transformation of Islam and the intellectual trends that we examined in chapter four occurred against an increasingly anarchic political backdrop. While the Arab Empire had enjoyed an almost-unbroken string of success since its founding in the early-seventh century, it abruptly entered a more challenging period immediately following the reign of Harun al-Rashid (r. 786-909). For the next century and a half, the caliphate experienced a sequence of debilitating civil wars, factionalism, and instability. As a result, the Abbasid Dynasty gradually saw its authority shrink until, in 945, it fell under the dominance of an Iranian dynasty, the Buyids (934-1055).

Worse followed. From the end of the eleventh century to the late-thirteenth century, a succession of Turkish nomadic groups, crusading Western knights, and, most dangerous of all, powerful Mongol warriors launched a series of sustained attacks on the Arab Empire. These invasions had enormous consequences for the Middle East. They produced a fundamental political reordering in the region, resulted in a legacy of misunderstanding over the nature and intent of the Crusades and, in the case of the Mongols, seemed for a time to constitute an existential challenge to Islamic rule.

The Withering of Abbasid Power

Civil War

Ironically, the decline of the Abbasid Dynasty began during the reign of the caliph widely viewed as the Arab Empire's greatest leader, Harun al-Rashid. Indeed, the idea that the caliphate was on the cusp of difficult times would have struck contemporaries as entirely implausible. After all, at the time of al-Rashid's death in 809, the empire was in an enviable position in nearly every regard. It had a strong, centralized system of administration, broad legitimacy throughout the Muslim world, and, perhaps most importantly, a powerful army centered on the *Abna*, the descendants of the Khurasani Arabs who had supported the Abbasid revolt against the Umayyad Dynasty. Harun al-Rashid had also left it with a treasury of enviable proportions owing to a steady stream of tax revenues from both the provinces and, most importantly, the Sawad: the fabulously rich agricultural land of southern Iraq that the Abbasids controlled directly.¹

In one critical way, however, al-Rashid had set the stage for the weakening of Abbasid power. As was typical of caliphs at that time, he had a number of sons by different mothers, and each could advance a claim to succeed him. In many ways, this was a wise practice. Infant mortality rates were staggeringly high in pre-modern times, so having multiple potential successors was a prudent hedge against the possibility that all legitimate claimants to the throne might pre-decease the reigning caliph—a development that could, in turn, set the stage for a disputed succession and attendant

¹ Hugh Kennedy, *When Baghdad Ruled the Muslim World: The Rise and Fall of Islam's Greatest Dynasty* (Cambridge, MA: Da Capo Press, 2004), 81–82, 130–32.

political instability. Al-Rashid's approach avoided that danger. In doing so, however, it ran a different risk: having a large number of legitimate successors threatened to open the door to an alternate kind of crisis in which half-brothers jockeyed for the succession or, worse, followed the caliph's death by engaging in outright civil war.²

Aware of these risks, Harun al-Rashid took two steps to ensure a smooth succession. First, he publicly declared that his son, al-Amin (r. 809-813), would succeed him as caliph. This decision was a prudent one that appeared to lay to rest the question of which son would accede to power upon al-Rashid's death. His second action was considerably less sensible, however. He followed by announcing that another son by a different mother, al-Mamun (r. 813-833), would concurrently become the governor of the wealthy, autonomous province of Khurasan and, critically, would become al-Amin's successor. Harun al-Rashid was not blind to the dangers inherent in this unorthodox arrangement. To ensure that al-Amin and al-Mamun abided by it, he required them to sign the Mecca Protocol in which they pledged before God that they would respect his wishes. This agreement may have satisfied the caliph, but it did little to assuage the concerns of those who doubted the wisdom of his convoluted succession scheme. Indeed, many astute observers predicted that his decision to name an heir to an heir would "cast dissension among his sons" and produce a civil war.³

They were not wrong. Though both al-Mamun and al-Amin appeared initially to be abiding by their father's wishes, competing factions soon began to push the brothers into conflict. The growing dissension revolved around control of the lucrative tax revenues of Khurasan. The *Abna* lived in Baghdad and supported al-Amin; they wanted the tax revenue of their ancestral homeland to continue to come to the capital as it had since the Abbasids had overthrown the Umayyads in 750. Opposed to them were the aristocrats of Khurasan; they sought greater autonomy from Baghdad and wished to retain the ample tax receipts that the province generated. Engaged in a zero-sum contest, these competing groups pushed al-Amin and al-Mamun to take their sides in the conflict over the revenue of Khurasan until they had succeeded in poisoning the brothers' relationship. Finally, in late 810, al-Amin irrevocably broke with al-Mamun by having his son, Musa, named as his successor and by burning the agreement that he and his brother had signed. The civil war had begun.⁴

The course of the conflict was very different than anyone had expected. Initially, both sides assumed that they were fighting to determine if al-Mamun would be able to retain control of Khurasan. The general Tahir ibn Husayn's (c. 775-822) stunning defeat of al-Amin's invading army in central Iran in 811 dramatically changed the complexion of the war, however. Thereafter, the two sides fought not over whether al-Mamun would continue to control Khurasan, but instead over the question of which brother would reign

² Hugh Kennedy, *The Prophet and the Age of the Caliphates*, Second Edition (New York: Routledge, 2004), 142–43.

³ Kennedy, *When Baghdad Ruled the Muslim World*, 68–71.

⁴ Chris Wickham, *The Inheritance of Rome: Illuminating the Dark Ages 400-1000* (New York: Penguin Books, 2009), 322–23.

as caliph. Still, even after his army's defeat, al-Amin appeared to be in a substantially stronger position than al-Mamun. He had much deeper financial and military resources than his brother, and he retained the loyalty of the *Abna*—the veterans who had constituted the core of the Abbasid army for the prior seven decades.⁵

Nonetheless, despite al-Amin's advantages, al-Mamun prevailed in the civil war. Three events tipped the contest in his favor. First, in a lightning campaign, Tahir seized control of the Sawad region and thus denied al-Amin what had been his main source of revenue. Second, rifts emerged between the *Abna* and al-Amin's other Arab supporters that paralyzed his effort to coordinate the resistance to his brother's army. Finally, growing awareness that al-Amin was responsible for the rupture and had ordered his subordinates to rip up the Mecca Protocol cost him nearly all his remaining popular support. As a result, by the time that Tahir's army surrounded Baghdad in August 812, al-Mamun had secured recognition as the rightful caliph everywhere but the capital.⁶

Still, the war dragged on. Indeed, it was only after a grueling, thirteen-month siege that Tahir's forces were finally able to breach the defenses and take control of the now-devastated city. In the process, al-Amin was killed while trying unsuccessfully to arrange his surrender.⁷

To a degree, al-Mamun's triumph turned out to be a pyrrhic one. It is true that he had defeated his brother and taken control of the throne; in the process, however, his conflict with al-Amin had grievously weakened the Abbasids' grip on the empire. Most obviously, the dynasty had suffered its first meaningful loss of territory when the governor of the distant North African province of Tunisia, Abdallah ibn Ibrahim (r. 812-817), took advantage of the chaos of the civil war to secure de facto independence for the Aghlabid Dynasty (812-909). More importantly, al-Mamun's inevitable decision to disband the *Abna* army had left the Abbasids without a reliable military force of their own. As a result, the new caliph had no choice but to enter into a relationship with the Tahir family to secure its continued military support. According to the arrangement they worked out, the Tahirids would rule Khurasan and much of the eastern part of the caliphate as autonomous, junior partners of the ruling dynasty; in exchange, they would provide the caliphate with the military force needed to maintain control of the empire.⁸

In the short term, this arrangement proved to be mutually beneficial. While Tahir assumed control of Khurasan and its abundant tax revenue, his capable son, Abd Allah (c. 798-c. 845) quickly restored caliphal authority over Syria, Egypt, and other parts of the empire that had broken away from the central government during the civil war. At the same time, however, dependence on the Tahirids for military support put the

⁵ Kennedy, *When Baghdad Ruled the Muslim World*, 90–95.

⁶ Kennedy, 100–102.

⁷ Kennedy, *The Prophet and the Age of the Caliphates*, 148–49.

⁸ Kennedy, *When Baghdad Ruled the Muslim World*, 202–3.

Abbasids in a very precarious position: without their own troops, they were vulnerable to being displaced in the future by their militarily powerful junior partners.⁹

Al-Mutasim

Al-Mamun's younger brother, al-Mutasim (r. 833-842), came up with an innovative and effective solution to this dilemma. In the decade following al-Amin's defeat, he began to purchase Turkish slaves and to organize them into an army. The Turks made excellent soldiers. Hailing from Central Asia, they were a tough, nomadic people who were highly skilled in the sophisticated military tactics of the steppes. It is important to understand that they were not slaves in the traditional sense. On the contrary, though legally considered to be the property of the caliph, they lived as free men of high status and earned an enviable salary.¹⁰

Disciplined, powerful, and, above all, highly loyal, al-Mutasim's small but rapidly growing army paid immediate dividends for the dynasty. It proved vital in helping al-Mamun to restore caliphal authority to the empire in the wake of the civil war. More importantly, as the only military formation that the Abbasids directly controlled, it gave the dynasty a force that could balance the Tahirids' powerful army—ensuring, in the process, that the ruling family would not find themselves pawns of their junior partners. Indeed, al-Mamun perceived the new army to be of such critical importance to the dynasty's independence that he felt compelled in the late 820s to name al-Mutasim rather than his own son, al-Abbas, as his successor.¹¹

The new force of Turkish slave soldiers played a large part in impelling al-Mutasim to make the fateful decision to relocate the seat of government soon after he became caliph in 833. The proximate cause of the move was the bitter resentment that the *Abna* felt toward the Turks who had replaced them in the caliph's army. Angry at their loss of social and economic status, they had begun to vent their frustration by assaulting and even murdering Turkish soldiers in the streets of the capital. Less immediate but equally important in spurring the move was the fact that Baghdad was already fully built out and thus lacked the room needed to house and train his new troops. Accordingly, in 835, al-Mutasim decided to move the army, palace, and bureaucracy to a new city, Samarra, located 130 kilometers north of Baghdad along the banks of the Tigris River.¹²

⁹ Ira M. Lapidus, *A History of Islamic Societies*, 2nd ed. (Cambridge: Cambridge University Press, 2002), 103–4; Kennedy, *When Baghdad Ruled the Muslim World*, 202–3.

¹⁰ Albert Hourani, *A History of the Arab Peoples* (Cambridge, Mass: Belknap Press of Harvard University Press, 2002), 116–17.

¹¹ Matthew S. Gordon, *The Breaking of a Thousand Swords: A History of the Turkish Military of Samarra* (Albany: State University of New York Press, 2000), 15–44.

¹² Gordon, 50–55.

At first, the shift to Samarra was an unalloyed success. The move ended the attacks on al-Mutasim's Turkish soldiers and provided him with ample land on which to house and train his new army. The benefits of the relocation were not limited to the military, moreover. On the contrary, moving the capital to Samarra paid substantial political dividends as well. Most notably, doing so gave al-Mutasim access to a new source of patronage that he could use to maintain the support of key aristocrats: developable land. Prior to the relocation, he had acquired at essentially no cost title to all of the property in and around Samarra, which, by virtue of its proximity to the new seat of power, was now of inestimable value. He then used it to secure the continued support of the elite by carefully doling it out to his most loyal supporters. In other words, as the historian Hugh Kennedy points out, the construction of the city of Samarra amounted to a "gigantic property speculation in which both government and its followers could expect to benefit."¹³

Anarchy in Samarra

The advantages inherent in the move to Samarra proved to be transitory, however. By the 860s, successive Abbasid caliphs found themselves virtual prisoners in their new capital. What had happened? What accounted for this dramatic change in circumstance? At root, the Abbasids' worsening situation in Samarra was a function of the seemingly inevitable intrusion of the slave-soldiers into politics. The trouble began in 861 when a group of Turkish leaders learned that Caliph al-Mutawakkil (r. 847-861) was planning to depose them and seize their wealth. In desperation, they moved to defend their position by murdering him at a drinking party and by replacing him with his son, whom they believed they could more easily control. This event was a watershed moment for the Abbasids. It initiated a nine-year period of anarchy in Samarra wherein successive caliphs, isolated from their supporters, could neither assert their control over the government nor arrest the growing power of the slave-soldiers.¹⁴

Unsurprisingly, the chaos did not remain limited to Samarra. Instead, it sparked a vicious cycle that bred yet further disorder. Instability in the capital allowed warlords and strongmen on the periphery of the empire to carve out independent states—in the process denying the central government the tax receipts it needed to meet the demands of the increasingly restive soldiers. Frustrated by their inability to get the money they felt they deserved, the Turks would respond by violently deposing the reigning caliph in favor of a candidate who they believed would be more pliable and therefore more likely to pay them. The rapid change in leadership would then further weaken the central government and permit even more regional leaders to secure autonomy, thus perpetuating the cycle.¹⁵

The result was anarchy on a grand scale. No fewer than four men sat on the throne between 861 and 870 as the Turks made and unmade caliphs with abandon.

¹³ Kennedy, *The Prophet and the Age of the Caliphates*, 162–63.

¹⁴ Gordon, *The Breaking of a Thousand Swords*, 88–90.

¹⁵ Kennedy, *The Prophet and the Age of the Caliphates*, 169–73.

Worse, all four died violently at the hands of their slave soldiers when they proved unable to come up with the revenue that the Turks demanded. Thus, by the late 860s, the Abbasids found themselves in dire circumstances. They had lost authority over large parts of the empire and had become prisoners of their angry and rebellious troops.¹⁶

The Abbasid Revival

That Abbasid rule continued after that point owed largely to the efforts of one man: prince al-Muwaffaq (842-891). Though he never formally held the title of caliph, he served as the de facto regent and the power behind the throne from 870 until his death in 891. In that capacity, he took several critical actions to restore the dynasty's position. First, he moved to revive Abbasid fortunes by placating the soldiers. Having a military background and a strong relationship with the army, al-Muwaffaq was able to end the anarchy in Samarra and bring the Turks back into the fold by reassuring them both that they would retain their position as the core of Abbasid military power and that the government would pay their salaries on a regular basis. The restoration of order in Samarra and subsequent relocation of the government back to Baghdad set the stage for his second action: the reestablishment of Abbasid authority over the empire's many rebellious provinces and breakaway states. In this effort, al-Muwaffaq enjoyed substantial success. In 870, the dynasty had controlled little more than northern and central Iraq; by the time of his death in 891, in contrast, it had regained control over western Iran, most of Syria, and southern Iraq.¹⁷

The Abbasids were never able to reassert the central government's authority over the rest of the empire, however. Taking advantage of the paralysis in Samarra during the 860s, rebels, outsiders, and, especially, regional governors had carved independent fiefdoms out of the empire's peripheral territory. In 868, for instance, the governor of Egypt, Ahmad ibn Tulun (r. 868-884), had established the independent Tulunid Dynasty (868-905); a few years later, Ismail Samani (r. 892-907) had followed suit by consolidating the existing Samanid Dynasty (819-1005) in the eastern province of Transoxiana. More troublingly, in 861, a rogue coppersmith named Yaqub al-Saffar (r. 861-879) had established the Saffarid Dynasty (861-1003) when he took control of the province of Sistan, located along the border between modern-day Iran and Afghanistan. Exploiting the instability in the capital, he moved quickly to acquire additional territory in the east. Most notably, al-Saffar was able to take possession of the Tahirid capital of Nishapur in 873—in the process ending the fifty-year long partnership between the Tahirids and the Abbasids.¹⁸

It is important to note that none of these states publicly declared that they were independent of the central government. Both regional leaders and caliphs alike saw value in upholding the fiction that the local dynasties ruled as subordinates of the Abbasids. Regional rulers such as the Samanids acquired legitimacy through this

¹⁶ Gordon, *The Breaking of a Thousand Swords*, 90–91.

¹⁷ Gordon, 141–43.

¹⁸ Hourani, *A History of the Arab Peoples*, 38.

arrangement, while the caliphs enhanced their much-diminished status by being able to claim authority over their nominal subordinates' territories.¹⁹

In reality, of course, caliphal power had diminished greatly. Claims to the contrary aside, the local dynasties that emerged in the late-ninth and early-tenth centuries were independent states in all but name. Indeed, real caliphal authority never again extended beyond Iraq, western Iran, and parts of Syria.²⁰

Problematically, moreover, the revival that al-Muwaffaq had spearheaded began to slow and then regress in the early tenth century. Fiscal issues accounted for this reversal of fortune. A combination of mismanagement, adverse environmental change, and the destruction of critical irrigation works in the civil wars and revolts of the ninth and early-tenth centuries had dramatically reduced agricultural productivity in the vital Sawad and thus the tax yield on which Abbasid rule depended. To offset the decline of this critical revenue stream, the government resorted to fiscal expediciencies such as tax farming: the already venerable practice of selling the right to collect state revenue in a province to private individuals. In the short run, this arrangement benefited the government by supplying the treasury with an upfront infusion of cash; since the practice encouraged the tax farmers to squeeze as much revenue out of the province as they could, however, it depressed economic activity in the long run and thus came at the cost of substantially diminished future tax receipts.²¹

By the 920s, the bill for the caliphate's fiscal imprudence had come due. Put simply, the regime lacked the revenue needed to maintain the cost of the palace and the army. The result was renewed chaos in the capital. Now compelled to get by with greatly diminishing resources, the central government experienced a new round of factional infighting and civil war capped by the murder of the feckless Caliph al-Muqtadir (r. 908-932). Nominal subordinates of the central government were quick to pounce. Taking advantage of the paralysis in Baghdad, provincial governors and others established de facto independent states in Syria, northern Iraq, and western Iran. By the early 930s, as a result, the caliph controlled little more than the capital and its immediate surroundings.²²

The Buyids

The end of caliphal autonomy finally came in 945. That year, the Buyids, a new and dynamic Shi'i dynasty that originated on the south shore of the Caspian Sea, took advantage of the instability in Baghdad to seize control of the capital. Though Buyid dominance of the caliphate was comparatively brief, their writ was substantial. Indeed, at its peak, the dynasty dominated the Muslim world—directly controlling Iraq, east and

¹⁹ Hugh Kennedy, *Caliphate: The History of an Idea* (New York: Basic Books, 2016), 87.

²⁰ Kennedy, 87.

²¹ Lapidus, *A History of Islamic Societies*, 106–11.

²² Kennedy, *Caliphate*, 87–88.

central Iran, and large parts of Syria, and exerting substantial influence in nearby regions.²³

Though its rule did not last very long, moreover, the Buyid Dynasty would have an enduring impact on other medieval Muslim states. Three of its governing practices would prove to be especially influential. First, aware that they lacked broad public support because of their adherence to Shi'ism, they opted not to depose the Abbasids but instead sought to legitimate their rule by pretending to govern in the name of the caliph. The Abbasids thus continued to reign as figureheads—as they would under future regimes—while a succession of Buyids wielded real power under the title *amir al-umara*, or Commander of the Commanders. Second, building on fiscal changes that had begun in the caliphate in the ninth century, the Buyid Dynasty put an end to the Abbasid tradition of having a central bureaucracy direct the affairs of the state and oversee the collection and distribution of taxes. Instead, it adopted a highly decentralized, tax farming arrangement wherein it divided its territory into semi-feudal land grants called *iqtas* that it then sold to military leaders. *Iqta* holders were obliged to provide military service to the state and to forward a reduced share of the tax revenue they collected to Baghdad; in exchange, they retained the bulk of the taxes that their *iqta* generated. Finally, the Buyids expanded and institutionalized the practice of recruiting slave soldiers that the Abbasids had begun.²⁴

The Buyids' semi-feudal, decentralized political system rapidly caught on. Other contemporary dynasties in the Middle East and Central Asia such as the Ghaznavids (977-1186) quickly adopted it, and it soon became the basic political structure that all subsequent medieval Muslim states would adopt. That it proved so influential is perhaps surprising in light of the fact that it had significant shortcomings. Three stand out. First, predictably, the decentralization of power produced internecine conflict among military leaders that helped to undermine the stability of the Buyid state. Second, the *iqtas* proved to have the same costs and benefits typical of other tax-farming systems: while they provided the central government with a short-term infusion of revenue, they did so at the expense of long-term income. Finally, because the *iqtas* were revocable, the system encouraged those who held them to pursue quick profits at the expense of long-term agricultural productivity—a change that weakened the economy in a number of areas and that accelerated the increasingly precipitous decline of the Iraqi economy that had begun in the ninth century.²⁵

The Fatimid Challenge

While the Buyids downplayed their Twelver beliefs to mollify their Sunni subjects, another Shi'i dynasty, the Fatimids, initially flaunted their religion and explicitly declared their intention to replace the Sunni Abbasid Caliphate with an Isma'ili Shi'i one. This new dynasty originated in the late-ninth century with a Syrian man named Ubayd Allah

²³ Kennedy, *The Prophet and the Age of the Caliphates*, 210–34.

²⁴ Kennedy, 210–34.

²⁵ Lapidus, *A History of Islamic Societies*, 114.

(r. 909-934). Claiming descent from Imam Muhammad ibn Isma‘il (c. 740-813)—and, thus, from Ali ibn Talib (r. 656-661) and the dynasty’s namesake, Muhammad’s daughter Fatima (605-632)—he declared publicly his intention to replace the Sunni caliphate with a Shi‘i one. Finding Syria inhospitable due to its close proximity to the Abbasids’ seat of power in Iraq, he sent emissaries to the fringes of the Muslim world to search for an appropriate place to start his revolution. He found it in Tunisia. The Berbers who lived in the interior of that region were both receptive to Isma‘ili Shiism and—as important—resentful of the ruling Aghlabid Dynasty. Accordingly, Ubayd Allah left Syria in the early-tenth century for Tunisia, where, after a brief campaign, his Berber troops succeeded in overthrowing the Aghlabids. He followed by assuming the title of Mahdi and by declaring the establishment of the Fatimid Caliphate.²⁶

Tunisia was a good starting point for Ubayd Allah, but it was too far away from the heart of the Muslim world to launch an effective challenge to the ruling caliphate. The Fatimids consequently set their sights on conquering the wealthy and populous province of Egypt, from which they could more effectively contest Abbasid dominance. Early efforts were unsuccessful, and for a time the dynasty appeared to be fated to remain little more than a bit player on the periphery of the caliphate. The Fatimids’ situation changed quickly in the 960s, however. A combination of political instability and famine had grievously weakened the Ikhshidid Dynasty (935-969) that had ruled Egypt since 935. Astutely sensing the province’s weakness, the Fatimid general Jawhar (?-992) seized Egypt in a nearly bloodless *coup de main* in 969. He followed his victory by ordering the construction of a new capital, Cairo. Four years later, the Fatimid caliph, al-Muizz (r. 953-975), relocated his government to the city.²⁷

The Fatimid challenge to the Abbasids was different in both kind and degree than that of the other regional powers that had emerged in the Muslim world. As we have seen, the dynasties that had established their independence in the ninth and tenth centuries had nearly all sought to legitimate their rule by maintaining the pretense that they governed in the name of the caliph in Baghdad. In contrast, the Fatimids forthrightly proclaimed that they intended to displace the Abbasids as the leaders of the Muslim world.²⁸

As a result of their astute economic policies, they appeared for a time to be poised achieve that ambitious goal. Most notably, a crash program of infrastructure construction and the suppression of pirates allowed the Fatimids to shift the western terminus of the lucrative, long-distance Indian Ocean trade from Iraq to Egypt—a change that produced a dramatic increase in tax receipts that the dynasty was in turn able to use to strengthen its position. The bulk of the new revenue went to the expansion of the army and navy, which gave the Fatimids the strength to seize control of the Hijaz and most of Syria. They did not spend all of their fiscal windfall on the

²⁶ Wickham, *The Inheritance of Rome*, 336–37.

²⁷ Marshall G. S. Hodgson, *The Venture of Islam, Volume 2: The Expansion of Islam in the Middle Periods* (Chicago: The University of Chicago Press, 1974), 23.

²⁸ Kennedy, *Caliphate*, 188–202.

military, however; instead, they devoted a substantial share of their newfound tax wealth to the goal of burnishing the dynasty's image in the Muslim world. Most notably, the Fatimids constructed the al-Azhar Mosque, a magnificent center of learning in Cairo, and moved to ensure the safety of the hajj pilgrims by paying off the Bedouin tribes that had been attacking the caravans traveling to Mecca. For a time, these initiatives were enormously successful. Indeed, by the year 1000, the Fatimids seemed to be well on their way to displacing the Abbasids as the rulers of a universal caliphate.²⁹

The dynasty's success in the tenth century did not continue in the eleventh, however. Instead, thanks to a combination of internal and external challenges, the Fatimids' efforts to supplant the Abbasids came to a standstill after the year 1000. Three issues combined to stymie the dynasty. First, the Fatimids found themselves limited by the enduring appeal and deep roots of Sunni Islam. Isma'ili missionaries did enjoy some success in spreading Shi'ism in Syria, but, on balance, they gained little traction among the Sunni majority that dominated the heart of the Islamic world. Ironically, they had even less success in Egypt. There, however, the failure to win converts was less a function Shi'ism's lack of appeal than of the Fatimids' desire to avoid taking actions that might spark dissension at home. Second, early in the eleventh century, the Fatimids adopted a slave-based military patterned on the system that the Buyids employed. The Turkish and Black-African slave soldiers they purchased fought well and ensured the dynasty's continued military success; however, they also undermined the regime from within—first when they began to engage in debilitating factional intrigue and then, in the 1060s and 1070s, when they waged outright civil war among themselves. Finally, beginning in the mid-eleventh century, the Fatimids began to face a new and serious challenge in the Muslim world in the form of the dynamic and powerful Turkish Seljuq Empire.³⁰

The Seljuq Turks

The militarily powerful Seljuq Dynasty would profoundly alter the Middle East despite the fact that it dominated the region for only a relatively brief period. It had its origins among the powerful Oghuz Turks who lived east of the Aral Sea. Under the command of the dynasty's namesake, the warlord Seljuq (?-c. 1009), a group of Oghuz Turks migrated south in the late-tenth century to Transoxiana where they adopted Sunni Islam. Under Seljuq's capable grandsons, Tughril (r. 1037-1063) and Chaghri (r. 1040-1060), the dynasty followed by conquering Khurasan and much of Iran before seizing Baghdad from the tottering Buyids in 1055. Continued expansion westward soon brought them into conflict with the formidable Fatimid Caliphate and the venerable Byzantine Empire. The Seljuqs had no more difficulty against those powers than they had against the Buyids, however. On the contrary, under Alp Arslan (r. 1063-1072) and his dynamic son, Malik Shah (r. 1072-1092), the Sunni Seljuqs drove the Fatimids out of Syria and wrested control of Anatolia from the Byzantine Empire. As a result, by the

²⁹ Kennedy, 188–202.

³⁰ Lapidus, *A History of Islamic Societies*, 286–87.

time of Malik Shah's death, the Seljuq Empire stood without question as the most powerful state in the Middle East.³¹

Seljuq success was not entirely without problems, however. The rapid expansion into the Muslim homeland created significant cultural friction between the Arabs and the Turkish newcomers. This situation may sound surprising in light of the fact that the Arabs had substantial experience interacting with Turkish slave soldiers. The enslaved Turks who served in the military had been relatively small numbers, however, and were generally taken as children; they thus had little connection with the culture into which they had been born and generally adapted to the customs of the Arab majority. Retaining the cultural practices of Turkish Central Asia and adhering to Islam in only the broadest sense, the Seljuq Turks were a different story altogether. In the eyes of educated, urban Arabs, the newcomers were violent, undisciplined boors who knew little of either Islamic law or the finer points of culture; as a result, Seljuq rule produced rumbles of discontent among the empire's Arab majority.³²

In response, the Seljuq Dynasty made a deliberate effort to ingratiate itself with its Arab subjects. This campaign had three components. First, taking a page from the Buyids, the dynasty sought to legitimate Turkish rule by retaining the Abbasids as figureheads. Tughril thus did not declare himself caliph but instead adopted the title of sultan; a derivative of the Arabic word for "authority," it indicated that the dynasty accepted the continuation of the caliphate while simultaneously making clear who held real power. Second, in an effort to win over the influential religious scholars, the Seljuqs signaled their embrace of education, culture, and Sunni orthodoxy by constructing and supporting religious colleges called madrasas. Finally, to overcome the Turks' reputation for crudeness and violence, the new rulers restored order and rebuilt damaged infrastructure. These initiatives were successful because, coupled with the Seljuqs' ability to establish direct control over most of the caliphate, they satisfied the Sunni Arabs' desire for continuity and orthodoxy and, importantly, seemed to suggest that the Seljuqs were in the process of reestablishing the universal empire that, since the ninth century, had existed only in name.³³

Ultimately, that conclusion proved to be inaccurate. The Seljuq family would continue to play a significant role in the Middle East for another century and a half following Malik Shah's rule, but it did not oversee the restoration of a genuinely unified caliphate. Instead, shortly after Malik Shah's death, the Seljuq state fragmented into a series of squabbling principalities. What caused the sudden splintering of the Seljuq Empire at a time when it faced little external pressure or internal unrest? Multiple factors played a role in its collapse, but the dominant one by far was the Turkish system of inheritance that the Seljuqs had brought with them from the steppes. Rather than practicing primogeniture wherein a single successor inherited the entirety of a political

³¹ Osman Aziz Basan, *The Great Seljuqs: A History* (Routledge, 2010), 53–54, 59–96.

³² Tamim Ansary, *Destiny Disrupted: A History of the World Through Islamic Eyes* (New York, NY: PublicAffairs, 2009), 127–30.

³³ Ansary, 127–32.

entity, the Turks followed a system of appanage in which families considered their holdings to be collectively held but individually administered. This arrangement functioned well enough when the dynasty was controlled by powerful figures who had the authority and respect needed to assign territories among their relations and to compel them to follow their commands. However, it did not work when the ruling family lacked a forceful leader. Such was the case following Malik Shah's death in 1092. When no dominant figure emerged to oversee the Seljuq realm, it quickly fractured among his brothers, sons, nephews, grandsons, and generals into a collection of petty, warring states.³⁴

The Caliphate Divided

The breakup of the Seljuq Empire at the end of the eleventh century marked the culmination of the process of devolution and division in the caliphate that had begun during the civil war between al-Amin and al-Mamun. Those changes were far-reaching and resulted in political, economic, and social environments that were dramatically different from those which had existed at the start of the ninth century. The political changes were the most obvious. While the idea of the universal caliphate endured, the *umma* was no longer unified; instead, it was divided into numerous competing entities including no fewer than three dynasties that claimed to be the legitimate rulers of the Muslim world: the Abbasids in Baghdad, the Fatimids in Cairo, and, after, 929, the revived Umayyads in Spain. Significant social changes paralleled the political decline of the caliphate. Most notably, the slave-recruited Turkish troops that al-Mutasim had introduced and that the Buyids had popularized had gradually displaced the Arab elite of bureaucrats, *Abna* soldiers, and landowners that had dominated the empire during the early years of the Abbasid Dynasty. Finally, civil war, instability, and the spread of the *iqta* system did grievous damage to large parts of the Middle Eastern economy with Iraq suffering the greatest decline. It was, in sum, a time of volatility and conflict—one that contrasted sharply with the stability, security, and prosperity that had existed at the end of Harun al-Rashid's reign.³⁵

The Crusades

That the Middle East had fragmented politically toward the end of the eleventh century was a stroke of good fortune for the crusading European knights who were about to mount a series of attacks on the region. Derived from the Latin word *crucesignatus*, meaning "signed with the cross," the term crusade denoted an armed, penitential pilgrimage organized and legitimated under the aegis of the Medieval Catholic Church. An innovation of the late-eleventh century, crusading was a broad phenomenon in both spatial and temporal terms. Indeed, over the next few centuries, Western European knights would undertake crusades in places as disparate as Eastern Europe, Spain, Italy, and the south of France. Throughout the period, however, the

³⁴ Lapidus, *A History of Islamic Societies*, 117–19.

³⁵ Lapidus, 108–11.

campaigns waged against the Muslims of the Middle East remained by far the most important crusades.³⁶

Origins

The Crusades were very much a product of the culture of Late Medieval Europe. They originated in the last decade of the eleventh century when a particular sequence of events, ideals, and needs came together to make them possible. While many of these factors were intrinsic to the culture of Western Europe, the proximate cause of the Crusades—the troubled situation in which Byzantium suddenly found itself in the last third of the eleventh century—was extrinsic. To understand the impetus for the Crusades, therefore, we must now briefly review the history of the Byzantine Empire in the centuries that followed the Arab Conquest.

For some time after the Arabs had irrupted from the desert, the Byzantines had confronted a remarkable combination of problems. Indeed, over the ensuing two centuries, they experienced what surely must have seemed to contemporaries like an endless series of political, economic, religious, and—especially—military crises. By the early ninth century, as a result, the once might Eastern Roman Empire had been reduced to a weak, regional power that controlled little more than Anatolia and a few scattered possessions in the Balkans.³⁷

Not long thereafter, however, the dark times passed, and the Byzantine state began to recover. Its revival was multifaceted. Militarily, the maturation of the Theme System—a decentralized, semi-feudal structure of administration and recruitment established in the seventh century—had produced a capable, well-motivated army composed of farmer-soldiers who safeguarded their local provinces. Highly effective at defense, the Theme System military gave the empire the ability to reassert control of the southern Balkans and to stabilize the frontier with the Abbasid Caliphate. Politically, meanwhile, the establishment of the Macedonian Dynasty (867-1056) initiated a nearly two-century-long period in which a sequence of capable emperors provided effective leadership in Constantinople. Economically, finally, the restoration of secure borders and the reestablishment of political stability set the stage for a period of sustained growth that dramatically improved the fiscal position of the state.³⁸

With the empire on a solid footing for the first time since the start of the seventh century, its leaders moved to recover some of the territory that their predecessors had lost. The founder of the Macedonian Dynasty, Basil I (r. 867-886), initiated this effort with his reconquest of southern Italy. Over the next century and a half, a succession of emperors followed suit by using powerful new, heavy-cavalry units called *tagmata* to

³⁶ Christopher Tyerman, *Crusades: A Very Short Introduction*, Very Short Introductions (Oxford: Oxford University Press, 2005), 12.

³⁷ Wickham, *The Inheritance of Rome*, 261–78.

³⁸ Wickham, 261–78.

regain control of Crete, Cyprus, southeast Asia Minor, and northern Syria—including the important city of Antioch.³⁹

Under Basil II (r. 976-1025), Byzantium's revival reached its apogee. His defeat of the Bulgars extended the empire's border to the Danube River for the first time in four centuries and gave Constantinople uncontested control of a belt of territory stretching from southern Italy in the west to the Euphrates River in the east. At home, meanwhile, his rule coincided with a time of economic growth, political stability, and religious harmony. Indeed, his reign capped what had been a remarkable comeback—one that left Byzantium more stable, unified, and secure than it had been at any point since Justinian the Great (r. 527-565) had sat on the throne in the sixth century.⁴⁰

The empire's revival came to an abrupt halt in 1071, however. At the Battle of Manzikert fought that year in eastern Anatolia, Alp Arslan's Turkish warriors not only defeated the Byzantine army but also captured Emperor Romanus IV Diogenes (r. 1068-1071). The engagement was neither a complete rout as earlier generations of historians had once described it, nor a defeat from which the empire could not recover; after all, it had weathered similar setbacks in the past. For two reasons, however, Manzikert turned out to be a fundamental turning point for the Byzantine state. First, the growing emphasis on the mobile, *tagmata* forces in the tenth and eleventh centuries had come at the expense of the Themes System army. By 1071, as a consequence, the empire no longer had the capable local military formations that had proven so effective at defending Byzantine territory in the eighth and ninth centuries and thus lacked the forces needed to prevent Seljuq troops from exploiting their victory. Second, the struggle to replace Romanus IV Diogenes failed to produce a quick winner; instead, it dragged on for a decade as Byzantine factions waged a vicious, zero-sum contest for control of the throne.⁴¹

This divisiveness was a gift to the Seljuqs. With the Byzantine elite focused so intently on the internal contest for power, the empire's borders—particularly its eastern frontiers—lay all-but undefended. Alp Arslan and his successor, Malik Shah, were well aware of the situation, and wasted no time in taking advantage of it to grab imperial territory. By the early 1090s, as result, the Seljuqs had taken control of nearly all of Anatolia—the territory that for centuries had constituted the Byzantine Empire's beating heart.⁴²

Despite the dire situation, the empire somehow managed to stage yet another comeback. It began with the accession of Alexius Komnenus (r. 1081-1118) to the throne in 1081. The victor in the struggle for power that had marked the 1070s, he was

³⁹ Warren Treadgold, *Byzantium and Its Army, 284-1081* (Stanford: Stanford University Press, 1998), 28–39.

⁴⁰ Robert Browning, *The Byzantine Empire* (Washington, D.C: Catholic University of America Press, 1992), 114–17.

⁴¹ Treadgold, *Byzantium and Its Army, 284-1081*, 39–42.

⁴² Basan, *The Great Seljuqs*, 78–88.

determined to restore the empire and to recover the territory it had lost. That task would not prove to be an easy one. By 1081, Byzantine weakness had drawn the attention of a number of predatory states; as a consequence, Alexius was compelled to deal with serious crises on nearly every front. Not only were the Seljuqs continuing to chip away at his territory in Anatolia, but Normans under Robert Guiscard (c. 1015-1085) and his son, Bohemund (c. 1054-1111) had launched an invasion of Greece the same year that he had become emperor. Worse, powerful Pecheneg nomads invaded the northern Balkans in 1090. The situation was grim to say the least.⁴³

Fortunately for the empire, Alexius was a peculiarly capable leader who had the *savoir faire* and cunning to see it through this difficult time. Working tirelessly, he moved to restore the army by revamping military recruitment and by reforming the system of tax collection, and he pursued well-designed diplomatic initiatives designed to secure allies for the now-weakened empire. These efforts were successful. With critical assistance from the Venetians, his restructured army managed to drive the Normans out of Greece in 1085. Meanwhile, after striking an alliance with another nomadic group, the Cumans, he decisively defeated the Pechenegs in 1091 and restored Byzantine control of the Balkans.⁴⁴

With the situation stabilized, Alexius next set his sights on regaining control of Anatolia. Conditions looked favorable for such an undertaking. As we have seen, the Seljuq Empire had split into a series of competing principalities and statelets controlled by rival members of the ruling dynasty following Malik Shah's death in 1092; as a result, a Byzantine campaign in Anatolia would face a sharply divided enemy rather than the unified one that Arp Arslan had commanded. Even in division, however, Seljuq forces in the region remained strong relative to the still-recovering Byzantine army. The empire would thus need help if it hoped to overcome the Turks and regain its lost territories. Alexius had already identified the ideal force to provide that assistance: the heavy, lance-armed knights of Western Europe. Though he had succeeded in defeating them, the emperor had nonetheless been immensely impressed by the power of those mounted warriors when he had fought them in Greece in the 1080s. In fact, he had come to conclude that a small force of such troops would give his army the extra punch it needed to beat the Seljuqs. Accordingly, he wrote a letter to Pope Urban II (r. 1088-1099) in 1095 formally requesting the pontiff's assistance in arranging for a contingent of knights to help his army regain the empire's lost territory in Anatolia.⁴⁵

The First Crusade

His request for aid from Western Europe found a warm reception in Rome. Urban II had been in contact with Alexius for several years and was increasingly excited about the prospects of sending a force of knights to the east. The pope's enthusiasm did not

⁴³ Browning, *The Byzantine Empire*, 157–61.

⁴⁴ Peter Frankopan, *The First Crusade: The Call from the East*, Reprint edition (Cambridge, MA: The Belknap Press of Harvard University Press, 2012), 35–38.

⁴⁵ Tyerman, *Crusades*, 38–39.

stem from a desire to help Alexius regain control of Anatolia, however. Rather, it emerged out of his belief that the emperor's appeal could be repurposed to justify a project of vastly greater scale and significance: the reconquest of the Holy Land from the Muslims. In other words, rather than taking Alexius's request for what it was—a comparatively modest appeal for military assistance from one medieval lord to another—Urban II instead chose to use it as a pretext to mobilize the forces of Western Europe for a religious war against the Muslims states in the Middle East.⁴⁶

Success would allow Urban II to achieve three of the eleventh-century Vatican's most cherished goals. First, commanding Christian nobles to wage war in the east would powerfully strengthen the papacy's hotly contested claim that the pope rather than the Holy Roman Emperor was the head of the Church. Second, assisting the Byzantine Empire might end the schism that in 1054 had divided the Church into Eastern Orthodox and Catholic branches on terms favorable to the papacy. Finally, sending a large army to the Middle East raised the possibility of achieving an objective that the papacy had, with growing enthusiasm, recently begun to contemplate: the return of Jerusalem—site of the Crucifixion and the Resurrection—to Christian control.⁴⁷

To win support for a holy war in the Middle East, Urban II went on a recruiting tour in 1095 capped by a massive gathering of nobles at Clermont in France. At the rally, he gave an impassioned address in which he called on the assembled aristocrats to organize a military campaign aimed at retaking control of Jerusalem. He was well aware that he was asking a great deal. Such an operation would come at substantial expense to those who participated in it and would take years to complete. Indeed, a similar proposal by one of his predecessors in the 1070s had met with a cool reception among nobles who were keenly aware of the costs that they would shoulder in undertaking such a complex and lengthy operation. Accordingly, Urban II carefully framed the expedition in a way that maximized its appeal to the aristocrats. Playing on the religious sentiment of the assembled lords, he characterized it not as a traditional military campaign, but instead as an armed, penitential pilgrimage. Critically, he also declared that those who participated would be absolved of all their sins.⁴⁸

This last provision was a peculiarly powerful inducement to the men gathered at Clermont. European nobles had long found themselves caught on the horns of a painful dilemma. On the one hand, the masculinity-driven social code to which they happily adhered impelled them to wage ceaseless warfare against each other; on the other, their deeply held religious beliefs led them to conclude that the violence that they committed in those conflicts would condemn them to eternal damnation. Urban II's request provided them with a way out of this predicament. Taking part in an armed

⁴⁶ Thomas Asbridge, *The Crusades: The Authoritative History of the War for the Holy Land* (New York: Ecco, 2010), 34–35.

⁴⁷ Thomas Asbridge, *The First Crusade: A New History* (New York: Oxford University Press, 2004), 15; Tyerman, *Crusades*, 21.

⁴⁸ Andrew Jotischky, *Crusading and the Crusader States* (New York: Pearson Longman, 2004), 23–36, 47–54.

pilgrimage to Jerusalem would not only absolve them of the many sins that they had committed but would permit them to do so in a way that was consistent with their class's social code. Accordingly, to shouts of "God wills it! God wills it!" thousands of the assembled lords at Clermont agreed to upend their lives by organizing and taking part in a crusade that they knew would involve an expensive, multiyear trip to the Holy Land from which many of them would never return.⁴⁹

Known to historians as the First Crusade, the operation got underway in the summer of 1096. Travelling overland, the first contingent reached the outskirts of Constantinople in the fall. With the steady arrival of further groups of knights over the winter and spring of 1097, the crusader army gradually swelled to between 50,000 and 70,000 soldiers—a substantial force by the standards of the day. It did not take long for the growing size of the crusader army encamped outside the Byzantine capital to alarm Alexius and his advisors. It was far larger than the modest contingent that the emperor had requested of Urban II and was dangerously close to the empire's seat of power; worse, many of its leaders were men who had sought to exploit Byzantine weakness by invading in the 1080s and were thus hardly figures that Alexius could trust. The emperor was a shrewd leader, however, and had little trouble identifying and exploiting the crusaders' needs. Taking advantage of the fact that they were desperately low on food and required transport across the Bosphorus, he compelled them to swear fealty to him as vassals—thus limiting their autonomy and ensuring that any former Byzantine territory they won would return to imperial control. The crusaders resented Alexius's demands but, in no position to refuse, reluctantly agreed to abide by his terms.⁵⁰

The military phase of the First Crusade began in earnest in the spring of 1097. After crossing the Bosphorus Strait in May, the combined Byzantine-crusader army decisively defeated a substantial Turkish force in Asia Minor—thereby permitting both the restoration of imperial control over western Anatolia and unmolested passage into Syria. The crusaders followed this victory by laying siege to Antioch in October 1097. After a grueling nine-month siege, they succeeded in taking the city in June 1098.⁵¹

Much needed, the victory was a major accomplishment for the crusaders. Coming after a difficult winter in which the knights endured food shortages and the bitter cold of northern Syria, the capture of the city lifted the crusaders' badly flagging morale. As important, their successful siege permitted them to regain their independence from Alexius. Using a misunderstanding between the crusaders and the Byzantine army over the conduct of the siege as a pretext, the leaders of the First Crusade reneged on the oath of fealty that they had taken to the emperor. As a result, the city did not retrocede

⁴⁹ Quoted in William Chester Jordan, *Europe in the High Middle Ages* (London: Penguin Books, 2002), 104–5; Jotischky, *Crusading and the Crusader States*, 31.

⁵⁰ Tyerman, *Crusades*, 22.

⁵¹ Frankopan, *The First Crusade*, 138–72.

to Byzantine control following its capture; instead, the ranking nobles of the First Crusade agreed to name Bohemond as the prince of Antioch.⁵²

The crusaders—Franks, as the Muslims called them—were now tantalizingly close to their final goal: Jerusalem. They wasted little time in pressing forward. Intent on occupying their objective as quickly as possible, they headed south in January 1099, and began to lay siege to the city on June 7. Taking it was not shaping up to be an easy task, however. Attrition, death, and the need to garrison captured territory had reduced the army that arrived outside Jerusalem to such a degree that the crusaders lacked sufficient strength to fully surround the city and thus could not reduce it through siege. In response, they decided to take it by storm. Launching their attack on July 15, they quickly breached the walls and poured into the city. There followed a gruesome massacre in which the crusaders killed thousands of Muslims and Jews. One eyewitness, Raymond of Aguilers, claimed that the slaughter was so enormous that it filled the city “with blood ‘up to their knees.’” Another firsthand observer, Fulcher of Chartres (c. 1059-1128) described a grisly scene in which crusaders burned piles of bodies in order to find gold coins that Muslims had swallowed in a vain effort to prevent the Franks from seizing their wealth.⁵³

The savagery described in these accounts has drawn a great deal of scrutiny in recent times. Many characterize the violence as extraordinary in both scale and intensity. For example, the writer Tamim Ansary calls the sack of Jerusalem “an orgy of bloodletting.” Such interpretations are not entirely inaccurate, but they overstate the extent and uniqueness of the violence in the Holy City. In reality, the crusaders’ actions were, while lamentable, neither extraordinary nor out of character with the times. Indeed, while the prevailing code of warfare to which both Christians and Muslims adhered entitled the victors to slaughter every resident of the city, current best estimates indicate that they only killed between 3,000 and 5,000 people. So what explains the lurid language that Aguilers and Fulcher used in recounting the events that they witnessed? Today, most historians agree that these descriptions were no more than the exaggerations typical of a less-literal time. As the historian Thomas Madden writes, “stories of the streets of Jerusalem coursing with knee-high rivers of blood were never meant to be taken seriously,”—a fact keenly understood by contemporaries, who, as Madden notes, “knew such a thing to be an impossibility.”⁵⁴

Over the next few years, the crusaders cemented their position in the Holy Land. Shortly after they took Jerusalem, they ensured their control of the region by defeating a

⁵² Jotischky, *Crusading and the Crusader States*, 58–59.

⁵³ Jordan, *Europe in the High Middle Ages*, 110; Raymond of Aguilers quote from Norman Housley, *Fighting for the Cross: Crusading for the Holy Land* (New Haven: Yale University Press, 2008), 218; Fulcher quote from Fulcher of Chartres, “The Capture of Jerusalem,” accessed January 17, 2020, <https://sourcebooks.fordham.edu/source/fulcher-cde.asp>.

⁵⁴ Ansary, *Destiny Disrupted*, 140; Thomas F. Madden, *The Concise History of the Crusades*, Third Student Edition (New York: Rowman & Littlefield Publishers, 2013), 38.

Fatimid army near Ascalon in southern Palestine—an engagement that is conventionally understood to mark the end of the First Crusade. Most of the remaining Franks thereafter returned to Europe to tend to their neglected landholdings. However, a steady stream of new arrivals offset their numbers and, with naval assistance from Venice, helped the crusaders who had remained in the Holy Land reduce the remaining Muslim-controlled coastal cities that lay between Jaffa and Antioch. The conquest of these ports established the basic structure of what Europeans had come to refer to as Outremer, or the “land overseas.” Thereafter, it occupied a thin belt of territory that extended from Ascalon in the south to the city of Edessa in the north, and consisted of four states: the County of Tripoli, the Principality of Antioch, the County of Edessa, and, most importantly, the Kingdom of Jerusalem.⁵⁵

The First Crusade had been a remarkably successful operation that had borne fruit beyond Urban II’s wildest dreams. The knights who had traveled east had, though operating in hostile territory thousands of miles from their base of support, managed to win three major and countless minor battles. In addition, they had succeeded in taking control of dozens of often-well-defended cities. Most importantly, the soldiers of the First Crusade had—despite long odds—achieved Urban II’s primary aim. Jerusalem and, with it, the Church of the Holy Sepulcher, were once more under Christian control.⁵⁶

Given the comparatively impoverished state of Western Europe, how were the knights of the First Crusade able to pull off such a sequence of feats? Three factors explain their success. First, the heavy knights that made up the elite nucleus of the crusader armies constituted an enormously powerful military force. Heavily armored and equipped with powerful lances, they were capable of mounting charges that were both ferocious and decisive—particularly against Turkish forces that had little experience fighting them. Second, the knights played their diplomatic cards very shrewdly. For example, the leaders of the First Crusade managed to lull the Fatimids into a dangerous complacency by engaging them in lengthy diplomatic talks that they terminated just weeks before their army arrived outside Jerusalem. Finally, as we have seen, the First Crusade arrived in the Middle East at a time when the region was peculiarly divided. The collapse of the Seljuq Empire following Malik Shah’s death in 1092 meant that the crusaders faced not a powerful, unified state, but instead a collection of divided, fractious principalities. These divisions largely ensured that the crusaders did not have to contend with an organized, Muslim coalition. In the rare instances in which they did face a larger combination of enemies, moreover, they often benefited from the emergence of divisions among their opponents. In a battle outside Antioch in 1098, for example, the crusaders were able to achieve victory in part because resentful vassals abandoned their patron, the Turkish leader Kerbogha (?-1102), in the heat of the fight.⁵⁷

⁵⁵ Jonathan Phillips, “The Latin East, 1098-1291,” in *The Oxford History of the Crusades*, ed. Jonathan Riley-Smith (Oxford: Oxford University Press, 1999), 111–13.

⁵⁶ Jotischky, *Crusading and the Crusader States*, 57–62.

⁵⁷ Carole Hillenbrand, *The Crusades: Islamic Perspectives* (Edinburgh: Edinburgh University Press, 1999), 58.

Contingency—luck—thus played a pivotal part in the First Crusades' success. Had its participants attacked a few years earlier or a few decades later rather than in the 1090s, they would assuredly have encountered a much more unified Middle East and would very likely have found the going far-more difficult if not impossible. There is irony in this situation, for the First Crusade succeeded for the same reason that the early Arab Conquest had. In both cases, the invaders triumphed not because they were the beneficiaries of divine providence, but instead because their timing was peculiarly fortunate.⁵⁸

Outremer

As we have seen, those who had taken part in the First Crusade had done so out of deeply held religious convictions that left little room for a tolerant approach to people of different faiths. Such views dominated the Franks' approach to the region during the early years of the Crusades. Upon taking control of the Holy Land, for example, the crusaders had sporadically massacred Muslims and had converted a number of religious shrines and mosques—including the Dome of the Rock and the al-Aqsa Mosque—into churches or administrative buildings. Within just a few years, however, the Franks who remained in the Crusader States no longer engaged in such actions and, in many cases, actively courted their non-Catholic subjects. What accounted for this shift away from intolerance?⁵⁹

Practical needs drove the change. The crusaders who remained in the region abandoned their earlier, intolerant approach not because they had a genuine change of heart, but instead because they grasped the perilousness of their situation. Weak in numbers—the Kingdom of Jerusalem was down to only three-hundred knights and fewer than two-thousand foot soldiers by early 1100—they came to conclude that they would not likely maintain control over the region's diverse population of Muslims, Jews, non-Catholic Christians, Druze, and Alawites if they continued to mistreat those people. They also keenly understood that the diplomatic structure of the region was unfavorable to them. Surrounded as they were by a number of powerful Turkish principalities, they realized that Outremer could retain its independence only if it succeeded in keeping those states divided. Accordingly, the Franks who remained in the Holy Land quickly shed the religious fervor of the First Crusade in favor of a more pragmatic approach to the region that centered on accommodation, discretion, and tolerance.⁶⁰

This pragmatism took several forms. Diplomatically, the Franks quickly accommodated themselves to the practice of entering into mutually beneficial, cross-religious alliances with nearby Turkish rulers aimed at maintaining the balance of power and thus at preventing any Muslim leader from establishing a powerful, unified state that could overwhelm Outremer. Most notably, the Kingdom of Jerusalem frequently allied with the Burid Dynasty of Damascus in order to prevent the powerful ruler of Aleppo

⁵⁸ Jotischky, *Crusading and the Crusader States*, 42.

⁵⁹ Phillips, "The Latin East, 1098-1291," 113.

⁶⁰ Phillips, 113–14.

from adding that rich and strategically located city to his domain. Meanwhile, to ensure domestic tranquility, they adopted a tolerant attitude toward their non-Catholic subjects that they patterned on the Muslim approach to the *dhimmi*. That is, in exchange for the payment of a poll tax, they left non-Catholic Christians, Muslims, and others free to practice their religion as they saw fit—an arrangement that guaranteed both domestic stability and the steady flow of tax revenues. Finally, the crusaders who remained in Outremer—and, especially, those born there—gradually embraced the multicultural cosmopolitanism of the region. They learned new languages, adopted regional forms of dress and food, often developed a sincere appreciation for other cultures, and, in many cases, formed genuine bonds with non-Catholics. One Muslim writer, Usama ibn Munqidh (1095-1188) claimed that he could tell at a glance if a Frank was a new arrival in the region based on whether that person had adopted local forms of dress, and he noted approvingly “that ‘there are some Franks who have settled in our land and taken to living like Muslims.’” Thus, within just a few short years, the Franks who had remained in the Holy Land had all but completely abandoned the intolerance and zeal that had driven them to crusade in favor of a pragmatism that, ironically, played no small part in ensuring the success of the Crusader States during the early twelfth century.⁶¹

Zengi

The states of Outremer also benefited from the fact that the response to the Crusades was initially muted. During the first third of the twelfth century, Muslim rulers made no effort whatsoever either to mount a meaningful counterattack or to organize a collective defense against the Franks. What accounted for their apparent complacency? Why did they not band together against the crusaders? To a substantial degree, the absence of a concerted Muslim response to the Crusades was a function of the same political and social fracturing that had facilitated the success of the First Crusade. Far more focused on gaining advantage against each other than on responding to the newcomers, the Turkish leaders of the principalities that adjoined Outremer generally paid little attention to the Crusader States save when they sought to enlist the Christians against their Muslim rivals.⁶²

There were, however, some who correctly grasped the intent of the crusaders. Most notably, a number of educated, urban Arab legal scholars had deduced the intent of the Franks from the start and had responded by reviving the concept of a jihad to defend *Dar al-Islam*. However, they lacked a relationship with or influence over the newly arrived and far-less sophisticated Turkish ruling class; accordingly, despite their vigorous efforts, they were unsuccessful in persuading the local princes to put down their differences and join together to form a common front against the Franks⁶³

⁶¹ Jotischky, *Crusading and the Crusader States*, 89–90; Quote from Madden, *The Concise History of the Crusades*, 47.

⁶² Madden, *The Concise History of the Crusades*, 46–48.

⁶³ Madden, 46–48.

Only with the emergence of the powerful Turkish ruler of Mosul and Aleppo, Zengi, (r. 1127-1146), would a coherent Muslim response to the Crusades begin to take shape. Zengi rose to prominence in 1144 when he seized the County of Edessa and added it to his growing state. His intentions in doing so are difficult to discern. On the one hand, his close relationship with the 'ulama' and adoption of their language of jihad suggests that he had religious motives. On the other, evidence indicates that he acted because Edessa's Frankish ruler had supported a rival Turkish dynasty against him. Regardless, the results of his victory were the same: thereafter, Zengi was able to credibly present himself as Islam's champion against the crusaders—a claim that won him growing support and legitimacy in the Muslim world.⁶⁴

While the fall of Edessa drew cheers in the Muslim world, it sent shockwaves through Outremer and Christian Europe alike. This response was merited. As we have seen, the Crusader States' foreign policy had relied on a divide-and-conquer strategy designed to prevent the emergence of a powerful, unified Muslim state adjacent to Outremer. Even with the addition of Edessa, Zengi's principality was as yet too small to constitute such a danger. As Frankish leaders grasped, however, his acquisition of Edessa put him in a position to assume control over the rest of Syria—an outcome that would upset the balance of power on which Outremer's security rested.⁶⁵

Accordingly, Pope Eugenius III (r. 1145-1153) responded to Zengi's victory in 1145 by calling for another armed pilgrimage—later known as the Second Crusade (1145-1149)—for the purpose of retaking Edessa. The response to his request for volunteers was substantial. Holy Roman Emperor Conrad III of Germany (r. 1138-1152) and King Louis VII (r. 1137-1180) of France both raised large armies and marched toward the Holy Land via Constantinople. Other groups of crusaders departed later via ship; en route, some of them helped King Alfonso I (r. 1139-1185) of Portugal conquer Lisbon from Spanish Muslims. Those who traveled by sea arrived in Outremer without difficulty; in contrast, those who took the overland route suffered crushing defeats at the hands of the Turks while trying to cross Anatolia. As a result, by the time Conrad III and Louis VII's battered columns joined those Franks who had traveled by ship to Outremer, the assembled troops of the Second Crusade lacked the strength to attack the powerful state that Zengi had assembled and thus could not recover Edessa.⁶⁶

Not wanting to return to Europe in defeat, they decided to abandon the original goal of the Second Crusade in favor of a new objective: the conquest of the less-well defended city of Damascus. It may have been a more realistic goal for the depleted forces of the Second Crusade, but the city's capture still proved to be well beyond the capacity of the assembled knights. Indeed, the effort to besiege Damascus ended in a humiliating retreat after only four days when the Franks realized that their army was in danger of being destroyed by superior Muslim forces. Thus, despite the commitment of

⁶⁴ Hillenbrand, *The Crusades*, 110–14; Paul M. Cobb, *The Race for Paradise: An Islamic History of the Crusades* (New York: Oxford University Press, 2014), 134–35.

⁶⁵ Madden, *The Concise History of the Crusades*, 53–54.

⁶⁶ Jotischky, *Crusading and the Crusader States*, 84–88.

substantial resources and the raising of significant armies, the Second Crusade ended in embarrassment and abject failure—unable either to retake Edessa or to shore up the Crusader States in any meaningful way.⁶⁷

Nur ad-Din

Meanwhile, a new leader had emerged to take charge of the struggle against the Franks. After an enslaved person murdered Zengi in 1146, his son, Nur ad-Din (r. 1146-1174), inherited the now-expanded principality of Aleppo. He proved to be an even more effective opponent of the crusaders than his father had been. Already well regarded by the growing number of religious scholars who sought a jihad against the Franks, Nur ad-Din had dramatically enhanced his anti-crusader bona fides by immediately going to war with Antioch and by decisively defeating a crusader army in 1149. Further success in the 1150s strengthened his position. Most notably, his seizure of the strategic city of Damascus in 1154 unified Muslim Syria under his control and demonstrated that he had both the will to wage a jihad against the Crusader States and the resources to sustain such a campaign. Accordingly, recruits flocked to his banner in the mid 1150s—especially after he declared that his central goal was the conquest of Jerusalem.⁶⁸

Nur ad-Din's victory at Damascus greatly alarmed the King of Jerusalem, Baldwin III (r. 1143-1163). As we have seen, the rulers of the Crusader States had long believed that they would be in mortal peril if a powerful Muslim leader managed to unify Syria. Zengi's acquisition of Edessa and, especially, Nur ad-Din's capture of Jerusalem marked significant setbacks in that effort. Now, the Franks faced the very situation that they had worked so hard to prevent—the establishment of a wealthy and powerful Muslim principality on the border of the Crusader States. Fortunately for Baldwin III and the other Crusade leaders, however, they were not alone in expressing alarm about the emergence of a formidable Muslim state in Syria. Byzantine Emperor Manuel I Komnenos (r. 1143-1180) shared their concerns and sought to work with them to contain Nur ad-Din. As a result, he struck an alliance with Baldwin III in the 1150s that succeeded in checking the Turkish leader's regional ambitions.⁶⁹

Nur ad-Din did not accept this diplomatic setback passively, however. Instead, he sought to offset it by building up his strength elsewhere in the region. The once-powerful Fatimid state appeared to offer the best opportunity for him to do so. By the 1160s, it had come to have a peculiar mix of strengths and weaknesses that made it a tempting target. On the one hand, Egypt's economy produced substantial tax revenue thanks to earlier improvements to its irrigation systems and to its dominance of the lucrative trade with India. On the other, worsening factionalism in the governing class and a succession

⁶⁷ Housley, *Fighting for the Cross*, 8–11.

⁶⁸ Jotischky, *Crusading and the Crusader States*, 111–13.

⁶⁹ Jotischky, 92.

of child caliphs had combined to progressively loosen the Fatimid Dynasty's grip on the province.⁷⁰

Accordingly, Nur ad-Din began to probe for ways in which he could exploit the political divisions in Cairo to his advantage. He got his chance in 1163. That year, the recently deposed Fatimid vezir, or minister, Shawar (?-1169) offered Nur ad-Din one-third of Egypt's tax revenue in exchange for helping him return to power. The Turkish leader pounced at the opportunity. He ordered one of his subordinates, the Kurdish general Shirkuh (?-1169), to march his army into Egypt and to help restore Shawar to the vezirate. Unfortunately for Nur ad-Din, his bid to assume a dominant position in Egypt failed thanks to the prompt actions of Baldwin III's successor, Amalric (r. 1163-1174). Well aware of the danger that Nur ad-Din's move posed to his kingdom, he intervened in Egypt and helped to defeat Shirkuh's expeditionary force. Amalric's victory proved short lived, however. A subsequent intervention by Shirkuh in Egypt in 1169 not only succeeded but did so beyond Nur ad-Din's wildest dreams. Rather than merely gaining control over part of the province or a share of its tax revenue as had been the goal in 1163, Shirkuh managed to secure for his Turkish lord the *entirety* of Egypt. It was a major coup. With Egypt added to Syria, Nur ad-Din now had more than adequate tax revenues and resources to defeat the Franks.⁷¹

Salah ad-Din

Ultimately, however, Nur ad-Din was not the ruler who would lead the attack on the Crusader States. Instead, leadership of the jihad against Outremer would fall to Shirkuh's nephew and lieutenant: Salah ad-Din (r. 1174-1193). Known as Saladin in the West, this ambitious and capable upstart spent most of the early part of his reign engaged in high-stakes political maneuvering—initially to gain ascendancy in Egypt and then to assume control of Nur ad-Din's empire. His ascent was rapid. He first established himself as a significant political figure when he succeeded his deceased uncle as Nur ad-Din's chief subordinate in Cairo in 1169; two years later, he consolidated his position in Egypt by abolishing the Fatimid Caliphate. To this point, his moves had been careful, incremental ones designed to enhance his power without alarming Nur ad-Din. In 1174, however, fortune provided him with the chance to reach for the brass ring. Nur ad-Din's death that year and the accession of his eleven-year-old son created a political vacuum that Salah ad-Din grasped he could exploit to take power. Moving aggressively, he travelled to Damascus where he married Nur ad-Din's widow—a common medieval Muslim political ploy—and assumed the role of regent. He followed by solidifying his control of Syria and, more importantly, by formally displacing Nur ad-Din's son as ruler—establishing, in the process, the Ayyubid Dynasty (1171-1260).⁷²

⁷⁰ Asbridge, *The First Crusade*, 265–72.

⁷¹ Cobb, *The Race for Paradise*, 160–63.

⁷² Hillenbrand, *The Crusades*, 171–72.

Salah ad-Din had long expressed a desire to lead a jihad against the Crusader States; with his territory now all-but surrounding Outremer, he was finally in a position to do so. Accordingly, after securing recruits by declaring holy war, he invaded the Kingdom of Jerusalem in 1187 with a huge army. Taking advantage of dissension among the Franks, he skillfully maneuvered the poorly-led crusaders into a position near the Sea of Galilee that lacked water. With the Franks suffering heat stroke and severe dehydration, the ensuing battle was a complete rout. Salah ad-Din's troops slaughtered thousands of crusaders, put the survivors to flight, and—worst of all in the eyes of the Christians—took possession of the True Cross. Quickly exploiting his victory, he followed by seizing Jerusalem and nearly the entirety of Outremer from the demoralized Franks. It was a dark day for the crusaders. Having been reduced to the ports of Tyre, Antioch, and Tripoli, Outremer was teetering on the brink of extinction.⁷³

The impending failure of the Crusader States and, especially, the loss of Jerusalem stunned and dismayed Europe. The news was so troubling, in fact, that Pope Urban III (r. 1185-1187) reputedly died of shock when told of the city's fall. The Church was not yet ready to give up the dream of a Christian-controlled Jerusalem, however. Instead, barely a week after Urban III's death, his successor called on Europe's kings to mount another crusade aimed at regaining control of the city. The response was tremendous. In 1190, French King Philip II (r. 1180-1223) and British monarch Richard the Lionheart (r. 1189-1199) both raised large armies and set out by sea for the Holy Land; concurrently, the Holy Roman Emperor, Frederick Barbarossa (r. 1155-1190), marched overland with an even-larger force. Massive in scale and enjoying the support of innovative new financial mechanisms such as the Saladin Tithe, the Third Crusade appeared to be unstoppable.⁷⁴

Unfortunately for the Franks, it proved to be anything but. The problems started when Frederick Barbarossa drowned while crossing a river in Asia Minor; absent his commanding presence, his demoralized army disintegrated. Meanwhile, though Richard and Philip II both arrived in the Middle East in 1191 without difficulty, they soon had a falling out stemming from Philip II's envy of Richard's popularity—a jealousy enhanced by the fact that Richard, though technically Philip II's vassal, was considerably more powerful than his putative liege lord. In response, the French king abandoned the field in a pique and returned to France with the bulk of his army. As a result, the Third Crusade appeared doomed to failure.⁷⁵

Despite seeing many of the crusaders return to France, Richard remained undaunted. A skilled and charismatic commander, he rallied his remaining troops and went on a successful offensive. Under his careful direction, his army retook the key port cities of Acre and Jaffa after which it savaged Salah ad-Din's army at Arsuf. For a time,

⁷³ Madden, *The Concise History of the Crusades*, 69–75.

⁷⁴ Jonathan Riley-Smith, *The Crusades: A Short History* (New Haven: Yale University Press, 1987), 84–87.

⁷⁵ Christopher Tyerman, *Fighting for Christendom: Holy Wars and the Crusades* (New York: Oxford University Press, 2004), 50–53.

as a result, Richard appeared poised to retake Jerusalem. Unfortunately for him, the victory at Arsuf marked the high point of his campaign. Thereafter, the two sides' asymmetric strengths produced a stalemate. Richard's superior generalship could not fully offset Salah ad-Din's larger army, while Salah ad-Din's bigger force could not entirely overcome the king's tactical gifts.⁷⁶

Ultimately, the stalemate did not last. Instead, a compromise emerged out of the two leaders' shared need to address pressing domestic political challenges. Richard feared that Philip II was exploiting his absence to weaken his position back home; Salah ad-Din, for his part, had watched the prestige he had acquired by taking Jerusalem in 1187 evaporate as a result of Richard's success against him. In 1192, as a result, the two men came to an agreement formalized in the Treaty of Ramla. By its terms, Richard recognized Muslim control of Jerusalem; in exchange, Salah ad-Din ceded the coastal strip from Jaffa to Antioch to the crusaders and agreed to grant Christian pilgrims free entry into Jerusalem.⁷⁷

Later Crusades

Richard had achieved a great deal, but his failure to regain control of Jerusalem left many Europeans feeling that the Third Crusade was a failure. Among them was the new pope, Innocent III (r. 1198-1216). Dismayed by the inability of the Third Crusade to retake the Holy City, he called for a new campaign, the Fourth Crusade, shortly after he acceded to the papacy. Launched with great fanfare, it not only failed to regain Jerusalem, but never even made it to the Holy Land. Instead, sniffing opportunity in the Byzantine Empire, the crusaders became involved in a complicated succession dispute in Constantinople in 1203 that ended the following year when they assaulted the city and established it as the capital of a new, Frankish-ruled state called the Latin Empire. It had a short and rocky life. Its Eastern Orthodox subjects bitterly resisted both Western rule and the imposition of Catholic doctrines and rites, and its rulers never controlled more than a small part of the Byzantine Empire's territory.⁷⁸

As a result, under the direction of the Palaeologus Dynasty, the Byzantine army was able to redeem Constantinople in 1261. For a time, it seemed that the empire was poised to achieve yet another recovery. Ultimately, however, the new rulers could not overcome Byzantium's many problems. Shorn of much of its territory and surrounded by predatory enemies eager to exploit its weaknesses, the empire thereafter survived as little more than a shell of its former self, until, as we shall see in chapter six, the cannon-equipped Ottoman Turks finally administered the coup de grace in 1453.⁷⁹

Meanwhile, the Western Europeans continued to mount crusades over the course of the thirteenth century. With the exception of the Sixth Crusade in the 1220s,

⁷⁶ Jotischky, *Crusading and the Crusader States*, 159–61.

⁷⁷ Jordan, *Europe in the High Middle Ages*, 174–75.

⁷⁸ Tyerman, *Crusades*, 35–38.

⁷⁹ Madden, *The Concise History of the Crusades*, 105.

however, which, through a diplomatic agreement, briefly regained control of Jerusalem, none proved successful. Largely focusing on Egypt—control of which would have simultaneously ensured the security of Outremer and made the crusaders fabulously wealthy—successive waves of Franks found themselves outmaneuvered by larger and better-led Ayyubid forces. Still, despite those failed efforts, the Crusader States themselves endured. Relying largely on savvy diplomacy and the lure of trade rather than on force of arms, the Franks managed to hang on to the territory that Richard had recaptured during the Third Crusade.⁸⁰

The Mongols

At about the same time as the Fourth Crusade, meanwhile, a vastly more dangerous threat to the Muslim world, the Mongol Empire, was quietly emerging northeast of China. The Mongols were a loose confederation of nomadic tribes that had long lived on the northern steppes of Asia. There, they herded livestock, raided the wealthy Chinese Empire of the Song Dynasty (960-1279), and warred with similar nomadic groups such as the Tatars. Like the other peoples of the region, the Mongols were militarily far more powerful than their relatively modest numbers would suggest. Toughened by the hard life of the Eurasian Steppe and well-practiced in sophisticated battle tactics as a result of the constant wars they fought with other nomads, they possessed a latent power constrained only by the fact that they remained divided among themselves into a number of competing tribes.⁸¹

Their disunity ended in 1206. That year, a powerful young war leader named Temujin (c. 1162-1227) succeeded in bringing all of the Mongol tribes together under his leadership. Taking the title Chingiss Khan, meaning Universal Leader, he followed by embarking on a remarkable—and unusually brutal—campaign of conquest. As the Great Khan, he first defeated nearby nomadic rivals such as the Tatars and the Tanguts and followed by incorporating them into his empire. He then wrested control of northern China from other nomadic groups who had long dominated that region. Spreading rapidly westward, his empire soon reached the frontier of Khwarazmia, the easternmost Muslim state in Central Asia.⁸²

Chingiss Khan initially had no hostile intentions toward the Muslim world. He viewed his new border with Khwarazmia not as a jumping off point for conquests, but rather as a valuable link to the Middle East's vibrant economy. Toward that end, he sent a caravan of merchants laden with eastern luxury goods on a trade mission into Khwarazmia in 1218 with instructions to negotiate a commercial agreement with its ruler, Shah Ala ad-Din Muhammad (r. 1200-1220). Muhammad was not receptive to Chingiss Khan's commercial initiative, however. Eager to demonstrate his toughness to the Great Khan, he responded by ordering his soldiers to plunder the caravan and to kill

⁸⁰ Tyerman, *Fighting for Christendom: Holy Wars and the Crusades*, 60–64.

⁸¹ Hodgson, *The Venture of Islam, Volume 2*, 286–87.

⁸² Peter Frankopan, *The Silk Roads: A New History of the World* (New York: Knopf, 2015), 154–56.

its 450 merchants. Worse, when Temujin sent three envoys to demand restitution, Shah Muhammad ordered one beheaded and the others shorn of their beards—a serious and calculated insult in the culture of medieval Central Asia.⁸³

This act enraged Chingiss Khan and led him to initiate the first phase of what would be a decades-long Mongol war with the Muslim world. Raising a huge force, he invaded Khwarazmia in 1219, defeated its army, and put Shah Muhammad to flight. Over the next few years, the vengeance-minded Mongols brutally sacked the cities of Balkh, Merv, Samarkand, and Nishapur as they pressed westward into the Islamic heartland. While we can reject the exaggerated claims of contemporary Muslim writers—the Mongols most certainly did not kill 1,750,000 people at Nishapur—the evidence is equally clear that they waged this war in a peculiarly ruthless and brutal fashion.⁸⁴

Temujin's death in 1227 led to a brief pause in the Mongol onslaught, but, by the early 1230s, they were back on the offensive. Having conquered Muslim Central Asia, they now focused on adding Iran to the great khan's empire. There, Mongol armies slaughtered huge numbers of people, destroyed cities, and even deliberately destroyed many of the *qanat*, the ancient, underground canals on which Iranian agricultural prosperity depended. Indeed, the devastation and loss of life in Iran was so enormous that the historian Ira Lapidus maintains that the Mongol invasion "amounted to a holocaust." The Mongols followed in 1243 by attacking Asia Minor, where they shattered the last significant Seljuq principality, the Sultanate of Rum. With that defeat, hope ebbed in the Islamic world that Muslim soldiers could ever defeat the Mongols. Just when the situation appeared grimmest, however, the Muslim Middle East caught a break: a succession dispute halted the onslaught and gave the Muslims a few years of desperately needed breathing space. As we shall see, that respite would prove pivotal.⁸⁵

With the succession finally sorted out in 1251, Mongke Khan (r. 1251-1259), the new Great Khan of the Mongols, moved to restart the conquests. He began by dividing his empire into four semiautonomous khanates that he distributed among his brothers. He gave one of them, the Ilkhanate, to his brother Hulagu (r. 1256-1265), along with instructions to complete the conquest of the Muslim world. Taking one-fifth of the Mongol army with him, Hulagu quickly set to work. He began in 1256 by reducing the Assassins' heretofore impregnable Alamut Castle—thus destroying their peculiar, non-contiguous state. He followed in 1258 by marching his army to Baghdad, where he hoped to compel al-Musta'sim (r. 1242-1258), the thirty-seventh and final Abbasid caliph, to surrender. When the latter refused to yield, Hulagu ordered his soldiers to storm the city. Possessing sophisticated siege equipment, they easily overwhelmed the defenders and took possession of Baghdad. They followed by sacking the city with a

⁸³ Ansary, *Destiny Disrupted*, 153.

⁸⁴ Steven R. Ward, *Immortal: A Military History of Iran and Its Armed Forces* (Washington, D.C: Georgetown University Press, 2009), 26.

⁸⁵ Ward, 26; Lapidus, *A History of Islamic Societies*, 226.

brutality that was peculiar even by the standards of the day. Mongol soldiers killed hundreds of thousands of people, razed buildings of inestimable beauty and historical significance, and destroyed the priceless contents of the city's many libraries by pitching books into the Tigris River. They also did enormous damage to the irrigation system of central Iraq, and thus ensured that the city would lack adequate food supplies to recover.⁸⁶

Perhaps most distressing to educated urban Muslims, Hulagu also ordered al-Musta'sim and his sons killed. Accounts differ as to how they were murdered, but a consensus of historians now agree that the Mongols most likely first wrapped them in carpets and then had horses trample them to death so as to avoid defiling the earth with royal blood. Regardless of the details, al-Musta'sim's execution sent shockwaves through the *umma*. Though the Abbasids had not exercised real power outside of Iraq for centuries, the dynasty nonetheless remained a powerful symbol of Muslim unity and power. As such, the end of the ruling house that had held the caliphate since 750 marked a distressing watershed for a Muslim world that increasingly felt it had lost its way.⁸⁷

Mamluks

Hulagu next set his sights on the only remaining independent Muslim state of any consequence: Egypt. Though vastly weaker than the Ilkhanate, Egypt was prepared to put up a vigorous resistance thanks to a recent change in leadership. In 1250, Turkish slave soldiers called Mamluks, a term that means, literally, "property" or "owned men," had taken advantage of instability in the Ayyubid court to seize power and establish the Mamluk Sultanate (1250-1517). Disciplined and well organized, they formed a powerful army that promised to valiantly resist the Mongols. Still, they faced long odds. Indeed, the sheer size of Hulagu's horde and its extensive battle experience suggested that the Mamluks would almost certainly go down in defeat as the two armies prepared for battle near Ayn Jalut in Palestine in 1260.⁸⁸

Ayn Jalut

Fortunately for the Mamluks, two factors combined to give them a shot at defeating the Mongols. First, not long before the battle, news had reached Hulagu that Monge Khan had died. Eager to take part in the deliberations at the *kurultai*—the meeting at which the Mongol chiefs would choose the next Great Khan—he returned to the capital of Karakorum with most of his troops. The Mongol army that continued toward Egypt thus only numbered about 10,000 fighters—still powerful, but a fraction of the force that had destroyed the caliphate in 1258. Second, Baybars (r. 1260-1277), the commander of the Mamluk army, was a peculiarly adept tactician and possessed an intimate familiarity with the region in which the battle was fought. The combination of

⁸⁶ Ansary, *Destiny Disrupted*, 153–56.

⁸⁷ Frankopan, *The Silk Roads*, 162–63.

⁸⁸ Hodgson, *The Venture of Islam, Volume 2*, 267–68, 417.

relatively even numbers and strong leadership proved enough to tip the balance. In the ensuing Battle of Ayn Jalut, the Mamluks scored a decisive victory. Egypt had been saved.⁸⁹

The results of the battle were enormously significant. First, and most obviously, it gave the victors time to consolidate and strengthen their regime. As a result, the Mamluk state would never again be as vulnerable to a Mongol attack as it had been on the eve of the Battle of Ayn Jalut. Second, and far more important, Baybars's victory ensured that independent Islamic rule would survive at a time when its future seemed very much in doubt.⁹⁰

The impact of the battle on Egypt's political structure was also far reaching. Just two months after the victory, Baybars and a group of co-conspirators arranged to have the reigning sultan, Qutuz (r. 1259-1260), assassinated. Not long thereafter, Baybars managed to navigate his way through the treacherous waters of Mamluk factional politics to succeed Qutuz as the new sultan. It was a bold move for the victor of Ayn Jalut, but it was also one that put him in a precarious position: if he did not quickly find some way to legitimate his rule, he would very likely find himself facing the same fate as his predecessor. In a stroke of good fortune, the arrival of an Abbasid prince, Abu al-'Abbas Ahmad, in Cairo in 1262 afforded him the opportunity to do so. In an elaborate ceremony, Baybars had him proclaimed Caliph al-Hakim I (r. 1262-1302); in exchange, the newly installed caliph urged Muslims in Egypt to acknowledge Baybars's rule as sultan, and thus accorded him the legitimacy he needed.⁹¹

In this way, the caliphate that the Mongols had destroyed in 1258 was resurrected in Cairo in 1262. Its rebirth was a symbolically significant event for Muslims—one that upheld the idea that the Islamic world remained unified under a single, universal ruler. In reality, however, nothing substantive had happened as a result of al-Hakim I's accession. The line of Abbasid caliphs who reigned in Cairo were powerless figures who lacked a court or even a retinue. Indeed, they continued to hold office only because the Mamluks saw utility in retaining the position. Lacking even the pretense of power, they were figureheads of the feeblest sort—a dynasty whose authority and responsibilities never rose above the ceremonial.⁹²

Mamluk Government and Economy

As we have seen, the use of slave soldiers had been commonplace in the Muslim world since al-Mutasim had first introduced the practice in the ninth century. It was under the Mamluk regime, however, that the arrangement reached its apotheosis. Superficially, the Egyptian system was similar to that of earlier dynasties that had relied on enslaved troops. Like those states, the Mamluk government replenished the ranks of

⁸⁹ Hillenbrand, *The Crusades*, 225–26.

⁹⁰ Asbridge, *The Crusades*, 620–28.

⁹¹ Kennedy, *Caliphate*, 248.

⁹² Kennedy, 248.

its army by purchasing Circassian or Turkish adolescents who had grown up in the harsh environment of the Eurasian Steppe. It then raised and educated them in barracks until they were ready to serve the regime. Here, however, the Mamluk system differed from earlier practices. While previous dynasties had restricted the use of slaves in state service to the army and had reserved the administration of government to free-born subjects, Mamluk Egypt instead relied upon enslaved people—and only enslaved people—to staff both its military and its bureaucracy. Even the children of the Mamluks were barred from joining the ruling class. Thus, the Mamluk system was a unique one in which nominally enslaved people rather than free-born ones controlled the government. It was also one that proved durable and effective, and that was, by the standards of the day, unusually meritocratic.⁹³

The Mamluk approach to the financing of government was considerably less innovative, however. Like earlier Muslim states such as the Buyid Dynasty, the Mamluks ruled through a decentralized, semi-feudal system based on *iqta* land grants. According to this arrangement, soldiers received the right to collect taxes over a defined territory; in exchange, they were obligated to provide military service to the state. To ensure a degree of control over what was an intrinsically decentralized system, the Mamluk government periodically redistributed or revoked *iqtas*. Initially, this practice proved reasonably effective at funding the military and government. As it had under the Buyids, however, the *iqta* system also imposed steep costs on Egypt's economy. The problem was one of incentives. Aware that they would receive tax revenue from a given land grant for only a brief period, *iqta* holders focused on squeezing as much wealth out of their peasant farmers as they could rather than on developing the infrastructure that would foster long-term growth; over time, as a result, Egypt's once bountiful agricultural sector—the heart of its economy—began to stagnate.⁹⁴

Islamic Political Recovery and the End of the Crusades

The Muslim political revival that began with the victory at Ayn Jalut was not confined to Egypt, but instead occurred throughout the Middle East. As we shall see in chapter six, new Muslim polities would emerge out of the wreckage of the Abbasid Caliphate and the Sultanate of Rum over the next few decades. Islamic rule even returned to the territory ruled by the Mongol Ilkhanid Dynasty. The religion's revival there began in 1295 when, for reasons of political expediency, the new khan, Mahmud Ghazan (r. 1295-1304), decided to abandon Buddhism in favor of Islam when he assumed the throne. Encouraged by the khan's conversion, nearly all of the Ilkhanid Mongols followed suit over the next decade. As a result, by the time of Ghazan's death, the Ilkhanate had completed a remarkable and shockingly fast transformation from

⁹³ Ansary, *Destiny Disrupted*, 157; Lapidus, *A History of Islamic Societies*, 292.

⁹⁴ Amina Elbendary, *Crowds and Sultans: Urban Protest in Late Medieval Egypt and Syria* (Cairo: American University in Cairo Press, 2015), 37; Lapidus, *A History of Islamic Societies*, 292.

existential threat to Islamic rule into what was, for all practical purposes, just another Muslim state.⁹⁵

By that point, meanwhile, Muslim principalities no longer needed to concern themselves with the Franks. Shortly after Ayn Jalut, Baybars launched a sustained effort aimed at eradicating the Crusader States. Planned and executed with great care, the campaign against the Franks bore quick results. The Mamluks took Haifa by storm in 1266, brutally sacked Antioch in 1268, and successfully assaulted Tripoli in 1289. In each case, they destroyed the port to prevent the Franks from using it to support future Crusades. The coup de grace came in 1291. That year, the Mamluks succeeded in breaching the mighty walls of Acre, the lone remaining crusader city. Thus, after nearly two centuries, Western Christian control of territory in the Levant had come to an end. The Crusades were over.⁹⁶

Legacy of the Crusades

How did two centuries of Christian control of the Levant affect the people involved? Put another way, what were the lasting consequences of the Crusades? In a word, their impact was complex—affecting Europe and the Middle East in very different ways in the short term while, in the long term, powerfully—if ahistorically—shaping present-day perceptions of relations between the Middle East and the West.

The consequences of the Crusades for the Middle East and the West differed substantially. Unsurprisingly, they had a largely positive short-term impact on Western Europe. Indeed, many groups profited from the Crusades. Among them were the merchants of the Italian city-states, who not only grew rich supplying Outremer but, ironically, also developed valuable new trade connections with Muslim states thanks to the Western presence in the Levant. As a result, even after the fall of Acre, they continued to prosper by supplying Europe with the imported eastern products and goods that Westerners had first encountered as a result of the Crusades. Monarchs also benefited. Taking advantage of the new authority and tax revenue that they had acquired to support crusading, a number of kings initiated multi-century-long campaigns to consolidate power that resulted in the emergence of the first genuinely centralized states in Europe since the fall of the Western Roman Empire in the fifth century. Finally, the Crusades dramatically boosted the power of the Church. For a time, in fact, the Crusades turned the papacy into the most powerful political institution in Western Europe.⁹⁷

In contrast, the Crusades produced little lasting political, economic, or cultural change in the Middle East and quickly faded from collective memory. While this situation may seem surprising today, it reflects the fact that the Crusades had at best a relatively marginal impact on the people of the region. The Franks never managed to

⁹⁵ Ansary, *Destiny Disrupted*, 157–58.

⁹⁶ Hillenbrand, *The Crusades*, 227.

⁹⁷ Asbridge, *The Crusades*, 657–70; Ansary, *Destiny Disrupted*, 149.

seize any of the region's leading cities, did not alter the social structure of the peoples they controlled, and remained confined to a small, albeit religiously significant, strip of land. The limits of their influence on the Middle East are particularly clear when contrasted to the enormous impact of the Mongol irruption. While the Ilkhanate destroyed the Abbasid Caliphate, demolished the irrigation systems on which Iraq's agricultural productivity depended, and contributed to the intellectual and spiritual retrenchment of the Muslim world, the crusaders main legacy was a series of imposing castles like Cracs des Chevaliers and, more significantly, a grievously weakened Byzantine Empire.⁹⁸

Perversely, however, while the Crusades themselves were of only limited significance, popular understandings of them—what the historian Christopher Tyerman calls “mythic memories”—continue to exert a powerful influence. Based on the ahistorical works of nineteenth-century European writers rather than on the Crusades themselves, these views are at best a caricature of the events they purport to describe. Nonetheless, they long ago attained the status of conventional wisdom—continuing to detrimentally influence modern perceptions of relations between the Muslim Middle East and the West and to fuel belief in the notion that the two cultures are locked in a millennia-long civilizational conflict.⁹⁹

Modern Western historical memories of the Crusades originated largely with the ideas of two influential, early-nineteenth century writers: Joseph-Francois Michaud (1767-1839) and Sir Walter Scott (1771-1832). An ardent French nationalist and proponent of imperialism, Michaud published an influential six-volume history of the Crusades, *Histoires des Croisades*, between 1812 and 1822. Removing the Crusades from their medieval context and stripping them of their penitential character, he recast them as glorious ventures in which brave French knights enthusiastically spread the virtues of France's superior culture to the people of the Middle East. Michaud thus transformed the Crusades into an anachronistic, proto-imperialist venture—in the process providing contemporary imperialists with the rhetorical ammunition needed to justify the acquisition of new colonies. Scott approached the Crusades in a very different but equally inaccurate way. A popular novelist, he wrote a series of books set during the time of the Crusades in which he developed two influential themes. First, he cast the Franks as brave but ignorant and uncultured louts who blundered violently into a peaceful and sophisticated Muslim world. Second, he described the Muslims—above all, Salah ad-Din—as people who were vastly more cultured and erudite than their crude and boorish European contemporaries.¹⁰⁰

To the lament of many modern scholars, Michaud's celebratory construction and Scott's more condemnatory one would establish the basis for contemporary Western historical memories of the Crusades. Michaud's positive interpretation cut multiple ways.

⁹⁸ Jonathan Riley-Smith, *The Crusades, Christianity, and Islam* (New York, NY: Columbia University Press, 2008), 70–71.

⁹⁹ Tyerman, *Crusades*, 136.

¹⁰⁰ Riley-Smith, *The Crusades, Christianity, and Islam*, 52–67.

On the one hand, it gave colonialists a language that they could use to justify imperialism; on the other, his positive account of the Frankish invasion of the Holy Land spurred neo-Marxist scholars to subsequently develop a highly critical—if also off base—interpretation of the Crusades as proto-imperialist ventures. Meanwhile, popular portrayals of the Crusades such as Ridley Scott’s film *Kingdom of Heaven* (2005)—which sharply contrasts violent, ignorant crusaders with urbane and cultured Muslims—reflect the enduring legacy of Sir Walter Scott’s critical interpretation of crusading.¹⁰¹

Ironically, contemporary understandings of the Crusades in the Middle East stem largely from the same sources. Prior to the late-nineteenth century, the Crusades were almost-completely forgotten in the Middle East. With the Franks long-since defeated and with the Ottoman Empire having shifted the border between the Christian and Muslim worlds deep into Central Europe there was little reason to pay attention to them; as a result, popular memories of the Frankish incursions faded rapidly in the Middle East. Beginning in the late-nineteenth century, however, the West reacquainted Muslims with the Crusades and, in the process, exposed them to the interpretations of Michaud and Scott that had come to predominate in Europe. The pivotal moment came at the very end of the nineteenth century. Inspired by Scott’s books, Kaiser Wilhelm II (r. 1888-1918) solemnly laid a wreath at Salah ad-Din’s tomb in Damascus in 1898—an act that many scholars argue reintroduced the Kurdish leader to a Muslim world that had all-but completely forgotten him.¹⁰²

Exposed to Western constructions of the Crusades and to surging Western imperialism, Muslim thinkers soon synthesized Scott and Michaud’s interpretations to form a new, critical understanding of the Crusades. It described the growing Western position in the Middle East in the late-nineteenth and early-twentieth centuries not as a product of the modern era, but instead as a continuation of the Crusades of the Middle Ages by the crusaders’ equally violent descendants. Lent credence by the inaccurate assumption that this interpretation had been Muslim conventional wisdom dating to the thirteenth century rather than a recent construct imported from Europe, this new interpretation soon became pervasive in the Middle East. It remains widespread today and has found particular appeal among antiimperialist Arab nationalists. As one writer declared in the 1930s, “the west is still waging crusading wars against Islam under the guise of political and economic imperialism.” More recently, jihadis have drawn similar conclusions. Sayyid Qutb (1906-1966), the so-called Godfather of Jihad, declared that “Western blood carries the spirit of the crusades within itself,” for instance, while Osama

¹⁰¹ Madden, *The Concise History of the Crusades*, 184–85; Riley-Smith, *The Crusades, Christianity, and Islam*, 66–67.

¹⁰² Riley-Smith, *The Crusades, Christianity, and Islam*, 62–66; For a persuasive rebuttal to the idea that Salah ad-Din was forgotten in the Muslim world see Diana Abouali, “Saladin’s Legacy in the Middle East before the Nineteenth Century,” *Crusades* 10 (2011): 175–89; Note, though, that Paul Cobb, otherwise supportive of Abouali’s argument, acknowledges that modern popular memories of him in the Middle East derive from European rather than Middle Eastern sources Cobb, *The Race for Paradise*, 5.

bin Laden (1957-2011) maintained that conflict between the West and the Muslim world in the early twenty-first century was merely the latest skirmish in a far-older civilizational conflict.¹⁰³

Thus, to the consternation of many scholars, the Muslim Middle East and the West alike frame relations today on the basis of inaccurate and ahistorical conceptions. Rather than locating the Crusades in the medieval culture to which they belonged, people instead cling to anachronistic constructions based on nineteenth-century polemics. At times, these “mythic memories,” to borrow Tyerman’s term again, lead people to the ludicrous, as when Islamists describe Israel as a latter-day Crusader State. At others, they fuel talk of a clash of civilizations—an idea that could prove dangerously self-fulfilling if it gains sufficient traction. More prosaically, historical memories rooted in the writings of Michaud and Scott impel people in both the West and the Muslim world to conclude incorrectly that contemporary conflicts between them lay rooted in a deep and unchangeable past, and thus fail to see that the origins of those issues lie firmly in the malleable present.¹⁰⁴

The Historical Debate

Historians do not merely take issue with popular understandings of the Crusades but also engage in vigorous debates among themselves over how the subject should be understood. The fact that the field continues to contest the meaning, origins, and consequences of the Crusades is hardly surprising given their importance to Europe’s internal development and their enduring significance in shaping popular perceptions of modern-day relations between the Middle East and the West. Two issues in particular have long animated historians of the Crusades: explaining what caused Europeans to launch them and assessing whether the Muslim leaders who fought against the Franks did so out of genuine religious conviction or out of a more cynical desire to win support against their rivals.

For much of the mid-twentieth century, most specialists argued that the Crusades were the product of economic forces. Writing often from a neo-Marxist perspective, they contended that class conflict and greed rather than idealism or devotion impelled people to go on crusade. For example, the French historian Georges Duby argued that crusading had its origins in the growing popularity of primogeniture among the nobility of eleventh-century Europe. In his view, the shift to that practice resulted in the creation of a class of disgruntled, landless younger sons whose unmet aspirations produced dissension and instability. The Crusades were a solution to this problem. They exported these unhappy and potentially dangerous young men to a place where they could acquire wealth and, more importantly—ensuring in the process continued stability in Europe. Other historians have cast the Crusades as proto-imperialist expeditions clothed in false piety that succeeded in enriching Europe at the expense of the Middle East. Amin Maalouf, for instance, has argued that the Crusades marked a watershed

¹⁰³ Riley-Smith, *The Crusades, Christianity, and Islam*, 71–78.

¹⁰⁴ Tyerman, *Crusades*, 136; Riley-Smith, *The Crusades, Christianity, and Islam*, 72–73.

moment for both the West and the Muslim world. For Europe, they were successful colonialist ventures that initiated a sustained period of growth; for the Middle East, in contrast, they marked the point at which the economy and culture began to stagnate.¹⁰⁵

Few historians today continue to argue that economic forces produced the Crusades. Drawing on new evidence that has shown that crusading was enormously expensive and that European nobles had to finance their participation by selling properties or mortgaging them at high interest rates, scholars now largely agree that people went on crusade out of a genuine and deeply felt sense of piety rather than out of a desire to get rich. In similar fashion, historians now largely reject the view that the Crusades simultaneously enriched Europe and impoverished the Middle East. In recent years, in fact, many scholars have flipped that interpretation on its head. The historian Thomas Madden, for one, not only maintains that the Crusades “constituted a massive drain on resources” in Europe, but that the economy of the Middle East continued to thrive long after the demise of Outremer.¹⁰⁶

Scholars have engaged in a parallel debate regarding the depth of Zengi, Nur ad-Din, and Salah ad-Din’s dedication to eradicating Outremer. Most modern scholars are skeptical that the destruction of the Crusader States was the primary objective of leaders like Zengi. Indeed, many who study the Crusades now contend that—jihadi rhetoric aside—the apparent hostility of those leaders to the Franks reflected more a desire to burnish their credentials in the Muslim world than a deep-seated commitment to holy war against the Crusader States. Noting that those leaders spent far more time warring with rival Turkish princes than fighting the Christians, for instance, the scholars Malcolm Lyons and D. E. P. Jackson argue that Zengi and Nur ad-Din had little commitment to driving the crusaders out of the region and used jihadi rhetoric largely to win support for their personal and dynastic goals. In Lyons and Jackson’s eyes, even Salah ad-Din’s reconquest of Jerusalem reflected such objectives. They maintain that while his capture of the city certainly contributed to the weakening of the Crusader States, it was aimed more at securing popular support for his campaigns against his Muslim rivals than at winning a holy war against the Franks.¹⁰⁷

Other scholars disagree with this line of reasoning. Challenging it directly, they argue that a leader’s public utterances carried great weight in the medieval Middle East and thus reveal that individual’s goals. While not denying that Salah ad-Din likely sought to expand his principality at the expense of his Muslim rivals, for example, the historian Carol Hillenbrand contends that his public embrace of jihad against the Crusader States was ultimately of far-greater significance than his personal and dynastic ambitions. As she writes, what was important was not his private motives, but instead “how his

¹⁰⁵ Amin Maalouf, *The Crusades Through Arab Eyes* (New York: Schocken Books, 1984), 264; Jotischky, *Crusading and the Crusader States*, 12.

¹⁰⁶ Frankopan, *The First Crusade*, 107; Madden, *The Concise History of the Crusades*, 189.

¹⁰⁷ Malcolm Lyons and D.E.P. Jackson, *Saladin: The Politics of the Holy War* (Cambridge: Cambridge University Press, 1982), passim.

contemporaries—his fellow military commanders, his personal advisers and the religious classes—saw him and reacted to him.”¹⁰⁸

Conclusion

With the end of the Crusades and, more importantly, the conversion of the Ilkhanid Mongols, the existential crisis that the Muslim Middle East had faced in the thirteenth century came to an end. What followed, however, was anything but an era of stability. The near-simultaneous fracturing of the Byzantine Empire in 1204, collapse of the Seljuq Sultanate of Rum in 1243, and outright destruction of the Abbasid Caliphate in 1258 produced a massive power vacuum in the Middle East that neither the Mamluks nor the Ilkhanid Dynasty—which itself fell apart in 1335—could fill. Instead, an obscure Turkish lord named Osman (r. 1299-1326) would establish a new state, the Ottoman Empire, in western Anatolia that would soon fill the political void. It is to the establishment of this new empire, and to the creation of a similar state in Iran, the Safavid Empire, that we shall now turn.

¹⁰⁸ Hillenbrand, *The Crusades*, 186.

Chapter Six: The Gunpowder Empires

Following the near-concurrent sack of Constantinople in 1204, collapse of the Seljuq Sultanate of Rum in 1243, and, especially, destruction of the caliphate in 1258, the Middle East appeared to have fallen into complete disunity. While the Mamluks controlled Egypt and Syria and a rump Byzantine state held sway in far-western Anatolia and Constantinople, the rest of the region lay hopelessly divided among a vast number of squabbling principalities, statelets, and tribal confederations. Indeed, by 1300, the stability and unity that had characterized the Middle East under the Abbasid Dynasty had become little more than a distant memory—one that seemed unlikely to ever again obtain in the region.

By the early sixteenth century, however, the situation in the Middle East had altered dramatically. By that point, two large states had emerged to divide the region between them: the powerful Safavid Empire that dominated Iran, and the mighty Ottoman Empire that controlled Anatolia, the Levant, Egypt, Iraq, and the Hejaz, as well as North Africa, Southeastern Europe, and the Black Sea littoral. What accounted for this dramatic geopolitical restructuring? What characteristics did these new empires have that permitted them to avoid the chronic instability that marked nearly all other contemporary states in the region?

In the mid-to-late twentieth century, the historians Marshall Hodgson and William McNeill provided a persuasive and widely accepted answer to these questions. These scholars located the success of these new states—as well as the similar Mughal Empire in India—in their adoption of an innovative system of interrelated administrative and military practices. Dubbing these states “gunpowder empires,” Hodgson and McNeill maintained that they were able to overcome the volatility that plagued other late-medieval Muslim states by centralizing political power, by creating and maintaining efficient bureaucracies able to raise substantial revenue, and, most importantly, by establishing regional monopolies over the use of firearms. Enormously powerful as a result of these innovative practices, these conquest states dominated the Muslim world and, in the case of the Ottoman Empire, constituted a serious and ongoing threat to the security and independence of Christian Europe in the early-modern period.¹

¹ William McNeill, “The Age of Gunpowder Empires, 1450-1800,” in *Islamic & European Expansion: The Forging of a Global Order*, ed. Michael Adas (Philadelphia: Temple University Press, 1993), 125–29; Marshall G. S. Hodgson, *The Venture of Islam, Volume 2: The Expansion of Islam in the Middle Periods* (Chicago: The University of Chicago Press, 1974), 405–6.

The Ottoman Rise to Power

Origins of the Ottoman Empire, 1299-1326

The Ottoman Empire emerged out of the anarchic situation that the sudden collapse of the political order in Asia Minor had created. The sack of Constantinople in 1204 had left the Byzantine Empire with inadequate resources to deal with its many enemies and thus rendered it too weak to defend its remaining territorial holdings in the far-western part of Anatolia; meanwhile, the demise of the Seljuq Sultanate of Rum in 1243 had resulted in the emergence of numerous, competing Turkish principalities called *beyliks* in central and eastern Asia Minor. Each of these small statelets took shape around a Turkish warlord who held the rank of *bey*—the lowest Turkish noble title. That the men who ruled these states possessed titles may suggest that the *beyliks* existed within a larger political arrangement or that their leaders enjoyed a degree of legitimacy conferred by a central government; such was most certainly not the case. On the contrary, this was a highly competitive and unstructured system wherein the power and status of a *bey* derived not from his title or from the legitimation that a respected authority might grant, but instead from his raw military talents and concomitant ability to attract followers. Those *beys* who could provide protection to peasants and merchants and who could lead frontier warriors called *akinjis* on successful raids gained subjects and saw their *beyliks* grow; those who could not suffered a decline in power as their followers deserted them in favor of their more effective rivals.²

The Ottoman Empire owed its start—and its name—to one of these *beys*, Osman I (r. 1299-1326). Osman I may have inherited a title and a tiny *beylik* from his father, but his success derived far more from his innate charisma, foresight, and remarkable leadership skills than from his patrimony. A military commander with the cunning to win battles and a lord with the wisdom to keep taxes low and predictable, he possessed all of the attributes needed to win supporters and territory at the expense of the less-successful and less-dynamic lords who controlled adjoining *beyliks*.³

As important, Osman I also had the savvy and talent to take full advantage of his principality's serendipitous location adjacent to the Byzantine Empire's remaining holdings in Asia Minor. The empire offered an almost perfect combination of wealth and weakness. On the one hand, its Anatolian possessions were endowed with rich pasturage, productive farmland, and small-but-important cities such as Bursa, Nicomedia and Nicaea. On the other, thanks to the more immediate threat that Western states such as Venice posed, Constantinople was compelled to deploy the bulk of its limited military forces to safeguard the empire's European territories and was thus unable to spare any more than a tiny fraction of its already small army for the defense of

² Douglas Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals* (Boulder: Westview Press, 2011), 34–35.

³ Caroline Finkel, *Osman's Dream: The History of the Ottoman Empire* (New York: Basic Books, 2005), 7–8.

its Anatolian holdings. As a result, those valuable possessions were isolated, vulnerable, and ripe for the plucking.⁴

Osman I built both his state and his reputation by preying on the empire. Grasping the weakness of its position in Anatolia, he began attacking it almost immediately after he took power. He began by boldly besieging the former imperial capital of Nicaea. While he was unsuccessful in taking the city, his victory over a Byzantine relief army in the Battle of Bapheus in 1302 earned him a reputation as a successful military commander among the Turks of Asia Minor and drew large numbers of *akinjis* to his banner.⁵

His growing success in Anatolia starkly revealed the dilemma in which the Byzantine Empire found itself. In desperation, Emperor Andronicus II Palaeologus (r. 1282-1328), responded by hiring a mercenary group, the Grand Catalan Company, to defend the territory in Anatolia that Osman I threatened. This move proved to be a poor one. While the mercenaries initially won several engagements against the Turks, they soon abandoned the war in favor of plundering the Greek peasants of the region. As a result, by the time the company had departed from Anatolia, many Byzantine subjects in the area had either fled across the strait to Constantinople or had come to accommodate themselves to Turkish rule.⁶

Osman I was quick to take advantage. Strengthened by his acquisition of new territory and subjects, He exploited the disarray that the Grand Catalan Company had created by launching annual raids into Byzantine territory. These attacks netted substantial loot, which drew yet more fighters to his cause. They in turn gave Osman I the strength first to gradually seize the Byzantine Empire's remaining rural land in western Asia Minor and then, around 1320, to lay siege to Bursa. Well-fortified and supplied, the city proved to be beyond the power of Osman I's warriors and managed to hold out through his death in 1326. Still, this setback aside, he had achieved a great deal. He had dramatically expanded his *beylik*, won a large following of *akinjis*, and, perhaps most importantly, had given his embryonic dynasty the imprimatur of success.⁷

Early Expansion, 1326-1402

Osman I's son and successor, Orhan (r. 1326-1362), proved to be every bit as capable and ambitious a leader as his father had been. Building on the stable foundation he inherited, Orhan skillfully transformed his *beylik* into a major power in western Anatolia. Like Osman I, he did so largely by exploiting Byzantine weakness. He began his reign by successfully concluding the siege of Bursa that Osman I had started. He followed by recruiting nearby Turkish *beys* to help him undertake an ambitious effort

⁴ Finkel, 7–8.

⁵ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 34–35.

⁶ Rudi Paul Lindner, *Nomads and Ottomans in Medieval Anatolia* (Bloomington: Research Institute for Inner Asian Studies, 1983), 25–27.

⁷ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 8, 12–13.

to overrun Byzantium's remaining territory in Asia Minor. This long and ultimately successful campaign was of enormous significance for Orhan. It not only resulted in the expansion of his state and its tax base, but, perhaps more importantly, it also elevated him above the nearby *beys* and thus gave him the prestige needed to start reducing those leaders to the status of Ottoman vassals.⁸

In doing so, Orhan pioneered what would become the dynasty's customary approach to acquiring territory. This process involved two steps. First, the Ottomans reduced defeated rivals or dependent principalities to vassalage. Then, after those states had accommodated themselves to Ottoman rule, the dynasty gradually assumed direct control of them. Minimizing the number of revolts during the early years of imperial oversight, this two-step progression proved to be a shrewd and highly successful strategy—one that the dynasty would employ to great effect over the next two centuries.⁹

Orhan had been enormously successful, but his territorial acquisitions had also produced a growing problem for his state. Put simply, the crude system of government that he had inherited was not appropriate for the changed circumstances of his *beylik*. Now encompassing a large territory of predominantly settled rather than nomadic people, his state required a far-more complex and sophisticated system of government than the one based on interpersonal relationships that he had inherited from his father. Accordingly, after he took Bursa, he sought to establish a simple bureaucracy to administer his state and, more importantly, to manage the collection of taxes. But who would organize and staff this bureaucracy? After all, neither the *akinjis* he surrounded himself with nor the peasants who constituted the majority of his subjects had the experience and knowledge needed to establish and run a complex administration.¹⁰

Unable to secure suitable people within his *beylik*, Orhan decided to look elsewhere for bureaucrats. He found them among the Sunni religious scholars of Anatolia. Promising that they would have the power to shape and manage his state, he succeeded in persuading a number of 'ulama' to relocate to Bursa. This move proved to be a smart one for the Ottoman leader. Drawing on centuries of well-developed Islamic political theory and practice, the religious scholars he recruited quickly established a crude-but-effective bureaucratic structure that was capable of providing him with the efficient administration and steady tax collection that his growing state required.¹¹

Perhaps unsurprisingly, the decision to rely on 'ulama' to run the government had important if-unintended consequences for the practice of religion in Orhan's state. To that point, the Ottomans had followed an unconventional brand of Islam that incorporated folk variants of Sunnism and Sufism into a syncretic blend of Christianity and Islam—a mix that made Orhan's *beylik* a welcoming place for people from a variety

⁸ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 35–36.

⁹ Streusand, 79–80.

¹⁰ Lindner, *Nomads and Ottomans in Medieval Anatolia*, 6–8.

¹¹ Lindner, 6–8.

of backgrounds. The embrace of heterodox religious practices began to ebb with the arrival of the 'ulama', however. Under their influence, Orhan and a number of his followers started to gradually abandon the loose approach to religion to which they had heretofore adhered in favor of an embrace of orthodox Sunnism. As we shall see, this shift would gradually accelerate under his successors.¹²

Orhan also began the process of transforming the Ottoman military from one that consisted exclusively of nomadic frontier warriors into a balanced one that included a substantial contingent of infantry. The key event that drove him to make this change was the siege of Bursa. Orhan learned from his early failed efforts to take the city that the *akinjis'* independence, lack of discipline and, most importantly, need to secure pasturage for their horses meant that he could not concentrate his forces for a long-enough duration to reduce fortified places through siege. Accordingly, he established the complementary *yaya* infantry to facilitate such operations. Based not on the nomads who continued to make up the bulk of his fighters but instead on settled recruits, this new force was able to maintain the tight blockades required to compel fortified places to capitulate. Increasingly important as Orhan shifted more and more to taking fortified urban centers, the *yaya* and other new types of troops would gradually supplant the nomads as the core of the more-and-more sophisticated and centrally directed Ottoman military.¹³

During the final decade of his reign, Orhan focused on a new and ambitious goal: expanding across the Dardanelles Strait into the Byzantine Empire's European territory. Ironically, his troops first crossed the strait not as invaders but instead as the invited allies of the Byzantine Emperor John V Cantacuzenus (r. 1347-1354), who was then locked in one of the many bitter civil wars that dominated Byzantium's internal affairs during much of the fourteenth century. Orhan's troops helped the emperor defeat his rival; in the process, however, they also seized control of the important port of Gallipoli on the European side of the Dardanelles Strait. Now in possession of a secure base in the Balkans, Orhan wasted no time in gobbling up the Byzantine Empire's European territory. Taking advantage of both the continued disarray in Constantinople and the sharp divisions that existed among the other Balkans powers, his troops engaged in a furious campaign of conquest that culminated in 1361 with the seizure of the important city of Edirne. The operation was another huge victory for Orhan. Now ruling substantial holdings in both Europe and Asia Minor, he had vaulted far ahead of the *beys* that had been his peers only a few years earlier. Accordingly, he assumed a new title that was more in keeping with his newfound stature and power: sultan.¹⁴

Orhan's son and successor, Murad I (r. 1362-1389), built on his father's success. He began his reign by relocating the capital from Bursa to Edirne so that he could more easily coordinate Ottoman efforts to expand into the weak and divided Balkans.

¹² Lindner, 6–8.

¹³ Lindner, 30–32.

¹⁴ Jason Goodwin, *Lords of the Horizons: A History of the Ottoman Empire* (New York: Henry Holt and Company, 1998), 14–15.

Concurrently, he dramatically accelerated Orhan's effort to remake the government from an informal one geared primarily to ruling a small nomadic principality into a highly centralized apparatus capable of administering a large, sedentary, multiethnic empire. He was quite successful in this effort. Indeed, more than any other Ottoman leader, it was Murad I who was responsible for establishing the basic structure of the Ottoman state.¹⁵

The new sultan also made a profoundly consequential addition to the Ottoman army. Intent on reducing the power of the Turkish vassals on whom Orhan had depended for much of his military force, Murad I established a new, tightly disciplined and highly loyal formation of infantry: the Janissaries. Following the pattern established by the Abbasid caliph al-Mutasim (r. 833-842) in the early ninth century, he recruited these soldiers from slaves. Called "new troops" or *yeni ceri*, from which the English term Janissary comes, these elite soldiers would be all-but unbeatable over the next three centuries and would prove critical in permitting the Ottomans to gradually end the independence of the other *beyliks*.¹⁶

In the immediate term, Murad I used the Janissaries to expand the empire in the Balkans. This campaign began with a great triumph over a coalition of Serbian and Bulgarian lords in 1371 on the Maritsa River. The victory permitted the sultan to take control of all of Macedonia and to turn the Serbian and Bulgarian nobles who had fought against him into Ottoman vassals. Two years later, he similarly reduced the Byzantine Empire to vassalage—surely a difficult pill for the heirs of the caesars to swallow. Finally, in 1389, he invaded Serbia to retaliate for the defeat of a small Ottoman army there a year earlier. Though Murad I died in the fighting, the Battle of Kosovo that marked the climax of the campaign proved to be a stunning Ottoman triumph and a battle of enormous significance. In the immediate term, the victory over a combined Serbian, Kosovan, and Bosnian army established the Ottomans as the now-unquestioned dominant power in the Balkans. The battle's long-term consequences were less tangible but equally important. As we shall see in chapter eight, nineteenth-century Serbian nationalists would reinterpret their forebears' defeat as a noble sacrifice against an unbeatable enemy and would, in so doing, transform the Battle of Kosovo into the central myth of modern nationalist conceptions of Serbian identity.¹⁷

Retrenchment and Recovery, 1402-1451

Murad I's successor, Bayezid I (r. 1389-1402), enjoyed a spectacular if ultimately unsuccessful reign as sultan. Dubbed Yidirim, or the lightning bolt, due to the quick pace of his conquests, he focused on rapidly and massively expanding the size of the empire. In the Balkans, he formally annexed Bulgaria as far as the Danube River,

¹⁵ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 17, 20.

¹⁶ Douglas A. Howard, *A History of the Ottoman Empire* (Cambridge: Cambridge University Press, 2017), 19.

¹⁷ Donald Quataert, *The Ottoman Empire, 1700-1922, Second Edition* (Cambridge: Cambridge University Press, 2005), 20.

initiated a long siege of Constantinople, and, at the Battle of Nicopolis in 1396, crushed a large army of crusaders sent to check his rising power. In Asia Minor, meanwhile, he waged a series of campaigns against rival Turkish states, gained territory through marriage, and, importantly, assumed direct control of a number of *beyliks* heretofore ruled by vassals. By 1400, as a result, the empire had become a major power—one that thoroughly dominated the northern Balkans and completely controlled western and central Anatolia.¹⁸

Impressive as Bayezid I's achievements were, however, they ultimately proved ephemeral. Ironically, it was his aggressive campaign of territorial acquisition and political consolidation that proved to be his undoing. By abandoning his predecessors' more cautious, two-step approach to expansion, he created enormous dissension among the recently deposed *beys* of Asia Minor. Intent on recovering their thrones, they appealed for assistance to Timur (r. 1370-1405)—known in the West as Tamerlane—the powerful sultan of the short-lived Timurid Dynasty that dominated Central Asia and Iran. Eager to cut down a rising rival, Timur was only too happy to oblige. In 1401, he lured Bayezid I into a confrontation by invading Anatolia. In the Battle of Ankara that took place the following year, he decisively defeated the Ottoman army—in the process, capturing and deposing Bayezid I.¹⁹

Despite the scale of his victory, Timur did not assume control of the Ottoman Empire or oversee its dismemberment. Instead, he imposed two significant changes designed to permanently weaken it. First, he reduced the power of the empire by restoring the Anatolian *beys* that Bayezid I had deposed. Second, he restructured the Ottoman government along traditional, nomadic lines by imposing on it the appanage model of collective rule to which the Mongols and Seljuq Turks had adhered. In keeping with that system, he divided the Ottoman territories among three of Bayezid I's sons, Sulëiman, Isa, and Mehmet I (r. 1413-1421).²⁰

The subsequent outbreak of a long, bitter civil war between the brothers gave the Christian rulers of the Balkans a much-needed respite. Not only did the Ottomans' internal struggle put the wars of conquest on hold, but it also spurred the brothers to adopt conciliatory diplomatic strategies toward the Christian states in hopes of dissuading them from intervening. Emblematic of their new approach were Sulëiman's entreaties. Eager to concentrate on defeating his brothers without fear of Western action against him, he freed the Byzantine Empire from vassalage and granted the Venetians and Genoese important commercial rights. These concessions were substantial, and they made plain that the balance of power in the region had shifted dramatically against the Ottomans. Indeed, as the fighting between the brothers neared its tenth anniversary, it appeared to Christian leaders such as Byzantine Emperor

¹⁸ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 25–28.

¹⁹ Howard, *A History of the Ottoman Empire*, 41.

²⁰ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 38.

Manuel II Palaeologus (r. 1391-1425) that the Ottoman threat had not merely crested but had begun to recede.²¹

They were mistaken. The empire did not continue to weaken, but instead experienced a rapid recovery following Mehmet I's victory in the civil war. His triumph over his brothers was a product of his shrewdness and patience as a leader. Rather than going for a knock-out blow early in the conflict, he smartly hung back and let his brothers weaken themselves fighting each other before he stepped in to pick up the pieces. As a result, he emerged victorious in 1413—the sole survivor and, from that point, the uncontested ruler of the Ottoman Empire.²²

He, and his son and successor, Murad II (r. 1421-1451), spent the next four decades overseeing the Ottoman recovery from Bayezid I's defeat. Each faced serious threats that complicated that effort. Between the end of the civil war in 1413 and his death in 1421, Mehmet I had to deal with the Venetian-supported revolt of yet another brother and had to put down a large uprising organized by the popular Sufi leader Sheikh Bedreddin (1359-1420). These were serious threats, and they greatly complicated his efforts to restore the dynasty to the position it had occupied on the eve of Timur's victory over Bayezid I. Mehmet I was up to the challenge, however. He was not only able to reestablish the dynasty's position in the Balkans but was also successful in reasserting substantial sultanic authority in Asia Minor.²³

Despite having to deal with a number of serious threats during his tenure as sultan, Murad II was even more successful than his father had been at restoring the dynasty's position. Those challenges were significant. He faced separate uprisings launched by an uncle and a brother, the continued efforts of the Anatolian *beys* to retain the autonomy that Timur had granted them, and Christian attacks aimed at weakening the empire's grip on the Balkans. Murad II addressed all of these threats swiftly and effectively. In the process, he restored full Ottoman control over Asia Minor for the first time since Timur's victory over Bayezid I. Meanwhile, his decisive triumph over a combined Hungarian-Polish army at Varna in 1444 ended Christian European efforts to weaken the empire's grip on the Balkans and fully secured its European borders. Thus, by the time of his death in 1451, he had completed the work that his father had started. Together, they had restored the empire to the position it had held in 1402 and had left it poised to become a dominant power in both Europe and the Middle East.²⁴

²¹ Steven Runciman, *The Fall of Constantinople 1453* (Cambridge: Cambridge University Press, 1965), 43.

²² Howard, *A History of the Ottoman Empire*, 41–42.

²³ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 38–39.

²⁴ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 33–38.

Ottoman Dominance, 1451-1566

Mehmet II

Before the Ottoman state could make that leap, however, it first needed to complete the destruction of the Byzantine Empire and, most importantly, to acquire its lone remaining city: Constantinople. Laying at the very heart of the Ottomans' expanding realm, the city was essential to the sultanate's continued success. Control of it would give the empire complete dominance of the lucrative trade between the Black Sea and the Mediterranean and would ease communications between its European and Asian holdings. Successfully reducing the city also promised to dramatically enhance the standing of the sultan. Muslim tradition had long attached eschatological significance to Constantinople, holding that the person who conquered it would be the Mahdi: the expected one who would usher in the end of history. The sultan who finally succeeded in taking it would thus not merely reap the material benefits attendant in the possession of such a fabulously well-located city but would also see his prestige and legitimacy dramatically enhanced.²⁵

No one was keener to take Constantinople than Murad II's son and successor, Mehmet II (r. 1451-1481). His interest in the city stemmed partly from his awareness of the substantial strategic and economic benefits that possession of it would bring. However, his desire to conquer it was largely a product of the tenuous position in which he found himself upon taking power. Only nineteen-years old when he acceded to the throne, he confronted whispered questions in the court and the army about his wisdom and capacity for leadership. Worse, many of the Turkish *beys* whom his father had only recently reduced to vassalage were scheming with the independent Turkish state of Karaman to regain their independence. These were serious challenges—ones that Mehmet II concluded he would be unable to deal with unless he quickly legitimated his rule through a great victory. Accordingly, over the objections of his powerful and long-serving grand vezir, Candarli Halil Pasha (?-1453), he ordered the Ottoman army to lay siege to the city beginning in April 1453.²⁶

Despite the fact that the sultanate's powerful military massively outnumbered the Byzantine defenders, taking the city was far from a foregone conclusion. Ringed by the massive, millennia-old triple Walls of Theodosius and able to bring in food by ship owing to its location on the Bosphorus Strait, Constantinople had famously withstood over a dozen sieges including recent ones undertaken by Bayezid I, Mehmet I, and Murad II. Undeterred, Mehmet II directed a well-thought-out campaign to reduce the city. First, he ordered the erection of a castle along the Bosphorus Strait called Rumeli Hisari—meaning “Strait Cutter”—from which cannons could interdict the shipment of food to the city. Second, and more importantly, he authorized the construction of a new type of cannon called a bombard that was designed to destroy the city's walls. Gunpowder

²⁵ Tamim Ansary, *Destiny Disrupted: A History of the World Through Islamic Eyes* (New York, NY: PublicAffairs, 2009), 175.

²⁶ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 45–49.

weapons were already centuries old, and smaller cannons had been commonplace in the region for decades. What made the bombards that Mehmet II had commissioned different was the fact that they were of an entirely unprecedented scale. Cast by a Hungarian metallurgist named Urban, the largest one had a twenty-eight-foot-long barrel and could fire a 1,300-pound granite projectile as far as a mile.²⁷

Still, even with these weapons, besieging the city was a risky bid on the sultan's part. In effect, he was gambling that he could do with his bombards what no previous attacker had managed to accomplish: smash down the Walls of Theodosius. Were he to succeed, he would have the legitimacy to deal with those who challenged him; were he to fail, he would instead almost certainly be deposed and killed.²⁸

Mehmet II's bet paid off. In systematic fashion, his bombards battered down a substantial section of the city's walls over the course of April and May. Finally, on May 29, the Ottoman army charged through the rubble and overwhelmed the outmatched and demoralized defenders. It was a momentous occasion. Not only did it mark the first time in history that cannons had succeeded in taking a major walled city, but it also dramatically raised Mehmet II's political standing. Known ever after as "the Conqueror," he now had the prestige needed to remake the empire and to complete the consolidation of power that his predecessors had begun.²⁹

He began with a concerted effort to transform Constantinople—or, as the Turks called it, Istanbul—into a suitable capital for a great empire. This task promised to be a difficult one. Severely depopulated by the outbreak of the Black Death in 1347 and impoverished by its economic decline during the terminal decades of the Byzantine Empire, Constantinople was as much a ruin as an urban center at the time of the Ottoman conquest. Undaunted, Mehmet II moved aggressively to restore it to its former splendor. He poured vast amounts of money into the city's reconstruction and swelled its population by successfully encouraging whole communities to relocate to it. He also moved to link the Ottoman Empire to the city and its glorious imperial past by turning the Hagia Sophia—its greatest cathedral and a literal edifice of Roman authority—into a mosque. Finally, he relocated his government to the newly constructed Topkapi Palace, a magnificent administrative center and royal residence in Constantinople that overlooked the Bosphorus Strait. These initiatives effected a startlingly rapid recovery. Within just a few short decades, the city was once again one of the world's great administrative, cultural, and commercial centers—one that both tied the empire together and gave it control over the main trade routes of the Black Sea and eastern Mediterranean.³⁰

²⁷ Goodwin, *Lords of the Horizons*, 32.

²⁸ Goodwin, 32.

²⁹ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 32.

³⁰ Marie-Janine Calic, *The Great Cauldron: A History of Southeastern Europe*, trans. Elizabeth Janik (Cambridge, MA: Harvard University Press, 2019), 72–73.

Concurrently, Mehmet II used the newfound prestige and wealth he had acquired through the conquest of Constantinople to complete the centralization of the Ottoman government—known thereafter as the Sublime Porte—that his predecessors had begun. He first acted by purging high-ranking officials whose fealty he doubted. Most notably, just three days after his troops had stormed Constantinople, he ordered his heretofore powerful grand vezir, Halil Pasha, executed. He followed by replacing unreliable officials with Christian converts to Islam who were, as such, dependent on the sultan and thus loyal. Indeed, of the six grand vezirs who succeeded Halil Pasha, five were Christian-born men who adopted Islam.³¹

Meanwhile, he also effected a series of important symbolic changes designed to enhance his and his successors' authority. The most important of these involved the transformation of the position of sultan from that of a first-among-equals who, like Osman, ate and socialized with his soldiers, into an aloof, near-divine figure who stood apart from his subjects as the embodiment of power and dignity. Accordingly, Mehmet II no longer ran or even attended the four-times-per-week imperial council meetings with his vezirs; instead, he observed them behind a curtained grill called the Eye of the Sultan.³²

Though he focused a great deal of attention on reforming his government, Mehmet II did not neglect the traditional Ottoman pursuit of territorial expansion. Here, too, he enjoyed enormous success. He assumed direct control over Serbia, reduced Bosnia and Herzegovina to vassalage, seized the remaining Byzantine successor states in Greece, and annexed the troublesome Turkish Karaman principality. He also successfully besieged the strategic Genoese trading center of Caffa on the Crimean Peninsula; in the process, he not only gained control of this valuable commercial city but also reduced the region's powerful Turkish Tatar nomads to vassals and turned the Black Sea into a virtual Ottoman lake. Finally, he set off alarm bells in the West when he briefly seized the city of Otranto on the heel of the Italian peninsula in 1480.³³

Selim I

For reasons we will explore presently, Mehmet II's successor, Bayezid II (r. 1481-1512), focused largely on consolidating his father's gains and did not acquire substantial territory. However, the next sultan, Selim I (r. 1512-1520), returned with a vengeance to the traditional Ottoman pursuit of expansion. His initial conquests came in response to a new menace that the empire confronted at the start of his reign: the establishment and growth of the powerful Safavid Dynasty in Iran. Its emergence had substantially and adversely altered the power dynamics on the Ottoman Empire's eastern frontier and posed a serious danger to the sultanate's security. The Safavid threat took two forms. First, the Iranian state possessed a powerful military that constituted a direct menace to the Ottoman Empire's control of eastern Asia Minor. Second, and more dangerously,

³¹ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 75–76.

³² Finkel, 67–80.

³³ Howard, *A History of the Ottoman Empire*, 75–76.

the Safavids articulated a populist Shi'i ideology that encouraged the nomadic Turkmen of far-eastern Anatolia to resist the sultanate's ongoing effort to impose direct control on them.³⁴

While Bayezid II had taken a cautious approach to the Safavid Empire, Selim I instead adopted an aggressive policy toward it aimed at ending the threat it posed to the Ottoman Empire once and for all. Accordingly, after brutally suppressing the Turkmen—sources claim that he executed as many as 40,000 prisoners—he invaded Iranian territory. It was a risky but ultimately successful maneuver. Owing substantially to the fact that the Safavids had not yet adopted gunpowder weapons, the Ottomans won a decisive victory in the Battle of Chaldiran in 1514; as a result, the Safavid military threat faded, and the Turkmen quickly—if grudgingly—came to accept Ottoman rule.³⁵

Emboldened, Selim I followed by moving against the Mamluks. Doing so proved far trickier than going to war with the Safavids. The issue was the Mamluks' religion. Many of the sultan's troops were deeply troubled about fighting fellow Sunnis and indicated that they would not take part in a campaign against them. Selim I dealt with this challenge by asserting that the Mamluks had been aiding the Shi'i Safavids and were thus not good Sunnis. Though the evidence for collaboration between the regimes in Egypt and Iran was wafer thin, his claim was sufficient to quell dissent in the ranks and to permit him to go forward with the invasion. The resulting campaign was an overwhelming success. Once more facing an enemy that was literally outgunned—the Mamluks had only recently and tentatively embraced firearms—Selim I's army seized Syria by winning a decisive victory at the Battle of Marj Dabik near Aleppo in 1516. A second triumph the following year finished the Mamluk state and permitted Selim I to add Egypt and the Hejaz to his now much-enlarged empire.³⁶

Selim I's military success against the Mamluks did not merely add territory and subjects or enhance his political standing but also profoundly altered the character and identity of the empire. It did so in two ways. First, the acquisition of Syria and Egypt transformed what had been a Muslim-ruled but Christian-majority empire into one in which Sunnis predominated. More importantly, the triumph over the Mamluks gave the Ottomans possession of Mecca and Medina and thus made the sultans the protectors of the Holy Places. Selim I and his successors were quick to grasp that their newfound status as the protectors of the Holy Cities promised to greatly enhance their legitimacy in the eyes of their now-majority Sunni subjects—but only if the dynasty adhered scrupulously to orthodox Sunni beliefs and practices. Accordingly, from that point forward, the Ottomans presented the empire as the leading Sunni Muslim state and acted to legitimate their rule through the promotion and embrace of traditional Sunnism.

³⁴ Gene R. Garthwaite, *The Persians* (Malden, MA: Blackwell Publishing, 2005), 161–65.

³⁵ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 102–6.

³⁶ Finkel, 106–9.

Thus, the shift from heterodoxy to orthodox Sunnism that had begun under Orhan was now complete.³⁷

Selim I's conquests also dramatically altered the Ottoman state's foreign policy. Control of Egypt, the Hijaz, and Syria turned what had been an empire of great regional significance in Anatolia and southeastern Europe into a truly global power—a three-continent behemoth that dominated all its frontiers. In response, the dynasty began to adopt a more-coherent, global strategy in its relations with other state.³⁸

Sulëiman the Magnificent

This global approach was evident in the reign of Selim I's successor, Sulëiman (r. 1520-1566). Known as the Magnificent in the West and as the Lawgiver in the Muslim world, Sulëiman pursued a complex, global grand strategy that saw the empire reach dizzying new heights. In Europe and the Mediterranean, he sought consolidation and security. To achieve those ends, he began his reign by taking two fortresses that had defied Mehmet II: the city of Belgrade, which had blocked the Ottomans from expanding into Central Europe, and Rhodes, an island that had long served as a base for piracy against Ottoman shipping. He followed his success at Belgrade with a crushing triumph at the Battle of Mohacs in 1526 that turned Hungary into an Ottoman-controlled buffer state. Three years later, he solidified his position in Central Europe by easily defeating Habsburg Emperor Charles V's (r. 1519-1556) attempt to retake Hungary and by briefly besieging Vienna, the capital of Habsburg Austria. Finally, taking advantage of an alliance he had struck with Charles V's dynastic rival, French King Francis I (r. 1515-1547), Sulëiman was able to complete the conquest of Algeria in the 1530s. Thus, by the time of his death in 1566, the Ottoman Empire had fully secured its position in both Central Europe and the Mediterranean.³⁹

Sulëiman's strategy in the Indian Ocean was less successful. In that region, he sought a largely defensive goal: the preservation of Ottoman commercial dominance from growing European competition. Direct Western involvement in the Indian Ocean was a recent and troubling phenomenon for the Ottomans. In 1498, the Portuguese mariner Vasco da Gama (c. 1460s-1524) had arrived in Calicut, India after a long, perilous voyage around the Cape of Good Hope—inaugurating, in the process, the first direct connection between the spices of South and Southeast Asia and the markets of Europe. It was a major commercial and nautical accomplishment. However, the Portuguese were not content merely to establish this link or even to hold a monopoly on the direct trade between Europe and the Indian Ocean. Instead, they followed da Gama's voyage with a shrewdly designed campaign to assume a dominant position over long-distance trade in the east. They seized control of key bottlenecks such as the Strait of Malacca in Southeast Asia, the Bab el-Mandem Strait at the opening to the Red Sea, and the Strait of Hormuz at the mouth of the Persian Gulf and then used their

³⁷ Calic, *The Great Cauldron*, 76–77.

³⁸ Calic, 76–77.

³⁹ Goodwin, *Lords of the Horizons*, 85–88.

command of those passages to deny competitors access to the Indian Ocean's sea lanes. It was a savvy if hard-ball strategy, and it helped to turn Portugal into a serious threat to the Ottoman Empire's commercial position.⁴⁰

In response, the sultan's government devised a strategy that sought to preserve the empire's continued commercial dominance. It centered on a complex, two-stage military campaign designed to eject the Portuguese from the Indian Ocean. Launched in the 1530s, its first phase was a success. Ottoman armies established control over Yemen as a preliminary step for an assault on Portugal's bases near the Bab el-Mandem Strait and took control of Iraq and the strategic port of Basra from the Safavids in preparation for a move aimed at driving the Portuguese from the Strait of Hormuz. The second phase was far-less successful, however. Despite launching coordinated assaults on Portuguese forts at the mouth of the Red Sea, at Hormuz, and even at key locations in India, Ottoman forces proved unable to overcome determined resistance. As a result, Sulëiman had to abandon the effort to force the Portuguese out of the region in favor of a diplomatic approach that ended with the two parties agreeing to respect each other's trade interests in the Indian Ocean.⁴¹

Despite his lack of success in countering Portugal's commercial threat, Sulëiman's reign is widely acknowledged as the high point of the Ottoman Dynasty. And what a zenith it was. Fiscally, the empire occupied an enviable position thanks to its robust agricultural economy, control of lucrative long-distance trade routes, and efficient system of tax collection. In the realm of religion, the dynasty enjoyed tremendous legitimacy and authority in the Muslim world owing to its embrace of orthodox Sunni Islam and its status as the protector of the Holy Place. Militarily, its mighty Janissary regiments were universally regarded as the most formidable formations in existence. Geopolitically, finally, the empire had no peer on the global stage. It controlled the Middle East and North Africa, had turned the eastern Mediterranean and Black Seas into Ottoman lakes, and had even assumed a position as a major player in the international order then taking shape in Europe. It was, in sum, an empire without equal—one that stood without question as the most powerful state on the planet.⁴²

The Historical Debate: The Ghazi Thesis

The above account of the Ottoman rise from frontier *beylik* to major power reflects a longstanding consensus among historians who have studied the empire. That is, in contrast to the debates over the emergence of Islam or the Arab Conquest, there have been no doubts raised about the chronology of the early Ottoman state or whether key figures described in the sources actually existed. Agreement on these fundamental points is not to suggest, however, that historians have held uniform views of the empire.

⁴⁰ Peter Frankopan, *The Silk Roads: A New History of the World* (New York: Knopf, 2015), 214–25; Finkel, *Osman's Dream: The History of the Ottoman Empire*, 128–29.

⁴¹ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 48–50; Finkel, *Osman's Dream: The History of the Ottoman Empire*, 128–29.

⁴² Ansary, *Destiny Disrupted*, 183.

On the contrary, scholars who focus on the Ottoman Empire have engaged in vigorous debates about a number of issues—particularly ones related to its nature and identity during its early years. What kind of state was it? Whom did the Ottomans permit to join its ruling class? What motivated the wars of expansion? Most importantly, what purpose did Osman I and his successors claim their sultanate was fulfilling?

For much of the mid-twentieth century, the Austrian-Jewish historian Paul Wittek's Ghazi Thesis dominated understandings of the empire. First articulated in a series of lectures given at the University of London in 1938 and later published as *The Rise of the Ottoman Empire*, the Ghazi Thesis took issue with the then-prevailing idea that the early empire was, at heart, ethnically Turkish in nature. Wittek rested his critique on a careful study of two key pieces of evidence from the early, ill-sourced years of the empire: a stone inscription in Bursa dating to 1337 and a work attributed to the fifteenth century Turkish poet Ahmedi (1334-1413), that characterized the early Ottomans not in ethnic terms as Turks but instead in religious ones as *ghazis*: Muslim frontier warriors who, motivated by religious zeal, fought ceaselessly to expand the realm of Islam. Based on this evidence and on corroborating secondary-source accounts written by religious scholars in the sixteenth century, Wittek concluded that ethnicity did not define the early Ottoman state; instead, religion—by which he meant the waging of holy war in the name of Islam—was the Ottoman's "dominant idea, the *raison d'être* of their state," and it remained so until just before the empire's demise in the early-twentieth century.⁴³

Wittek's Ghazi Thesis dominated interpretations of the early Ottoman Empire for four decades. Indeed, it was only after his death in 1978 that scholars began to seriously question both his conclusions and the evidence on which they rested. The historian Rudi Paul Lindner was one of the first to do so. Focusing on the action of the early Ottomans rather than on what legal scholars wrote about them in the sixteenth century, he observed that they routinely fought against neighboring Muslims, practiced heterodox rather than orthodox Islam, welcomed the participation of Greek Christian soldiers and commanders on their raids, and refrained from forcing conquered people to convert—actions that were hardly consistent with the behavior of the sort of hardened zealots that Wittek had described. Then, having dispensed with the idea that religion defined the early Ottoman state, Lindner followed by articulating a new interpretation of their motives based on anthropological theories about the nature of tribal groupings. Drawing on recent research that holds that tribes are a function of shared interests rather than ethnicity or common bloodlines, he argued that the Ottomans were an inclusive tribal group consisting of both Muslims and Christians bound together by a shared interest in raiding and territorial expansion for the purpose of obtaining loot. In his view, it was only after the 'ulama' that Orhan had recruited began to promote the

⁴³ Paul Wittek, *The Rise of the Ottoman Empire: Studies in the History of Turkey, Thirteenth-Fifteenth Centuries*, Royal Asiatic Society Books (New York: Routledge, 2012), 35.

ghazi ideal that the Ottomans abandoned their early religious inclusivity in favor of an embrace of holy war aimed at the territorial expansion of Islam.⁴⁴

While Colin Imber has recently contended that the absence of evidence renders the early Ottoman Empire a “black hole” to scholars seeking to uncover its true nature, other historians have instead amplified Lindner’s view that the sultanate was remarkably inclusive. Drawing attention to the astonishing number of Byzantine and Balkan nobles who held prominent positions in the state, for example, the historian Heath Lowry argues that the early Ottoman Empire is best understood as a highly meritocratic, “predatory confederacy open to all rather than as an Islamic-based” state dedicated to holy war. In his view, it was the demand for manpower that sparked this inclusive approach. Needing soldiers for their profitable raids and wars of expansion, the Ottomans willingly accepted Christian nobles into their military; eager to gain loot and slaves, the latter were only too happy to join. As a result, the Ottoman state developed an elite that was an amalgam of Muslim Turks and Greek Christians. Lowry concludes that it was only much later, when the dynasty had begun to see value in using orthodox Sunni Islam to legitimate its rule following the acquisition of the Holy Places, that Muslim writers—“projecting an Islamic past back in time”—reimagine the early Ottomans as *ghazi* warriors.⁴⁵

Sources of Ottoman Success

Regardless of the motives that animated its expansion, the speed of the dynasty’s rise is astonishing. A crude, frontier *beylik* in 1300, the Ottoman confederacy had become an expansive and powerful empire with complex institutions and sophisticated political ideologies by the time Mehmet II conquered Constantinople. What accounted for this remarkable transformation? What characteristics and circumstances set the Ottomans apart from rival states? Most importantly, why did the Ottomans and not some other group come to dominate the Balkans and Middle East in the Early Modern Period?

Contingency

To a significant degree, Ottoman success was a matter of good fortune. With the near-simultaneous fracturing of the Byzantine Empire, shattering of the Seljuq Sultanate, and demise of the Abbasid Caliphate, much of the Middle East had become a vast power vacuum. As a result, rather than standing in the shadow of a nearby great power, Osman I’s tiny *beylik* existed in a world of similarly small states—precisely the sort of environment in which a charismatic and able military leader could thrive. The dynasty also enjoyed the added good fortune of possessing a *beylik* that bordered a soft

⁴⁴ Lindner, *Nomads and Ottomans in Medieval Anatolia*, 1–38.

⁴⁵ Imber quote from Colin Imber, “The Legend of Osman Gazi,” in *The Ottoman Emirate: (1300-1389)*, ed. Elizabeth A. Zachariadou (Heraklion, Greece: Crete University Press, 1993), 75; Lowry quote from Heath W. Lowry, *The Nature of the Early Ottoman State* (Albany: State University of New York Press, 2003), 3, 134.

frontier in the form of the weak but wealthy Byzantine Empire rather than ruling a territory surrounded entirely by other aggressive Turkish *beys*. As we have seen, raiding and seizing territory from the Byzantines was both profitable and comparatively easy and allowed Osman and his successors to build the kind of military reputation that would attract growing numbers of frontier warriors to the Ottoman banner. Thus, just as had been the case for the Arabs who had conquered the Middle East in the seventh century and for the knights of the First Crusade who had taken Jerusalem in the eleventh century, luck played a substantial part in the early empire's success.

Luck can only go so far in explaining Ottoman success, however. Many of the contemporary Turkish *beyliks* were considerably larger and fielded many more soldiers than did Osman I's tiny state in 1300. Likewise, his *beylik* was hardly the only one with the good fortune to share a border with the vulnerable Byzantine Empire. Luck was thus an important precondition for the dynasty's astonishing rise, but it cannot explain why Osman I and his descendants rather than a different Turkish dynasty came to dominate the region and, thus, cannot alone account for Ottoman success. Ultimately, answering the question of why the empire rose so quickly from such humble origins requires an examination of its internal structures, for it was the ideologies, practices, and evolving institutions of the Ottoman ruling class that ultimately explain the dynasty's remarkable achievements.

Strong Leadership

One of the most obvious sources of Ottoman success was the empire's unusually effective leadership during its early centuries. As we have seen, from its founding at the turn of the fourteenth century through the end of Sulëiman's reign in the mid-sixteenth century, the sultanate benefited from a remarkable string of peculiarly effective rulers. Nearly all the men who served as sultan during that period—the Ottomans did not permit women to hold the throne—were wise and charismatic rulers who provided able administration and effective military command. As the writer Jason Goodwin notes, the nicknames they earned pay testament to their leadership abilities: the Grim, the Conqueror, the Lawgiver, and the Great are hardly the monikers of weak leaders. Moreover, the sheer length of their tenures in office—the first ten sultans enjoyed average reigns of a remarkable twenty-seven years—lent a degree of stability that was invaluable to the early Ottoman state.⁴⁶

How did the dynasty come to have such a remarkable succession of capable and far-sighted rulers? Chance once again surely played a part, but far more important was the dynasty's rejection of primogeniture in favor of a brutal, but effective practice known as fratricidal succession—a system that the Ottomans adopted to ensure, in the historian Donald Quataert's words, the "survival of the fittest, not eldest son." It was the anarchic interlude that followed Timur's defeat of Bayezid I in 1402 that pushed the dynasty to embrace fratricidal succession. As we saw earlier, Timur had imposed the traditional, Turko-Mongolian appanage system of shared family governance on the

⁴⁶ Goodwin, *Lords of the Horizons*, 167–68.

Ottomans following his victory—an arrangement that, unsurprisingly, produced a decade-long civil war. The victor in that conflict, Mehmet I, was determined to ensure that the dynasty avoided a repeat of that debilitating conflict; accordingly, not long after he had defeated his brothers, he embraced the practice of fratricidal succession as a necessary if unpleasant way of ensuring a stable succession.⁴⁷

The arrangement had two provisions. First, based on the prevailing view that each of the sultan's sons possessed an equal claim to the throne and was thus a possible future monarch, it stipulated that all of the heirs should learn how to rule effectively by first serving as provincial governors. Second, upon the death of the reigning sultan, it called for the sons to engage in a brief struggle for power. That contest would end when one of them made his way to the capital, secured the backing of the bureaucracy and military, and assumed the title of sultan. Once in power, the new ruler would then ensure that he faced no future challenge to his authority by having any surviving brothers executed.⁴⁸

The Ottomans rigorously adhered to this system for the entirety of the fifteenth and sixteenth centuries. At times, the results were shocking. For example, Mehmet II's first act as sultan was to command one of his assistants to suffocate his fifteen-month-old brother. Some of his successors went even further. Selim I had two of his brothers killed upon acceding to the throne while Murad III (r. 1574-1595) had all five of his brothers—four of whom were infants—executed. However, it was Mehmet III (r. 1595-1603) who set the dubious record of having the most siblings killed: upon taking power, he ordered his royal executioners to strangle all nineteen of his brothers.⁴⁹

The system of fratricidal succession may have been cruel, but it was also highly effective. Indeed, it provided four substantial benefits that, collectively, constituted the primary reason that the empire enjoyed a succession of strong leaders during the fifteenth and sixteenth centuries. First, having heirs gain experience by serving as provincial governors worked as intended: with great consistency, those who became sultan possessed the administrative and military experience needed to rule effectively. Second, fratricidal succession's Darwinian nature generally resulted in the most capable son emerging victorious, and, thus ensured that the strongest candidate became the new sultan. Third, the arrangement's winner-take-all principle proved effective in helping the empire avoid both the decentralizing tendencies of the traditional Turco-Mongolian appanage system and the debilitating strife that accompanied internal conflicts such as the civil war that had followed Bayezid I's reign. This last concern loomed large for Ottoman political theorists, who defended the system of fratricidal succession as a practical necessity. As one thinker later wrote about the murder of Mehmet II's infant

⁴⁷ Quote from Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 90–91; Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 72.

⁴⁸ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 91.

⁴⁹ Franz Babinger, *Mehmed the Conqueror and His Time* (Princeton, N.J.: Princeton University Press, 1992), 65; Colin Imber, *The Ottoman Empire, 1300-1650: The Structure of Power*, Second edition (New York: Palgrave Macmillan, 2009), 95–101.

brother, “it was . . . the better course to root up the sapling of mischief, before it put forth leaves and branches.” Finally, the system ensured that a pretender did not survive to become a leader around whom disaffected elements could rally against the reigning sultan.⁵⁰

The one instance during this period in which fratricidal succession failed to eliminate a rival claimant to the throne made clear the system’s benefits. It occurred when Bayezid II’s defeated brother, Cem (1459-1495), managed to evade the sultan’s executioners by fleeing into exile in Western Europe. For much of Bayezid II’s reign, Cem’s presence in Europe constituted a genuine danger to the stability of the empire. The issue was his ability to rally disaffected elements to his cause. Aware that Cem enjoyed broad support among the recently dispossessed Anatolian *beys*, his Western hosts used the threat of helping him raise a revolt in Asia Minor as leverage to successfully block Bayezid II’s efforts to expand at their expense. Fortunately, for Bayezid II, Cem proved to be only a transitory menace and never produced the mass uprising that the sultan feared. Nonetheless, his mere existence had succeeded in handcuffing the empire for a substantial period—a fact attested to by Bayezid II’s failure to expand the sultanate.⁵¹

Thereafter, and for the next century, the system of fratricidal succession did what the Ottomans had designed it to do. It ensured that the sultanate would not be hamstrung by a pretender like Cem. Likewise, it all but guaranteed that the empire would benefit from the rule of highly effective leaders like Selim I and Sulëiman. Thus, brutal as it was, the system of fratricidal succession functioned as intended: it made certain that the Ottoman Empire enjoyed a stable succession—something that could not be said about many contemporary states.⁵²

Efficient Administration

The early sultans proved successful not merely because they were capable leaders, but also because they oversaw a peculiarly effective central bureaucracy. Loosely patterned after the now-venerable Muslim tradition of slave-recruited militaries, the imperial administration of the fifteenth and sixteenth centuries was a uniquely Ottoman innovation that had no parallel in any other state. Staffed by men who were utterly loyal to the dynasty, it gave the sultans a remarkably efficient and sophisticated instrument through which they could implement their policies.

Initially, it did not vary appreciably from longstanding Islamic structures and practices. The religious scholars that Orhan had recruited to run his expanding government were well versed in the Muslim administrative tradition that dated to the Abbasid Dynasty, and the bureaucracy that they established reflected that system.

⁵⁰ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 72–73; Quote from Imber, *The Ottoman Empire, 1300-1650*, 96.

⁵¹ Finkel, *Osman’s Dream: The History of the Ottoman Empire*, 82–85.

⁵² Goodwin, *Lords of the Horizons*, 155–57.

Through the late-fourteenth century, as a result, the Ottoman system of administration did not differ substantially from those of other contemporary Muslim states.⁵³

Beginning in the early-fifteenth century, however, the Ottoman state broke radically with Muslim tradition with the introduction of a new system of recruitment called the *devshirme*. Meaning literally “collection” but typically referred to in the West as the “levy of boys,” the practice likely originated in the early-fifteenth century when the Ottomans moved to create new military and administrative structures aimed at freeing the dynasty from a dependence on the Turkish nobles. The resulting *devshirme* system that they developed to achieve that end was a highly innovative one that centered on the enslavement, conversion, and education of Christian boys. Every three years, a caravan of tribute officers traveled from settlement to settlement in Greece and the Balkans seeking suitable candidates for service to the sultan among the Christians of the region. After assembling the boys in each village, the officers selected only the brightest and most capable among them at a rate of about one in forty. They then brought the boys to Constantinople, where they had them converted to Islam—a violation of Islamic law justified on grounds of *raison d'état*—and divided them into two groups. The more physically capable ones were selected to serve in the famed Janissary corps, which we will treat in more detail in a moment, while the ones who demonstrated the highest intellectual capacity were reserved for the bureaucracy. Both groups then received years of physical and intellectual education in the empire’s rigorous palace schools.⁵⁴

The *devshirme* has for a long time equally fascinated and repulsed Westerners. The separation of children from their parents, their enslavement, and their forced conversion offend modern sensibilities and seem to suggest that the Ottomans were peculiarly ruthless. Such views are ahistorical and lie rooted in a facile understanding of the world from which the boys came. In fact, the new system offered substantial benefits to the people recruited through it and was far less arbitrary than it appears at first blush. Such was certainly the case with regard to the boys’ conversion to Islam. Those selected in the *devshirme* were uneducated peasants living in isolated villages where people practiced folk variants of Christianity that bore little resemblance to Orthodox Christian doctrine; conversion to Islam was thus a case more of introducing them to organized religion than of compelling them to abandon deeply held beliefs.⁵⁵

Selection through the *devshirme* also offered opportunities to the boys that were wildly beyond a Balkan peasant’s most outlandish dreams. Joining the bureaucracy or the janissary corps meant going from being one of the tax-paying masses, or *raya*, to becoming a member of the *askeri*, or ruling class, and involved trading the parochial world of the village for the cosmopolitan culture of the court and capital city. Like al-Mutasim’s slave soldiers, moreover, the *devshirme*-recruited slaves, or *qul*, were, while

⁵³ Lindner, *Nomads and Ottomans in Medieval Anatolia*, 6–8.

⁵⁴ Quotes from Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 82–83; Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 30–31.

⁵⁵ Goodwin, *Lords of the Horizons*, 56–62.

technically the property of the sultan, hardly chattel slaves as commonly understood in the West. Instead, they were respected and powerful people who drew a salary and enjoyed access to the levers of power. Recruitment through the *devshirme* thus not only came at little cost to the boys, but also opened unparalleled opportunities to them.⁵⁶

If the *devshirme* system was good to its recruits, it proved even more beneficial to the sultans. They derived three critical advantages from it. First, they benefited from the fact that the bureaucrats were utterly devoted to them. Dependent entirely upon the Ottoman rulers for their position and subject to dismissal or worse should they displease their patron, the bureaucrats were—in sharp contrast to the independently minded Turkish nobles—completely loyal to the dynasty. Second, as the boys lacked powerful family connections and as their children, born Muslim, could not succeed them in government service, the ability of a bureaucrat to rise in the administration was dependent almost entirely on his ability. In other words, the sultans' *devshirme* system produced a central administration that was unusually meritocratic and thus highly competent and efficient.⁵⁷

Finally, the *devshirme* system provided the dynasty with a way out of a predicament that stemmed from its growing embrace of Islam beginning in the fifteenth century. As we have seen, the Ottoman state had gladly and profitably employed Christians from the very start of the empire; indeed, its willingness to do so had been a key source of its strength. However, the dynasty's increasing inclination to legitimate its rule in orthodox Sunni terms had begun to make the practice of employing Christian bureaucrats and, especially, soldiers increasingly hard to justify. The sultans thus found themselves caught on the horns of a dilemma: if they continued to rely on the Christians that had proven so essential to the dynasty's early success, they would undermine the legitimacy that their embrace of Sunnism had provided; if they instead refused to employ those subjects, they would risk losing a vital source of highly effective soldiers and bureaucrats.⁵⁸

The *devshirme* system provided a way to resolve this predicament. It allowed the dynasty to continue to use Christian-born subjects in the military and bureaucracy without raising doubts about the sultans' religious *bona fides* or their status as the protectors of orthodox Sunnism. The results were nothing if not ironic. By formalizing a system that put Christian-born people in charge of the government and the military, the *devshirme* system ensured that the conquered gradually came to displace Muslim-born Turks as the backbone of the ruling class.⁵⁹

⁵⁶ Goodwin, 56–62.

⁵⁷ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 82–83, 93–94.

⁵⁸ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 30–31.

⁵⁹ Quataert, 30–31.

The Ottoman Military

The empire's powerful army was perhaps the most obvious contributor to its rapid rise to dominance. From the very beginning, when Osman's military abilities had attracted *akinji* frontier warriors to his banner, the dynasty had enjoyed a steadily widening quantitative and qualitative advantage over its enemies that consistently gave it the edge needed to expand at their expense. That gap continued to gradually widen over the course of the next hundred years. By the early-fifteenth century, however, the empire's military advantage no longer rested primarily on the prowess of its leaders or their ability to attract nomadic fighters. Instead, from that point until the end of the sixteenth century, it stemmed chiefly from the empire's embrace of innovative military practices including—most notably—its precocious adoption of gunpowder weapons and attendant development of groundbreaking new tactics that facilitated their effective use in battle.⁶⁰

Even before their wholesale embrace of firearms, Ottoman leaders had aggressively pioneered the development and financing of new, highly effective types of soldiers. Indeed, much of the dynasty's early success had stemmed from its ability to quickly replace the fearsome-but-unreliable *akinji* raiders on whom Osman I had relied with more dependable and disciplined troops. The most important of those new soldiers during the dynasty's early years were the *sipahis*. Though far-less famous than the gun-armed Janissaries, the *sipahi* horsemen were a key part of the Ottoman military from the time they displaced the *akinji* as the core of the army in the 1360s through the end of the sixteenth century. Able to move quickly thanks to their light armor, they could outmaneuver or outrun heavier cavalry such as the knights that the European states fielded while still packing a formidable punch. Indeed, using the powerful Turkish recurved bow, the typical *sipahi* could successfully strike an enemy horseman two-hundred-and-fifty meters away an astonishing one time in four. This combination of mobility and firepower was all-but unmatched in the fifteenth and sixteenth centuries, and, for a time, it gave the Ottomans a substantial military edge that allowed them to consistently win pitched battles.⁶¹

What really set the *sipahis* apart, however, was not their tactical abilities but instead the innovative administrative arrangement—the *timar* system—that the dynasty developed to raise and manage them. With obvious antecedents in the older *iqta* structure of land grants, the new arrangement derived its name from a unit of agricultural land or pasturage known as a *timar*. Individual *sipahi* received a share of the revenue generated by a standard *timar* for their upkeep, while district governors known as *sanjakbeys* and provincial governors called *beylerbeys* enjoyed a portion of the income produced by larger territories in accordance with their greater power and responsibilities. Importantly, to vouchsafe against recipients establishing roots that might result in them acquiring hereditary control over their grants, the Sublime Porte regularly rotated *sipahis* and other officials to new *timars*.⁶²

⁶⁰ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 82–87.

⁶¹ Streusand, 84.

⁶² Streusand, 82–87, 98–99; Quataert, *The Ottoman Empire, 1700–1922, Second Edition*, 29.

This novel new organizational arrangement allowed the empire to sidestep a common problem that premodern states experienced in financing their militaries. Lacking the administrative capacity to tax agricultural land directly, such regimes could only raise large forces by ceding substantial independence to powerful and frequently unreliable local subordinates. As a result, they often found themselves facing the dilemma of having to choose between maintaining a high degree of central political control or possessing a large army. The *timar* system was a successful response to this predicament—a case of the Ottomans having their cake and eating it too. By devolving the power to raise taxes to local leaders while simultaneously retaining control of the land, it permitted the dynasty to raise a large military force even as it continued to maintain central direction of the provinces.⁶³

While the *sipahis* remained an important part of the Ottoman military through the end of the sixteenth century, they increasingly found themselves eclipsed by the Janissary foot soldiers after 1440. Recruited through the *devshirme* system and thus, technically, in thrall to the sultan, the Janissaries may appear at first glance to be figures of relatively low rank in the Ottoman hierarchy. They could not marry, were barred from taking up a trade, and had to live in the barracks for their entire adult lives. Such an impression is entirely misleading, however. In reality, the Janissaries were—like the bureaucrats—high-status members of the *askeri* who enjoyed a unique position in the government.⁶⁴ What accounted for their power and prestige? At root, their privileged position stemmed from their adoption of firearms in the mid-fifteenth century. Guns had existed for some time, but, to that point, no one had yet found a way to overcome their intrinsic shortcomings as battlefield weapons. Those limits were substantial. While they packed a powerful punch, firearms were also heavy, slow to reload, and remarkably inaccurate. Perhaps most importantly, they did not integrate effectively with longstanding battle tactics that centered on bows and melee weapons. Through the early-fifteenth century, as a result, they remained more of a novelty than a mainstay in field engagements.⁶⁵ The Ottomans were the first to devise new tactics that could render early guns effective battlefield weapons. They determined that soldiers marching in a tightly organized formation and firing simultaneously could not only overcome the weaknesses intrinsic to early firearms, but could, if sufficiently disciplined, enjoy a decisive advantage over armies using traditional weapons. There was one snag, however. Firearms required not merely well-disciplined troops, but also ones willing to abandon millennia of military practices centered on traditional weapons in favor of radically new tactics and arms. For reasons we will explore in a moment, most soldiers were reluctant to do so. Here, the Janissaries were a notable exception. Beholden to the sultan and trained from an early age to follow orders without question, they broke ranks with their comrades in arms and willingly took up the new weapons in the 1440s.⁶⁶

The results were revolutionary. Wearing uniforms—itsself an innovation at the time—and marching into battle at a pace set by an accompanying military band, the Janissaries

⁶³ Gábor Ágoston, *Guns for the Sultan: Military Power and the Weapons Industry in the Ottoman Empire* (New York: Cambridge University Press, 2005), 23.

⁶⁴ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 30–31.

⁶⁵ Quataert, 30–31.

⁶⁶ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 83.

were the first real professional army since the Romans. Grim, disciplined, and full of pride, they gave the Ottomans an incalculable advantage and proved to be the difference in pivotal battles such as Varna, Chaldiran, and Marj Dabik. They were no flash in the pan, moreover. On the contrary, the Janissaries would retain their commanding lead in gunpowder weapons and related tactics for well over a century. Indeed, only with the maturation of a series of changes to warfare in the sixteenth and early-seventeenth centuries—so far reaching that the historian Geoffrey Parker refers to them as the Military Revolution—would the empire finally begin to face comparably powerful rivals along its European borders.⁶⁷

Fiscal Strength

The early Ottoman Empire also enjoyed the enormous advantage of being in a consistently strong fiscal position. Despite the vast cost of the bureaucracy, standing army, and palace, the Sublime Porte only occasionally found itself facing fiscal shortfalls and, in sharp contrast to nearly all contemporary states, frequently ran a surplus. What explained this unusual situation? That is, what accounted for the Ottoman government's peculiarly favorable budgetary circumstances?

Three broad factors explain the empire's fiscal good fortune. First, the Sublime Porte benefited from a robust, expanding economy during its first three centuries. To a substantial degree, this growth was a function of the order and security that the Ottomans brought to the heretofore fractious regions of Anatolia and the Balkans. The sultan's powerful government quickly ended the constant warfare and brigandage endemic to the region and freed the people it ruled from the confiscatory taxes that the local nobles had imposed following the Fourth Crusade and the collapse of Byzantine authority. The results were impressive. Once again able to keep a large share of any additional wealth that they produced, farmers responded by putting new land under the plow and by intensifying the cultivation of existing fields, while merchants and craftspeople reacted by aggressively investing in commercial ventures that were once again profitable. The result was an economic and demographic boom over the course of the fifteenth and sixteenth centuries—one that produced a concomitant rise in the tax receipts that the central government collected.⁶⁸

Second, the Sublime Porte enjoyed substantial revenue derived from the duties it assessed on trade. From early on, the empire enjoyed steadily rising customs receipts thanks to its fortuitous location between the West and the sources of the Eastern goods that Europeans craved. The Ottomans began to tap into this revenue stream in the late-fourteenth century when Bursa supplanted the Byzantine city of Trebizond as the primary entrepot for Asian goods being transshipped to Europe. Later, the conquest of

⁶⁷ Geoffrey Parker, *The Military Revolution: Military Innovation and the Rise of the West, 1500-1800*, Second Edition (Cambridge: Cambridge University Press, 1996), 24; Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 83.

⁶⁸ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 28; Goodwin, *Lords of the Horizons*, 118.

Constantinople cemented Ottoman control of the western end of the Silk Road, while the extension of the imperial writ to the northern Red Sea in the 1510s gave the sultanate a lock on the traditional maritime trade routes that linked the Mediterranean region to India. As a result, over the course of the fifteenth and early-sixteenth centuries, the empire enjoyed steadily rising customs receipts from goods transshipped through the empire.⁶⁹

Even the decline in trade that stemmed from the establishment of a direct maritime connection between Europe and the Indian Ocean barely dented the amount of customs revenue that the sultan's government received. Not only did the volume of trade that went around the Cape of Good Hope remain comparatively small relative to the amount that went through the sultan's territory through the end of the sixteenth century, but the establishment of profitable sugar plantations in Egypt and the growing involvement of Ottoman merchants in the shipment of Yemeni coffee to Europe generated new streams of tax revenue that more than offset any decline in customs revenue that may have resulted from the direct trade with Europe.⁷⁰

Finally, the very act of conquest itself dramatically improved the fiscal health of the central government. That imperial expansion generated revenue may seem counterintuitive at first blush. After all, the army that acquired those territories consumed the vast bulk of the state's revenue. Still, even when accounting for the cost of the sultan's military—and that expense was substantial—the empire's nearly unbroken string of victories between 1299 and 1566 proved enormously profitable for the Sublime Porte. The key was the way that the Ottomans divided the loot acquired in conquest. Longstanding Islamic custom decreed that the sultan received a hefty one-fifth of all booty taken in war. Coupled with the Ottoman military's near-constant success, this provision ensured that the state received enormous and regular infusions of wealth—massive surpluses that more than offset the cost of the wars of conquest.⁷¹

Thus, the Sublime Porte only rarely faced fiscal challenges during its rise to power. Instead, the general growth in economic activity that stemmed from the reimposition of order, the steady increase in customs duties that accompanied the empire's growing dominance of vital long-distance trade routes, and the regular influx of enormous quantities of loot ensured that the Ottomans consistently received more than adequate revenue to cover the empire's substantial expenses during its rise to dominance. It was, to say the least, an unusual and highly enviable position.

⁶⁹ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 105–6.

⁷⁰ Jane Hathaway, "The Ottomans and the Yemeni Coffee Trade," *Oriente Moderno* 25 (86), no. 1 (2006): 161–71; Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 105–6.

⁷¹ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 176.

The Circle of Justice

While a strong central government, favored fiscal position, and powerful military made possible the acquisition of an empire, it was the Ottomans' tolerant approach to the subject peoples they governed that permitted them to retain it. The dynasty's mild rule reflected its adherence to an older Islamic ideology of just government called the Circle of Justice. This doctrine held that if the monarch was a beneficent ruler who kept taxes reasonable and predictable and who maintained order, the peasants would thrive, and the treasury would obtain adequate tax revenue each year. In turn, those strong tax receipts would underwrite a powerful and effective army that would complete the Circle of Justice by ensuring that the ruler's position remained secure and unchallenged. If, in contrast, the sultan did not rule justly, then the peasants would fail to prosper, tax receipts would gradually slip, the army would inevitably weaken, and, eventually, the ruler would be overthrown. Thus, just rule ensured not merely the well-being of society, but also the continued success of the ruling dynasty.⁷²

Ottoman rulers well understood the implications of the Circle of Justice, and they approached the administration of their empire in accordance with its dictates. For their Muslim subjects, doing so was straightforward: the government merely needed to keep taxes at a reasonable rate and to ensure the stability and order required for shari'a law to function. The empire's substantial non-Muslim population—it remained majority Christian through the 1510s and possessed nearly as many Christians as Muslims even after Selim I's conquest of the Middle East—was a different matter, however. How could the Ottoman dynasty provide its non-Muslims subjects with the just rule on which prosperity and the well-being of the state depended?⁷³

It did so by expanding and codifying the empire's longstanding system of religious toleration. Established by Mehmet II and later formalized as the *Millet* System, this arrangement extended formal recognition to religious minorities by granting them the power to administer their own legal, educational, and charitable affairs under the aegis of their leading ecclesiastical figure. For example, the Patriarch of Constantinople oversaw the administration of the Greek Orthodox community. His office provided the Greek community with education through church-run schools, maintained social order within the community by overseeing issues such as marriage and inheritance through its ecclesiastical courts, and aided poor Greeks through religiously based charitable organizations. At sharp variance with the structure of nearly all modern states, this system worked remarkably well for both the confessional groups and the Ottoman government. By ceding autonomy to the different religious communities, it not only satisfied their needs and desires, but—in keeping with the Circle of Justice—also

⁷² Linda T. Darling, *A History of Social Justice and Political Power in the Middle East: The Circle of Justice From Mesopotamia to Globalization* (New York: Routledge, 2013), 2–14.

⁷³ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 103–4.

ensured that they enjoyed the just rule on which the fiscal health of the state depended.⁷⁴

The sultanate's tolerance of the empire's Jewish community was particularly notable and proved highly beneficial to both parties. Ottoman acceptance of Jewish people came in two forms. First, the empire accorded them official status as a sanctioned religious community. Beginning with Mehmet II, the government recognized the Grand Rabbi of Constantinople as the head of the Jewish community and gave his office authority over its legal, educational, religious, and charitable affairs. As a result, Jewish subjects of the empire enjoyed legal protection, educational opportunities, and—should they fall on hard times—access to social welfare benefits through institutions that their community controlled. Second, the Ottoman government did not merely tolerate the existing Jewish community in the empire, but also gladly welcomed Jewish refugees fleeing discrimination in Christian Europe. Most notably, when King Ferdinand II (r. 1479-1516) of Aragon issued the Alhambra Decree in 1492 requiring Sephardi Jews to either convert to Christianity or leave Spain, Bayezid II instructed his officials to welcome them in the Ottoman Empire. Settling largely in Salonika, the Sephardi Jews brought with them profitable industries that soon, in keeping with the Circle of Justice, paid the sultan back for his tolerance by generating substantial tax revenue for the state. Bayezid II was well aware of the benefits he was accruing thanks to this policy. Gleeful about his gain, he reputedly mocked Ferdinand II for kicking Jewish people out of Aragon, declaring that “[h]e is impoverishing his country and enriching my kingdom.”⁷⁵

Cultural Openness

The Ottoman state also demonstrated a remarkable willingness to adopt new and effective cultural practices regardless of their provenance. Indeed, in sharp contrast to its more chauvinistic and prejudiced contemporaries, the dynasty unabashedly embraced ideas, methods, and even people from other cultures as well as disruptive new technological innovations. Doing so greatly strengthened the Ottoman Empire and played no small part in facilitating its rapid expansion.

The Ottomans' cultural openness is most evident in three ways. First, the dynasty was unusually receptive to having non-Turks and non-Muslims hold important leadership positions. As we have seen, during its early centuries, the Ottomans did not make membership in the *askeri* contingent on being Turkish or even following the Islamic faith. Instead, it gladly welcomed substantial numbers of Jewish people and Christians into what the historian Ira Lapidus calls “a condominium of elites.” Thus, a Jewish refugee from Spain was an important advisor to Sulëiman and later became the Duke of Naxos. Similarly, a converted nephew of the final Byzantine emperor became one of Mehmet II's grand vezirs while a second served as the powerful Ottoman governor of the Balkans. Indeed, Greeks—both Christians and converts to Islam alike—

⁷⁴ Calic, *The Great Cauldron*, 90–92.

⁷⁵ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 88.

were so common in the administration of the early empire that the historian Dimitri Kitsikis characterizes the Ottoman state of that period as a “Turkish-Greek Empire.”⁷⁶

Second, the Ottomans readily adopted practices and methods from other cultures—including, notably, non-Muslim ones. For example, while the *timar* system had obvious antecedents in the semi-feudal *iqtas* that the Buyids and many subsequent Muslim states had adopted, they also drew on a similar late-Byzantine fiscal arrangement called a *pronoia*. Ottoman architecture—particularly the design of monumental mosques such as the Blue Mosque that the regime had sponsored to legitimate its rule—likewise made clear the empire’s openness to foreign ideas. Those structures may have adhered to longstanding Muslim architectural traditions, but they also borrowed in obvious ways from the design of grand Byzantine buildings such as the Hagia Sophia.⁷⁷

Finally, Ottoman openness to new ideas and technologies was central to its precocious adoption of gunpowder weapons. Firearms were not new when the Janissaries began to use them. On the contrary, so-called hand cannons had existed since the late-thirteenth century. For two reasons, however, they had not yet gained widespread acceptance among the armies of Europe and the Middle East. First, as we have seen, military leaders found them to be difficult to integrate into established battlefield tactics that centered on traditional weapons such as bows, swords, and lances and were thus reluctant to use them in large numbers. This hesitance was important, but the second reason was far-more critical in explaining the limited adoption of gunpowder weapons in the fourteenth and early-fifteenth centuries: broadly speaking, soldiers in most armies perceived them to be incompatible with longstanding masculine military ideologies and traditions. Nearly everywhere, both the rank-and-file and their leaders believed that killing at a distance using arms as crude as guns violated the warrior code of bravery to which they adhered and preferred to fight in hand-to-hand combat using traditional melee weapons. Thus, restrained by the seeming incompatibility of gunpowder weapons with existing tactics and by the soldiers’ reluctance to use them, nearly all military forces in both Europe and the Middle East had only incompletely adopted guns by the early-fifteenth century.⁷⁸

The Ottomans were not burdened with such hidebound views, however. On the contrary, they were happy to adopt effective new technologies even if they upset existing practices and values. Indeed, once they had devised tactics for their effective use, the Ottomans quickly embraced gunpowder weapons and began to employ them on a mass scale. As a result, for a brief but critical period, the sultan’s forces enjoyed a

⁷⁶ Lapidus quote from Ira M. Lapidus, *A History of Islamic Societies*, 2nd ed. (Cambridge: Cambridge University Press, 2002), 264; Kitsikis quote from Lowry, *The Nature of the Early Ottoman State*, 12; Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 67.

⁷⁷ Lapidus, *A History of Islamic Societies*, 263; Mark Bartusis, *The Late Byzantine Army: Arms and Society, 1204-1453* (Philadelphia: University of Pennsylvania Press, 1992).

⁷⁸ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 30.

near-monopoly on the use of massed firearms. That advantage proved to be decisive. As we have seen, the empire's deployment of siege cannons was the deciding factor during Mehmet II's successful siege of Constantinople, while the use of gunpowder weapons against Safavid and Mamluk armies—which had vigorously resisted the adoption of firearms—had made possible the victories at Chaldiran and Marj Dabik that had allowed the sultanate to seize vast swaths of territory in the Arab Middle East.⁷⁹

The Safavid Empire

While the Ottoman Empire was the dominant state in the Middle East during the Early Modern Period, it faced for a time a powerful rival in the region in the form of the Safavid Dynasty (1501-1722). Controlling the Iranian plateau for more than two centuries, the Safavid state bore many similarities to the Ottoman Empire. Both reached their apogee as a result of their adoption of gunpowder weapons, and both used longstanding Muslim administrative practices to centralize power. As a consequence, historians have typically treated the Ottoman and Safavid states as a pair of similar, gunpowder-based empires. This characterization is not without merit. As we shall see, those states did indeed resemble each other in a number of critical ways. At the same time, however, many of their commonalities were surface ones that obscured fundamental differences in their political systems, methods of military recruitment, and, especially, religious policies.

Origins

Those differences stretched back to the two dynasties' sharply contrasting origin stories. Unlike the Ottomans, the Safavids began not as a military or political group but instead as a populist, Sufi religious order that found appeal with some of the Turkish tribes of eastern Anatolia and western Iran. Established at the turn of the fourteenth century by its namesake, Safi ad-Din (c. 1252-1334), the family at first oversaw a fairly typical mystical, Sufi order. It did not retain that orientation for long, however. On the contrary, like a number of similar orders in the region, the Safavids' theological position grew more extreme and heterodox in response to the political disorder and the collapse of Sunni institutions that occurred following the demise of the Ilkhanate and the Seljuq Sultanate of Rum. Indeed, within just a few decades, it had transformed into a militant organization that embraced *ghuluww*: a term that connoted an extremist, nomadic opposition to centralized authority and a belief that certain charismatic leaders possessed a direct link to the divine.⁸⁰

The pivotal moment in that transformation came under Safi al-Din's great grandson, Junayd (r. 1447-1460). Facing a challenge from one of his uncles who enjoyed the backing of the powerful Qaraqyunlu, or Black Sheep confederation of Turkish tribes, Junayd abandoned the Safavids' heretofore narrow focus on spiritual

⁷⁹ Lord Kinross, *Ottoman Centuries* (New York: Harper Perennial, 1977), 52; Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 30.

⁸⁰ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 159–61.

concerns in favor of an aggressive approach that would fundamentally restructure the order. Advancing messianic claims to justify his new role, he assumed the political and military leadership of the tribes that adhered to his teachings and forged an alliance with the Qaraqyunlus' rivals, the Aqqyunlu, or White Sheep confederation of Turkish tribes. These moves proved consequential. They not only resulted in the defeat of Junayd's uncle, but, more fundamentally, also completed the Safavid order's metamorphosis from a mystical organization into a militant movement driven by *ghuluww* ideology.⁸¹

Further dramatic changes followed under his son and successor, Haydar, (r. 1460-1488). Not long after taking power, he abandoned Sunnism in favor of Shi'ism and began to assert that the Safavid family descended from the Prophet Muhammad through the line of Ali and Fatimah. These were bold moves—ones that would have significant consequences. In the immediate term, they inspired a number of Haydar's more-devoted supporters to follow his lead and adopt Shi'ism. To demonstrate their commitment to their new sect, they began to wear a distinctive, twelve-folded red turban—each fold representing one of the twelve Imams—called the Qizilbash, a term used thereafter to denote the tribes that supported the Safavid order. The long run consequences of Haydar's conversion were far-more significant. They led directly to the people of Iran adopting Shi'ism—a change of enormous significance that would have important consequences stretching to the present day.⁸²

Ironically, Haydar's success would lead directly to a serious setback for the Safavid Dynasty. The issue was the growing appeal of his revolutionary mix of Shi'i and *ghulat* ideology among the Turkmen of eastern Anatolia and western Iran. Fearing that the Safavid leader's surging appeal would reduce him to vassalage, the Aqqyunlu leader Yaqub (r. 1478-1490), broke with Haydar in 1488 and went to war with the Qizilbash. The conflict went disastrously for the Safavids. Not only did the Qizilbash tribes suffer a serious defeat at the hands of Yaqub's soldiers, but Haydar himself died in the fighting. Taking advantage of the ensuing chaos, the Aqqyunlu followed by systematically killing or imprisoning all of the remaining members of the family save for Haydar's seven-year-old son, Ismail (r. 1501-1524), who managed to escape into hiding in 1494. It was a major setback for the Safavids—one that appeared to have doomed the dynasty to irrelevance.⁸³

Shah Ismail

Instead, it achieved a remarkable revival under the youthful Ismail. Reemerging in 1499 at the age of twelve, the charismatic Safavid leader rallied a confederation of Qizilbash and disaffected Aqqyunlu tribes to his banner. Eschewing moderation, he declared himself to be Ali reincarnate, the Hidden Imam, and a descendant of the Seventh Imam. Under his messianic leadership, the Qizilbash seized the Aqqyunlu

⁸¹ Lapidus, *A History of Islamic Societies*, 233–34.

⁸² Garthwaite, *The Persians*, 160–61.

⁸³ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 143.

capital of Tabriz in 1501—which Ismail made his own—and then defeated its army in 1503. Having eliminated their most powerful regional opponent, the Qizilbash followed with a lightning campaign that resulted in Ismail assuming control of the entirety of the Iranian plateau. He followed by declaring himself ruler of all Persia. In so doing, he did not assume a conventional title such as sultan. Instead, seeking to legitimate his rule by connecting his state to Persia’s distinct identity and history and—perhaps—to obscure his Turkish background, Ismail adopted the ancient Persian title of shah.⁸⁴

Flush with success, he next moved to extend his influence among the tribal Turkmen of Ottoman-controlled eastern Anatolia. Wielding the Safavids’ revolutionary mix of Shi’i and *ghulat* ideas as a subversive, ideological cudgel, he sent missionaries into the region and made contact with key Turkish leaders in hopes of detaching the tribes from the Ottoman Empire. While his message found a welcome reception among the many Turkmen who wished to regain their autonomy, his initiative ultimately ended in disaster. Determined to maintain his grip over eastern Asia Minor, Sultan Selim I responded—as we have seen—by invading the Safavid Empire. Ismail and his Qizilbash soldiers met the Ottoman army at Chaldiran in the far-northwest part of the Iranian plateau. Genuinely believing that he was the Hidden Imam, the shah was blithely unconcerned that Selim I’s army enjoyed a substantial numerical advantage or that its troops possessed firearms while his army instead relied on bows, swords, and lances. Indeed, Ismail was so confident that he enjoyed divine favor that he personally led a cavalry charge against the center of the Ottoman line. The results were disastrous. Inflicting heavy casualties, the massed fire of the Ottoman cannons and gun-armed Janissaries decimated his horsemen and turned the Battle of Chaldiran into a bloody rout. With Ismail’s army all but destroyed, the victors followed by briefly occupying Tabriz before retiring from Safavid territory.⁸⁵

The Qizilbash Interlude

Selim I’s withdrawal rescued Ismail but could not obscure the fact that Chaldiran had been a crushing setback—one that fundamentally altered the political balance in Persia. The most immediate result of the defeat was the shattering of Ismail’s confidence. Depressed by his failure, he never again led an army in battle and instead spent the remaining decade of his reign drinking and otherwise indulging himself in the pleasures of his position. His withdrawal from leadership would have serious ramifications for the Safavids. No longer believing after Chaldiran that Ismail was divinely ordained, the dynasty’s onetime Qizilbash supporters took advantage of his distraction to assert their autonomy. Controlling valuable pasturage that permitted them to field a large force of horsemen, they quickly assumed a dominant position over Iran and its central government that they would retain for the balance of the sixteenth century. As a result, Ismail’s successors found themselves in a difficult situation—too

⁸⁴ Lapidus, *A History of Islamic Societies*, 234; Howard, *A History of the Ottoman Empire*, 86; Ansary, *Destiny Disrupted*, 186.

⁸⁵ Roger Savory, *Iran Under the Safavids* (Cambridge: Cambridge University Press, 2007), 39–46.

weak either to defend the country from Ottoman depredations or to curb the power of the Qizilbash.⁸⁶

Shah Abbas I

It was only at the very end of the sixteenth century that a new and dynamic leader, Ismail's great grandson Abbas I (r. 1588-1629), would be in a position to restore the dynasty's position. That he was able to do so was surprising in light of the trying circumstances that he faced at the start of his reign. In the years immediately preceding his accession to the throne, Ottoman forces had exploited his father's weak rule to hive off substantial territory along Iran's western border. Worse, Qizilbash factions had begun to engage in destabilizing civil conflict and to intervene in the affairs of the central government. This last challenge was especially dangerous for the Safavids. The tribes not only killed the queen in 1579 and the grand vezir in 1583 but had also taken to meddling in the succession. Indeed, Abbas I himself had become shah at the behest of a Qizilbash leader, Murshid Quli Khan Ustajlu, who hoped to rule from behind the throne.⁸⁷

The sixteen-year-old monarch quickly proved that he was up to these challenges. Taking advantage of worsening factionalism among the Qizilbash, he first acted by lining up the political support needed to have Murshid Quli Khan executed. Having secured his autonomy in this way, he followed by centralizing power in his hands. Key to this effort was his creation of a royal monopoly over the production and export of silk, Iran's most valuable product. It was a characteristically shrewd move on his part—one that provided him with the revenue needed both to develop an efficient central administration and to turn the firearm-equipped slave army that his grandfather, Tahmsap (r. 1524-1576), had established into a force powerful enough to check the Qizilbash.⁸⁸

Abbas I followed by relocating the capital to the city of Isfahan. Doing so came with two substantial advantages. First, the move put his government in a central location from which it could more easily control Iran. Second, and far more important, relocating the capital made possible the construction of a new palace complex of monumental architecture that could serve as the symbolic centerpiece of a renewed and strengthened Safavid state. The results were—and are—impressive. Arranged around the *Naqsh-e Jahan*, a massive and beautiful new *maidan*, or square, that measures one hundred and sixty meters by five hundred meters, the complex of structures he built was stunning. It included the Ali Qapu Palace, the Sheikh Lotfollah Mosque, and, above all,

⁸⁶ Garthwaite, *The Persians*, 167–69; Savory, *Iran Under the Safavids*, 64–65.

⁸⁷ Abbas Amanat, *Iran: A Modern History* (New Haven: Yale University Press, 2017), 72–75.

⁸⁸ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 150–53.

the magnificent Shah Mosque and its attendant madrasas. Befitting his state's dependence on revenue derived from trade, it also featured a new bazaar.⁸⁹

The palace complex and the reconstructed city of Isfahan served a number of ends. The buildings accommodated the administrators on whom Abbas I's centralizing plan depended, while the new neighborhoods in Isfahan supported the fiscal key to his state: the silk workers and their workshops. Meanwhile, the grandness of the architecture and the brilliance and orderliness of Isfahan proclaimed the glory of Abbas I and the Safavids and suggested that the stability embodied in the orderly, planned city of Isfahan depended on the continuation of the dynasty. Isfahan thus constituted not only the physical center of Abbas I's powerful new administrative state, but also the symbolic core of his effort to legitimate Safavid rule.⁹⁰

Abbas I's domestic accomplishments soon led to success abroad. His powerful new army recovered Tabriz in 1603 and Armenia in 1604 and decisively defeated an Ottoman counterattack in 1605. After a lull, it scored a number of further victories. In 1622, it succeeded both in retaking the important eastern city of Qandahar from the Mughal Empire and, in conjunction with the British East India Company, in ejecting the Portuguese from the Island of Hormuz. This last triumph was especially significant. It gave the delighted shah control over both the entrance to the Persian Gulf and the trade that passed through it. Finally, in 1624, his armies not only seized Baghdad but also managed to hold it in the face of a concerted Ottoman counterattack.⁹¹

Thus, by the time he died in 1629, Shah Abbas I had fully restored Safavid rule. He had ended Qizilbash dominance of the government, restored control of the border, and affirmed the idea that a powerful monarch should rule over Iran. Importantly, he had also replaced tribal rule with a strong, centralized bureaucracy and turned the firearm-equipped slave army that his grandfather had pioneered into a powerful force comparable to the Ottoman Janissaries.⁹²

Dissolution

Unfortunately for the Safavids, however, the dynasty was unable to maintain the bureaucratic and military system that Abbas I had established. Two issues combined to prevent it from doing so. First, none of Abbas I's successors possessed his creativity, charisma, or leadership skills. Indeed, they were, for the most part, his polar opposites. For example, the penultimate shah, Safi II (r. 1666-1694), spent most of his time drinking and pursuing sexual conquests and thus provided little guidance during what was a difficult period. Remarkably, his successor and brother, Sultan Husayn (r. 1694-1722), was even worse. Thanks to his excessive alcohol consumption and, ironically,

⁸⁹ Garthwaite, *The Persians*, 175–76.

⁹⁰ Amanat, *Iran*, 82–91.

⁹¹ Garthwaite, *The Persians*, 174.

⁹² Lapidus, *A History of Islamic Societies*, 241.

the intense piety that monopolized his time, he failed to provide any effective direction of the state whatsoever.⁹³

The second issue, the empire's worsening fiscal situation, proved to be even more problematic. Importing more goods than it exported, Persia experienced steady specie drain and inflation in the decades following Abbas I's reign. Exacerbated by the decision to end the state monopoly on silk exports, the worsening trade balance resulted in the steady decline of tax receipts. Making matters worse, the dynasty failed to maintain the tenuous control that Abbas I had established over Persia's agricultural regions; as a result, successive shah's saw those lands—and the tax revenue that they generated—revert to the increasingly restive and independent tribes.⁹⁴

The results were catastrophic for the state. With tax receipts steadily falling, Abbas I's successors had little choice but to let his bureaucracy and his once-vaunted military gradually wither. Relative peace along Persia's borders obscured the decline of Safavid power for some time. By the 1710s, however, the dynasty could no longer disguise the feebleness of its position, and foreign powers began to peel off its outlying territory. The coup de grace finally came in 1722 when the tribal Ghalzay Afghans seized Isfahan and compelled the final Safavid shah to abdicate—thus bringing the dynasty to an ignominious end. The demise of Safavid rule was a real blow to Persia and precipitated hard times for the country's people. Indeed, as we shall see in a later chapter, there would follow costly invasions, seven decades of civil war, and precipitous economic decline before the country was finally able reunify.⁹⁵

Persian Shi'ism

Concurrent with their efforts to create a powerful, centralized state, the Safavids successfully imposed Shi'ism on a people that had heretofore largely adhered to Sunnism. The dynasty had done so for the purpose of legitimating its rule, and, accordingly, had retained tight direction over the religious establishment. That dominance did not last, however. Taking advantage of Safavid weakness, Shi'i religious leaders would gradually carve out a fully autonomous position and would achieve vast, independent authority over lay Shi'a. Together, these developments would powerfully influence Iran's culture and the subsequent course of its history.

Shi'ism was a core element in the dynasty's early efforts to secure legitimacy in Iran. Initially, the Safavids promoted a heterodox variant that held that Shah Ismail was the Hidden Imam and the *Murshid-i Kamil*, or perfect master, who had to be obeyed by his followers without question. For a time, this claim earned the dynasty the unquestioned fealty of his Qizilbash supporters. As we have seen, however, Ismail's stunning defeat at Chaldiran in 1514 demolished the carefully developed myths that he and his supporters had developed during his rise to power. The battle was thus a

⁹³ Savory, *Iran Under the Safavids*, 226–54.

⁹⁴ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 155–56.

⁹⁵ Savory, *Iran Under the Safavids*, 226–54.

double blow to the Safavid Dynasty—one that cost it not merely territory and prestige, but, more importantly, the religious legitimacy on which its rule ultimately rested.⁹⁶

The dynasty's response to this setback was shrewd and effective. The Safavids quickly discarded Ismail's claim to be the *Murshid-i Kamil* and began to throw their support behind orthodox Twelver Shi'ism in hopes of using it to legitimate their continued rule. As part of that effort, the government brought Shi'i 'ulama' from Bahrain and Lebanon to Iran over the next few decades and tasked them with converting Iranians to Twelver Shi'ism and with establishing a state-dominated religious hierarchy. Financially dependent on the government, the Shi'i 'ulama' worked diligently to achieve these ends. They dutifully upheld Safavid legitimacy, preached that the people should obey the dynasty, accepted successive shahs' claims to be the Hidden Imam incarnate, and acted to bring Shi'ism to the people of Iran. Their efforts were successful. Abetted by the state's vigorous persecution of Sunnis, the 'ulama' succeeded in converting nearly the entire Iranian population to Shi'ism by the early-seventeenth century.⁹⁷

The close relationship between the dynasty and the Shi'i 'ulama' did not endure for long, however. Instead, as the dynasty weakened following Abbas I's reign, the religious scholars became less inclined to slavishly support the Safavids. This change took place very gradually. It started with the clerics reasserting their commitment to the traditional view that the Hidden Imam would only reemerge from Occultation at the end of time. This position was a clear repudiation of one of the fundamental justifications for Safavid rule: that the shahs were themselves the Hidden Imam incarnate. Later, in the seventeenth and eighteenth centuries, the religious scholars substantially broadened and deepened their challenge to royal dominance of the clergy. Central to this effort was the development of a new claim that it was not the shahs but instead the *mujtahids*, the most learned 'ulama' and the precursors of the ayatollahs, who were the legitimate heads of the Shi'i community until the Hidden Imam returned—a position that directly challenged the monarch's authority. While that idea gained broad acceptance among the clergy, a minority advanced an even more extreme position: that the shah should not hold authority over the *mujtahids* but should instead serve as the instrument through which those clerics effected their rulings.⁹⁸

The 'ulama' went much further in asserting their independence in the eighteenth century. Political and economic conditions were key to the success of this effort. Politically, the absence of a central government following the collapse of the Safavid Dynasty permitted the 'ulama' to turn the autonomy that they had won following Abbas I's reign into outright independence. As a result, in contrast to their counterparts in the Ottoman Empire who received their salaries from the state and were thus dependent on it, the Shi'i clerics of Iran had achieved complete autonomy from state authority by the middle of the eighteenth century. Economically, meanwhile, their ability

⁹⁶ Lapidus, *A History of Islamic Societies*, 241–42.

⁹⁷ Lapidus, 242.

⁹⁸ Garthwaite, *The Persians*, 186–89; Nikki Keddie, *Modern Iran: Roots and Results of Revolution* (New Haven: Yale University Press, 2006), 17–20.

to exploit agricultural land they either owned outright or controlled through a *waqf* (pl. *awqaf*)—an inalienable piece of endowed property that supports a religious institution or charitable cause—ensured that they would remain financially independent even if royal authority returned to Iran.⁹⁹

The clerics had not merely secured their financial and political autonomy during the eighteenth century, moreover, but had also dramatically increased their religious, moral, and legal sway over Iran's Shi'i majority. Their ability to win broad recognition of two significant theological innovations accounted for this enhancement of their authority. First, they gained widespread acknowledgement of their claim that the *mujtahids* were sufficiently learned and wise that they could exercise *ijtihad*; that is, they secured acceptance of the idea that the leading clerics were so well grounded and accomplished in *fiqh*, or jurisprudence, that they were capable of rendering new interpretations of the Shi'i shari'a code in the name of the Hidden Imam. Second, the 'ulama' gained popular adherence to their position that lay Shi'a, understood to be incapable of grasping the hidden, esoteric core of the faith, had to follow—or emulate—a *mujtahid* and to abide by his rulings.¹⁰⁰

Thus, by the time that Iran reunified at the end of the eighteenth century, the clergy had established themselves as a potent, independent force. Their recognized right to engage in *ijtihad*, their autonomy from the state, their success in advancing the claim that the clerics were the leaders of the Shi'i community during the Occultation, and their insistence that lay people had to emulate one of the *mujtahids* afforded them substantial—if still latent—political power. In other words, the door was now open for them to assert themselves in political affairs if they wished. For the time being, they refrained from crossing that threshold. Instead, they remained focused on issues of religion and the law while regarding questions of politics as the prerogative of the shah. As we shall see in a later chapter, however, their restraint had its limits, and internal disputes arising from the West's growing dominance of Iran in the nineteenth century would eventually spur them to intrude into political affairs.¹⁰¹

A Gunpowder Empire?

At first blush, the Safavid Empire seemed to have been very similar to its Ottoman rival. Like its Turkish contemporary, the Iranian dynasty built its regime on the core elements of a gunpowder state: a highly centralized government based on traditional Muslim administrative methods and a centrally directed, firearm-equipped army. In two fundamental ways, however, the Safavid Empire broke with the gunpowder model that the Ottoman Empire embodied—indeed, its differences with the Ottoman state were so significant that they raise questions about whether it in fact met the definition of a gunpowder state.

⁹⁹ Amanat, *Iran*, 202; Lapidus, *A History of Islamic Societies*, 469–70.

¹⁰⁰ Keddie, *Modern Iran*, 20.

¹⁰¹ Lapidus, *A History of Islamic Societies*, 470, 475–76.

First, unlike Osman and his descendants, the Safavids were unable to fully centralize political power in their hands. In part, this failing was fiscal in nature. As impressive as Abbas I's monopoly over silk exports had been, it could not supply his government with all of the revenue it needed to pay for his army. As a consequence, the Safavids relied heavily on *tuyul*—a fiscal device similar to the Ottoman *timar* or the Buyid *iqta*—designed to sustain soldiers using land grants rather than scare tax revenue. This dependence was not in itself significant; after all, the Ottoman sultans raised a large part of their army in a similar fashion. However, in contrast to the Sublime Porte's ability to prevent *timar* holders from making their land grants hereditary, Abbas I's government lacked the administrative capacity to stop the possessors of *tuyuls* from assuming ownership of the land that the state had assigned them; as a result, while the Ottomans retained tight control over the *timar* holders, the Safavid state proved unable to prevent the gradual devolution of power to those who held *tuyuls*.¹⁰²

Second, the Safavids differed substantially from the Ottomans in their adoption of gunpowder weapons. While their Turkish rivals acquired the vast bulk of their empire as a direct result of their precocious embrace of firearms, the Safavids instead gained nearly all of their territory *before* they began to use guns. Indeed, they only reluctantly adopted firearms after their disastrous defeat at Chaldiran had demonstrated the effectiveness of such weapons.¹⁰³

These are significant divergences. Indeed, they are so substantial that they suggest that the substantive differences between the two empires outweigh their apparent surface similarities. As a result, many scholars today conclude that the Safavid state was a gunpowder empire only under the most generous definition of the term and only during the latter half of its existence.¹⁰⁴

Legacy

The Safavid state may not have been a full-fledged gunpowder empire, but it did play a vital role in shaping Iran's future development. The dynasty established the rough outlines of the country's modern borders and left it with a tradition—albeit a weak one—of centralized, bureaucratic government. Above all, the Safavids established Twelver Shi'ism as the dominant religion in Iran. This last development would turn out to be particularly momentous. It not only resulted in Shi'i Islam becoming a central component of the country's national identity, but it also left Iran with a powerful, organized, and, above all, independent religious establishment—one that was poised to exploit its strong position to exert a powerful voice in political affairs.¹⁰⁵

¹⁰² Lapidus, 237.

¹⁰³ Amanat, *Iran*, 56–59.

¹⁰⁴ McNeil, "The Age of Gunpowder Empires, 1450-1800," 131.

¹⁰⁵ Garthwaite, *The Persians*, 157.

Conclusion

Gunpowder weapons and the political centralization and efficient bureaucracies they required permitted the Ottomans and, to a lesser degree, the Safavids to remake the map of the Middle East. Gone were the unstable principalities, *beyliks*, and tribal confederations of the late-medieval period. In their place, there instead stood highly centralized empires that were able to maintain control over large territories using sophisticated, firearm-equipped military forces.

For a time, these regimes were mighty. In particular, as we have seen, the Ottoman Empire stood without question as the most powerful state on Earth for much of the fifteenth and sixteenth centuries. By the end of that period, however, the empire's dominance would begin to slip in the face of an increasingly powerful and assertive West. The result would be a peculiarly trying period for the dynasty—one, as we shall see in the next chapter, that would compel it to make significant changes to the empire's political, economic, social, and military systems.

Chapter Seven: Ottoman Crisis and Adaptation, 1566-1768

The century that followed the reign of Sulëiman the Magnificent (r. 1520-1566) was a trying one for the Ottoman government. During that period, the empire faced a combination of vexing political, economic, and military problems—ones that seemed at times to be beyond the capacity of its existing institutions and practices to resolve. However, these issues did not, as historians once maintained, prove insurmountable for the empire or doom it to a gradual degeneration ending in its final dissolution in 1922. Instead, benefitting from a government that remained dynamic and flexible, the empire successfully responded to the crises of the late-sixteenth and seventeenth centuries by transforming its institutions, modifying its political practices, and reconstructing its fiscal systems. These adaptations were complex and at times imperfect, but they were also successful and ensured that the empire continued to thrive well into the eighteenth century.

The Ottoman ‘Time of Troubles,’ 1574-1656

The challenges that the sultanate faced were not immediately apparent under Sulëiman the Magnificent’s immediate successors, Selim II (r. 1566-1574) and Murad III (r. 1574-1595). Indeed, during their reigns, the Ottoman Empire appeared to be as strong as ever. Domestically, its economy thrived, and its government continued to enjoy an unusual and envious degree of political stability. In the realm of foreign affairs, it dominated its neighbors as it had in the past and remained the expansionary, conquest state that it had been for nearly three centuries. Most notably, between 1578 and 1590, the sultanate inflicted a series of significant defeats on its Persian rival and gained control of strategic territory in the Caucasus including the Safavid capital of Tabriz. While the empire was unable to make comparable gains in the west at that time, it nonetheless continued to raid Austrian territory regularly and posed a constant menace to the Habsburg capital of Vienna that lay just a few hundred kilometers from the border.¹

For a brief moment in 1571, however, the empire’s fortunes did appear to have taken a turn for the worse. That year, the Holy League, an alliance of the Papacy, Venice, and Spain, destroyed nearly the entire Ottoman fleet at the Battle of Lepanto in Greece’s Gulf of Patras. It was a decisive victory for the league—one so total that many European observers concluded that the tide had finally turned in their long struggle against the empire. They were mistaken. Fiscally strong, the Sublime Porte was able to make good the navy’s losses within a matter of months and thus ensured that the Christians could not turn their victory into permanent control of the seas around Greece. Lepanto thus had no immediate adverse consequences for the empire and did nothing to weaken its grip on the Eastern Mediterranean. Indeed, thanks to its conquest of

¹ Colin Imber, *The Ottoman Empire, 1300-1650: The Structure of Power*, Second edition (New York: Palgrave Macmillan, 2009), 56–58.

Cyprus from the Venetians in the same year as its defeat at Lepanto, the Ottoman Empire actually improved its strategic position in the region.²

Beginning with the reign of Mehmet III (r. 1595-1603), however, the Ottomans experienced six decades of domestic and international problems that contemporaries referred to as the “time of troubles.” This characterization was apt, for it was without question a difficult period for the empire both at home and abroad. Internationally, the sultan’s armies found themselves bogged down in a series of long, costly, and frustrating wars. Domestically, meanwhile, the empire faced a combination of debilitating including political factionalism, destabilizing succession issues, and fiscally draining rebellions in Anatolia. These were serious challenges—ones that left it far weaker than it had been at the time of Selim II’s accession in 1566.³

The Long War, 1593-1606

The “time of troubles” began with the Long War that the empire fought with Austria from 1593-1606. It was a frustrating conflict for the Sublime Porte—one that followed a markedly different course than had earlier wars with Christian European states. Rather than achieving easy victories over an outmatched and outgunned foe as it had in the past, the sultan’s army instead found itself fought to a standstill by a potent, tactically sophisticated, and well-armed opponent. Indeed, the Ottomans won just a single field battle during the thirteen-year-long war, the Battle of Mezokeresztes in 1596, and did so only because Habsburg troops, having achieved a breakthrough, had lost cohesion when presented with the opportunity to loot the Ottoman camp. Nearly as bad, the conflict consisted for the most part not of decisive field engagements, but instead of long, drawn out, and costly sieges against a new, and nearly invulnerable type of fortification: the star fort. With walls too thick to be battered down by artillery, the new fortresses that the Austrian military engineers constructed prevented the Ottomans from achieving the sort of quick victory that had been the hallmark of the sultan’s military during the prior two centuries. Thus, unlike earlier wars with the Europeans that had been quickly and easily won, the Long War turned into an expensive stalemate—one that sapped morale even as it put great pressure on the treasury.⁴

Driven to exhaustion and near their fiscal breaking points, the Habsburgs and Ottomans agreed in 1606 to engage in talks along their common border for the purpose of bringing the conflict to an end. The result was the Peace of Zsitvatorok, which ended the war on the principle of *uti possidetis*, meaning that each party retained the territory that it controlled at the conclusion of the fighting. At first glance, the agreement appeared to be a favorable one for the Sublime Porte. It did not require the empire to cede any land, left it in possession of Hungary and the Danubian Principalities of

² Mark Greengrass, *Christendom Destroyed: Europe 1517-1648* (New York: Penguin, 2015), 505–6.

³ Imber, *The Ottoman Empire, 1300-1650*, 66.

⁴ Caroline Finkel, *Osman’s Dream: The History of the Ottoman Empire* (New York: Basic Books, 2005), 174–75.

Wallachia and Moldavia, and even awarded the sultan two border fortresses. Nonetheless, European and Ottoman diplomats alike understood that the settlement was a significant symbolic defeat—one that reflected a seismic shift in the balance between the empire and the European states. To that point, treaties involving the Ottoman Empire had always been negotiated in Istanbul and presented as acts of sultanic grace between highly unequal parties. Agreed to instead at the border and explicitly acknowledging for the first time that the Habsburg emperor was of equivalent rank to the sultan, the Peace of Zsitvatorok announced in unambiguous language that the Ottoman Empire was no longer the dominant force it had been under Sulëiman the Magnificent.⁵

War with the Safavids

In the meantime, the already stretched Ottoman military had to fight a second war—this time with the Safavids. For Shah Abbas I (r. 1588-1629), this conflict was one of revenge. Smarting from the concessions he had been forced to make to the Ottomans in 1590, the Persian leader had spent the final decade of the sixteenth century busily reorganizing his army and reequipping it with a large force of cannons. By the early 1600s, as a result, he was well positioned to take advantage of the Ottomans' preoccupation with the Habsburg Empire to start a war aimed at recovering the provinces he had been compelled to cede a decade earlier. He made his move in 1603. Striking boldly, his new military launched a coordinated series of attacks that quickly overran Ottoman forces in western Iran and the Caucasus. The Safavid invasion was a serious blow to the Sublime Porte. It not only cost the empire large amounts of valuable territory but also threatened to unhinge the entire Ottoman position in the east. In response, Sultan Ahmed III (r. 1603-1617) ordered a powerful counterattack in 1605 aimed at recovering Tabriz. Unfortunately for him, the offensive collapsed spectacularly when Shah Abbas I destroyed the invading Ottoman forces at Sufiyan. Thereafter, the war settled into another costly stalemate until it finally concluded with the Treaty of Serav in 1618.⁶

The stunning victories that Abbas I's reformed army won in the early years of the war made manifest the extent to which Ottoman military dominance had evaporated by the turn of the seventeenth century. Indeed, Safavid success at that time amounted to nothing less than a complete turnaround in the balance that had obtained between the two empires for nearly a century. Since the Battle of Chaldiran in 1514, Ottoman supremacy in the east had been so total that successive shahs had had little choice but to retreat when faced with the prospects of a field battle with the sultan's army. Shah Abbas I's riposte thus came as a great shock to the Sublime Porte and showed emphatically that Ottoman forces no longer held the upper hand along the eastern frontier. The Treaty of Serav reflected this tectonic shift in the military balance. By its terms, the Ottomans agreed to give back all the provinces they had won from Shah

⁵ Finkel, 175.

⁶ Abbas Amanat, *Iran: A Modern History* (New Haven: Yale University Press, 2017), 79–82.

Abbas I in 1590 and to accept a return to the frontier that had existed prior to 1578—a huge concession for an empire that had never before yielded more than token amounts of territory.⁷

The *Jelali* Revolt

The outbreak of a serious rebellion in Anatolia in 1596 significantly complicated the Ottoman Empire's ability to wage its simultaneous wars against the Safavid and Habsburg Empires. Known as the *Jelali* Revolt, the uprising had its roots in the grievances of key provincial groups. It began when the central government seized *timars* from *sipahi* who had deserted in battle in 1596; angry about losing their income, the dispossessed cavalymen responded by revolting and engaging in brigandage. Irregular, firearm-equipped peasants known as *sekban* soon joined them. These men had served in the Ottoman-Safavid War of 1578-1590 and expected to become members of the ruling elite at its conclusion; instead, the government had unceremoniously demobilized them in a cost-cutting move. Incensed, they engaged in local banditry until the *sipahi* revolt presented them with the opportunity to form a common front with the dispossessed cavalymen through which both groups could secure concessions from Istanbul. It was the union of these disaffected elements that turned the *Jelali* revolt into a serious threat to Ottoman control of Anatolia in the late 1590s—particularly after the governor-general of Karaman, who had been tasked with suppressing the rebellion, instead assumed leadership of it.⁸

As a result, by the early years of the seventeenth century, the situation in Anatolia had become grim for the government. Now well led and sophisticated, the *Jelali* rebels ran riot through the region. They occupied large swaths of Asia Minor, defeated a succession of armies sent against them, and besieged and extorted important cities such as Ankara. Crucially, they even briefly took control of the strategic passes that linked Syria and Asia Minor and thus threatened the sultan's control of the Arab Middle East.⁹

The central government's position in Anatolia only began to improve following the conclusion of the Long War. Using troops freed up as a result of the Peace of Zsitvatorok, Grand Vezir Kuyuju Murad Pasha (r. 1606-1611) directed a massive offensive designed to finally bring the *Jelali* to heel. Even then the going was tough. Only through the skillful use of divide-and-conquer tactics and the liberal dispensation of bribes—he bought off key rebel leaders with the promise of government offices—was he finally able to bring the revolt to a conclusion in 1610.¹⁰

⁷ Imber, *The Ottoman Empire, 1300-1650*, 62–63.

⁸ Douglas Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals* (Boulder: Westview Press, 2011), 124.

⁹ Imber, *The Ottoman Empire, 1300-1650*, 65.

¹⁰ Imber, 65.

Succession Issues

If the *Jelali* Revolt and the costly wars with the Safavid and Habsburg Empires were not challenging enough, the Ottomans also confronted growing political instability in the capital. These problems first emerged in 1589 when the Janissaries revolted over a cut in their pay. Murad III suppressed their rebellion easily enough, and, to all outward appearances, seemed to have restored order to his government. In fact, he had done nothing of the sort. On the contrary, while the Ottoman state appeared to be as stable as ever during the ensuing decades, it had in reality begun to splinter into competing factions that struggled with one another for control of the levers of power and, pivotally, for influence over the sultan.¹¹

The latent conflicts between the cliques finally burst into the open in 1617 following the death of Sultan Ahmed I (r. 1603-1617). One faction headed by the sheikh *ül-Islam*, the highest-ranked Muslim cleric in the empire, arranged to pass over Ahmed I's son, Osman II (r. 1618-1622), in favor of his more-malleable, though also mentally unbalanced brother, Mustafa I (r. 1617-1618, 1622-1623). Mustafa I's enthronement was portentous for the empire in two ways. First, it broke with the Ottoman precedent of having sons succeed their fathers as sultan. Second, and more importantly, it marked a dramatic escalation in the involvement of the bureaucratic and military factions in the selection of the ruler.¹²

The result was a period of severe political volatility in Istanbul. Just three months after Mustafa I assumed the throne, the Janissaries deposed him on the justifiable grounds that he was mentally unstable and replaced him with Osman II. The new sultan's reign did not mark a return to normalcy, however. Instead, dispensing with the obsequiousness that they had shown to previous rulers, competing cliques within the *askeri*—most egregiously, the Janissaries—schemed to control Osman II and challenged his prerogative to direct policy. By 1620, the exasperated sultan had had enough. Furious about the Janissaries' heavy-handed involvement in politics, he began plotting in secret to replace them with more politically reliable musketeers recruited from the peasants of Anatolia. This maneuver was a shrewd one and might have succeeded had the Janissaries remained unaware of his efforts. Unfortunately for Osman II, however, the well-connected soldiers sniffed out his plan before he could put it into effect. Moving quickly to defend their status and position, they dethroned and executed him in 1622. They followed by briefly returning the mentally disturbed Mustafa I to the throne before once more deposing him in favor of another of Ahmed I's sons, Murad IV (r. 1623-1640).¹³

¹¹ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 177–78.

¹² Finkel, 177–78, 196–97.

¹³ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 55. Douglas A. Howard, *A History of the Ottoman Empire* (Cambridge: Cambridge University Press, 2017), 146–47.

The empire rapidly descended into chaos following Osman II's murder. Taking advantage of their military power and the fact that their barracks were adjacent to the palace, the Janissaries demanded and received a substantial accession donation—functionally a bribe—from the new sultan to ensure their continued loyalty; they followed by further destabilizing the already-shaky political system by deposing a succession of grand vezirs. These actions benefited the Janissaries in the short term, but they were enormously unpopular and produced a backlash in the provinces. The governor of Erzurum, Mehmet Pasha (1576-1634), soon gave focus to that anger. Declaring that he sought to avenge the murder of Osman II, he raised a revolt that quickly secured control of large parts of Asia Minor including the important cities of Sivas, Bursa and Ankara. To the delight of many people, he also ordered the execution of any Janissaries found in the territory he controlled on grounds that they were complicit in Osman II's murder. Short of funds as a result of the accession bonus given to the Janissaries, the central government could offer no meaningful response to this new challenge.¹⁴

The crisis the Ottomans faced soon worsened dramatically, moreover. Taking advantage of the paralysis in the Sublime Porte and the unrest in Asia Minor, Shah Abbas I invaded Ottoman Iraq in 1623 and laid siege to Baghdad. To his delight, his soldiers took possession of the city early the next year. It was a shocking blow to Ottoman prestige—one that brought into clear focus the problems that the central government's paralysis had created.¹⁵

Murad IV—Recovery

Surprisingly, given that he was only twelve when he took the throne, the new sultan, Murad IV, was able to provide the effective leadership needed to restore internal order and the empire's territorial integrity. That he would be successful in doing so was far from obvious during the early years of his reign. While he had been able to quickly bring Mehmet Pasha to heel, his effort to retake Baghdad in the early 1630s had ended in defeat. Worse, his subsequent decision to sack the popular grand vezir who had commanded the failed campaign in Iraq, Husrev Pasha (?-1631), had incited the Janissaries to revolt once again. Thus, by the early 1630s, the situation in Istanbul appeared to have reverted to the chaos that had obtained in the early 1620s. Indeed, most observers had come to conclude that the sultan was a weak leader—one who was powerless either to prevent the empire from descending further into disorder and instability, or, more immediately, from avoiding the same fate as Osman II.¹⁶

In this case, however, appearances were deceiving. Murad IV in fact proved to be a remarkably effective ruler who—aided, as we shall see, by his capable mother—defly outmaneuvered the Janissaries as the first step in a successful campaign designed to reassert sultanic control over the empire. His effort to deal with the soldiers' uprising was smartly planned and executed. He initially responded with a shrewd

¹⁴ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 202–5.

¹⁵ Gene R. Garthwaite, *The Persians* (Malden, MA: Blackwell Publishing, 2005), 174.

¹⁶ Imber, *The Ottoman Empire, 1300-1650*, 69–71.

tactical retreat in which he appeared solicitous and agreed to meet their main demands. Concluding from the episode that the sultan would not challenge them, the mollified Janissaries stood down—just as Murad IV had hoped they would. Taking advantage of the breathing space he had created for himself, he followed by quietly securing allies from among those factions that resented the soldiers' dominance of the government. Then, having solidified his position in this way, he had the rebellion's leaders arrested and executed.¹⁷

He followed with a flurry of initiatives designed to restore internal order and to reassert the empire's dominance of its borders. He pursued a series of reforms to the system of taxation, moved to impose stringent new moral regulations, and acted to subdue brigands who had exploited the anarchy in the capital to ravage Anatolia. He also personally led a campaign against the Safavids in 1638 aimed at reconquering Baghdad. Murad IV was successful in all of these initiatives. As a result, he was not only able to cement his control of the government, but also to end the instability that had dogged the Sublime Porte since Ahmed I's death in 1617. Indeed, by the time his reign came to a close in 1640, he had so thoroughly reestablished order that his subjects hailed him as the restorer of the empire.¹⁸

Ibrahim the Mad

Initially, the recovery that Murad IV had achieved continued under his successor, Ibrahim, (r. 1640-1648). It did so thanks not to the skilled leadership of the sultan—who was already demonstrating signs of mental instability—but instead to the tactful and effective guidance of his grand vezir, Kemankesh Mustafa (1592-1644). Sidelining Ibrahim, Kemankesh Mustafa was able to keep the factions in check and to maintain the empire's international position. As a result, the first four years of Ibrahim's tenure passed uneventfully.¹⁹

A pair of unrelated events in 1644 shattered the calm, however, and ushered in another difficult period for the empire. First, a faction headed by Ibrahim's charismatic, spiritualist advisor, the charlatan Jinji Hoja (?- 1648), used its influence with the sultan to have Kemankesh Mustafa executed—thus paving the way for it to assume control of the levers of power. Interested primarily in enriching themselves, the men provided little effective leadership and imposed no check on the sultan's increasingly erratic actions. Second, the empire started a war that year that quickly transformed into a costly quagmire. The conflict began when Maltese corsairs seized a ship carrying a high-ranking Ottoman official. Briefly stopping at the Venetian-controlled Island of Crete, the pirates issued a ransom demand before quickly spiriting their captive to Malta. The sultan and his advisors were furious. Believing that the slight could not go unpunished, they mobilized the fleet and army for a retaliatory attack. They did not target Malta,

¹⁷ Imber, 69–71.

¹⁸ Imber, 69–71.

¹⁹ Leslie P. Peirce, *The Imperial Harem: Women and Sovereignty in the Ottoman Empire* (New York: Oxford University Press, 1993), 246.

however; instead, using the apparent complicity of the Venetians in the incident as a pretext, they invaded the valuable and strategically located island of Crete. Catching the defenders by surprise, the Ottoman army quickly overran all of the island save for the heavily fortified port city of Candia.²⁰

The fortunes of war soon turned against the Ottomans, however. Venice may not have been the power it once was, but it still possessed a formidable fleet that could ferry a steady stream of supplies to the garrison in Candia; as a result, the city was able to hold out for an astonishing twenty-five years despite being besieged by a huge Ottoman army. More damagingly, in 1647 and 1648, the Venetians used their powerful navy to blockade the critical Dardanelles Strait that connected the capital to the Aegean Sea. The closure of the waterway was a huge blow to the empire. It not only prevented Ottoman soldiers on Crete from receiving desperately needed supplies, but also inflicted serious economic hardship and food shortages on the people of Istanbul.²¹

Meanwhile, at the same time that the empire was suffering setbacks against the Venetians, Ibrahim's sanity began to deteriorate rapidly. The decline in his mental health came in the context of his effort to meet a pressing dynastic need. Social and political order in the empire depended on the unifying symbol of the Ottoman family; as Ibrahim was the sole remaining Ottoman male, however, the dynasty on which stability rested was in danger of ending. Accordingly, to ensure that the line did not die out, his advisors pressed him to make siring sons a high priority. Sexual relations were a pleasant obligation for Ibrahim at first; after 1644, though, they evolved from a dynastic necessity into an all-consuming personal fetish. Ignoring his official duties, the sultan soon spent all of his time in the harem—the part of a Muslim household reserved for women and children—in the company of a growing roster of concubines recruited to meet his particular sexual tastes. Given his mental-health issues, the sultan's absence from the divan would likely have been more of a help than a hindrance to the Ottoman state. However, his desire to lavish costly gifts on his sexual partners—he insisted on carpeting the palace of one of his favorites with fabulously expensive sable furs—posed growing problems because they could only be met by raising taxes on the empire's already overburdened subjects.²²

Those levies contributed to a spectacular crash in support for Ibrahim among key groups. Coming at a time of blockade-induced food shortages and lurid rumors of the sultan's sexual excesses, the imposition of new taxes to pay for the luxuries he showered on his concubines embittered the residents of the capital. Meanwhile, the sultan's fumbling rule and his government's inability to bring the war with Venice to a successful conclusion infuriated the Janissaries. Fed up, the soldiers determined that Ibrahim had to go. Accordingly, they secured a fatwa, or legal ruling, in August 1648 from the sheikh ül-Islam authorizing them to depose and kill the increasingly

²⁰ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 223–27.

²¹ Finkel, 227. Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 56–57.

²² Imber, *The Ottoman Empire, 1300-1650*, 70–72.

unbalanced sultan. They followed by arranging for his six-year-old son, Mehmet IV (r. 1648-1687), to succeed him.²³

While a great relief to the people of the capital, the murder of Ibrahim did not result in a return to the stability of Murad IV's reign. Instead, the empire experienced eight more years of upheaval in what seemed like a replay of the anarchic 1620s. Mounting resentment of the Janissaries' dominance of the government resulted in the outbreak of yet another rebellion in Asia Minor while stepped-up factional infighting in the capital left the central government incapable of providing effective rule. Even the streets of the capital experienced disorder. Angry over new taxes and a costly debasement of the currency, the residents of the city repeatedly vented their anger by rioting.²⁴

By 1656, the crisis had reached a boiling point. Factionalism and instability in the government had metastasized to the point that the grand vezirate had become a revolving door: indeed, no fewer than seven different men occupied the post between May 1655 and October 1656 including one, Zurnazen Mustafa Pasha, who served for a mere four hours. Meanwhile, the soldiers chose this difficult period to launch yet another revolt—one that, predictably, further paralyzed the central government. Finally, and most shockingly, Venice inflicted a crushing defeat on the Ottoman fleet in the summer of 1656 that permitted it to seize the islands of Lemnos and Bozcaada and, with them, control of the Dardanelles Strait. Times were very dark indeed for the empire.²⁵

Sources of Ottoman Weakness

The Venetian victory in 1656 made clear to observers both inside and outside the Ottoman Empire that the sultanate was no longer the dominant power it had once been. What accounted for its reversal of fortune? How did a state so overwhelmingly strong in 1566 that it could dominate all its frontiers find itself facing defeat at the hands of tiny Venice not even a century later? Paradoxically, the difficulties that the Ottomans experience during the first half of the seventeenth century were related to the peculiar characteristics that had facilitated the empire's meteoric rise. Put simply, many of the strengths that had fueled the sultanate's stunning growth and transformation during its early years had instead become relative or absolute liabilities that undermined it during the seventeenth century.

Economic and Fiscal Issues

The first area of strength that transformed into a weakness was the economy. As noted in chapter six, the early Ottoman Empire had benefited from several centuries of

²³ Jason Goodwin, *Lords of the Horizons: A History of the Ottoman Empire* (New York: Henry Holt and Company, 1998), 226–27.

²⁴ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 236–46.

²⁵ Howard, *A History of the Ottoman Empire*, 168; Finkel, *Osman's Dream: The History of the Ottoman Empire*, 247–53.

steady economic expansion. Its rising population, the development of new trade routes, the cultivation of new land, and the monetization of exchange in rural areas—that is, the shift from barter to cash-based transactions—had combined to produce substantial growth. Following Sulëiman the Magnificent’s reign, however, the empire’s economy began to stagnate. The rise in population slowed, trade plateaued, and, as the marginal utility of new land brought under the plow steadily decreased, the growth in agricultural productivity stalled.²⁶

Making matters worse, the Ottoman economy began to experience at that time what scholars refer to as the Price Revolution: a general period of inflation in western Eurasia caused by the influx into Europe of large quantities of gold and silver from the Americas. While the effect of the sudden rise in the supply of precious metal on the empire was substantially more limited than its impact on Western Europe—and far-less consequential than historians once believed—it nonetheless fueled a steady rise in prices that further contributed to the dampening of commercial activity in the sultanate.²⁷

Unsurprisingly, the Ottoman government began to experience significant fiscal shortfalls at this time. As we have seen, these deficits had little precedent. Through the early 1580s, the Sublime Porte had regularly enjoyed substantial budget surpluses; thanks to the sultanate’s subsequent economic stagnation, however, the tax receipts that had grown in the sixteenth century now stopped rising. Worse, since the empire levied taxes on a nominal basis—meaning that its assessments did not rise along with inflation—the Price Revolution meant that existing tax revenue declined in real terms.²⁸

Two other issues further exacerbated the Ottoman Empire’s worsening fiscal situation. Beginning in the early-seventeenth century, the growing oceanic exchange between Asia and Europe that had started with Vasco da Gama’s (c. 1460s-1524) arrival in Calicut in 1498 finally began to eclipse the Silk Road trade that passed through the Ottoman Empire. As a result, the customs revenue that the overland trade had heretofore provided began a long, gradual decline. Meanwhile, in contrast to the earlier wars of expansion that had generated a steady inflow of loot to the treasury and had thus largely paid for themselves, the conflicts of the late-sixteenth century constituted a significant drain on the exchequer. The conflict with the Safavids, the Long War with the Habsburgs, and, especially, the naval contest with the Holy League from 1570-1573 were all extremely costly ventures that rapidly consumed the treasury’s surplus even as they failed to generate compensatory revenue in the form of loot.²⁹

Thanks to the empire’s worsening fiscal situation, meanwhile, an existing but heretofore manageable peculiarity of the Ottoman fiscal calendar began to pose a serious, if periodic, challenge to the exchequer in the late-sixteenth century. This

²⁶ Sevket Pamuk, *A Monetary History of the Ottoman Empire* (Cambridge: Cambridge University Press, 2000), 125–32.

²⁷ Pamuk, 125–32.

²⁸ Pamuk, 128, 138–39.

²⁹ Pamuk, 131–32, 139.

problem was a function of an unusual disparity between the treasury's receipt and disbursement of funds. It collected taxes on a solar calendar—a necessity in an agricultural economy—but paid salaries to its troops and officials on a quarterly basis according to the Muslim lunar calendar. As the lunar calendar is eleven days shorter than a solar year, the Ottoman treasury thus had to pay roughly thirty-three years of salary every thirty-two years; in other words, every eighth year, it had to come up with a fifth quarter's worth of salaries. During the sultanate's astonishing rise, this issue had been a manageable one for the treasury. Finding the extra funds every eighth year had simply not been an issue when the sultan's army was bringing in vast extra revenue in the form of loot. As we have seen, however, the sultanate had stopped growing following the reign of Sulëiman the Magnificent; beginning in the late-sixteenth century, as a result, the octennial need to find the excess funds to pay an extra quarter's worth of salaries would constitute a serious challenge for the increasingly strapped treasury.³⁰

Aggravated by the long and expensive Ottoman-Safavid War of 1578-1590 and by the lingering costs of the new fleet it had to construct after the Battle of Lepanto, the bill finally came due in the mid 1580s. Desperate to secure enough coinage to meet its expenses, the treasury responded by instituting a draconian, 44 percent debasement of the currency, the silver *akce* coin, in 1585-1586. This devaluation gave the government the funds it needed to meet its obligations, but it also produced significant civil unrest and a destabilizing surge in inflation that hit soldiers and bureaucrats with particular severity. More importantly, the dramatic depreciation of the currency set in motion a vicious fiscal-economic cycle. Thereafter, debasement and monetary instability produced inflation that depressed the economy; the resulting decline in economic activity in turn reduced real tax receipts, which then exacerbated the fiscal crisis and impelled the treasury to pursue further debasements.³¹

Thus, by the early-seventeenth century, the empire's fiscal and commercial position had grown bleak. Its once-expanding economy had stalled, and its treasury officials grappled with steadily worsening budget deficits that they only belatedly grasped could not be resolved through further debasements. It was a difficult situation for the sultanate—one made worse by virtue of the fact that both the economy and the state's fiscal standing had so recently been among its core strengths.

Weak Sultans

The empire might have had greater success resolving the fiscal crisis of the late-sixteenth and early-seventeenth centuries had it continued to enjoy strong leadership. Unfortunately, the succession of highly capable rulers that began with Osman I (r. 1299-1326) ended following the death of Sulëiman the Magnificent. Indeed, after a brief stretch of merely lackluster sultans, the empire would experience a nearly unbroken,

³⁰ Halil Sahillioglu, "Sivis Year Crises in the Ottoman Empire," in *Studies in the Economic History of the Middle East: From the Rise of Islam to the Present Day*, ed. M. A. Cook (New York: Routledge, 1970), 230–49.

³¹ Pamuk, *A Monetary History of the Ottoman Empire*, 135, 141, 147–48.

century-long string of weak rulers who were unable to provide effective direction of the Sublime Porte.³²

On the surface, two factors explain the decline in sultanic leadership. First, as we have seen, two of the sultans who reigned during the next century, Mustafa I and Ibrahim, suffered from severe mental instability; as such, they proved incapable of directing the affairs of state. Second, nearly all of the others who reigned during the seventeenth century came to power as children. Mehmet IV was seven when he became sultan, for example, while Osman II and Ahmed I were each thirteen at the time they acceded to the throne. Unsurprisingly, the child sultans had a difficult time asserting themselves over the factions that had come to dominate the bureaucracy and court. Murad IV, of course, had broken this pattern and had enjoyed a successful reign despite becoming sultan at the age of nine; however, he was very much the exception. Indeed, Murad IV was the *only* genuinely effective ruler between Mehmet III's death in 1603 and Sulëiman II's (r. 1687-1691) accession to the throne more than eighty-years later.³³

Youth may have been a contributing factor in the decline in sultanic leadership—again, Murad IV excepted—but a fundamental change in the succession process in the early-seventeenth century proved far-more consequential. Beginning with the accession of Ahmed I in 1603, the *askeri* adopted a new system for selecting sultans that ended the practice of fratricidal succession that had worked so effectively since the late-fourteenth century to ensure that the empire enjoyed strong leadership. Where the old system had involved a Darwinian contest among the heirs that ended with the winner executing his fraternal rivals, the new arrangement instead spared the sultan's brothers by giving the bureaucrats and courtiers the power to select the new leader. To ensure that the passed-over heirs did not possess powerbases from which they could challenge the recently installed ruler, the new system also ended the practice of having the sultan's sons gain administrative experience by serving as provincial governors. Instead, all of the heirs, including the next sultan, led pampered lives of ease and pleasure in the luxurious, unreal atmosphere of the Gilded Cage—the name given to the imperial harem—where they acquired no sense of the real world and easily fell under the influence of powerful courtiers.³⁴

Why did the Ottomans adopt this new method of succession if it failed to produce leaders of the same caliber as the fratricidal system? They did so for two reasons. First, at only thirteen-years of age, Ahmed I had not yet fathered any heirs; as such, killing his brother Mustafa I ran the risk of ending the dynasty and sparking a period of anarchy and crisis. Second, popular revulsion at the practice of royal fratricide had begun to surge in the late-sixteenth century. The line of five, child-sized coffins wheeled out of Topkapi Palace on the day that Murad III acceded to the sultanate, and, even more so,

³² Finkel, *Osman's Dream: The History of the Ottoman Empire*, 165–66.

³³ Finkel, 165–66.

³⁴ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 91–93.

the train of nineteen caskets that emerged when Mehmet III took the throne shocked the people of Istanbul and compelled the *askeri* to adopt the new system.³⁵

While the new arrangement may have been more humane, it was also much less effective at producing capable rulers than the fratricidal system it replaced. Raised in the luxurious, fantastical environment of the Gilded Cage, the sultans of the seventeenth century were ill-equipped to respond effectively to the increasingly serious challenges confronting the empire. Bereft of practical experience or even a realistic grasp of the world in which they lived, they were figureheads rather than actual rulers—the creatures of competing factions of bureaucrats, ‘ulama’, Janissaries, eunuchs, and courtiers rather than leaders in their own right. Thus, thanks to the abandonment of the system of fratricidal succession—the sultanate—another of the empire’s early sources of strength, had become a significant source of weakness.³⁶

Military Problems

As the Long War and the Ottoman-Safavid War of 1603-1618 made clear, even the empire’s vaunted military, its most obvious strength during its rise to power, had lost its edge. This change marked an enormous shift in the military balance. During the conflicts of the early- and mid-sixteenth century, the sultan’s army had easily dominated its enemies on every front. By the turn of the century, in contrast, the balance between the Ottoman military and the armies of its rivals had shifted to such a degree that Habsburg forces had been able to fight the sultan’s soldiers to a draw while Shah Abbas I’s Safavid troops had outright trounced them. What had happened during the late-sixteenth century? Why was the Ottoman military no longer able to defeat its enemies as it had in the past?

First, to be clear, the change in the military balance between the empire and its rivals was not indicative of a general decline in the efficacy of the sultan’s army. The Ottoman military of the early-seventeenth century remained a potent, well-led, and well-supplied force. Indeed, it was every bit as powerful as it had been when Selim II had triumphed so spectacularly over the Safavids at Chaldiran in 1514 or when Sulëiman the Magnificent had crushed the Hungarians in 1526.³⁷

Rather, the empire experienced a *relative* decline in military power—due not to its shortcomings or lack of discipline but instead to significant improvements in the armies of its European and Persian rivals. As earlier noted, Shah Abbas I had modernized the Safavid military in the 1590s in part by dramatically increasing the number of cannons it employed; by the early-seventeenth century, as a result, his troops were able to go toe-to-toe with Ottoman forces. European armies improved to an even greater degree relative to the Ottomans due to the Military Revolution, a series of

³⁵ Imber, *The Ottoman Empire, 1300-1650*, 95–101.

³⁶ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 92.

³⁷ Virginia H. Aksan, *Ottoman Wars, 1700-1870* (Harlow, England: Routledge, 2007), 45.

technological and organizational changes among European armies that dramatically altered the way wars were fought. It involved four major changes: the deployment of much-larger armies, the use of powerful new weapons such as muskets and field guns, the adoption of innovative and devastating new tactics such as volley fire—a technique that permitted a formation to shoot continuously—and the construction of new, star-shaped, *trace italienne*-style bastion forts, which, as the Ottoman military discovered during the Long War, were nearly impervious to artillery fire. These were dramatic changes. They not only allowed European militaries to close the gap with the Ottoman army by the turn of the seventeenth century but gave them the ability thereafter to fight it on an even footing. Thus, while the sultan's armies were as strong as ever, they no longer enjoyed the relative advantage that the empire's early adoption of gunpowder weapons had previously afforded.³⁸

That said, two critical components of the Ottoman military, the *sipahis* and the Janissaries, did decline in effectiveness between the early-sixteenth and the start of the eighteenth centuries. The *sipahis* were the first to lose their edge. The bow-armed provincial cavalry had gradually become obsolete over the course of the sixteenth century as a result of the growing adoption and increasingly sophisticated employment of gunpowder weapons; however, the continued success of Ottoman armies masked their decline. Indeed, it was only during the Long War that their military ineffectiveness—particularly in the face of massed formations of musket-armed infantry—fully manifested itself and made clear that they were rapidly becoming an anachronism on the battlefield.³⁹

The Janissary corps' decline was more gradual. Its devolution began in 1566 when the Janissaries won the right to marry and to produce legal heirs. Eager to ensure that their children inherited their *askeri* status, they followed by winning modifications to the traditional system of recruitment that permitted their sons to join the corps. These changes gradually transformed the Janissaries over the next century. Indeed, by the time Sultan Ahmed III formally abolished the *devshirme* in 1703, the corps had long since completed its shift from a meritocratic institution beholden to the sultan into an hereditary one dedicated to furthering its parochial interests.⁴⁰

The debasement-induced inflation that began in the late-sixteenth century accelerated the Janissary corps' decline. As we have seen, the series of debasements that the treasury had undertaken beginning in the late-sixteenth century had brutally eroded the purchasing power of the Janissaries' salaries. In response, the soldiers began to supplement their pay by taking jobs as porters or butchers or by opening small businesses. Ostensibly, the Janissaries were employed in those positions on a part-time basis for the purpose of offsetting the decline in their wages. In reality, their side jobs

³⁸ Geoffrey Parker, *The Military Revolution: Military Innovation and the Rise of the West, 1500-1800*, Second Edition (Cambridge: Cambridge University Press, 1996), 6–24.

³⁹ Imber, *The Ottoman Empire, 1300-1650*, 292–93.

⁴⁰ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 45.

were full-time obligations that left scant time for them to train or otherwise fulfill their official duties.⁴¹

The end of the *devshirme* system and the Janissaries' adoption of wage-earning work proved doubly harmful to the empire. First, the transformation of the Janissaries from elite soldiers into full-time workers led to a dramatic decline in their military effectiveness between 1600 and 1750. Abandoning the training and discipline that had been so critical to their earlier success in order to work their civilian jobs, the Janissaries devolved into a completely ineffective military force and eventually stopped going on campaign altogether. Paradoxically, meanwhile, even as the corps' military capabilities declined, its political influence reached new heights. No longer bound to the sultan as they had been when they were slaves recruited through the *devshirme* system, they were now in a position to use their monopoly over military power in the capital to raise and overthrow sultans at will. This they did with abandon. Between 1618 and 1730, either alone or in conjunction with other cliques, they deposed Mustafa I—twice— Mehmet IV, Mustafa II (r. 1695-1703), and Ahmed III; meanwhile, they had not merely dethroned Osman II and Ibrahim the Mad but had also executed them.⁴²

In recent decades, scholars have significantly complicated this understanding of the Janissaries. While still acknowledging the corps' declining martial abilities and its destructive meddling in the succession, they have shown that it also served an important social role in the capital as the populist protector of the urban common people. Two peculiar characteristics allowed it to do so. First, in contrast to nearly all other groups, the Janissaries were members of both the *askeri* elite and the urban working class; as such, they were in a unique position to mediate between the Sublime Porte and their lower-class brethren. Second, they held a monopoly on military power in the capital and were thus well positioned to compel the sultan's government to alter policies perceived as harmful to urban workers. The corps began to assume its role as a populist intermediary body early in the eighteenth century. Thereafter, it frequently exploited its dominant military position in the capital to compel the Sublime Porte to abolish laws and policies that impinged on the welfare of the urban lower class— particularly those that harmed the interests of the artisan guilds that employed many of the Janissaries.⁴³

The corps' new role as an intermediary body may have been a socially critical one, but it could not offset the fact that the Janissaries and their *sipahi* comrades no longer dominated the battlefield as they had during the Ottoman rise to power. Indeed, by the end of the seventeenth century, the decline of those forces had become obvious. The *sipahis* troopers who had once nimbly outmaneuvered heavier cavalry and routed infantry during the empire's early years were now almost completely obsolete. More dramatically, the Janissary corps had not merely lost its military dominance but had, in the words of the historian Donald Quataert, devolved from being "the terror of its foreign

⁴¹ Quataert, 45.

⁴² Finkel, *Osman's Dream: The History of the Ottoman Empire*, 202, 251.

⁴³ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 47–48.

foes to the terror of the sultans.” In short, two more institutions that had been vital to the dynasty’s early success had, by the eighteenth century, instead become liabilities.⁴⁴

The Capitulations

Meanwhile, though they had little immediate impact, a series of bilateral trade agreements called Capitulations that the Sublime Porte entered into with several European states beginning in the late-sixteenth century would have important long-term ramifications for the empire. Based on earlier Byzantine commercial treaties, the Capitulations provided merchants from signatory states with special privileges while operating in the sultanate. Thanks to those treaties, they enjoyed extraterritoriality, or immunity from Ottoman laws, and exemption from the payment of local taxes.⁴⁵

The Capitulations served the Sublime Porte’s interests in two ways. First, they provided a means of locating European traders within the broader social, political, and economic structure of the empire. As non-Ottoman subjects, merchants from the capitulatory states were neither entitled to the benefits of Ottoman law nor permitted to access the courts of the empire’s recognized socioreligious communities; they thus lacked legal standing and protection while operating in the sultanate. The Capitulations solved this problem by creating what were in essence new *millets* headed by the monarchs of the European states from which the merchants came. Second, the agreements were an attractive concession that the Sublime Porte could grant to European states with which it hoped to conclude diplomatic agreements. Indeed, the Ottoman state granted the first Capitulation to France in 1569 as a reward for allying with the empire against their common Habsburg enemy. For similar reasons, the Sublime Porte entered into a capitulatory agreement with Britain in 1579 that opened trade in the empire to the royally chartered Levant Company.⁴⁶

At the time they were signed, the Capitulations did not appear to be terribly significant politically or even economically. The volume of trade was small, and the empire was much stronger than the capitulatory states. As such, the Sublime Porte could easily abrogate the agreements at no cost if it wished. Unbeknownst to Ottoman officials in the late-sixteenth and early-seventeenth centuries, however, the Capitulations posed a latent, long-term threat to the empire’s political independence and economic health. Were the balance between the sultanate and the European kingdoms to reverse, those states would not only be able to use the Capitulations to assume a dominant position in the empire’s economy but would have the military strength to prevent the Sublime Porte from revoking the agreements. In other words, the

⁴⁴ Quataert, 45.

⁴⁵ Howard, *A History of the Ottoman Empire*, 166.

⁴⁶ Donald Quataert, *The Ottoman Empire, 1700-1922, Second Edition* (Cambridge: Cambridge University Press, 2005), 78–79.

Capitulations constituted a ticking time bomb for the empire, one slated to explode if and when—paradoxically—the Sublime Porte was too weak to end them.⁴⁷

The Historical Debate: Decline or Adaptation?

For a long time, the “Ottoman decline thesis”—the view that the Ottoman Empire experienced a gradual, inexorable degeneration that began with the conclusion of Sulëiman the Magnificent’s reign and that ended with the sultanate’s dissolution in 1922—dominated the historical literature. While Western scholars of the mid-twentieth century are most commonly associated with this interpretation, it is in fact a very old idea that originated in the empire itself. Indeed, the people who first advanced it were the seventeenth-century writers of *Nasihatname*, or Advice of King, literature—political commentaries designed to influence how the Ottoman Dynasty governed. Coming largely from the bureaucratic class, these authors argued that the “time of troubles” was a direct product of the ineffective leadership of the sultans who succeeded Sulëiman the Magnificent. Interpreting the world through the lens of the Circle of Justice, they maintained that the rulers of the late-sixteenth and early-seventeenth centuries were degenerate, pleasure-seeking men who had failed to rule justly or to maintain order; as a result, the peasants suffered, tax receipts fell, and the strength of the military waned. How could the empire regain its equilibrium? The *Nasihatname* writers argued that the answer was obvious: it could do so only through the return of effective, moral leadership at the top.⁴⁸

Initially, the *Nasihatname* concept of Ottoman decline was not widely known. Written by a modest number of officials for an equivalently small audience of sultans and top leaders, it remained confined largely to the bureaucratic class. However, the idea that the empire was in decline gradually spread—first to a wider audience within the Ottoman world and then, in the eighteenth and nineteenth centuries, to Europe. There, with the Ottoman Empire’s relative weakness as a backdrop, it soon began to shape Western interpretations of the sultanate’s troubles.⁴⁹

It was not until the mid-twentieth century, however, that the Ottoman decline thesis fully took shape. Seeking to explain the empire’s apparent failure to modernize, Western scholars writing after World War II drew heavily on the *Nasihatname* literature to argue that a shift from strong to weak sultans beginning in the late-sixteenth century led ineluctably to the decentralization of political power and to the broader degeneration of the empire. The most influential of these writers was the Princeton historian Bernard Lewis. He contended that the lack of effective leadership in the late-sixteenth and early-

⁴⁷ Quataert, 78–79.

⁴⁸ Jane Hathaway and with contributions by Karl Barbir, *The Arab Lands under Ottoman Rule: 1516-1800* (New York: Routledge, 2013), 8; Douglas A. Howard, “Genre and Myth in the Ottoman Advice for Kings Literature,” in *The Early Modern Ottomans: Remapping the Empire*, ed. Virginia H. Aksan and Daniel Goffman (New York: Cambridge University Press, 2007), 150.

⁴⁹ Howard, “Genre and Myth in the Ottoman Advice for Kings Literature,” 150.

seventeenth centuries resulted in the gradual and irreversible decay of key structures and institutions within the empire. The bureaucracy slid into venality and corruption, the military lost its discipline, the local elites carved the provinces into semiautonomous fiefdoms, and the economy gradually stagnated under the weight of corrupt tax collectors until, unable to defend its territory from the increasingly dynamic and modern West, the empire finally collapsed in 1922. Thus, to proponents of the Ottoman decline thesis such as Lewis, the empire's demise was the inevitable result of the military, economic, bureaucratic, and even cultural decadence that had begun with the shift to weak sultanic leadership in the late-sixteenth and early-seventeenth centuries.⁵⁰

While the Ottoman decline thesis dominated understandings of the empire in the 1950s and 1960s, it began to face mounting criticism from a new generation of scholars after 1970. They started by critiquing the logic that underlay it. Most pointedly, they noted that the idea that decline in the late-sixteenth and early-seventeenth centuries led directly to the empire's collapse in the 1920s was a teleological fallacy: a flawed form of reasoning that interprets all past events as the causes of an inevitable final outcome. In their eyes, such thinking ignored the many obvious ways in which the Ottoman situation improved after 1650, and, more problematically, denied agency to people living in later times.⁵¹

Those scholars followed by reexamining the primary sources of the late-sixteenth and early-seventeenth centuries. Here, they concluded that the evidence simply did not sustain the idea that the empire had degenerated. Most notably, they found that the writers who composed the *Nasihatname* literature were not offering unbiased or even realistic assessments of the issues that the empire confronted during the "time of troubles." Rather, their arguments reflected either the broader complaints of the bureaucratic class, or, more frequently, their own personal grievances; as such, their claims of imperial decay cannot be taken at face value. Likewise, these scholars saw success rather than weakness or degeneracy in the military's performance in the conflicts of the early-seventeenth century. They acknowledged that the empire's inability to achieve a decisive victory over the Habsburgs, the Safavids, and the *Jelali* demonstrated that its overwhelming dominance of the battlefield had become a thing of the past. At the same time, however, they also contended that the very fact that the sultanate could wage war simultaneously on three fronts was indicative of the Ottoman state's ongoing strength and resilience rather than a sign of its stagnation or weakness.⁵²

By the 1980s, historians arguing along these lines had fully repudiated the Ottoman decline thesis. While acknowledging that the empire experienced a period of instability in the late-sixteenth and early-seventeenth centuries as it transitioned to a new, decentralized political system, they also persuasively showed that the devolution

⁵⁰ Bernard Lewis, "Some Reflections on the Decline of the Ottoman Empire," *Studia Islamica*, no. 9 (1958): 111–27.

⁵¹ Howard, "Genre and Myth in the Ottoman Advice for Kings Literature," 137–66.

⁵² Howard, 137–66.

of power was a successful response to the military, economic, and diplomatic challenges that the empire confronted during its “time of troubles” rather than a function of decadence or weakness. In other words, the institutions and practices that had proven so effective in the fifteenth century did not degenerate but instead evolved as part of a successful response to the changing circumstances that confronted the empire beginning in the late-sixteenth century. Thus, by the 1990s, a new historical consensus had emerged that characterized the period from the death of Sulëiman to the mid-eighteenth century not as a time of decline but instead as a period of successful adaptation to the empire’s changed circumstance.⁵³

Ottoman Adaptation

What did this transformation encompass and when did it occur? First, to be clear, it was not a lagging response to the challenges that began to emerge in the late-sixteenth century. Instead, while the evolution of the empire peaked between the 1650s and 1690s, the process of adaptation started almost immediately after the problems that confronted the Ottoman state began to manifest themselves. Second, the empire’s evolution was far reaching and impacted nearly all aspects of the Ottoman system. The distribution of political power in the capital, the governance of the provinces, the collection and disbursement of taxes, and the structure of the military all changed in response to the empire’s new circumstances. The result was a very different state than the one that Mehmet II or Sulëiman the Magnificent had ruled—one that was far-better suited to the conditions of the seventeenth and eighteenth centuries.

Military Reforms

Unsurprisingly, the empire first acted to reform its army. It did so as a direct result of the frustration it experienced fighting the Long War. Unaware of the impact that the Military Revolution was having on the armies of the Christian European states, Ottoman generals had assumed at the start of the conflict that their soldiers would easily defeat Habsburg forces just as they had so many times in the past. Instead, to their shock, the army found itself bogged down in a long, costly, and frustrating stalemate against a vastly improved enemy. Most alarmingly, the Austrian military had not merely closed the gap that had long existed between it and the sultan’s army but had actually surpassed Ottoman forces in several key areas including the quantity and quality of field guns deployed and the use of advanced tactics such as volley fire.⁵⁴

In response, the Sublime Porte undertook a series of changes designed to restore the Ottoman military’s qualitative superiority. Though they touched on everything from logistics to grand strategy, the reformers focused their efforts on effecting changes to three key areas. First, they sought to bring Ottoman battlefield tactics and weapons in line with contemporary Western practices. As a result, the army quickly adopted

⁵³ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 99–100; Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 121.

⁵⁴ Imber, *The Ottoman Empire, 1300-1650*, 291–92.

European-style field guns and muskets as well as the volley-fire system that the Austrians had used with such effect in the Long War. Second, they moved to match the European states' shift to larger militaries. Accordingly, at great expense to the treasury, the reformers increased the size of the central standing army from 20,000 soldiers in 1567 to 37,000 at the conclusion of the war with the Habsburgs. Finally, they did not merely expand the army but also dramatically altered its composition. Most notably, they increased substantially the proportion that consisted of musket-armed Janissaries and *sekban* and severely reduced the share composed of the increasingly obsolete provincial *sipahis*.⁵⁵

The military reforms undertaken in the early-seventeenth century proved quite successful. The quick adoption of the strategies, tactics, and weapons of the Military Revolution substantially improved the Ottoman army's capabilities and ensured that it did not fall behind its European peers. Indeed, Virginia Aksan, one of the leading scholars of Ottoman military affairs, points out that the reforms ensured that the sultan's army remained roughly equivalent in strength and capability to the forces of its two primary rivals—Russia and Austria—at least through the turn of the eighteenth century. In other words, the sultan's military did not continuously and ineluctably weaken beginning in the seventeenth century as proponents of the Ottoman decline thesis once argued, but instead successfully evolved to meet the needs and challenges of a different time.⁵⁶

Fiscal Changes

While these reforms ensured that the Ottoman military enjoyed at least parity with its European rivals in the seventeenth century, they also contributed to the fiscal difficulties that bedeviled the empire at that time. Most obviously, the expansion of the army imposed new costs on the exchequer; after all, more troops inevitably meant greater expenditures. The increased size of the military was only part of the story, however. Indeed, the modifications that the Sublime Porte made to the army's composition likely played a more important part in increasing the fiscal burden than did the numerical expansion of the military. Put simply, shifting from a cavalry-heavy army to one in which infantry predominated was not merely an exercise in swapping one set of troops for another. Instead, it required the sultanate to make fundamental changes in the way that it raised and funded its military forces. Most notably, it involved replacing the provincial *sipahis*, whose expenses had been met through the *timar* system, with musket-armed Janissaries and *sekban*, whose costs were instead borne by the treasury.⁵⁷

⁵⁵ Aksan, *Ottoman Wars, 1700-1870*, 48–49; Imber, *The Ottoman Empire, 1300-1650*, 291–94; Gabor Agostono, "Firearms and Military Adaptation: The Ottomans and the European Military Revolution, 1450-1800," *Journal of World History* 25, no. 1 (March 2014): 97.

⁵⁶ Aksan, *Ottoman Wars, 1700-1870*, 45, 67–70.

⁵⁷ Imber, *The Ottoman Empire, 1300-1650*, 198.

In response, the central government began to gradually replace the *timars* that had sustained the provincial *sipahis* with *iltizam*, or tax farms, overseen by tax farmers called *multezim*. Common in Europe in the Early Modern Period and with antecedents in the Muslim world stretching back to the Buyid Dynasty and the Abbasid Caliphate, tax farming was a decentralized and privatized revenue system in which the state auctioned the right to collect taxes to private individuals. The Ottoman system was fairly typical of tax-farming regimes in the Early Modern Period. The winning bidder to an *iltizam* was obliged to pay the government an initial lump sum and then to forward to the treasury periodic tax payments raised during the three-year term of the contract; any excess revenue raised was the *multezim's* to keep. The Sublime Porte began the shift to tax farming in 1596 when it seized and converted into *iltizam* the *timars* of those cavalymen who had deserted during the Battle of Mezokeresztes. Thereafter, as the share of infantry in the army steadily rose, the government gradually dispossessed more and more *sipahi* from their *timars* in order to create the additional revenue-producing *iltizam* needed to pay for the new foot soldiers.⁵⁸

Tax farming was a mixed blessing for the empire. In the immediate term, it proved able to supply the Ottoman government with revenue desperately needed to pay for the reformed and expanded army and thus helped to ease the fiscal crisis. At the same time, however, the short duration of the *iltizam* contracts encouraged the *multezim* to impose crushingly high tax rates and to use brutal methods to collect revenue in order to maximize short-term profits. As such, the system discouraged the economic growth on which future tax-revenue depended.⁵⁹

More importantly, substantial as they were, the tax receipts that the *iltizam* generated proved insufficient by themselves to offset the Sublime Porte's myriad fiscal problems. The central government responded in two ways—one modest and one substantial. First, in 1677, bureaucrats arrived at a way of resolving the budgetary issue created by the asymmetry between the solar and lunar calendar; thereafter, the treasury no longer had to deal with the crisis that arose every eight years when it needed to secure the extra revenue required to pay an additional quarter-year's-worth of salaries. Second, and more significantly, sometime between the start of the Long War and 1620, the government made permanent the *avariz* tax: a special assessment imposed during time of conflict to pay for the extraordinary costs of maintaining an army in the field. Bringing in substantial income for the state, the *avariz* went a long way toward easing the fiscal crisis that had plagued the empire since the late-sixteenth century.⁶⁰

Provincial Decentralization

The fiscal reforms of the seventeenth century did not merely benefit the treasury's bottom line but also fostered a transformation in the structure of provincial

⁵⁸ Imber, 198–201.

⁵⁹ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 126–27.

⁶⁰ Howard, *A History of the Ottoman Empire*, 184; Imber, *The Ottoman Empire, 1300-1650*, 242–43.

government and a related relaxation of the central state's control of the periphery. These changes marked a substantial break with past practice. Keenly aware that centrifugal political forces had grievously weakened earlier empires, the Sublime Porte had traditionally worked to maintain a tight grip over the provinces. For example, it had periodically rotated governors to new positions in order to prevent them from establishing patronage networks with local notables that might permit them to rule autonomously. More importantly, it had maintained close ties with the semi-feudal *sipahis*, whose strength in the provinces acted as a check on local officials and ensured that they could not raise illegal taxes or hire mercenary armies composed of demobilized *sekban*.⁶¹

For two reasons, however, those restraints weakened and then collapsed in the early-seventeenth century. First, growing factionalism in the capital not only diverted attention from the provinces but also impelled officials in the central government to grant concessions to provincial governors in order to win their support against rival cliques. Second, and more importantly, the demise of the *sipahis* meant that the Sublime Porte no longer had a military force in the provinces to curb the ambitions of the governors and local notables. Increasingly, as a result, provincial officials were able to negotiate new arrangements with the central government that accorded them greater autonomy. Initially, the accommodations that they secured were limited ones such as the right to remain in their posts permanently and thus to establish patronage networks. Soon, however, they were able to win more far-reaching concessions including the power to impose new taxes, the legal sanction to create private armies, and, ultimately, the right to transform their territories into permanent, hereditary fiefdoms over which the central government could exert only limited control. By the mid-seventeenth century, these changes had become irreversible. Thereafter, local autonomy rather than central control characterized the administration of the provinces.⁶²

It is important to note, however, that the growing independence of the provinces neither weakened the empire nor constituted a sign of its degeneration. Ties between center and periphery remained strong, and the new system in no way interfered with the empire's ability to function. Indeed, provincial governors did not typically have antagonistic relationships with the court, but instead retained close patronage connections with officials in the capital. They also continued to forward tax revenue to the exchequer and contributed troops from their private armies for service in the empire's wars. In other words, the hierarchical system of provincial administration that had existed in the fifteenth and sixteenth centuries had not degenerated into a weaker, decentralized system as proponents of the Ottoman decline thesis had once argued; instead, it had transformed into a cooperative arrangement—one that that functioned effectively over the next two centuries.⁶³

⁶¹ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 86.

⁶² Ira M. Lapidus, *A History of Islamic Societies*, 2nd ed. (Cambridge: Cambridge University Press, 2002), 277–78.

⁶³ Hathaway and Barbir, *The Arab Lands under Ottoman Rule*, 8–9.

Decentralization in the Capital

With the growing weakness of the sultans, power devolved in the central government as well. As we saw earlier, the succession of child rulers and mentally unbalanced sultans who reigned in the first half of the seventeenth century had created a power vacuum in the court. In response, shifting factions composed of harem eunuchs, queen mothers, Janissaries, vezirs, ‘ulama’, and powerful bureaucrats had assumed a steadily increasing share of power. While these groups often competed with each other—sometimes violently—they increasingly worked cooperatively, and soon came to provide the empire with what the historian Donald Quataert argues was an effective system of “collective leadership.”⁶⁴

What of the sultans? How did they fit into this new structure of shared rule? Some, such as Murad IV and Mehmet IV, participated in the running of the empire, albeit more as first-among-equals than as absolute monarchs on the model of Mehmet II (r. 1451-1481). These leaders were very much the exception, however. For the most part, the sultans of the seventeenth century reigned rather than ruled—serving as unifying symbols that legitimated the collective system rather than as executives who actually directed the administration of the empire.⁶⁵

As we have seen, proponents of the Ottoman decline thesis once argued that the shift from strong, sultanic rule to factional governance during the early-seventeenth century marked the moment that the empire began its inexorable degeneration. According to this school of thought, the early Ottoman state had prospered and expanded under the effective leadership of powerful sultans such as Mehmed II, Selim II and Sulëiman the Magnificent. After 1566, in contrast, it endured the rule of a rotating cast of venal and corrupt courtiers, pashas, and vezirs. More interested in accumulating personal power and wealth than in providing wise leadership, the new ruling class failed to respond effectively to foreign threats, ignored the empire’s gradual economic, diplomatic, and military degeneration, and set the sultanate on a path that led, ineluctably, to its dissolution.⁶⁶

There is a grain of truth to this argument. The early period of collective rule was indeed an anarchic time—one marked by violent riots in the capital, political paralysis in the Sublime Porte, and rebellions in the provinces that, collectively, left the Ottoman government poorly equipped to deal with external challenges. The empire’s enemies were quick to exploit the instability. Most notably, it was the chaos that followed the deposition and execution of Osman II that gave Shah Abbas I the opportunity to seize control of Baghdad in 1624.⁶⁷

⁶⁴ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 34.

⁶⁵ Quataert, 34.

⁶⁶ Lewis, “Some Reflections on the Decline of the Ottoman Empire,” 111–27.

⁶⁷ Finkel, *Osman’s Dream: The History of the Ottoman Empire*, 205–6.

In recent decades, however, scholars who argue that the empire transformed rather than degenerated have persuasively rebutted this interpretation. They contend that the Ottoman decline thesis vastly overstates the role that the early sultans had played in the government, unrealistically assumes that the empire's fortunes were entirely a function of the competence of its top ruler, and, most importantly, ignores the many successes that the Sublime Porte enjoyed after the "time of troubles." Indeed, directly contradicting the Ottoman decline thesis, they maintain that the system of collective rule that took effect in the early-seventeenth century provided capable governance. In particular, they argue that the empire was especially well led during two successive periods: The Sultanate of Women from the 1620s to the 1650s and the Koprulu Era from 1656 to 1703.⁶⁸

The Sultanate of Women

During the Sultanate of Women, the *valide sultans*, or queen mothers, provided strong leadership in what proved to be a peculiarly rocky period for the Ottoman Empire. The preeminent force in the government from the death of Osman II in 1622 until her assassination in 1651, Kosem Sultan (1689-1651), mother of Murad IV and Ibrahim, was the most important of the *valide sultans*. She proved pivotal in seeing the empire through a peculiarly difficult period. Serving as regent from 1623 to 1632, she was instrumental in helping Murad IV to put down a Janissary revolt and to restore central control of the sultanate; later, she played a key role in bringing Ibrahim's disastrous reign to a close.⁶⁹

Turhan Sultan (1627-1683), the mother of Mehmet IV, offered similarly effective direction of the empire during hard times. Serving as regent and co-ruler from 1651 to 1656, she succeeded in maintaining some semblance of order in the capital—no small task given the anarchy that prevailed at the time—and overaw the reconstruction of many of the empire's critical defensive bastions. Undoubtedly, however, her most important action came in 1656. Confronting defeat at Venetian hands, she made the bold decision that year to elevate an obscure and elderly bureaucrat named Koprulu Mehmet Pasha (1575-1661), to the position of grand vezir—a choice that not only demonstrated her wisdom and decision-making abilities but that also launched the highly successful Koprulu Era.⁷⁰

The Koprulu Era

Like the *valide sultans*, the members and dependents of the Koprulu family who controlled the grand vezirate from 1656 to 1703 provided wise and effective rule. That they were able to do so was critically important for the empire, for the Koprulu era

⁶⁸ Hathaway and Barbir, *The Arab Lands under Ottoman Rule*, 8–9.

⁶⁹ Peirce, *The Imperial Harem*, 253–56; Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 75–76.

⁷⁰ Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 75–76; Peirce, *The Imperial Harem*, 253–56.

began during a peculiarly dark time. The issue was the ongoing war with Venice. As we have seen, the Venetians had followed their victory over the Ottoman fleet in 1656 by reinstating their blockade of the Dardanelles Strait; as a result, Istanbul experienced serious food shortages that produced periods of civil unrest among the city's people as well as destabilizing revolts among the soldiers. Upon becoming grand vezir, Mehmet Pasha moved aggressively to restore the situation. He began by organizing a naval campaign in 1657 that succeeded in defeating the Venetian fleet and in reopening the strait to commerce. This victory proved pivotal for Mehmet Pasha. By dramatically enhancing his stature, it gave him the political capital needed to arrest and execute the leaders of the rebellious soldiers and to purge the bureaucracy of officials from rival cliques.⁷¹

He followed those moves with a series of further successes. He put down a rebellious vassal in Transylvania, defeated the Habsburgs, and efficiently and quickly reconstructed the capital after a devastating fire and thereby further boosted his reputation and his grip on power. By the time he passed away in 1661, he had resolved the crises of the 1650s and had left the empire on its most stable footing since Murad IV's death in 1640. Indeed, his direction of the sultanate had been so effective that he was able to compel Mehmet IV to name his son, Koprulu Fazil Ahmed Pasha (1635-1676), as the empire's next grand vezir—an action that institutionalized Koprulu control of the Sublime Porte for the next three decades.⁷²

Strong leadership continued under Fazil Ahmed Pasha and his successor as grand vezir, his adopted brother, Kara Mustafa Pasha (1634-1683). Indeed, the two men built so effectively on Mehmet Pasha's successes that the empire was able to wage wars of conquest for the first time in nearly a century. Fazil Ahmed Pasha began the Koprulu era of expansion by finally completing the conquest of Crete in 1669 and by taking Podolia from the Polish-Lithuanian Commonwealth in 1672. Kara Mustafa continued the campaign of expansion when he became grand vezir in 1676. Most notably, he secured suzerainty over the semi-nomadic Cossacks of Right Bank Ukraine—thereby shoring up the Ottoman position along the border with the increasingly aggressive Russian Empire. These victories were substantial and earned the Koprulu great acclaim. Indeed, as contemporaries approvingly noted, the empire had never before controlled as much territory as it did under Kara Mustafa.⁷³

The grand vezirs of the Koprulu Era enjoyed similar success implementing much-needed domestic reforms. They restored discipline to the military, struck absentee troops from the rolls, and began to offset the declining efficacy of the Janissaries by making greater use of Albanian, Circassian, and, especially, Tatar auxiliaries. Most importantly, successive Koprulu vezirs implemented sweeping fiscal reforms that put the treasury on a solid footing for the first time since the 1570s. Central to that effort was the issuance in 1688 of a new silver coin, the *kurus*, to replace the now completely

⁷¹ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 253–64.

⁷² Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 76, 58.

⁷³ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 273–83.

depreciated *akce*. Stable in value, the *kurus* quickly brought an end to the inflation that had plagued the empire for much of the prior century and thereby helped spark what the historian Douglas Howard characterizes as “a long period of prosperity in Ottoman society.”⁷⁴

Malikane

The Koprulu also oversaw a far-reaching reform of the tax system in the 1690s that would have enormous ramifications both for the central government’s fiscal situation and for the administration of the provinces. Mounting dissatisfaction with the *iltizam* form of tax collection prompted this move. By the end of the seventeenth century, treasury officials had become well aware that the three-year terms of the *iltizam* contracts were encouraging the *multezim* to seek short-term gain at the expense of the empire’s long-term economic well-being. They also understood that the *iltizam* system had never provided the state with the consistent tax payments it needed in peacetime let alone during its frequent foreign conflicts. Such was the case during the long and costly War of the Holy League in the closing decades of the seventeenth century, which we shall explore presently. Indeed, it was the need for additional funds to pay for that conflict that would serve as the catalyst for a complete overhaul of the tax system.⁷⁵

The centerpiece of that reform program was a comprehensive change to the system of tax farming that went into effect in 1695. That year, the treasury began to replace the *iltizam* system with a new form of tax farming called *malikane* that it hoped would deliver the revenue needed to finance the war. The *malikane* tax farms were superficially similar to the *iltizam* ones that they replaced. They required the winning bidder to pay an initial lump sum and thereafter to forward a share of the annual tax receipts they collected to the treasury. The *malikane* system differed from the *iltizam* arrangement in one significant way, however: The leases on them were for the lifetime of the holder rather than for three years.⁷⁶

That change made little difference in the immediate term. In the long run, however, the *malikane* not only provided the steady collection of tax revenue the empire needed but also indirectly benefitted the exchequer by helping to boost the broader Ottoman economy. The key here was the duration of the *malikane* contracts. Lasting for the life of the holder, the agreements encouraged the *multezim* to make investments in infrastructure and to set taxes at rates designed to encourage long-term economic growth rather than to generate quick profits. As a consequence, the new system helped to establish the conditions that would see the empire experience a steady agricultural

⁷⁴ Lapidus, *A History of Islamic Societies*, 277; Howard, *A History of the Ottoman Empire*, 170–72, 186.

⁷⁵ Finkel, *Osman’s Dream: The History of the Ottoman Empire*, 326.

⁷⁶ Howard, *A History of the Ottoman Empire*, 184–85.

and commercial expansion over the course of the eighteenth century and a concomitant rise in tax receipts.⁷⁷

Advocates of the Ottoman decline thesis once maintained that the shift to lifetime tax farming marked another critical inflection point in the empire's decline. Characterizing the adoption of the *malikane* system as a desperate grab for short-term cash, they argued that it not only came at the expense of long-term tax receipts but also had the unintended consequence of accelerating the ongoing process of decentralization that allowed the more distant parts of the empire to become fiscally and politically autonomous. To proponents of the Ottoman decline thesis, in other words, it incurred a double cost to the Sublime Porte—failing to provide the treasury with adequate revenue even as it weakened the government's control of peripheral provinces.⁷⁸

More recently, scholars who argue that the empire experienced transformation rather than decline have persuasively challenged this interpretation. While acknowledging the centrifugal political implications of the new arrangement, they contend that by maintaining control of the bidding process and by regulating the collection of revenue, officials in the capital ensured that the state retained a significant degree of authority over the provinces. In addition, they show that the attraction of lifetime tax farms—they were quite lucrative for the lease holders—and the Sublime Porte's continued monopoly over the bidding process combined to strengthen patronage links between local notables and officials in Istanbul. Coupled with the provincial governors' desire for formal recognition—only the sultan could grant them the legitimacy they needed—the *malikane* system thus ensured that center and periphery would remain bound together for the next century and that the provincial notables would continue both to forward tax revenue to the treasury and to provide troops for the defense of the empire. In other words, while the new system of tax collection and the new relationship between the Sublime Porte and the provinces differed substantially from past practice, they did not weaken the power of the central government as scholars once argued.⁷⁹

Wars of Contraction?

The War of the Holy League, 1683-1699

Koprulu rule was not entirely successful, however. Indeed, the family's ability to restore order after the "time of troubles" must be balanced against the disastrous war into which it dragged the empire in 1683. That conflict was a product of Kara Mustafa's desire to continue the renewed drive to expand the empire. Seeking to take advantage

⁷⁷ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 325–26.

⁷⁸ Lewis, "Some Reflections on the Decline of the Ottoman Empire," 111–27.

⁷⁹ Albert Hourani, *A History of the Arab Peoples* (Cambridge, Mass: Belknap Press of Harvard University Press, 2002), 252; Howard, *A History of the Ottoman Empire*, 184–86; Hathaway and Barbir, *The Arab Lands under Ottoman Rule*, 8–9.

of a revolt by Protestant dissidents in the Habsburg-controlled part of Hungary, he invaded Austria in 1683. He was not content merely to take a few border fortresses as the Ottomans might have done in the recent past, however. Instead, seeking greater glory, he marched his army deep into Austrian territory and laid siege to Emperor Leopold I's (r. 1658-1705) capital, the city of Vienna. Vastly outnumbering the defenders, the Ottomans appeared poised in the late summer of 1683 to achieve the empire's greatest victory since the reign of Sulëiman the Magnificent. Instead, the campaign ended in disaster. On September 12, a Polish-Habsburg relief force under the command of Polish King John Sobieski (r. 1674-1696) decisively routed the Ottoman army and broke the siege. The battle was a major setback for the sultan's military and a fatal one for Kara Mustafa: blaming his grand vezir for the defeat, Mehmet IV had him executed on Christmas Day, 1683.⁸⁰

The situation rapidly deteriorated from that point. Following the victory at Vienna, the Habsburg Empire, the Polish-Lithuanian Commonwealth, Venice, and Russia formed a powerful alliance against the Ottomans called the Holy League. Bereft of Kara Mustafa's effective leadership and gripped by political instability—thirteen men served as grand vezir between 1683 and the conflict's conclusion in 1699—the empire found itself outnumbered and outmatched. Habsburg forces repeatedly pummeled the Ottoman army and won a victory of substantial strategic and symbolic significance at Mohacs in 1687, the site of Sulëiman the Magnificent's greatest triumph. Concurrently, Venice took possession of the Morea in Greece and Russian troops seized the fortress of Azov at the mouth of the Don River—the linchpin of the empire's defensive line north of the Black Sea. The loss of that bastion was a serious blow to the Ottomans and made clear just how tenuous the empire's position had become by the late 1690s. Indeed, the fall of the strategic fortress was so shocking that it finally compelled the Sublime Porte to sue for peace.⁸¹

The Treaty of Karlowitz, 1699

The Treaty of Karlowitz that emerged out of the subsequent negotiations made clear the extent of the Ottoman Empire's defeat. Based on the principle of *uti possidetis*—meaning that each party retained the territory that it controlled at the conclusion of the fighting—the agreement cost the sultanate substantial amounts of territory. By its terms, Istanbul ceded Hungary and Transylvania to the Habsburg Empire, Podolia to the Polish-Lithuanian Commonwealth, and the Morea and the Dalmatian coast to Venice. A separate treaty signed with Russia in 1700 recognized the loss of the fortress of Azov and—in a stipulation pregnant with symbolic significance—acknowledged that the Russia tsar, once clearly the Ottoman sultan's diplomatic subordinate, was now of equivalent rank.⁸²

⁸⁰ Marie-Janine Calic, *The Great Cauldron: A History of Southeastern Europe*, trans. Elizabeth Janik (Cambridge, MA: Harvard University Press, 2019), 122–23.

⁸¹ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 289–318.

⁸² Streusand, *Islamic Gunpowder Empires: Ottomans, Safavids, and Mughals*, 60–61.

Proponents of the Ottoman decline thesis once cast the War of the Holy League and the treaty of Karlowitz as the first clear, outward manifestations of the empire's accelerating internal decay. They argued that a century of weak leadership, venal corruption, military indiscipline, and provincial decentralization had weakened the empire and left its army incapable of fighting off the militaries of its increasingly dynamic and sophisticated European rivals. As a result, Ottoman forces endured a string of defeats in the war that ended only when the Sublime Porte accepted a peace agreement that cost the empire most of its territory in Central Europe.⁸³

In recent years, historians critical of the decline paradigm have persuasively countered this interpretation of the conflict. While acknowledging the extent of the Ottoman defeat, they dispute the view that the empire lost the war as a result of a broader process of cultural, political, economic, and military degeneration. Instead, they maintain that the Ottoman army was every bit as effective as those of its enemies and contend that it fought well during the conflict. The problem was not one of quality, they continue, but instead of quantity. At war with a broad alliance of enemies on a series of fronts that stretched from Hungary in the west to the Don River in the east, the Ottoman army found itself overwhelmed and outgunned by a numerically superior coalition. In other words, these historians conclude that the empire lost the war not because of internal decline but instead because its soldiers were badly outnumbered.⁸⁴

Continued Military Success

Those scholars also point to the continued effectiveness of Ottoman military forces in the decades following the War of the Holy League as evidence that the sultan's military had not declined relative to those of its main rivals. For example, they note that Ottoman forces succeeded in ending the Fourth Russo-Ottoman War in 1711 by surrounding and compelling the surrender of a Russian army under the command of Tsar Peter the Great (r. 1682-1725)—a victory that briefly returned the critical fortress of Azov to Ottoman control. Led by the Crimean Tatars—who had emerged as the sultan's best troops—the Ottoman military likewise performed competently in a conflict the empire fought with Russia and Austria between 1736 and 1739. While the sultan's troops had at best mixed success against the tsar's armies in that war, it routinely battered Habsburg forces in Central Europe. Thus, the wars of the early-eighteenth century made clear that while the Ottoman army had long since lost the commanding edge that it had enjoyed in the days of Selim II and Sulëiman the Magnificent, it nonetheless showed no signs of decline or degeneration and remained comparable in capability to the militaries of its Habsburg and Russian rivals through the middle of the eighteenth century.⁸⁵

⁸³ For example, see Goodwin, *Lords of the Horizons*, 228–36.

⁸⁴ Gabor Agostono, "Firearms and Military Adaptation: The Ottomans and the European Military Revolution, 1450-1800," 200–202.

⁸⁵ Aksan, *Ottoman Wars, 1700-1870*, 83–118.

Conclusion

As the ongoing success of its military demonstrated, the Ottoman state had—after a difficult start—succeeded in addressing the challenges of the “time of troubles.” It had neither degenerated as scholars once claimed nor clung blindly to the institutions and practices that had proven so successful in its rise. Instead, it had undertaken a sweeping reform program that, while far from perfect, nonetheless resolved many of the empire’s problems and ensured that it continued to thrive well into the eighteenth century.

After that point, however, the Ottoman Empire found itself in an astonishingly different and far-more dangerous world. With the onset of the Commercial and Industrial Revolutions, the Western states that the Ottomans had once dominated were undergoing an unprecedented economic transformation that was rapidly rendering them vastly stronger than the empire. By the early-nineteenth century, the gap in power between the sultanate and the European states had become so great and so dangerous that the Sublime Porte felt compelled once again to embark on a far-reaching reform program. This time, however, it was not acting alone. Throughout the empire, officials and subjects alike debated how Muslim states could best respond to the growing threat emanating from Europe. It is to the challenge of the West and to the Ottoman Empire’s manifold responses that we shall now turn.

Chapter Eight: Western Intrusions, Eastern Responses, 1768-1878

In the late-eighteenth century, the Ottoman Empire began to face a new set of challenges far greater than any it had yet confronted. The successful transformation that the sultanate had undergone in the seventeenth century had provided it with an imperfect, if-nonetheless effective system of government, but did not endow it with the economic structures and social practices needed to keep pace with the commercially dynamic states of Europe. Starting in 1768, as a result, the Ottomans found themselves increasingly unable to fend off the aggressive Western powers or to prevent the success of national liberation movements in the Balkans; above all, they could no longer keep the predatory Russian Empire from periodically seizing large swaths of strategic territory. The result was what Western leaders euphemistically referred to as the “Eastern Question”: the problem that Ottoman weakness posed to the balance of power on which European peace rested. For the next century and a half, the Eastern Question was a front-burner issue—a constant menace to European security and stability that occupied Western diplomats who were intent on preventing volatility in the Ottoman Empire from sparking a general war.

While the imbalance in power between the West and the empire compelled European leaders to engage in seemingly constant crisis management, it prompted Muslims to undertake a searching reappraisal aimed at understanding what had gone wrong and at determining how they could best respond to the new challenges that the empire faced. Bureaucrats, intellectuals, provincial governors, and religious scholars alike offered a number of competing explanations for the empire’s weakness and prescribed a variety of often-conflicting reform proposals aimed at redressing the yawning gap in power that had opened between the Muslim world and the European states. Those efforts ultimately failed to arrest the West’s growing dominance or to prevent the loss of further territory; they did, however, succeed in creating a legacy of reform that would powerfully shape political and social views well into the twentieth century.

The European Powers and the Ottoman Empire

The Rise of the West

The source of the Eastern Question and the impetus for the reform efforts that emerged in the nineteenth century was the menace posed by the rise of the West. Between 1500 and 1800, Europe underwent a series of significant changes that opened a widening economic and military gap between it and other parts of the world. Its transformation began with the Commercial Revolution: the emergence of new trade routes, industries, and financial networks in the years following the discovery and exploitation of the New World that resulted in large parts of Europe shifting to a proto-capitalist economic system. Concurrently, the West underwent the Scientific Revolution. This new method of studying the material world and understanding humanity’s relationship to it broke with millennia of received wisdom and made new and vastly

more sophisticated technological breakthroughs possible. Most importantly, thanks to the Scientific and Commercial Revolutions, Europe entered the Industrial Revolution in the eighteenth century, which produced a period of sustained economic growth vastly greater than had been the case at any point in the past.¹

The West also drew strength from the emergence of two revolutionary new ideologies in the seventeenth and eighteenth centuries: liberalism and nationalism. These new belief systems constituted a direct challenge to longstanding political theories and to the the dynastic system that had predominated in Europe for centuries. Liberalism called for the replacement of hereditary privilege, the divine right of kings, and arbitrary rule with written constitutions, the consent of the governed, civil rights, and the rule of law. Meanwhile, the related ideology of nationalism demanded that culturally homogenous groups of people—those who shared a language and history—should no longer be incorporated into dynastic empires overseen by unchecked kings but should instead rule themselves as independent, sovereign states.

On the surface, nationalism and liberalism seemed to impose substantial new constraints on the authority of Western rulers. After all, those ideologies explicitly challenged the monarchs' claims to absolute rule and called for an end to their monopoly over foreign relations. Paradoxically, however, nationalism and liberalism did not weaken central governments but instead enhanced their power. Liberalism made this contradiction clear from the very start. As political theorists and government ministers slowly realized, the British monarchy's acceptance of Parliament's power to set tax rates and to have a voice in foreign affairs had resulted not in the legislature curbing the executive's conduct of foreign relations but instead in it enthusiastically funding and supporting the British state's many expensive wars. Nationalism similarly strengthened the power of central governments, at least those that ruled culturally homogenous populations. Most notably, the revolutionary French state's use of nationalist appeals in the *Levée en Masse* 1793 had permitted it to quickly raise an army of one-and-a-half million conscripts to defend the country from Britain and Habsburg Austria—a mobilization on a scale that would have been inconceivable only a few years earlier.²

The Russian Threat

Ironically, the most dangerous threat to the Ottomans, the Russian Empire, similarly lagged behind Western Europe. A staunch opponent of liberalism to the bitter end, it did not begin to industrialize until the late-nineteenth century and thus did not possess a dynamic economy similar to those of the Western European states. Nonetheless, thanks to Peter the Great's (r. 1682-1725) military reforms and to its participation in both the War of the Austrian Succession (1740-1748) and the Seven

¹ Tim Blanning, *The Pursuit of Glory: The Five Revolutions That Made Modern Europe: 1648-1815* (London: Penguin Books, 2008), 93–141.

² William Doyle, *The French Revolution: A Very Short Introduction* (Oxford: Oxford University Press, 2001), 56.

Years' War (1756-1763), it fielded a fearsome military—one that was equipped with modern Western weapons and drilled in the latest tactics. Making matters worse for the Sublime Porte, St. Petersburg had a powerful desire to expand at Ottoman expense. This aspiration was, at root, ideological. As far back as Ivan III (r. 1462-1505), Russian monarchs had described their empire as the “Third Rome”: the rightful heir to the Byzantine Empire and, importantly, its former territory. It was a claim that Russian rulers took very seriously. Indeed, Ivan III had lent legitimacy to it not only by marrying a Byzantine princess, but also by styling himself tsar, or caesar, in imitation of the Byzantine emperors of old.³

Thereafter, successive tsars pressed Russia's claim to be the heir of the Byzantine Empire and to be entitled, by right, to its territory. For a long time, of course, that belief remained no more than an aspiration. Faced with a still-powerful Ottoman state, successive Russian monarchs found themselves unable to realize their dream of succeeding the Byzantine Empire. Indeed, it was only under the forbidding and strong-willed Catherine II (r. 1762-1796) that Russia began to transform that claim from ambition into reality. Determined to make real the pretension that the tsars were the legitimate successors to the Byzantine emperors, she pursued a “southern strategy” aimed at securing control of the Ottoman-held warm-water ports on the northern shore of the Black Sea from which Russian forces could subsequently threaten Constantinople and the straits. It was a bold program; astonishingly, it was also one that Catherine II succeeded in achieving.⁴

Ottoman Weakness

Russo-Ottoman War, 1768-1774

She began to pursue her southern strategy with the Russo-Ottoman War of 1768-1774. From the start, the conflict went disastrously for the Ottomans. Having sat out the War of the Austrian Succession and the Seven Years' War, the sultanate had failed to remain abreast of the rapid evolution of European weapons and tactics that had occurred since 1740. As a result, Catherine II's powerful and well-led troops easily thrashed the Ottomans' much-larger armies and handily occupied both the Crimean Peninsula and the strategic Danubian Principalities of Moldavia and Wallachia. Indeed, by the time the war ended in 1774, the Ottoman Empire's heretofore-sophisticated supply system lay in shambles and its army was on the verge of collapse.⁵

³ Tamim Ansary, *Destiny Disrupted: A History of the World Through Islamic Eyes* (New York, NY: PublicAffairs, 2009), 235.

⁴ Virginia H. Aksan, *Ottoman Wars, 1700-1870* (Harlow, England: Routledge, 2007), 129.

⁵ Caroline Finkel, *Osman's Dream: The History of the Ottoman Empire* (New York: Basic Books, 2005), 373–79.

Treaty of Kucuk-Kaynarja, 1774

At first blush, the Treaty of Kucuk-Kaynarja that concluded the war seemed to salvage at the negotiating table what the Ottoman military had lost on the battlefield. The key was Western intervention. Keen to maintain the balance of power, European states led by the Habsburg Empire compelled Catherine II to give up the main territorial prizes—Wallachia and Moldavia—that her armies had won. Indeed, European diplomats managed to limit Russia's tangible gains to relatively modest ones: a large indemnity, a few key fortifications and ports on the Black Sea coast, and some border regions in the north Caucasus. Thus, to the casual observer, the Ottomans appeared to have dodged a bullet.⁶

Such an assessment is inaccurate. The agreement was in fact a disastrous one for the Sublime Porte that fully reflected the scale of the empire's military defeats from 1768-1774. Two provisions were especially damaging. First, the treaty granted Russia the right to protect Orthodox Christians in the sultan's domain; seemingly innocuous at the time, this stipulation would serve as the justification that St. Petersburg would use to intervene in Ottoman affairs. Second, the agreement formally ended the sultan's sovereignty over the Crimean Khanate and made the Tatars independent. This concession would prove to be a painful one for the Sublime Porte. It not only deprived the military of its best troops but also marked the first time that the empire had been compelled to cede territory with a predominantly Muslim population. Worse, as was widely understood at the time, Tatar independence was only a temporary condition destined to lead—as it did in 1783—to Russian annexation of the strategic Crimean Peninsula.⁷

The war of 1768-1774 and the treaty that ended it thus constituted a critical turning point for the empire. As a result of them, the Sublime Porte had lost its most dependable soldiers, its monopoly over the Black Sea, and its freedom from foreign interference in its internal affairs. Perhaps worse, that conflict—and another, only slightly less-disastrous war with Russia from 1787-1792—had revealed that the sultan's military no longer stood on an even footing with the armies of the European powers.⁸

Napoleon in Egypt, 1798-1801

The Ottoman Empire's relative weakness was certainly clear to the ambitious French general, Napoleon Bonaparte (r. 1799-1814, 1815). Seeking to exploit the qualitative gap that had opened between the sultan's military and those of the European states, the young general had begun plotting in the mid 1790s to seize territory in the Middle East from which he could interdict commerce between Britain and its economically vital colony of India. Those plans came to fruition in 1798 when he invaded and occupied Ottoman Egypt. Initially, the operation was a cakewalk for his

⁶ Aksan, *Ottoman Wars, 1700-1870*, 158–60.

⁷ Aksan, 158–60.

⁸ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 377–87.

disciplined soldiers. They easily took control of the province—demonstrating in the process that the divide between the military capabilities of the Western European states and those of the sultan’s army had widened substantially since the Ottoman Empire’s recent defeat at Russian hands. Most notably, at the Battle of the Pyramids outside Cairo in July 1798, Napoleon’s soldiers used their superior artillery, better tactics, and bayonet-equipped guns to decisively rout an Ottoman army in an engagement that came at the astonishingly modest cost of only thirty French lives.⁹

Ultimately, however, Napoleon’s invasion did not prove successful. The destruction of his fleet by a British naval force in the Battle of the Nile in August 1798 rendered his position in Egypt untenable and thus dashed his grandiose plan to bring England to its knees by severing its communications with India. Eager to shed himself of a losing venture, Napoleon followed the defeat by passing command of his army in Egypt to one of his subordinates and returning to Paris. Eventually, an Anglo-Ottoman army—in which the disciplined British troops did the heavy lifting—defeated what remained of his expeditionary force at Alexandria in 1801 and compelled Paris to abandon the Egyptian operation.¹⁰

The “Eastern Question”

The ease with which the French took Egypt confirmed what the Russian victories in the late-eighteenth century had suggested: that the Ottoman Empire was no longer a first-rate military power. Accordingly, attuned as they were to the strengths and weaknesses of the world’s states, European leaders thereafter factored the increasingly vast gap in military and economic power that had emerged between the Ottoman Empire and the West into their diplomatic calculations. Indeed, it was at that time that the reality of Ottoman weakness and the threat it posed to the maintenance of peace in Europe—the Eastern Question—first began to emerge as a central concern in European diplomatic circles.¹¹

Views of the Eastern Question differed. For Russian tsars and their ministers, the problems of the Ottoman Empire smelled of opportunity. They hoped to exploit its weakness in order to replace sultanic rule in the Balkans with dependent, Russian-dominated states. More importantly, they dreamed of finally achieving their long-held goal of taking control of Constantinople and the straits and thereby giving Russia both the prestige of possessing the former Byzantine capital and the tangible benefits of at last enjoying access to uninhibited warm-water ports. In contrast, diplomats from other states viewed the Eastern Question as a serious issue that threatened to undermine the broader, post-Napoleonic European peace. Leaders in Britain, France, and Austria in particular understood that the problems of the Ottoman Empire created a dangerous

⁹ Juan Cole, *Napoleon’s Egypt: Invading the Middle East*, First edition (New York: Palgrave Macmillan, 2007), 3, 13–14, 66–69.

¹⁰ Cole, 108–10, 204.

¹¹ Erik J. Zürcher, *Turkey: A Modern History*, Third Edition (London: I.B. Tauris, 2004), 38–39.

power vacuum; should any state—especially Russia—exploit its weakness to seize territory in the Balkans or take control of the straits, the European balance of power on which peace and security rested would be fatally undermined and the Continent would be plunged into a costly war.¹²

The leaders of those states proved better at perceiving the risks inherent in the Eastern Question than at managing them, however. Divided among themselves and unable to resist taking advantage of the Ottoman Empire, they never found a satisfactory way to resolve the Eastern Question. Instead, left to fester, it increasingly strained relations between the Great Powers and helped to bring about the First World War. That conflict lay in the distant future at the time that French troops left Egypt, however. In the meantime, the centrality of the Eastern Question in European diplomacy waxed and waned—rising when problems arose in the Balkans or when Russia sought to expand at Ottoman expense and falling when other crises took precedence.¹³

Nationalist Revolts

In keeping with this pattern, the Eastern Question remained on the back burner during the cataclysmic Wars of Napoleon (1803-1815) even as it grew substantially more complicated due to the appearance of nationalist revolts in the European part of the Ottoman Empire. The first of these occurred in Serbia. Emerging out of a local conflict among Ottoman officials in 1804 and drawing strength from socioeconomic grievances that had little to do with Serbian identity, the revolt only belatedly acquired nationalist overtones. It ended in 1816, when the Sublime Porte, bowing in part to Russian pressure, granted autonomy to the province.¹⁴

A similar nationalist rebellion broke out in Greece in 1821. In two ways, it was reminiscent of the recently resolved Serbian revolt. First, it too arose out of a conflict within the Ottoman ruling class and only belatedly transformed into a nationalist movement. Second, like the Serbian uprising, it succeeded as a result of foreign involvement—in this case, the intervention of Russia, Britain, and France. Their motives for becoming involved varied. St. Petersburg interceded in hopes of turning Greece into a protectorate; in contrast, London and Paris intervened partly to curb the tsar's ambition and partly to satisfy the demands of their increasingly pro-Greek publics.¹⁵

¹² Ansary, *Destiny Disrupted*, 235–36; Marie-Janine Calic, *The Great Cauldron: A History of Southeastern Europe*, trans. Elizabeth Janik (Cambridge, MA: Harvard University Press, 2019), 229–30.

¹³ Donald Quataert, *The Ottoman Empire, 1700-1922, Second Edition* (Cambridge: Cambridge University Press, 2005), 56.

¹⁴ Zürcher, *Turkey*, 31; Finkel, *Osman's Dream: The History of the Ottoman Empire*, 403–5.

¹⁵ Richard J. Evans, *The Pursuit of Power: Europe 1815-1914* (New York, New York: Penguin Books, 2016), 53–59.

The support of these states proved critical in helping Greece gain independence. In 1824, the powerful ruler of Ottoman Egypt, Pasha Mehmet Ali (r. 1805-1848), had deployed his army and navy against the rebels at the Sublime Porte's request; in short order, his disciplined forces defeated the insurgents and restored sultanic control over Greece. Perversely, however, his intervention proved to be too successful. Faced with the prospect of an outright Greek defeat, the people of France and Britain pressured their governments into joining with Russia to rescue the rebellion. Accordingly, in 1827, an Anglo-Russian-French naval force destroyed the combined Egyptian-Ottoman fleet at Navarino. The battle proved to be a turning point. With the European states now firmly committed to the rebels, the Sublime Porte concluded that victory was no longer possible. In 1829, as a result, it agreed to grant Greece its independence.¹⁶

The Historical Debate: Nationalism

Understandings of nationalism and, especially, the origins of the Greek and Serbian nationalist movements, have changed markedly in recent decades. Prior to World War II, most scholars argued that nationalism was the natural precondition for the nation-state. According to this view, each nation, or ethnically-defined group of people, is a fixed or essential entity, meaning that it has always existed and that its fundamental characteristics—its culture, beliefs, and language—are unchanging. When faced with oppression, a people that does not enjoy self-determination will, because of its shared sense of nationalism, seek independence in order to form a nation-state; in other words, nation precedes state. Thus, according to this perspective, the independence drives that occurred in the Balkans in the nineteenth century were the natural and inevitable product of the innate nationalism of the peoples who lived there. Having suffered under the yoke of oppressive Ottoman rule for centuries and able to look back to a time when their forebears had enjoyed independence, the various peoples of the region took advantage of the empire's growing weakness to rise up and regain their long-lost and much-cherished independence.¹⁷

After 1960, a new generation of historians and political scientists began to challenge this conception of nationalism. They reject outright the idea that nations are essential and timeless. Instead, these scholars argue that each is a cultural construct that originates in a particular time and set of circumstances. They also contend that nationalist campaigns are not typically mass movements. On the contrary, they are almost always the products of small groups of elites who organize them in order to achieve political or economic ends that they could not attain under outside rule. Only after they have successfully achieved independence and established a state do those groups set about creating a sense of national identity—a task they undertake for the purpose of securing the loyalty of the people who live within the new country's borders. In other words, this interpretation contends that a widely held sense of national identity does not precede the creation of the state; instead, it emerges *after* the state's

¹⁶ Richard J. Evans, *The Pursuit of Power: Europe 1815-1914* (New York, New York: Penguin Books, 2016), 53–59.

¹⁷ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 70.

establishment, and does so as a result of a concerted effort on the part of powerful interests to establish a shared sense of membership in what the influential political-scientist Benedict Anderson calls the “imagined political community” of the nation.¹⁸

Based on this conception of nationalism, scholars of the Ottoman Empire like Donald Quataert now dismiss the older idea that surging nationalist sentiments caused by purported Ottoman misrule produced the independence movements that rocked the Balkans after 1800. Drawing on Anderson’s ideas, they argue that through the mid-nineteenth century, nationalism and national identity remained foreign concepts in a region in which occupation and religion rather than ethnicity or even language defined identity. Thus, scholars arguing from this perspective contend that nationalism did not spark the emergence of independence drives in the Balkans; instead, small groups of local elites created and led those movements in order to achieve narrow political and economic goals. Later, after they had won independence from the sultanate, those elites began to promote new, nationalist-based identities in order to secure the loyalty of the people that they ruled. In other words, this interpretation holds that it was only after elites had achieved independence that they developed new national symbols and repurposed historical events into collective, nationalist myths and thus created the “imagined political community” on which a shared sense of national identity rests—most notably when Serbian nationalists transformed the Battle of Kosovo (1389) from an ordinary military defeat into a glorious if doomed, lost-cause struggle against foreign domination.¹⁹

French and British Support

In conjunction with these revolts, renewed Russian aggression put the Eastern Question back on the front burner in the late 1820s. Indeed, St. Petersburg was not shy at that time about acting to achieve its goals vis-à-vis the Ottoman Empire. Taking advantage of the Greek revolt, the tsar had followed the Battle of Navarino by going to war with the now-nearly defenseless empire. Russia handily won the war and secured significant concessions through the Treaty of Adrianople that ended it. Later, after Russian troops protected Istanbul from the army of the rebellious Mehmet Ali in 1833—a stunning case of the fox guarding the hen house—the tsar’s government compelled Sultan Mahmud II (r. 1808-1839) to sign the Treaty of Hunkar Iskelesi, which, while nominally a defensive agreement between equals, was, in reality, a one-sided deal that promised to reduce the Ottoman Empire to little more than a Russian protectorate.²⁰

These actions greatly alarmed the British and French governments. Much of their concern lay rooted in the diplomatic risks that they perceived in tsarist dominance of the Ottoman Empire. Both Paris and London were acutely aware that Russian control of the

¹⁸ Benedict Anderson, *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, Revised edition (London: Verso, 2006), passim; quote from page 6.

¹⁹ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 188–90; Calic, *The Great Cauldron*, 217–18; Quote from Anderson, *Imagined Communities*, 6.

²⁰ Zürcher, *Turkey*, 34–37.

straits or the Balkans would pose a grievous danger to the balance of power in Europe and likely produce a devastating general war among the Great Powers. For that reason alone, they were determined to limit St. Petersburg's influence over the Sublime Porte and, especially, to prevent it from acquiring Constantinople or the empire's territory in the Balkans.²¹

London had a second, more self-interested geostrategic reason to oppose the growth of Russian influence in the Ottoman Empire: the threat that tsarist dominance of the Middle East would pose to the security of its economically vital Indian colony. Concerns about Russian expansion to the south and the implicit danger its growing influence in Central Asia posed to the Subcontinent dominated London's foreign policy in the nineteenth century. The result was the "Great Game": the economic, political, and diplomatic struggle between London and St. Petersburg for influence over the regions that lay between India and Russia. Unsurprisingly, the Ottoman Empire loomed large in Britain's approach to the Great Game. If Russia was able to exploit the Sublime Porte's weakness to gain control of ports on the Persian Gulf, its ships would be in a position to interdict the vital maritime communications with India on which the security and economic might of the British Empire rested. Conversely, if London could preserve the Ottoman Empire intact, it could function as a critical buffer state that could help to contain Russian influence.²²

Paris and London's concerns about Russian influence over the Sublime Porte also stemmed from the substantial economic stake that they had acquired in the Ottoman Empire. Their commercial preeminence in the sultanate was largely a function of the capitulatory treaties that they had signed with the Sublime Porte in the late-sixteenth century. Thanks to the exemption from Ottoman taxes and customs duties that those agreements afforded them, British and French merchants had gradually assumed control of the empire's lucrative international trade. Their economic dominance of the Ottoman Empire grew even more substantial starting in the late 1830s. In a treaty we will discuss in more detail presently, the Sublime Porte agreed in 1838 to transform the Capitulations into free-trade agreements that established import duties at rates so low that the empire could no longer protect domestic industries from European competition. Over the course of the nineteenth century, as a result, the Ottoman Empire gradually became a neo-colony of France and Britain—serving as a source of cheap raw materials and a market for their factory-produced finished goods.²³

Meanwhile, as a result of a provision in the capitulations, Ottoman Christians had begun to purchase certificates called *berats* from European consulates that gave them the same legal status as Western merchants. Originally intended to be a limited

²¹ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 56.

²² Edward Ingram, "Great Britain's Great Game: An Introduction," *The International History Review* 2, no. 2 (1980): 160–71.

²³ Carter V. Findley, *Turkey, Islam, Nationalism, and Modernity* (New Haven: Yale University Press, 2010), 139–41; Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 78, 126–29.

perquisite reserved for a small number of diplomatic translators, the system came to be massively abused in the nineteenth century when hundreds-of-thousands of Ottoman subjects obtained *berats* and the tax advantages that they conferred. As a result, European-affiliated merchants, particularly those whose *berats* linked them to Britain and France, came to dominate the coastal trade in the empire, while European consulates, entitled to oversee the legal affairs of their states' *berat* holders, assumed sovereign powers within the empire.²⁴

Economically, in other words, London and Paris had a great deal at stake in the continued independence of the sultanate. Were St. Petersburg to solidify its dominant position in the empire, it could replace the capitulations with protective tariffs—a move that the Sublime Porte was too weak to make on its own. In turn, those Russian-enforced tariff barriers would likely shut British and French merchants out of the Ottoman Empire and, in so doing, deny their factories both the raw materials they needed for the production of finished goods and the overseas markets necessary for the disposition of their surplus production.

Thus, for a combination of strategic and economic reasons, Britain and France gradually shifted their approach to the Ottoman Empire after the Napoleonic Wars. Determined both to retain their privileged commercial position and to prevent St. Petersburg from upsetting the prevailing diplomatic order, they committed themselves to preserving the empire's territorial integrity and to limiting Russian influence over the Sublime Porte. British Foreign Minister Lord Palmerston (1784-1865) neatly summarized this new approach shortly after the signing of the Treaty of Hunkar Iskelesi. Henceforth, he declared, it was British policy that “[t]he Ottoman Empire was to be preserved, supported, reformed, and strengthened.”²⁵

The interplay between Russian aggression and French and British support for the sultanate would establish a recurring, cyclical pattern in the Ottoman Empire's diplomatic affairs. Each cycle would begin with the sultanate suffering a serious military setback either at the hands of Russian-backed nationalist groups or tsarist armies and then being forced to sign a peace agreement in which it either ceded substantial territory or made significant concessions to St. Petersburg. Alarmed by the threat both to their narrow economic interests and to the broader balance of power, Britain and France would then respond by joining with other concerned states such as the Habsburg Empire to compel Russia to agree to a reduction—but not a full rollback—of the concessions it had secured from Istanbul. In doing so, they would ensure the survival of a smaller, but-still coherent and independent Ottoman state even as they acted to appease popular opinion at home by pressing the Sublime Porte to implement reforms designed to improve the rights of the empire's non-Muslims. This 'six-step forward, five-step back' pattern first began to take effect shortly after the conclusion of the Greek War of Independence when London compelled Russia to accept more-limited

²⁴ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 129; Bernard Lewis, *The Middle East: A Brief History of the Last 2,000 Years* (New York: Scribner, 1995), 295.

²⁵ Quoted in Aksan, *Ottoman Wars, 1700-1870*, 375.

gains in the Treaty of Adrianople than its military success over the Ottomans merited; it would remain in effect for the remainder of the nineteenth century.²⁶

The Crimean War, 1853-1856

There was, however, one major exception to this pattern: the Crimean War of 1853-1856. In that instance, rather than acting to reduce the tsar's gains after his armies had defeated the sultan's forces, Britain and France instead went to war with Russia in order to prevent it from assuming de facto control of the Ottoman Empire. The conflict had complex origins. Nominally, it was the byproduct of France's successful campaign to have the sultan transfer control of the Christian Holy Sites in Palestine from the Orthodox Church to the Catholic Church. In reality, the war stemmed from Tsar Nicholas I's (r. 1825-1855) attempt to use this religious dispute as a pretext for pressuring the Sublime Porte into granting his government sovereign power over the empire's vast Orthodox Christian population—a concession that would have reduced the Ottoman state to little more than a Russian appendage had it gone into effect. News of the tsar's demand hit London and Paris like a bombshell. Already alarmed by Nicholas I's earlier suggestion that London and St. Petersburg should partition the Ottoman Empire—or, as he famously called it, “the sick man of Europe”—the French and British governments responded by joining with the Sublime Porte to declare war on Russia. Fought almost exclusively on the peninsula for which it was named, the Crimean War differed from earlier conflicts between the Ottoman and Russian Empires in that it was the tsar's forces rather than the sultan's that were on the defensive. Costly to all its participants, the struggle finally came to an end in 1856 when the fall of the key port of Sevastopol compelled Nicholas I's successor, Tsar Alexander II (r. 1855-1881), to sue for peace.²⁷

The Treaty of Paris that formally ended the war seemed on the surface to be a significant diplomatic victory for the Ottomans. Thanks to Britain and France's desire to keep the Russians out of the Middle East, it contained terms that appeared to secure the empire from further tsarist depredations. It barred Russia from operating or maintaining naval forces on the Black Sea, compelled it to renounce any right to intervene on behalf of Orthodox Christians in the empire, and even required it to cede a small amount of territory in Bessarabia to the sultanate. More importantly, the treaty gave the Ottoman Empire a new degree of international legitimacy by welcoming it into the Concert of Europe: the diplomatic structure that the victors in the Napoleonic Wars had created to maintain peace and order in nineteenth-century Europe.²⁸

In the longer term, however, the Crimean War did little to arrest the Ottoman Empire's growing relative weakness or to ensure its security. Two problems soon

²⁶ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 56; Zürcher, *Turkey*, 34–37.

²⁷ M. Şükrü Hanioglu, *A Brief History of the Late Ottoman Empire* (Princeton: Princeton University Press, 2008), 79–81; Aksan, *Ottoman Wars, 1700-1870*, 438–45.

²⁸ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 458.

manifested themselves. First, waging the war had required the Ottoman treasury to take out substantial international loans that it soon discovered were beyond its capacity to service. Second, London and Paris lacked the strength to ensure that Russia continued to abide by the terms of the Treaty of Paris—particularly after the Franco-Prussian War of 1870-1871. Indeed, it was France's defeat in that conflict that permitted St. Petersburg to unilaterally abrogate the provision banning it from maintaining naval forces on the Black Sea. Thus, in the long run, the Crimean War not only resulted in the Ottoman Empire becoming mired in a gradually worsening fiscal crisis that left it at the mercy of its foreign creditors but also failed to offer it any real long-term security against the Russian threat.²⁹

Pan-Slavism

Despite its defeat in the Crimean War, meanwhile, Russia had not abandoned its efforts to assume a dominant position in the Balkans or to control the straits. In fact, it remained as committed as ever to exploiting Ottoman weakness to its advantage. With military threats, diplomatic pressure, and proposals to partition the empire having backfired so spectacularly in 1853, however, St. Petersburg felt compelled to pursue a subtler strategy toward the Ottoman Empire after 1856. That new approach centered on using the increasingly popular ideology of Pan-Slavism as a stalking horse for the advancement of Russian interests. A variation of nationalism, Pan-Slavism had emerged in the mid-nineteenth century as a purported ideology of national liberation. It held that the Slavic peoples of the Balkans were groaning under the weight of oppressive and alien Ottoman rule and that they could best achieve their independence through collective action.³⁰

As such, Pan-Slavism appeared to provide St. Petersburg with a powerful ideological tool it could use to advance its designs on Ottoman territory. The key was its indirect approach. Supported by a well-orchestrated Russian propaganda campaign and by substantial covert assistance, it had the potential to nurture nationalist movements and, eventually, independent states among the millions of Slavs who lived in the Ottoman Empire's Balkans territory without directly implicating St. Petersburg. As the most powerful Slavic state and as the self-proclaimed protector of the Slavic people, Russia would then be in a position to gradually assume suzerainty over these new nation-states. In other words, Pan-Slavism promised to achieve through indirect means that which Russia had been unable to secure through intimidation or outright conquest.³¹

²⁹ Norman Rich, *Great Power Diplomacy: 1814-1914* (Boston: McGraw-Hill, 1991), 119–21; Sevkett Pamuk, *A Monetary History of the Ottoman Empire* (Cambridge: Cambridge University Press, 2000), 213–14.

³⁰ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 479–80.

³¹ Evans, *The Pursuit of Power*, 671–73.

The Russo-Ottoman War of 1877-1878

Russia's new approach began to pay dividends in the 1870s. Coming against the backdrop of the Ottoman government's default on its international debt obligations—a topic we shall explore further in a moment—nationalist uprisings broke out among the Slavs of Bosnia-Herzegovina in 1875 and spread to Bulgaria in April 1876. Taking advantage of the crisis, the autonomous principality of Serbia stepped up the pressure on the Sublime Porte by declaring war two months later. The tsar's ministers were gleeful. With much of the Balkans in revolt, Russia appeared to have a golden opportunity to achieve its goal of assuming a dominant position in the region. Still, past experience suggested that the tsar's government needed to move cautiously. Accordingly, wary of repeating the aggressive approach that had sparked the Crimean War, it refrained from taking actions that might risk provoking Paris and London and instead waited to see how the situation developed.³²

It did not need to wait long. Lurid press reports about the “Bulgarian Horrors”—a series of bloody attacks in which ill-disciplined Muslim irregulars massacred thousands of Bulgarian Christians—had poisoned the British and French publics against the Ottoman Empire; as a result, London and Paris were unable to provide Istanbul with diplomatic support during the crisis. Now free to act, Tsar Alexander II responded by throwing caution to the wind and declaring war on the Ottoman Empire in early 1877.³³

Unsurprisingly, the fighting went disastrously for the Ottomans. Despite vigorous early resistance by the sultan's army, superior Russian forces won a series of decisive victories on the frontier and then advanced rapidly into the empire. As a result, the Sublime Porte was compelled to sue for peace in January 1878. Signed just two months later, the Treaty of San Stefano that formally ended the war reflected both Russia's battlefield victories and the Ottoman Empire's diplomatic isolation. By its terms, the sultan ceded substantial territory in the Caucasus and agreed to the creation of a huge, independent Bulgarian state that stretched from the Black Sea to the Aegean coast—thereby dividing the Ottoman Empire's remaining Balkan possessions in half. More importantly, as it was well understood that Bulgaria was going to be a Russian protectorate, the treaty effectively made St. Petersburg the dominant power in the Balkans and, pivotally, gave it unimpeded access to the Mediterranean for the first time.³⁴

The Treaty of Berlin, 1878

In securing those gains, however, Russia had once more badly overplayed its hand. Its sudden expansion into the Balkans posed a serious challenge to the balance of power on which European peace rested while its newfound access to the

³² Zürcher, *Turkey*, 71–75.

³³ Zürcher, 71–75.

³⁴ Douglas A. Howard, *A History of the Ottoman Empire* (Cambridge: Cambridge University Press, 2017), 270–71.

Mediterranean threatened London's vital communications with India via the recently opened Suez Canal. As a consequence, the earlier pattern wherein Britain and the other powers intervened to prevent Russia from dismembering the Ottoman Empire reasserted itself—this time in the form of the Congress of Berlin, an international meeting that took place in the summer of 1878 in the capital of the recently unified German Empire. Dominated by states such as Britain that wanted to rollback Russia's recent achievements, the conference replaced the Treaty of San Stefano with a new accord, the Treaty of Berlin, that undid many of the gains that St. Petersburg had secured earlier that year. It focused largely on Bulgaria. The Treaty of Berlin dramatically reduced Bulgaria in size, made it an autonomous part of the Ottoman Empire rather than an independent state, and—to the great relief of the British government—stripped it of territory along the Aegean Sea.³⁵

These revisions were substantial, but, even with them, the events of 1877 and 1878 constituted a crushing setback for the Sublime Porte. That is, while the Treaty of Berlin may have eliminated the worst aspects of the Treaty of San Stefano, it also contained new and revised provisions that grievously weakened the empire. It affirmed Russia's territorial acquisitions in the Caucasus, allowed the Habsburg Empire to occupy Bosnia-Herzegovina, imposed a stiff indemnity on the Ottoman Empire, and, most importantly, granted independence to the heretofore autonomous principalities of Montenegro, Serbia, and Romania. Meanwhile, in separate agreements, Britain took possession of Cyprus in exchange for a promise to guarantee the empire's Middle Eastern possessions, and France acquired title to the nominally-Ottoman province of Tunisia. These losses—especially those in the Balkans—were enormous. Indeed, while half the population of the empire had lived in its European provinces prior to the crisis, thereafter, the share that lived west of the straits fell to a mere 20 percent.³⁶

The Ottoman Public Debt Administration, 1881

The Congress of Berlin had one other vital consequence: it compelled the sultanate to accept the establishment of the Ottoman Public Debt Administration (OPDA) in 1881 in exchange for a significant restructuring of the Sublime Porte's international financial obligations. Set up in 1875 after the sultanate stopped servicing its foreign loans, the OPDA was a vast bureaucracy managed and staffed not by Ottoman officials but instead by Westerners appointed by the Sublime Porte's international creditors, who, with a mandate to ensure the continued payment of the bonds on which the empire had defaulted, controlled a significant share of the sultanate's tax revenue. Efficiently run, its impact on the empire was decidedly mixed. On the one hand, it succeeded over the next two decades in restoring the Sublime Porte's creditworthiness on international money markets and in reducing substantially its foreign debt. On the other, by granting its international creditors first claim on the empire's tax revenues, it ceded an enormous degree of sovereignty to the West and

³⁵ Evans, *The Pursuit of Power*, 677.

³⁶ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 54; Evans, *The Pursuit of Power*, 677.

turned the empire into a de facto financial colony of the European powers. As important, the OPDA starved the Ottoman state of the revenue it needed to construct desperately needed infrastructure and to support economic development and thus ensured that it remained locked in a position of dependency in relation to the industrial West.³⁷

The Eastern Response

The challenge that the increasingly powerful Western states posed to the Ottoman Empire shocked the sultan's subjects and sparked a great deal of soul searching both within the empire and in the broader Muslim world. The power of the Russian army and, especially, the discipline, tactical skill, and advanced weapons of the French and British troops that had fought in Egypt during the Napoleonic Era had made abundantly clear to even the most insular of observers that the Ottoman Empire and the broader Muslim world had fallen distressingly far behind the West in technological and military terms. It was a sobering realization—one that raised a number of troubling questions. How had the heretofore weak and impoverished West overtaken the once-mighty Ottoman Empire? How could the march of history—hitherto understood to be unfolding according to a divinely-ordained plan in which the Muslim world steadily outpaced other cultures—have gone so distressingly off course? Most pertinently, how could Muslims best respond to the challenge that the increasingly powerful and aggressive West posed?

These questions produced an enormous outpouring of thought, debate, and action over the course of the nineteenth century. Indeed, a wide variety of movements—some top-down, others populist—would emerge at that time seeking to provide answers to these questions and a path forward for the empire and the Muslim world. The conclusions that they drew were varied, often contradictory, and, ultimately, unsuccessful in arresting the sultanate's continued decline relative to the West. Nonetheless, by establishing patterns that later political and social movements would adopt, the ideas developed at that time would ultimately prove to be of enormous long-term significance for the development of the Middle East.

Westernizing Reforms in the Ottoman Empire: The Tanzimat Era, 1826-1876

Objectives

The most important of the reform efforts in the nineteenth century was the Tanzimat program that occurred under the direction of the Ottoman state. Meaning literally “reorganization,” the Tanzimat was a top-down initiative designed to westernize the empire through the adoption of European economic, social, and political practices. Though ultimately unsuccessful in staving off the West, the Tanzimat reforms would

³⁷ Pamuk, *A Monetary History of the Ottoman Empire*, 213–16.

nonetheless powerfully influence subsequent reform efforts in the Middle East and were thus of enormous consequence for the region's development.³⁸

The Tanzimat program had antecedents in an earlier reform effort launched in response to the crushing defeats the Ottoman military had suffered at the hands of Russia in 1768-1774 and in 1787-1792. Alarmed by these reverses, reform-minded officials in the late-eighteenth and early-nineteenth centuries determined that the empire had to begin fielding modern armies organized and equipped along European lines if it wished to avoid further territorial losses. As the reformers soon realized, however, doing so was not simply a matter of changing the armed forces' composition or its equipment. Instead, the sultanate could only maintain a modern military if it had a sophisticated and growing economy—one able both to provide soldiers with advanced weapons and to sustain the complex logistical systems on which contemporary warfare depended.³⁹

Increasingly, they also grasped that the modernization of the empire's military and economic superstructure would fail absent a more fundamental, root-and-branch reform of Ottoman society. Such was the lesson that they drew from Western Europe's recent history. It showed that the extension of civil rights and the institutionalization of social and political equality—at least for men—were necessary preconditions for a rapidly growing economy; after all, the countries that had extended such rights to their people had also been the first to industrialize.⁴⁰

Those legal guarantees were also prerequisites for the creation of a mass military composed of motivated soldiers. This last point was the paradoxical lesson of the *Levée en Masse*. By guaranteeing rights to its citizens and thus ceding power to them, the government of France had secured the loyalty of the French people and ideologically bound them to the survival of the republican state; as a result, it had been able to raise and field the huge armies of highly motivated soldiers needed to defend the revolution in the mid 1790s. The success of the *Levée en Masse* dramatically shaped the thinking of Ottoman reformers. As the historian Donald Quataert writes, in their eyes, “[t]he lesson was clear: universal conscription meant vastly enhanced military and political strength. But, to render such conscription palatable, the state had to grant universal rights (to males).”⁴¹

Instituting universal rights would prove very tricky in the Ottoman Empire, however, because such a legal arrangement ran counter to the unequal group—or corporate—structure of its society. It did so in two ways. First, according to the prevailing Ottoman system, the sultan's subjects enjoyed varying rights and privileges depending on the group to which they belonged. The most fundamental distinction was religious: Muslims enjoyed greater status than did non-Muslims. Even among the sultan's Muslim subjects, however, rights differed greatly. For example, while all

³⁸ Howard, *A History of the Ottoman Empire*, 230–33.

³⁹ Howard, 230–33.

⁴⁰ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 67.

⁴¹ Quataert, 67.

Muslims paid lower taxes than did Christian and Jewish subjects, some, most notably the Janissaries and certain guild members, were exempt from paying taxes altogether. Second, individuals in this system did not typically interact directly with the central state; instead, intermediary groups—be they *millet*s, guilds, religious orders, etc.—mediated between their members and the central government. For example, a Greek Christian would pay taxes to a local tax farmer and would abide by a legal system administered and enforced by the Orthodox Church. The individual in question would thus have no direct dealings with the central government; instead, the *multezim* to whom that person remitted taxes and the Greek Orthodox ecclesiastical officials who oversaw the legal system under which he or she lived would indirectly connect them to the Sublime Porte.⁴²

Since these intermediary groups accorded their members disparate privileges, instituting the universal rights that reformers hoped would bind people to the state required, as a first step, that the Sublime Porte abolish those bodies or eliminate their legal standing. Once it had done so, the state could then build direct relationships with its people—particularly with the sultan’s heretofore socially and legally inferior non-Muslim subjects—and grant them the civil liberties that the reformers believed would ensure their loyalty. Eliminating the empire’s many corporate bodies would not be an easy task, however. Those groups were deeply embedded in the structure of Ottoman society and had leaders that would staunchly oppose any change that threatened to reduce their prerogatives.⁴³

Military Reforms—the “Auspicious Incident,” 1826

One group in particular promised to bitterly resist the reform efforts: the Janissaries. A serious problem for the Sublime Porte for some time, the Janissaries had become a millstone around the neck of the Ottoman Empire by the turn of the nineteenth century. They were not only militarily ineffective, but, worse, they had fueled instability by deposing sultans and by using their location adjoining the palace to repeatedly block any meaningful efforts to reform the government or the military. Indeed, they would play a leading part in crushing an effort to create a new, modern corps called the *Nizam-ı Cedid*, or New Order Army, that reform-minded officials hoped would give the empire a force that could fight European armies on an equal footing.⁴⁴

Launched in the 1790s under the direction of Sultan Selim III (r. 1789-1807), the New Order Army was an experimental force that fought using contemporary European weapons, tactics, and discipline. Though small, it quickly proved its worth in battle against Napoleon’s elite troops in 1799. In fact, Selim III and his ministers were so pleased with its performance against the French that they ordered the New Order Army’s expansion. Thus, to the delight of the reformers, the empire at last seemed on track to modernize its military. Unfortunately for those officials, however, the Janissaries

⁴² Quataert, 63–65.

⁴³ Quataert, 45–46, 136–37.

⁴⁴ Hanioglu, *A Brief History of the Late Ottoman Empire*, 44.

did not share their enthusiasm for the *Nizam-I Cedid*. They had grudgingly tolerated the New Order Army when it had been a small, experimental force, but they feared that its growth would cost them their dominant political position and thus were determined to prevent its expansion. Accordingly, in conjunction with other conservative forces, the Janissaries rose up in rebellion in 1807. They deposed and killed Selim III and followed by compelling the new sultan, Mustafa IV (r. 1807-1808), to disband the New Order Army.⁴⁵

The successful uprising demoralized the reformers. Having seen firsthand what Western soldiers were capable of, they believed that the empire could only stave off its enemies if it too possessed a modern army such as the *Nizam-I Cedid*. In their eyes, therefore, the Janissary revolt had not merely set the Ottoman military back for decades but had also enormously complicated the Sublime Porte's efforts to navigate a rapidly worsening diplomatic environment. The uprising did come with a silver lining, however: it clarified the initial step that the reformers needed to take if they wished to achieve a broad restructuring of the Ottoman government. That is, the events surrounding Selim III's deposition demonstrated that the first move in any consequential reform effort would have to be the elimination of the Janissary corps; until it had been eradicated, no meaningful restructuring of the empire was possible.⁴⁶

Its destruction finally came to pass in the mid 1820s as a result of careful planning on the part of Sultan Mahmud II (r. 1808-1839) and his advisors. Aware that the qualitative gap between the armies of the Ottoman Empire and those of the Western powers was continuing to widen, Mahmud II had for some time hoped to restart Selim III's military reform program. Recalling what had happened to his predecessor, however, he grasped that he needed to move very carefully lest he incite the Janissaries and their supporters. Accordingly, the new military modernization initiative that he announced in 1826 called not for the creation of a new, competing force like the *Nizam-I Cedid* but instead for the gradual restructuring of the corps itself along European lines. In other words, it seemed to be a compromise proposal that promised to finally provide the empire with a modern army while simultaneously permitting the Janissaries to retain their privileged political position. Even this moderate proposal did not sit well with the corps, however. Complaining that the new training regimen was contrary to their traditional methods, the Janissaries responded by once again rising up in revolt. It seemed like an inauspicious beginning to the reform program. Indeed, to the casual observer, it appeared that Mahmud II's effort to modernize the military had begun unravelling even before it started.⁴⁷

The sultan and his advisors were unfazed by the Janissaries' latest rebellion, however. Aware that the soldiers were likely to resist the restructuring program, the reformers had quietly laid the groundwork for a military operation against them should

⁴⁵ Ali Yaycioglu, *Partners of the Empire: The Crisis of the Ottoman Order in the Age of Revolutions* (Stanford: Stanford University Press, 2016), 38–64.

⁴⁶ Yaycioglu, 38–64.

⁴⁷ Aksan, *Ottoman Wars, 1700-1870*, 318–22.

the need arise. Well in advance, they had rallied key groups and institutions to the sultan's cause in order to ensure that he had support for a move against the corps; more importantly, before announcing the reform proposal, the sultans' advisors had deployed to the palace a substantial force of dependable soldiers that had been secretly trained and equipped with modern European weapons. Then, using the revolt as a pretext, the reformers ordered the loyal soldiers to mount a sudden assault on the corps' barracks. It was an enormous success. Making liberal use of modern cannons, the sultan's troops killed 6,000 Janissaries in a mere twenty-one minutes and reduced their barracks—the physical manifestation of the Janissaries' power—to a smoldering ruin. By nightfall, the “Auspicious Incident,” as the assault soon came to be known, had completely eradicated the corps. The road to reform now lay open.⁴⁸

The reformers moved quickly to take advantage. Unsurprisingly, they began by aggressively pushing forward with their military modernization scheme. Over the next few decades, they dramatically increase the size of the central government's army and rebuilt it along European lines. As a result, Ottoman soldiers were soon wearing Western-style uniforms, training in accordance with the most up-to-date military doctrines, and going into battle equipped with modern weapons.⁴⁹

Mahmud II and his advisors had only gotten started, moreover. Determined to push forward with the more fundamental changes that they believed the empire needed, they followed the destruction of the Janissaries by launching a series of social, economic, and political reforms. These included greater central government control over the provinces, a census conducted in 1830-1831 that permitted the treasury to assert control over the collection of taxes, and the reorganization of the bureaucracy along modern, European lines through the creation of new departments including the ministries of finance, foreign relations, and interior.⁵⁰

They also instituted far-reaching changes to the laws that governed how people could dress—a move that constituted the first step in the elimination of social distinctions between groups. Traditionally, the government had enforced a complex series of sumptuary laws that visibly differentiated people according to their rank and status. These required the sultan's subjects to wear clothing that indicated their religious affiliation and class status and included a series of tortuously complex rules about the type, size, and color of turbans that each rank of Ottoman official could wear. Seeking to move the empire toward a system of greater equality, Mahmud II ordered that the existing sumptuary laws be superseded in 1829 with new regulations that imposed a common dress code designed to obscure social and political hierarchies and to undermine group identity. Central to this effort was the replacement of the myriad

⁴⁸ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 433–38.

⁴⁹ Findley, *Turkey, Islam, Nationalism, and Modernity*, 39–43.

⁵⁰ Findley, 39–43.

turbans that officials had worn with a new type of hat, the fez, that men of any rank or religion could wear.⁵¹

Rose Chamber Edict, 1839

The reform effort took a major step forward in 1839 when Sultan Abdulmecid I (r. 1839-1861) issued the Rose Chamber Edict. This proclamation formally inaugurated the Tanzimat Era by promising a comprehensive array of reforms. These included the replacement of tax farming with a fair system of direct taxation, the promulgation of laws guaranteeing the security of life and property, the establishment of an equitable system of military conscription, and the creation of an efficient bureaucracy. Importantly, the promised provisions were universal in nature and thus bound the government to legally and socially emancipate the sultan's millions of non-Muslim subjects. The Rose Chamber Edict did not itself carry the force of law; instead, it merely affirmed that the government would soon pass such legislation. Nonetheless, by committing the state to a sweeping program of reform aimed at erasing the legal privileges enjoyed by the different intermediary bodies, the proclamation marked a watershed moment in the empire's history and ushered in a period of rapid change.⁵²

Under the direction of the central bureaucracy, which emerged as the dominant political force during the Tanzimat Era, the Sublime Porte followed with a series of reforms in the 1840s and early 1850s aimed at remaking the empire's social, political, and economic systems along the lines articulated in the Rose Chamber Edict. In the social realm, the government instituted mixed commercial courts for hearing cases involving people from different confessional groups, drew up plans for a system of compulsory, secular education centered on a curriculum that stressed loyalty to the state, and replaced the term *dhimmi* with the designation "non-Muslim Ottoman"—thereby blurring the previously sharp legal distinction between Muslims and non-Muslims. In economics, meanwhile, it sought to effect greater growth by introducing paper money, by stabilizing the value of the currency, by adopting the practice of state budgeting, and by enacting a new commercial code based on French law that included provisions for bankruptcy and partnerships. Politically, finally, it adopted a national flag and a national anthem as unifying symbols, established the first post offices in the empire, and sought to exert greater central control over the provinces.⁵³

Hatti-i Humayun, 1856

Issued in 1856, a second statement of intent extended and clarified the Rose Chamber Edict. Though the reformers enthusiastically supported it, the announcement's genesis lay not in their hopes for further restructuring but instead in Britain's insistence

⁵¹ Donald Quataert, "Clothing Laws, State, and Society in the Ottoman Empire, 1720-1829," *International Journal of Middle East Studies* 29, no. 3 (1997): passim.

⁵² Aksan, *Ottoman Wars, 1700-1870*, 403-4.

⁵³ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 453-54; Quote from Hanioglu, *A Brief History of the Late Ottoman Empire*, 86-91.

on the issuance of a more concrete elaboration of the new rights that the sultan's subjects were to enjoy. Two issues drove London to call for a new declaration. First, the queen's government was keenly aware that continued discrimination along religious lines abetted St. Petersburg's ability to use its claim to be the protector of Orthodox Christians to stir up instability in the Balkans. Second, as in other Western states, popular opinion in Britain was increasingly dissatisfied with the continued second-class status of Ottoman Christians and had begun to demand that the British government pressure the Sublime Porte to move forward with the reforms it had promised in 1839. Accordingly, shortly before the Crimean War came to an end, London made both admission into the Concert of Europe and continued British diplomatic support of the empire contingent on the issuance of a new statement that would articulate the rights that the Rose Chamber Edit had promised to non-Muslims.⁵⁴

In response, the reformers who dominated the Ottoman bureaucracy had the sultan issue a follow-up edict in 1856, the *Hatti-i Humayun*, that focused on articulating the new legal and social position that the empire's Christian and Jewish minorities would henceforth enjoy. It guaranteed religious freedom, explicitly opened civil-service employment and schools to non-Muslims, promised the abolition of the death penalty for Muslims who converted to Christianity or Judaism, ended the *jizya*, and, most notably, made men of all confessional groups rather than just Muslims subject to conscription. As important, the edict marked the point at which the sultanate replaced its traditional embrace of Sunni Islam as a means of legitimating the government with Ottomanism, a new ideology designed to counter the centrifugal force of nationalism through the promotion of loyalty to the Ottoman state on the basis of universal equality.⁵⁵

The *Hatti-i Humayun* marked another important step in the empire's reform efforts. However, it is important to understand that many of its provisions were not as sweeping than they may have appeared at first blush. The ban on the death penalty for apostasy on the part of Muslims was symbolically important, for example, but of little practical consideration in light of the fact that the Ottoman government and religious hierarchy had long tolerated conversion to Judaism or Christianity.⁵⁶ More importantly, the elimination of the *jizya* and the expansion of the draft to include non-Muslims turned out to be largely cosmetic. It is true that the edict expanded conscription to encompass Christians and Jews; however, the *Hatti-i Humayun* also included a provision permitting non-Muslims to pay an annual exemption tax that released them from the obligation to serve. Both the Ottoman government and the empire's religious minorities embraced this loophole. Keen to avoid serving in the sultan's army, non-Muslims were only too happy to pay the exemption tax. Though securing large numbers of committed recruits was a key aim of the reform program, meanwhile, the Ottoman government was perpetually in such desperate financial straits that it was more than willing to forego soldiers in exchange for tax revenue. In practical terms, in other words, the *Hatti-i*

⁵⁴ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 457–62.

⁵⁵ Finkel, 457–62, 475.

⁵⁶ Conversion to Shi'i Islam was a different matter altogether, however.

Humayun resulted in the perpetuation of the *jizya* system—the payment of a tax in lieu of military service—albeit as a voluntary arrangement and under a new name.⁵⁷

The limits of the *Hatti-i Humayun* were also evident in the fact that many Ottoman subjects saw little improvement in their rights or legal standing as a result of the reform effort. This reality was most apparent with regard to Muslim women. While they did gain access to education thanks to the *Hatti-i Humayun*, they saw no liberalization to their sartorial code and continued to face legal action when dressed in a manner that the state deemed too revealing. Worse, the gradual adoption of secular laws based on Western codes came at the cost of many of the property rights that Muslim women had traditionally enjoyed under Islamic law.⁵⁸

Still, these important caveats aside, the *Hatti-i Humayun* and the other Tanzimat reforms were of enormous consequence thanks to the fact that they mounted a direct challenge to the corporate structure of Ottoman society. The embrace of universal, individual rights and the effort to establish a direct connection between the state and its subjects squarely undermined the authority of the guilds, *millets*, and religious orders. Likewise, the related efforts to eliminate tax farming and to assert central control over outlying parts of the empire weakened the position of the local notables and provincial governors who had long mediated between the Sublime Porte and the people of the provinces. Finally, the guarantee of religious freedom that was inherent in the emancipation of Jewish people and Christians and the elimination of Islam's central social and political position in the empire eroded the privileged position that Muslims had long enjoyed and undercut the power of the 'ulama'.⁵⁹

Opposition

Predictably, the Tanzimat reforms generated hostile reactions from those who lost status and privilege. Muslims were the most vocal in complaining about them. Ordinary Muslims protested the relative decline that the changes brought to their social standing while the 'ulama' criticized the gradual displacement of Islamic law by secular legal codes and condemned the idea of social equality between Muslims and *dhimmi*. As one conservative religious scholar declared shortly after the issuance of the 1856 edict, “[f]or Muslims this is a day to weep and moan.” Muslims were not alone in complaining about the *Hatti-i Humayun*, however. Provincial governors and rural notables also bitterly fought what they correctly saw as a threat to their autonomy and wealth. Perhaps surprisingly, even the leaders of the *millets*, bodies whose members clearly stood to gain, contested the Tanzimat. What explains their hostility? Why did they oppose a reform effort that sought to enhance the rights of the empire's non-Muslims? They did so for reasons of naked self-interest. Jealous of their authority, they demanded that the state refrain from directly granting new rights to individuals and called on it to instead apply the new privileges through the *millets*—an arrangement that

⁵⁷ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 460–61.

⁵⁸ Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 66–67.

⁵⁹ Quataert, 66–68.

would have permitted them to maintain the power that they wielded as intermediaries between the central government and the communities they had long overseen.⁶⁰

Assessment

In the face of this opposition, how effective was the reform program at restructuring the empire along Western lines and at building a new sense of Ottomanist identity? While it certainly had its successes, in the end the results were at best mixed. On the one hand, the reform effort had succeeded in destroying the Janissaries, had dramatically improved the military, and had modernized the central bureaucracy. Of pivotal importance, it had also secured British support against both the internal threat that Mehmet Ali posed—a situation that we shall explore presently—and the external challenge that Russia presented. On the other hand, however, many of the reforms failed to produce the intended results owing to a combination of poor execution and exogenous factors. For example, the effort to provide universal, compulsory education stumbled due to the lack of qualified teachers. Similarly, attempts to foster the development of a modern economy foundered in the face of the Capitulations and the low import duties that the European powers had compelled the empire to accept in exchange for their diplomatic support.⁶¹

Most importantly, the Tanzimat movement had failed spectacularly in achieving that which had emerged as the empire's most pressing reform goal: winning the allegiance of its non-Muslim population. Despite the government's efforts, neither the reform program nor the Ottomanist ideology that the reformers had promoted with such vigor found traction with the sultan's non-Muslim subjects. Instead, the vast majority were either unmoved by the Sublime Porte's efforts to win their loyalty or had fallen under the influence of one of the nationalist ideologies that the new, ethnically defined states of the Balkans had begun promoting in hopes of winning the loyalties of their co-nationalists who still lived under the sultan's rule. By the 1870s, in other words, no promise to rationalize the tax code, to reform the legal system, or even to grant universal rights and equality to all subjects could persuade the vast majority of the empire's remaining Christian subjects to express a sense of loyalty toward an empire that many were ambivalent about and that a growing number wished above all else to leave. The response to the Sublime Porte's revocation of the provision permitting non-Muslims to purchase exemptions from conscription in 1908 made the failure of both the reformers' integrationist program and their promotion of Ottomanist ideology emphatically clear. Rather than embracing their newfound rights or accepting the responsibilities that went with them, tens of thousands of the sultan's Christian and Jewish subjects responded to the prospect of serving in the Ottoman military by promptly emigrating to the Americas.⁶²

⁶⁰ Hanioglu, *A Brief History of the Late Ottoman Empire*, 74–75.

⁶¹ Hanioglu, 92, 102.

⁶² Quataert, *The Ottoman Empire, 1700-1922, Second Edition*, 67.

The Young Ottomans

Mounting frustration with the limited success and undemocratic nature of the top-down Tanzimat program led to the emergence a competing reform effort known as the Young Ottomans. Originating in the 1860s, this new movement was composed of a loose coalition of liberal Muslim intellectuals, journalists, and writers. Disappointed by the inability of the Tanzimat reforms to secure the empire, they came to a seemingly simple but also hugely influential idea: that Islam and constitutional democracy were complementary and could be integrated. This they proceeded to do. They merged the ideas of Western Enlightenment thinkers such as Jean-Jacques Rousseau (1712-1778) and Montesquieu (1689-1755) with Islam's social-justice values and the Qur'an's embrace of *shura*, or consultation, which they interpreted as a divine endorsement of Islamic democracy. The result was a blueprint for a government that was both democratic and Islamic—one, the Young Ottomans believed, that could both provide the just rule that the people of the empire sought and demonstrate to the predatory West that the sultanate was a modern state capable of self-rule rather than a premodern one requiring neo-colonial tutelage.⁶³

The Young Ottomans did not need to wait long to see their dream of a constitutional order take form. A succession of problems in the mid-1870s undermined the legitimacy of the existing government and opened the door to the kind of sweeping political changes that they proposed. The first crisis was a fiscal one. After a serious drought in Anatolia and the onset of a global depression in 1873, tax receipts fell to the point that the Ottoman treasury could no longer service the empire's massive debt obligations; as a result, to the shock of European governments and domestic investors alike, the Sublime Porte defaulted on its debt in October 1875. Meanwhile, as we have seen, the outbreak of nationalist revolts in the Balkans in 1875 and 1876 plunged the Ottoman Empire's European territories into a complex diplomatic and military crisis that Russia seemed well positioned to exploit. Finally, Sultan Abdulaziz (r. 1861-1876) further complicated what was an already difficult situation by attempting to use the Balkan revolts as a pretext to end the Tanzimat and to reassert sultanic absolutism. Coming on the heels of the other crises, Abdulaziz's brazen move led the reformers in the bureaucracy to conclude that they had to take extreme actions. Accordingly, they joined forces with the Young Ottomans to launch a coup in early 1876 that put Abdulaziz's brother, Abdülhamit II (r. 1876-1909), on the throne.

Now firmly in control of the state, the Young Ottomans and their allies wasted little time in remaking the empire along liberal, democratic lines. They started by drafting and putting in place a constitution that reflected both Islamic values and recent Western political science. Based on Belgium's constitution, it transformed the empire into a constitutional monarchy—albeit one that lodged substantial power in the sultan's hands.

⁶³ Şerif Mardin, *The Genesis of Young Ottoman Thought: A Study in the Modernization of Turkish Political Ideas* (Princeton: Princeton University Press, 1962), passim.; Ira M. Lapidus, *A History of Islamic Societies*, 2nd ed. (Cambridge: Cambridge University Press, 2002), 496; Findley, *Turkey, Islam, Nationalism, and Modernity*, 104–6.

With the increasingly grim Balkans crisis as a backdrop, they followed by holding elections for a bicameral parliament formally known as the General Assembly; to great fanfare, it began meeting in March 1877. Thus, despite the manifold crises that enveloped the empire in the mid 1870s, the constitution appeared to be well on its way to becoming institutionalized by early 1878.⁶⁴

Unfortunately for the reformers, Abdülhamit II was a staunch opponent of constitutional democracy. Indeed, he had begun plotting a return to absolutist rule from the moment he took the throne. It did not take long for him to put his plan into action. Armed with sweeping emergency powers accorded him—ironically—by the very constitution he sought to overthrow, the sultan took advantage of the war that had broken out with Russia to prorogue the assembly and to suspend the constitution in February 1878. Firmly in control of the capital, he followed by intimidating the Tanzimat reformers and the Young Ottomans into accepting what amount to a self-coup, or *autogolpe*. Thus, in less than a year's time, the Ottoman Empire's Islamic-based, liberal constitutional government—the dream of the Young Ottomans—had been replaced by a return to sultanic despotism.⁶⁵

Despite their spectacular short-term failure, the Young Ottomans proved to be enormously influential in the long run both in the capital and in the broader Middle East. Indeed, while their ideas played no role whatsoever in shaping the structure of the government or the spirit of the laws under Abdülhamit II, their vision of an Islamic democracy would remain the starting point for most political thinking inside the empire for decades thereafter. Most notably, it would inform the political views articulated by an important underground resistance movement called the Committee for Union and Progress (CUP) that we will examine in the next chapter. More broadly, the Young Ottoman synthesis of European political ideas and Islamic values would inspire the emergence of similar movements in other parts of the Middle East—most notably in Iran.⁶⁶

Westernizing Reforms in Egypt

The Tanzimat bureaucrats and the Young Ottomans were not the only top-down Westernizing reformers in the empire in the nineteenth century. In fact, they were not even the first. Several decades earlier, a separate modernization effort had begun in the autonomous Ottoman province of Egypt, first under the direction of its Albanian-born governor, Mehmet Ali and later under his grandson, Ismail (r. 1863-1879). For a time, their restructuring effort appeared to be enormously successful and even seemed to outdo the central government's Tanzimat modernization program. Mehmet Ali and Ismail's achievements ultimately proved to be ephemeral, however. Not only did their reform effort flounder in the face of concerted European opposition, but, worse, it paved

⁶⁴ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 488–90.

⁶⁵ Finkel, 488–90.

⁶⁶ Hanioglu, *A Brief History of the Late Ottoman Empire*, 104.

the way for Britain to take control of Egypt on an ostensibly temporary but ultimately permanent basis in 1882.

Mehmet Ali

Mehmet Ali's meteoric rise occurred in the aftermath of Napoleon's invasion of Egypt in 1798. Deployed to the province as an officer in 1801, he took part in the Anglo-Ottoman military operations that ended the French occupation. Sensing opportunity in the power vacuum that followed the evacuation of British troops in 1803, he arranged to remain in Egypt. Making skillful use of both his considerable diplomatic talents and the 6,000-man military force he commanded, he followed by mounting a successful bid to compel the powerful Mamluk landowners to accept him as the province's new ruler. He followed in 1805 by securing Selim III's reluctant acknowledgement of him as governor. Finally, six years later, he removed the last meaningful restraint on his rule when he had his soldiers massacre the Mamluks while they were attending a ceremony in the capital.⁶⁷

With the Mamluks out of the way, Mehmet Ali was now free to pursue his goals unhindered. They were ambitious. He wanted nothing less than to break with the Ottoman Empire and establish a powerful, dynastic state centered on Egypt. A shrewd man, he understood that the task he had set himself was a daunting one. Achieving independence from the sultan and maintaining it in the face of the challenge that the rising European powers posed would require a strong army and navy. In turn, such a military would demand an efficient central bureaucracy and a strong economy—one that was not merely able to generate the tax revenue that the state needed but that was also sophisticated enough to produce the uniforms, supplies, and state-of-the-art weapons that the army and navy required. In other words, Mehmet Ali could only achieve his political objectives if he first modernized and industrialized Egypt.⁶⁸

He moved aggressively to do so from the moment he eliminated the Mamluks. He began by seeking new sources of revenue to fund his ambitious program. He boosted government tax receipts by replacing the *iltizam* tax farms with direct levies and by imposing taxes on Egypt's many *awqaf*—despite the fact that land endowed for the purpose of supporting a religious institution or charity traditionally enjoyed tax-exempt status. He followed by establishing state monopolies over the sale of key agricultural products including wheat, sugar, and, of special importance, the high-quality, long-staple cotton that Egypt produced in abundance and that Europe's textile mills voraciously demanded. To maximize the state's take, he bought these crops from peasants at fixed, low prices and then sold them internationally at higher, market-rate ones. These initiatives—particularly the export monopolies—were successful. Thanks in

⁶⁷ Eugene Rogan, *The Arabs: A History – Revised and Updated Edition* (New York: Basic Books, 2012), 66–69.

⁶⁸ Khaled Fahmy, *All the Pasha's Men* (Cambridge: Cambridge University Press, 1997), 9–13.

part to the high price of grain during the Napoleonic Wars, Mehmet Ali's government was soon collecting substantial amounts of revenue.⁶⁹

He used that income to support an aggressive industrialization program. He established workshops that could manufacture goods that the province had previously imported like rope, metal items, and chemicals, and he hired European experts who could establish the schools, technical institutions, and military academies needed to produce the skilled workers and soldiers on whom his ambitions depended. He also tapped the state's new revenue sources to construct the workshops and shipyards required to make the uniforms, tents, muskets, cannons, warships that his military would need to engage in modern warfare. Finally, and most notably, he lavished funds on a new textile industry that he hoped would be able to compete with the more mature ones in Britain, Belgium, and France. Boosted by governmental support and cheap, state-subsidized cotton, it expanded rapidly. Indeed, by 1840, the Egyptian textile industry was producing more than 1.2 million items of calico per year.⁷⁰

Mehmet Ali's aggressive economic development program soon permitted him to field the modern military that he had long coveted. It quickly proved its worth. At the sultan's request, he deployed troops to the Hijaz from 1811 to 1818 to reconquer Mecca and Medina from the Wahhabis, a puritanical Arabian religious group that had seized the Holy Places in 1806. Well-armed and disciplined, his soldiers soon drove the Wahhabis back into the desert; as a reward, the sultan permitted him to add the Hijaz to the territory that he governed. Not long thereafter, he further expanded his fledgling empire by invading and annexing Sudan.⁷¹

Changes in the Egyptian military's system of recruitment soon made his army even more powerful. Beginning in 1822, Mehmet Ali abandoned his earlier reliance on traditional slave soldiers in favor of conscripting Egyptian peasants into a new, French-trained army patterned after Selim III's *Nizam-I Cedid*. It was a fearsome and effective force. Equipped with modern weapons, his disciplined army of indigenous Egyptian soldiers was, without question, the most powerful military force in the Middle East.⁷²

Indeed, it was so much more capable than the as-yet reformed Ottoman military that Sultan Mahmud II felt compelled in 1824 to offer Mehmet Ali's son, Ibrahim (r. 1848), control of Cyprus, Crete, and the Morea in exchange for Egyptian assistance in suppressing the Greek Revolt. In the short term, the intervention proved successful. Under Ibrahim's effective-if-brutal command, Mehmet Ali's powerful military made short

⁶⁹ Khaled Fahmy, "The Era of Muhammad 'Ali Pasha," in *The Cambridge History of Egypt. Volume 2, Modern Egypt, from 1517 to the End of the Twentieth Century*, ed. M. W. Daly (Cambridge: Cambridge University Press, 1998), 147–50.

⁷⁰ Aksan, *Ottoman Wars, 1700-1870*, 308–13; David Landes, *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor: Why Some Are So Rich and Some Are So Poor* (New York: W. W. Norton & Company, 1999), 403–8.

⁷¹ Aksan, *Ottoman Wars, 1700-1870*, 308–13.

⁷² Aksan, 308–13.

work of the Greek rebels and quickly returned the region to Ottoman rule. The very success of the Egyptian army proved to be its undoing, however. Ibrahim's victory over the Greeks alarmed Russia, France, and England, which, for different reasons, supported Greek independence. The result, as noted earlier, was Western intervention in the conflict on the Greek side and, ultimately, the Battle of Navarino in 1827 that cost Mehmet Ali the modern navy that he had so painstakingly constructed.⁷³

Navarino soon produced a serious crisis in the Ottoman Empire. Bitter over the destruction of his fleet, Mehmet Ali demanded that the sultan compensate him for his losses by granting him control of Syria. Mahmud II had become leery of his nominal subordinate's rising power, however, and was determined to prevent him from further expanding his territory. Wagering that the Greek campaign had weakened the Egyptian army to the point that the Ottoman military could defeat it, he consequently rebuffed Mehmet Ali's demand. Furious, the latter responded by ordering Ibrahim to invade Syria in 1831. The ensuing conflict between the sultan and his putative vassal quickly proved that Mahmud II had miscalculated badly. With the Egyptian army having long since fully recovered from the fighting in Greece, Ibrahim quickly overran Syria before pressing on with an invasion of Asia Minor. There followed several months of fruitless negotiations after which Ibrahim decisively crushed a much-larger Ottoman force near Konya in December 1832. The battle dramatically raised the stakes of the conflict. Indeed, it not only put the integrity of the Ottoman Empire into question but also transformed what had been an internal conflict into a full-blown international crisis.⁷⁴

With Britain paralyzed at that time by domestic political issues, Russia took the lead in resolving what came to be known as the First Egyptian Crisis. Believing that a weak Ottoman Empire was more conducive to its territorial ambitions than a dynamic and powerful Egyptian state, St. Petersburg deployed troops to protect Istanbul and compelled Mehmet Ali to order Ibrahim to withdraw from central Anatolia. The war thus ended with the sultan still firmly in power and with nearly all of Asia Minor remaining under direct imperial control. Still, the settlement was a highly favorable one for Mehmet Ali. In exchange for halting his advance on the capital, the agreement he had negotiated with St. Petersburg added Syria, Crete, and the Anatolian province of Adana to his burgeoning empire-within-an-empire.⁷⁵

Neither Mehmet Ali nor the Ottoman government were satisfied with the terms of the Russian-brokered agreement, however, and their competing efforts to alter it would produce another confrontation, the Second Egyptian Crisis, in 1839. Mehmet Ali moved first. Intent on securing a complete break with the empire, he offered £3 million to the sultan in May 1838 in exchange for full independence. Unwilling to countenance what would have been the loss of nearly all of the Ottoman Empire's Arab territories, Mahmud II declined his nominal subordinate's offer. Instead, three months later, his government moved to counter the Egyptian ruler's efforts to divide the sultanate by

⁷³ Evans, *The Pursuit of Power*, 58.

⁷⁴ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 443–45.

⁷⁵ Zürcher, *Turkey*, 36–37.

agreeing to the Treaty of Balta Liman with London. On the surface, the treaty appeared to be little more than a one-sided agreement that came at great cost to the sultanate. A monument to free trade, it capped customs duties and banned monopolies throughout the empire and thus fully opened Ottoman markets to British merchants.⁷⁶

On a deeper level, however, the Treaty of Balta Liman was a shrewd move on the Sublime Porte's part designed to contain and greatly weaken Mehmet Ali's position. It promised to do so in two ways. First, since Egypt remained legally a part of the Ottoman Empire, the treaty abolished the export monopolies on which Mehmet Ali's power rested and thus greatly eroded the military threat that he posed to continued Ottoman rule. Second, since British merchants would only benefit from the end of Mehmet Ali's export monopolies if Egypt formally remained part of the Ottoman Empire, the agreement ensured that London would continue to back the Sublime Porte's efforts to prevent the Egyptian ruler from gaining independence.⁷⁷

Indeed, now confident that it could count on Britain's support, Mahmud II's government moved to force a crisis by attacking Syria in 1839. Militarily, the invasion was an utter disaster for the empire. Disciplined and better equipped, a much-smaller Egyptian force thoroughly routed the Ottoman army in June 1839. Diplomatically, however, it was a rousing success. Just as Ottoman officials had hoped, concerns about the possible dismemberment of the empire pushed the British government to intervene to end the crisis. With the support of several other European states, it gave Mehmet Ali an ultimatum in the summer of 1840 thinly disguised as a choice. If he agreed to accept the sultan's nominal authority, his position would become hereditary and he would retain control of southern Syria; however, if he instead pressed on with the war, Britain would compel him to immediately return all of Syria to sultanic rule.⁷⁸

Mehmet Ali badly miscalculated when faced with this diktat. Determined to achieve real independence rather than autonomy, he defied London's demand and continued his war with the sultan. The vast gap between Egypt's military power and that of the European states quickly made clear the futility of this decision. The British government promptly responded by ordering the Royal Navy to sever the supply lines on which Ibrahim's troops depended for food and ammunition—thus exposing them to defeat by Ottoman forces. Mehmet Ali had no answer to this action. Caught in a no-win situation, he reluctantly accepted defeat and agreed to pull his soldiers out of Syria.⁷⁹

The Egyptian ruler had suffered a huge diplomatic setback in the Second Egyptian Crisis. European opposition to the fragmentation of the Ottoman Empire had cut short his bid to achieve independence, while his intransigence had cost him control of Syria. In the end, all he received as a result of his successful military campaigns against the sultan was formal recognition of his family's hereditary control of Egypt.

⁷⁶ Fahmy, "The Era of Muhammad 'Ali Pasha," 174; Rogan, *The Arabs*, 80.

⁷⁷ Fahmy, "The Era of Muhammad 'Ali Pasha," 174.

⁷⁸ Rogan, *The Arabs*, 79–81.

⁷⁹ Rogan, 79–81.

While valuable, even this concession masked a much-more substantial defeat. As earlier noted, continued Ottoman suzerainty over Egypt meant that Mehmet Ali and his successors had to abide by the terms of the empire's foreign agreements including, most importantly, the Treaty of Balta Liman. As such, he was compelled to dismantle the state monopolies and to drastically lower the export duties that had constituted the financial foundation on which his army, his program of industrialization, and, ultimately, his power, rested. Thus, in the wake of the Second Egyptian Crisis, Mehmet Ali's modernization program lurched to a halt.⁸⁰

Scholars have long been divided over the efficacy of his top-down development plan. Some argue that his forced-industrialization program—particularly the textile industry that he built—produced impressive results. For example, the historian Jean Batou notes that Egypt at one point had the fifth most cotton spindles in the world and concludes that “its [industrial] achievements were not so different from those of certain Western European regions.” More broadly, scholars such as Batou also maintain that Mehmet Ali's development scheme was on track to industrialize Egypt and that it only failed because the European states denied it the right to erect protective tariffs or to subsidize its domestic production. Others are considerably less impressed with Mehmet Ali's achievements, however. Most notably, the historian David Landes contends that the Egyptian textile industry's reliance on expensive animal labor rather than coal or other power sources suggests that the province did not experience a genuine industrial revolution in the 1820s and 1830s. More broadly, he argues that the failure of Mehmet Ali's industrialization plan stemmed from internal shortcomings rather than from European efforts to kill a potential competitor in its cradle and that it would have failed regardless of Western actions.⁸¹

Whether Egypt was doomed from the start or undermined by the industrial states, the outcome was the same: Mehmet Ali's plan to turn his state into a military and industrial power ground to a halt in the wake of the Second Egyptian Crisis. Indeed, within just a few years, Egypt's powerful military had shrunk to a shell of its former self. More importantly, its once-rapidly growing, proto-industrial economy had been transformed into a colonial appendage of Britain that supplied its factories with raw materials and served as a market for its finished goods.⁸²

Khedive Ismail

Mehmet Ali's dream of turning Egypt into a modern state was not completely dead, however. Elevated to the rank of khedive, a Persian title analogous to the term 'viceroy,' his grandson Ismail launched a second, top-down effort to transform the

⁸⁰ Rogan, 81–83.

⁸¹ Jean Batou, “Muhammad-'Ali's Egypt. A Command Economy in the Nineteenth Century?,” in *Between Development and Underdevelopment: The Precocious Attempts at Industrialization of the Periphery, 1800-1870*, ed. Jean Batou (Geneva: Librairie Droz, 1991), 181; Landes, *The Wealth and Poverty of Nations*, 405–6.

⁸² Fahmy, “The Era of Muhammad 'Ali Pasha,” 175–79.

province into an economic power beginning in the early 1860s. His achievements were briefly dazzling and appeared for a time to have put Egypt on a path that would soon see it join the ranks of the world's economically developed states. In the end, however, his ambitious program proved to be a spectacular and costly failure. Mismanagement and, more importantly, the staggering foreign debt on which his modernization scheme rested would not only prevent Egypt from transforming into an industrial power but would also pave the way for Britain to turn it into a de facto colony in 1882.

Ismail was a man of enormous ambition. Though he did not seek to make Egypt a great military power as his grandfather had, he shared his predecessor's desire to dramatically modernize its economy and to achieve political independence from both Istanbul and the increasingly rapacious European powers. Accordingly, he embarked on a massive, state-led development program. It was nothing if not ambitious. To increase agricultural exports, he ordered the government to construct numerous new port facilities and to dig thousands of miles of irrigation canals. Meanwhile, to end the country's dependence on low-margin agricultural exports, he expanded Egypt's rail network, built huge numbers of schools, and subsidized industries like sugar refining, paper milling, and cotton processing. Finally, to give the country the sheen of modernity, he invested huge sums into a project to transform Cairo into a European-style city. He constructed wide, straight boulevards, outfitted fashionable neighborhoods with gas lighting, and laid out a series of new parks and public gardens. He also ordered the erection of a new palace, an opera house, and a national theater.⁸³

These were important projects to Ismail, but they paled in comparison to the centerpiece of his modernization scheme: The Suez Canal. The waterway was the fulfillment of an age-old dream to connect the Mediterranean Sea and the Red Sea at the point where Egypt proper meets the Sinai Peninsula. The mid-nineteenth century seemed a propitious time to finally realize that vision. With the globe becoming more and more economically connected thanks to the industrialized states' seemingly insatiable demand for raw materials, the proposed canal promised to be well utilized and immensely profitable. Egypt did not possess the capital and technical knowledge to construct the waterway itself, however. Accordingly, Ismail's predecessor, Said (r. 1854-1863), had entered into a joint venture with a French consortium under the direction of the famed canal builder Ferdinand de Lesseps (1805-1894) to dig the canal. By the terms of the agreement, the French would supply capital, equipment, and technical knowledge in exchange for a 56 percent share of the canal while the Egyptian government would provide land and *corvée* labor—uncompensated work by Egyptian peasants—and would retain the remaining 44 percent stake. With this agreement in place, construction began under de Lesseps's direction in 1856. Completed in 1869, the canal contributed enormously to the growth in world trade that occurred in the late-nineteenth century. Almost overnight, it massively cut the cost to ship goods between

⁸³ Dominic Green, *Three Empires on the Nile: The Victorian Jihad, 1869-1899* (New York: Free Press, 2007), 14–17; Rogan, *The Arabs*, 100–101; Peter Stearns, *The Industrial Revolution in World History*, 3rd edition (Boulder: Westview Press, 2007), 97.

Europe and South, Southeast, and East Asia and opened the east coast of Africa to European imperial penetration and control.⁸⁴

It also gave Ismail the opportunity to present himself as a modernizer and global leader. He did so most famously in 1869 when he celebrated the canal's completion by hosting the cultural and political elite of Europe at a lavish series of parties, ceremonies, fireworks displays, and feasts highlighted by a performance of Giuseppe Verdi's opera *Rigoletto* and by a ceremonial procession of ships headed by the French Empress Eugénie's (1826-1920) royal yacht. It was an impressive series of events and seemed to suggest that, under Ismail's leadership, Egypt was poised to join the ranks of the world's developed states.⁸⁵

While the khedive's economic-development program appeared enormously impressive for a time, it ultimately rested on fiscal quicksand and thus proved untenable. Three factors contributed to the budgetary problems that Egypt began to experience in the late 1860s. First, corrupt and ill-suited to overseeing the flood of initiatives that the khedive rushed to undertake, the Egyptian government proved incapable of either completing projects in a timely manner or keeping them under budget. Second, Ismail himself was an infamous spendthrift who lived lavishly and who spent extravagantly on edifices such as the Opera House and the Egyptian Museum in Cairo that were designed not to be financially self-sustaining but instead to boost his international standing.⁸⁶

These factors were surely important, but it was ultimately the means by which Ismail had financed his economic-development program that pushed his government to the fiscal brink. As we have seen, the Treaty of Balta Liman had barred Egypt from imposing export duties or from operating state-owned agricultural monopolies. As such, the treaty left Ismail with no option but to underwrite his aggressive modernization scheme by securing loans on international money markets using the tax proceeds of future cotton crops as collateral. At first, the terms were very good. Thanks to the North's blockade of the Confederacy during the American Civil War (1861-1865), the value of Egypt's cotton crop soared, and Ismail was able to borrow at very low interest rates. However, after Southern exports began to flood the market following the Union victory in 1865, cotton prices—and with them, Egypt's tax receipts—collapsed. Thereafter, Ismail could only pay his existing debt by taking on new loans at ever-higher rates of interest. By 1870, the situation had become unsustainable. With one-third of the state's revenue committed to servicing the debt, the Egyptian government and Ismail's economic modernization program were teetering on the brink of bankruptcy.⁸⁷

⁸⁴ Rogan, *The Arabs*, 99–100.

⁸⁵ Green, *Three Empires on the Nile*, 1–5.

⁸⁶ F. Robert Hunter, *Egypt Under the Khedives: 1805-1879: From Household Government to Modern Bureaucracy* (Cairo: American University in Cairo Press, 1999), 38–40.

⁸⁷ Hunter, 38–40.

The khedive responded by taking drastic actions to keep Egypt solvent. While well-intentioned, these efforts reflected short-term thinking and merely served to push his government into an even deeper fiscal hole. Such was certainly the case with the *Muqabala* Law that went into effect in 1872. It offered landowners the opportunity to receive a permanent, 50 percent reduction in their tax assessment in exchange for an immediate payment of six-years' worth of tax revenue. Unsurprisingly, the *Muqabala* Law proved to be very popular with those wealthy enough to take advantage of it, and it generated a quick infusion of funds in its first few years. With equal predictability, however, it severely depressed the collection of taxes in subsequent years and thus deepened rather than eased the growing debt crisis.⁸⁸

Thereafter, the situation progressively worsened. By the mid 1870s, pressure from Egypt's creditors compelled Ismail to sell Egypt's 44 percent stake in the Suez Canal to the British government. This transaction once again brought a quick injection of cash, but, like the *Muqabala* Law, it did so at the cost of long-term revenue and thus could only slow rather than arrest Egypt's ongoing slide into insolvency. Finally, in 1876, the bill came due. No longer able to meet its obligations, Egypt called a moratorium on the payment of its international debts—in effect announcing that it was bankrupt.⁸⁹

At that point, the British and French governments intervened to ensure that their bondholders were made whole. Initially, London and Paris did so by imposing on the Egyptian government what was known as the system of Dual Control—an arrangement designed to ensure that it collected sufficient revenue to service its debts. By the terms of this understanding, a British economist would serve as Egypt's minister of finance and a Frenchman as minister of public works. Controlling Egypt's budget, they were tasked with ensuring that it could fund its international debt, now consolidated at a rate of 7 percent. But how would they come up with the additional revenue to meet that obligation? They did so by imposing a biting austerity regime on the Egyptian government that saw wages and positions in both the bureaucracy and army savagely cut.⁹⁰

Ismail initially accepted this arrangement. Dependent on the European states for trade and financing, he had little choice but to comply. Chafing at the loss of power, however, he soon soured on the system of Dual Control and sought to exploit popular bitterness about the European-imposed austerity program to alter the terms of the agreement. He made his move in 1879 when he dissolved the cabinet and formed a new one that did not include any European ministers. He simultaneously demanded that his creditors lower the terms of Egypt's consolidated debt to 5 percent. These moves went over like a lead balloon in Europe. Unwilling to accept either Ismail's challenge to their authority or a further reduction in the interest paid to French and British holders of

⁸⁸ Rogan, *The Arabs*, 120–21.

⁸⁹ Rogan, 120–21.

⁹⁰ F. Robert Hunter, "Egypt Under Muhammad 'Ali's Successors," in *The Cambridge History of Egypt. Volume 2, Modern Egypt, from 1517 to the End of the Twentieth Century*, ed. M. W. Daly (Cambridge: Cambridge University Press, 1998), 195–97.

Egypt's debt, London and Paris refused to negotiate. Instead, they responded by pressuring Sultan Abdülhamit II to remove the khedive. Desperate at that moment for British and French support against Russia, the Ottoman ruler was quick to comply. Accordingly, the sultan deposed Ismail in 1879 in favor of the khedive's son, Tawfiq (r. 1879-1892). Tawfiq was merely a figurehead, however. As the speedy removal of his father had made clear, real power in Egypt thereafter lay not in Egyptian hands, but instead in those of the French and British ministers who once again controlled the cabinet and the budget.⁹¹

The 'Urabi Revolt, 1881-1882

Even then, Egyptians continued to resist Western dominance. This time, however, opposition did not come from the political elite but instead from the officers in the Egyptian army. Furious that the austerity measures that the Dual Control government had imposed had cut their wages in half in order to satisfy the interests of European bondholders, a group of officers led by the revolt's namesake, Colonel Ahmad 'Urabi (1841-1911), organized a series of popular, proto-nationalist demonstrations in 1881 under the slogan "Egypt for the Egyptians" aimed at compelling Tawfiq to accept a list of reforms. Their demands included the formation of a new cabinet, the approval of a new constitution, and the reconvening of the Assembly of Delegates—a heretofore weak legislative body established in 1866 that Ismail had prorogued in 1879. Facing broad popular support for the revolt, the cautious Tawfiq had no choice but to comply. As a result, by February 1882, Egypt had a constitution, an empowered legislature, and a reform-minded cabinet in which 'Urabi himself served as minister of war.⁹²

The events of 1881 and early 1882 alarmed Britain and France. 'Urabi's presence in the government raised the prospect that Egypt might stop meeting its debt obligations under the Dual Control system and even the possibility that it might unilaterally abolish privileges such as the Capitulations that Europeans had long enjoyed in the province. Unwilling to countenance such changes, the French and British governments responded by sending a joint fleet to Alexandria for the purpose of intimidating the Egyptians into backing down. The deployment failed to achieve that end, however. Instead, the arrival of the warships merely served to enrage the Egyptian people and to intensify the impasse between Cairo and the European states. By early July, as a result, Egypt had become a tinderbox—a country destined to burst into flames in the presence of the tiniest of sparks.⁹³

One was not long in coming. A fight between a Briton and an Egyptian in Alexandria in early July quickly transformed into a bloody riot in which over 125

⁹¹ Hunter, 195–97.

⁹² Rogan, *The Arabs*, 123–28.

⁹³ Donald Reid, "The 'Urabi Revolution and the British Conquest," in *The Cambridge History of Egypt. Volume 2, Modern Egypt, from 1517 to the End of the Twentieth Century*, ed. M. W. Daly (Cambridge: Cambridge University Press, 1998), 231–37.

Egyptians and 50 foreigners died. The British fleet—but not the French one, which had received orders from Paris to refrain from using force—responded to the incident by bombarding Alexandria on July 11, 1882, causing heavy loss of life. Three days later, British troops landed with orders to restore order and protect Western property. Though London claimed that its occupation was “temporary,” it made clear the following year that it intended to remain in Egypt indefinitely by sending Evelyn Baring (1841-1917), the future Lord Cromer, to assume the position of consul general—a title that only slightly disguised the fact that, he, rather than Tawfiq, was Egypt’s real ruler.⁹⁴

Ismail’s modernization attempt thus ended in disaster for Egypt. Economically, it had not only proven to be every bit as unsuccessful as his grandfather’s industrialization scheme but, worse, had given rapacious Western imperialists the opportunity to reduce Egypt to economic servitude. Thereafter, foreign businesses and governments owned a huge and growing share of the province’s economy including its greatest asset, the Suez Canal, and they opted to repatriate their considerable profits to Europe rather than to reinvest them in Egypt’s development. The political impact was even worse. The province may have technically remained an autonomous part of the Ottoman Empire, but, after 1882, it was in reality a British colony in all but name—one that served both as the primary artery connecting the metropole to its prized colony of India and as the nexus of London’s system of imperial defenses.

Populist Reform Movements

Not all of the Muslim reform efforts of this period were top-down ones such as those that occurred in Egypt and the Ottoman capital. On the contrary, the nineteenth century was a time of vigorous debates among intellectuals, religious leaders, and others over how Muslims could best resolve the problems of their time—particularly the challenge posed by the West—and reclaim what they saw as the Islamic world’s rightful place as the globe’s preeminent power. The two most influential of these reform movements drew starkly different conclusions about how the Islamic world could right itself. The Islamic modernist followers of Jamal al-Din al-Afghani (1838-1897) argued that Muslims should respond to the threat of European dominance by adopting Western technologies and science while retaining the values, beliefs, and embrace of rationalism that were characteristic of early Islam. In contrast, the puritanical adherents of Wahhabism, argued instead that Muslims could resolve their problems only by reembracing the conservative, canonical understanding of the values and practices of the early *umma*.

⁹⁴ Donald Reid, 231–37; M W Daly, “The British Occupation, 1882-1922,” in *The Cambridge History of Egypt. Volume 2, Modern Egypt, from 1517 to the End of the Twentieth Century*, ed. M. W. Daly (Cambridge: Cambridge University Press, 1998), 240–41.

Wahhabism

Wahabism predated substantial European penetration of the Middle East but would come to offer an influential framework for understanding the source of Muslim weakness and for responding to the Western challenge. Its founder and namesake, Muhammad ibn Abd al-Wahhab (1703-1792), developed its broad outlines while a religious student in Medina. His teachers there had exposed him to the writings of the thirteenth-and-fourteenth century Sunni theologian, Ibn Taymiyyah (1263-1328), who had argued that Muslims could regain God's blessing in the wake of the Mongol conquest by returning to the core values and practices of the early community of believers and by purging those who refused to do so. Ibn Taymiyyah's ideas would powerfully influence Abd al-Wahhab and shape the austere and dogmatic interpretation of Sunnism that he subsequently developed.⁹⁵

Abd al-Wahhab's new conception of Islam revolved around three broad, interrelated points. First, stressing *tawhid* or the oneness of God, it held that the only acceptable way to practice Islam was by focusing one's devotion exclusively on Allah. The inclusion of *any* intermediary figure or practice in the act of worship—be it the intercession of the Shi'i Imams, the popular practice of praying at the tombs of Muslim saints such as Muhammad, the veneration of Sufi *pirs*, or even the observance of many religious holidays—was to engage in idolatry or polytheism and was therefore *shirk*. Second, like Ibn Taymiyyah, Wahhabism championed the older ideology of Salafism: the idealization of the first few generations of Muslims and the belief that Islam could reform itself and end the problems of the present day only through a return to the practices of the early *umma*. Accordingly, it held that all additions to the practice of Islam made after the seventh century had to be ruthlessly excised. Finally, drawing explicitly on the ideas of Ibn Taymiyyah and indirectly on Kharijism, Abd al-Wahhab's conception of Islam maintained that God required all believers to engage in jihad against Muslims who had corrupted the religion. Those who resisted the imposition of Wahhabism's strict rules, those such as Shi'i and Sufis who had added new practices to the faith, and those who practiced *shirk* were *kafirs*, or apostates, whom the *umma* was obliged to *takfir*—or excommunicate—and kill.⁹⁶

At first, Abd al-Wahhab's doctrine alienated many people; indeed, he was forced to flee his hometown of Al-'Uyana, in eastern Arabia after he had a woman stoned to death for adultery. Soon, however, his message of religious renewal through a return to the practices of the past gained a small-but-highly committed following. Among them was Muhammad Ibn Saud (r. 1726-1765), the ambitious chief of a small tribe that controlled the oasis town of Dariyah in the inhospitable central region of Arabia called the *Najd*. Perceiving the benefits that could accrue from an alliance, the two men agreed to an arrangement in 1744 in which Ibn Saud recognized Abd al-Wahhab's

⁹⁵ John L. Esposito, *Islam: The Straight Path*, Fourth Edition (New York: Oxford University Press, 2011), 144–46.

⁹⁶ Reza Aslan, *No God but God: The Origins, Evolution, and Future of Islam*, Updated Edition (Random House, 2011), 246–47.

religious authority while Abd al-Wahhab acknowledged Ibn Saud's secular power. It would prove to be a momentous agreement—one that at last gave Abd al-Wahhab the opportunity to put his religious teachings into practice.⁹⁷

The Wahhabi-Saudi alliance soon flourished. Under the skilled military leadership of Ibn Saud and his successor, Abdulaziz (r. 1765-1803), the zealous and disciplined Wahhabi fighters quickly won control of central Arabia. They followed by launching religiously motivated raids into Ottoman territory. In 1802, Abdulaziz's zealous and uncompromising soldiers attacked Karbala during the festival of Ashura. Brutally sacking the city, they killed more than 2,000 Shi'i worshippers and destroyed the Imam Husayn Shrine—the burial place of the Third Imam. More spectacularly, Saudi forces seized Mecca and Medina from the sultan four years later. There, they banned the consumption of coffee and tobacco, ordered all books other than the Qur'an burned, and compelled men to wear beards and women to seclude themselves. They even destroyed the tombs of Muhammad and other prominent early Muslims to prevent people from worshipping at them.⁹⁸

Wahhabi control of Mecca and Medina posed a serious threat to the Ottoman government. Not only did it prevent the treasury from receiving substantial and much-needed tax revenues from the Hijaz, but, more importantly, it made a mockery of the Ottomans' claim to be the protectors of the Holy Places and thus undermined the legitimacy on which the dynasty's rule rested. Unsurprisingly, the Sublime Porte took this situation very seriously. Accordingly, as we saw earlier, it ordered Mehmet Ali to recapture the Hijaz and to put an end the Wahhabi threat. Though resistance was fierce, the Egyptian ruler succeeded in doing so. His troops not only restored the Holy Places to Ottoman control in 1811 but, after further hard fighting, managed to bring the Saudi state to heel in 1818.⁹⁹

Despite its military defeat, however, Wahhabism remained a potent ideology. It persevered under the protection of a revived Saudi state, and maintained a small but committed following of Muslims who believed that the umma had lost its way. Later, its austere, Salafist conception of Islam, previously shunned by most believers, would begin to acquire a degree of theological legitimacy. Indeed, after the turn of the twentieth century, growing numbers of Sunni Muslims would come to view it not merely as a way to purify and revive Islam but also as a prescription for how—through scrupulous adherence to the canonical understanding of the traditions of the early *umma*—believers could reform their society and meet the growing challenge that European imperialism posed to the Muslim world. Wahhabism thus offered a very different model for addressing the European threat than that advanced by the top-down,

⁹⁷ Madawi al-Rasheed, *A History of Saudi Arabia*, Second Edition (Cambridge: Cambridge University Press, 2010), 17–18.

⁹⁸ Sayed Khatab, *Understanding Islamic Fundamentalism: The Theological and Ideological Basis of Al-Qa'ida's Political Tactics* (Cairo: The American University in Cairo Press, 2011), 74–75; Aslan, *No God but God (Updated Edition)*, 248.

⁹⁹ al-Rasheed, *A History of Saudi Arabia*, 23.

Westernizing reformers in Istanbul and Cairo—one that stressed that the Muslim world could parry the West not by imitating it but instead by regaining God's trust through a return to the practices of the *salaf*.¹⁰⁰

Islamic Modernism

Salafism also shaped the ideas of the Islamic Modernist movement that emerged in the late-nineteenth century. Like the Wahhabi, Islamic modernists such as the Egyptian Muhammad Abduh (1849-1905) were critical of those who called for Muslims to abandon their cultural traditions in favor of the wholesale adoption of Western ideas and values. They also agreed that Muslims needed to revive the practices and principles of the first few generations of believers. However, while the Wahhabis wanted to do so in order to reestablish a society based on the conservative religious scholars' canonical understanding of the early *umma*, the modernists instead sought to return to the values of the first Muslims for the purpose of *freeing* the contemporary Islamic world from what they saw as the suffocating tyranny of the superstitious and close-minded 'ulama'.¹⁰¹

Islamic Modernism's emergence as a powerful force in the Middle East stemmed largely from the work of the Iranian-born writer and political activist Jamal al-Din al-Afghani. His early life gave no hint of his later radicalism. He received a typical Shi'i education, followed the conventional path of an affluent Iranian, and held unremarkable views. His understanding of the world changed with great suddenness, however, when he witnessed Britain's brutal response to the anti-imperialist Sepoy Rebellion of 1857 while a student in India. Shocked by the violence with which the British suppressed the revolt, he experienced a dramatic and sudden transformation into a committed, lifelong foe of imperialism—a change that led him to think deeply about how the Muslim world could most effectively respond to the rising threat that the West posed.¹⁰²

He began by first seeking to understand how Islamic civilization had fallen so far behind the West. To al-Afghani, the fact that Europe had managed to vault ahead was puzzling in light of the Muslim world's onetime enormous lead in commerce, science, and philosophical thinking. During the Arab Empire's golden age, he noted, Islam's intrinsic rationalism and embrace of reason had made the Caliphate the most scientifically and intellectually advanced civilization on earth. Indeed, modern Western science and technology—however impressive it may have been—ultimately rested on earlier Muslim breakthroughs. So what had happened? How had the Islamic world fallen behind?¹⁰³

In al-Afghani's view, the answer was clear. He argued that the key moment in the Muslim world's intellectual decline had occurred in the thirteenth century when the

¹⁰⁰ Ansary, *Destiny Disrupted*, 254–57.

¹⁰¹ Aslan, *No God but God (Updated Edition)*, 237–38.

¹⁰² Ansary, *Destiny Disrupted*, 261–62.

¹⁰³ Esposito, *Islam*, 154–56.

religious scholars had exploited the crisis atmosphere that the Mongol irruption had created to entrench themselves as the unquestioned arbiters of Islam. Emphasizing *taqlid*—the uncritical acceptance of existing views—and calling for the closing of the gates of *ijtihad*, the superstitious religious scholars had effectively barred any effort to question their stifling, canonical understanding of scripture or religious law. Having secured a monopoly over how people understood the Qur'an and the shari'a code—and thus over what kind of thinking was permissible—they followed by imposing rigid limits on scientific and intellectual inquiry that left little room for further discoveries. As a result, al-Afghani concluded, Islamic civilization had become an intellectual backwater that remained stuck in a Medieval mindset while the once intellectually inferior Europeans had surged ahead.¹⁰⁴

But how could the Muslim world end the monopoly that the 'ulama' had established over understandings of the shari'a code and the Qur'an? Al-Afghani's answer lay, ironically, in going back to the values of the first three generations of believers—the source of the conservative religious scholars' control of Islam. Doing so, he contended, would permit Muslims to retake control of the past from the 'ulama' and to purify Islam of the religious scholars' superstitious and anti-intellectual beliefs, thus permitting people to once again use the religion's intrinsic rationalism to interpret the shari'a code.¹⁰⁵

Al-Afghani's leading disciple, the Egyptian legal scholar Muhammad Abduh, was particularly vociferous in arguing for a return to the Islam of the Salaf as a way of breaking the dominance of the religious scholars. To Abduh, the religion of the first Muslims was not the rigid or fixed system that the 'ulama' described, but instead a dynamic set of vital moral and intellectual guidelines that people could revisit and reinterpret to meet the needs of each age. By retaking control of the past from the religious scholars, he continued, Muslims could reopen the gates of *ijtihad* and thus engage in a much-needed critical reinterpretation of the sources of religious law—the *qiyas*, *ijma*, sunna, and even the Qur'an—aimed at bringing the shari'a code in line with the needs of the contemporary world. Free to reason once again, Muslims would thereafter achieve the scientific and technological breakthroughs that had in earlier times been the hallmark of Islamic civilization and, in so doing, would erase the gap in power that had opened between the West and the Middle East.¹⁰⁶

The Islamic Modernists did not reserve their barbs exclusively for the 'ulama', however. Instead, they were also highly critical of advocates of Westernization such as the Tanzimat reformers, who, as we have seen, called not just for the adoption of

¹⁰⁴ Christopher De Bellaigue, *The Islamic Enlightenment: The Struggle Between Faith and Reason, 1798 to Modern Times* (Liveright, 2018), 202–3; Aslan, *No God but God (Updated Edition)*, 235.

¹⁰⁵ Esposito, *Islam*, 154–56.

¹⁰⁶ Albert Hourani, *A History of the Arab Peoples* (Cambridge, Mass: Belknap Press of Harvard University Press, 2002), 307–8; Aslan, *No God but God (Updated Edition)*, 236–38.

Western technologies, but also for the wholesale restructuring of Islamic societies along European lines. The Islamic Modernists agreed that Muslims could and should adopt European technologies and build on Western science. Indeed, they were forthright in arguing that Muslims could counter the Western threat only by adopting European weapons, commercial methods, and factories. However, the modernists vigorously disagreed with the notion that the Islamic world also had to adopt European social arrangements and values. On the contrary, they insisted that Muslims needed to fit Western technologies and science strictly within the context of traditional Islamic values such as social justice. To do otherwise was to court disaster, as the calamitous efforts to remake Egypt along Western lines had made all too clear.¹⁰⁷

The Islamic modernists applied similar reasoning to political issues. Most notably, like the Young Ottomans, Abduh sought to develop a form of democracy that was rooted in Islamic values rather than in the Western ideas that had informed the Tanzimat reformers' program. Toward that end, he recast *ijma*—the idea of consensus that the 'ulama' had for so long used to uphold their monopoly over religious questions—to mean popular sovereignty, and he joined the Young Ottomans in reconceiving *shura* to mean self-rule; in so doing, he gave Muslims the foundation on which they could construct an indigenous system of representative government—one rooted in Islamic values rather than in the Western Enlightenment.¹⁰⁸

Conclusion

Reform efforts of the nineteenth century such as Islamic Modernism cast long shadows in the twentieth and twenty-first centuries. Many later movements directly or indirectly found inspiration in their ideas, and some even drew specifically on their programs. For instance, al-Afghani's ideas were instrumental in shaping the views of Hasan al-Bana (1906-1949), the founder of the Egyptian Muslim Brotherhood, while Abd al-Wahhab's puritanical understanding of Islam provided the ideological underpinnings of the Saudi state that would win control of Arabia in the early-twentieth century and shaped the views and the ideas and practices of twenty-first century Jihadi groups such as the Islamic State of Iraq and the Levant (ISIL). Meanwhile, the Tanzimat reformers' top-down approach to modernization would influence the efforts of Arab nationalist such as Egyptian President Gamal Abdel Nasser (r. 1954-1970) to remake their countries after World War II.

While they may have been influential on later social and political movements, however, the reform efforts of the nineteenth century ultimately failed to achieve the goals that their proponents had sought. They had proven unable either to end the imbalance between the Muslim Middle East and the industrializing West or to arrest the Islamic world's continued relative decline. Indeed, despite the efforts of the reformers, European exploitation of the Ottoman Empire accelerated in the early-twentieth century. Ultimately, the inability of the nineteenth-century reform movements to halt the decline

¹⁰⁷ Ansary, *Destiny Disrupted*, 267.

¹⁰⁸ Lapidus, *A History of Islamic Societies*, 516–17; Ansary, *Destiny Disrupted*, 263–65.

of the empire dimmed their appeal and opened the door to the emergence of new, secular-nationalist movements such as Turkism and, later, Arab nationalism that would displace the earlier reform efforts as the dominant modes of resistance to continued Western encroachment in the Middle East. It is to the appearance of these new ideologies and to the global conflict that would nurture them that we shall next turn.

Chapter Nine: The Twilight of the Ottoman Empire, 1878-1916

The four-decade period that followed Sultan Abdülhamit II's (r. 1876-1909) suspension of the constitution was a time of dizzying change during which the Ottoman Empire struggled to secure its territorial integrity in the face of threats from within and without. Paradoxes abounded. The government formally abandoned the Tanzimat program in favor of a return to traditional Ottoman practices while also quietly pressing forward with a remarkably similar modernization program. Likewise, the liberal revolution that the sultan's subjects jubilantly cheered in 1908 seemed for a time to herald a new, democratic dawn but instead paved the way for a secretive cabal to transform the empire into a one-party state that was even more autocratic than that of the recently deposed sultan. Perhaps most obviously, while the Ottoman military endured demoralizing setbacks at the hands of Italy and the small, impoverished nations of the Balkans in the early 1910s, it followed by winning a number of significant victories over the world's most powerful empire just a few years later.

There was nothing paradoxical about the effects of nationalism on the Ottoman Empire, however. Despite the Tanzimat reformers' promotion of the supranational concept of Ottomanism, nationalism increasingly eclipsed all other constructions of identity in the region in the late-nineteenth and early-twentieth centuries. Indeed, at an accelerating pace, the sultan's subjects gradually abandoned not only the supranational identities that the government promoted in favor of new nationalist identities rooted in shared ethnicity and language. For many, the reason was simple: in their eyes, nationalism was a liberating ideology—one that provided a pathway to a better future outside of an empire that they had come to view as alien. Problematically, however, nationalism had also acquired a dark side. Once seen as a partner to liberalism, it increasingly brought new ethnic rivalries, chauvinisms, and sinister hatreds to the region. The result would be violence, ethnic cleansing, and genocide on a heretofore unimaginable scale.

The Hamidian Age, 1878-1908

Known as the Hamidian Age, the era of Abdülhamit II was emblematic of the contradictions of the late Ottoman Empire. Outwardly, his reign was a time of retrenchment—a period in which the empire demonstratively repudiated modernity in favor of a return to the traditional values and practices of the Ottoman past. It was a very different story beneath the surface, however. There, in direct contrast to the official line emanating from the central government, Abdülhamit II oversaw an efflorescence of innovation and reform that differed little from the social, political, and economic restructuring promoted by the reformers and Young Ottomans whom he had so unceremoniously shoved aside.

From the time that he suspended the constitution in 1878, Abdülhamit II had made emphatically clear that he intended to abandon the Tanzimat. Publicly disdainful of both democracy and the idea of restructuring along Western lines, he joined

conservatives in arguing that the empire could best respond to the challenge that Europe posed not by modernizing but instead by returning to the traditions and values of the Ottoman golden age. Accordingly, armed with sweeping emergency powers that—ironically—the liberal constitution had accorded him, Abdülhamit II effected a dramatic, reactionary shift in the distribution of power in the Sublime Porte that sidelined the Tanzimat reformers and their Young Ottoman allies. Thereafter, it would be the sultan and his conservative courtiers rather than the liberal bureaucrats who would direct the Ottoman state.¹

Now firmly in control, the new sultan acted aggressively to pursue his reactionary agenda. He began by devising a new conception of identity designed to retain the loyalty of the bulk of his remaining subjects. Acknowledging that Ottomanism had failed to secure the loyalty of the sultanate's Christian subjects, he formally jettisoned that ideology and, with it, the reformers' goal of retaining the allegiance of *all* people living in the empire. In its place, he promoted Pan-Islamism: a narrower, though still supranational, identity designed to provide the ideological glue needed to bind his Sunni subjects to the state on the basis of a common religious identity—an approach that appeared to have a great deal to recommend it now that Sunni Muslims constituted the overwhelming majority of the empire's population.²

Central to his effort to promote Pan-Islamism was the dynasty's longstanding claim that the reigning sultan was also the caliph and thus the head of the global Sunni community. Abdülhamit II believed that this title gave him enormous authority, and that he could use it to achieve two critical goals. First, by emphasizing his standing as Muhammad's successor, he was confident that he could ensure the continued loyalty of his Sunni subjects and forestall the emergence of nationalist movements among them. Second, he believed that publicly promoting his position as caliph would win the religious allegiance of the hundreds of millions of Muslims who lived in the French, British, and Russian Empires—loyalty that he believed would bolster the sultanate by giving him valuable leverage over those states in any future conflict.³

He followed by undertaking a series of high-profile actions designed to burnish his standing among the world's Sunni Muslims. Most notably, aware that the caliph was the traditional protector of the Holy Places, he ordered the construction of the Hijaz railway to facilitate the movement of people to Mecca during the hajj and he lavished fifteen times as much revenue on the province as it generated in taxes to ensure the pilgrims' ongoing health, safety, and comfort. These were significant expenses; given

¹ Donald Quataert, *The Ottoman Empire, 1700-1922, Second Edition* (Cambridge: Cambridge University Press, 2005), 64–65.

² Sean McMeekin, *The Ottoman Endgame: War, Revolution, and the Making of the Modern Middle East, 1908-1923* (New York: Penguin Books, 2015), 29–30; Caroline Finkel, *Osman's Dream: The History of the Ottoman Empire* (New York: Basic Books, 2005), 488–501.

³ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 488–501.

the potential domestic and international benefits, however, they were, so far as the sultan was concerned, well worth the cost.⁴

Abdülhamit II's embrace of his role as caliph and criticism of the Tanzimat reformers earned him a reputation then and later as a reactionary who waged a fierce, if doomed, rearguard action against modernity. A closer look reveals, however, that his call for a return to traditional Ottoman and Islamic practices was a largely rhetorical one that disguised both a genuine embrace of modernization and a deep commitment to the spirit if perhaps not the letter of the Tanzimat reformers' efforts to restructure the Ottoman Empire. His actions made this clear. As sultan, he promoted the rationalization of provincial administration, doubled the number of upper-elementary schools, modernized the army with German help, and aggressively pushed economic development schemes. Most notably, he oversaw the construction of the first significant railways in the empire—the very symbol of modernity in the nineteenth century—and signed an agreement in 1899 granting Germany a concession to build a line linking the capital to Baghdad. Indeed, he was so committed to the development of the rail network that he once urged an advisor to “[j]ust get it built; run it over my back if you must.”⁵

Even Abdülhamit II's promotion of his status as caliph did not constitute a case of retrenchment. On the contrary, it was an example of innovation and a clear break with longstanding dynastic practice. As we saw in chapter six, the Ottomans had claimed to be the caliphs as far back as Selim II's (r. 1512-1520) conquest of Egypt in 1517. From the beginning, however, the dynasty's use of the title had been more notional than real. Successive sultans held it largely as an honorific much in the way that they had also claimed to be the *Kayser-i Rum* or Emperor of the Romans. In other words, as is so often the case with self-conscious efforts to revive long-dead practices, Abdülhamit II's assertion that he was the rightful leader of all the world's Sunnis was less a return to an older tradition than a novel invention—one designed, in this case, to dramatically expand sultanic power.⁶

Political thinking and social attitudes outside of the governing elite also followed currents that ran counter to Abdülhamit II's avowed embrace of tradition. Politically, the sultan's prorogation of the assembly and suspension of the constitution were unpopular and sparked anger rather than mute acceptance. In the wake of those actions, many would come to join new, underground opposition organizations that called for liberalizing the government and for effecting a top-down restructuring of Ottoman society along lines similar to Japan's Meiji Restoration. Meanwhile, in spite of the sultan's promotion of traditional values, social practices continued to evolve dramatically during the Hamidian Age. Three changes stood out. First, as the large number of photographs taken of women with uncovered faces suggests, a growing feminist movement was beginning to challenge longstanding attitudes towards veiling at the very moment that

⁴ Finkel, 488–501.

⁵ Carter V. Findley, *Turkey, Islam, Nationalism, and Modernity* (New Haven: Yale University Press, 2010), 146–58.

⁶ Findley, 146–58.

the sultan was publicly calling for a return to traditional Muslim practices. Second, the technical impossibility of separating people by gender on popular new technologies such as ferries and streetcars made a mockery of Abdülhamit II's efforts to enforce Islamic customs such as segregation along gender lines. Finally, changes to urban social mores during this period reflected the middle class's newfound openness to modern, European cultural practices. Most notably, it was during the Hamidian Age that the middle class abandoned the traditional custom of sitting on pillows in favor of the Western practice of using chairs, tables, and sofas. Thus, despite Abdülhamit II's rhetorical opposition to reform and social innovation, many of the ideas implicit in the Tanzimat reform effort continued to gain traction in the late-nineteenth and early-twentieth centuries—often at the direct, if quiet, urging of the sultan.⁷

The Second Constitutional Period, 1908-1913

Abdülhamit II's long reign finally came to an end early in the twentieth century. To the joy of the growing number of Ottoman subjects who supported democracy, a revolution that took place in 1908 restored both representative government and the constitution that the sultan had suspended in 1878. The result was a brief era of democratic rule known as the Second Constitutional Period—a time that began with the optimistic belief that the empire was at last poised to become a modern state that enjoyed a thriving economy and, importantly, the respect of the European countries. The euphoria that had greeted the revolution quickly faded, however. The new government proved no more capable of guaranteeing the empire's borders in the face of continued European predation than had the Hamidian regime it supplanted and thus soon began hemorrhaging popular support. Ultimately more interested in aggrandizing power than in institutionalizing the democratic constitutionalism that so many wanted, it responded by solidifying its control over the levers of power. As a result, the Ottoman Empire emerged from the Second Constitutional Period not with the representative government that many had wished for, but instead with an undemocratic state dominated at every level by an autocratic and secretive ruling party.

The Young Turks

Organized opposition to Abdülhamit II first emerged in 1889 when a group of military medical students established a secret, revolutionary society called the Committee of Union and Progress (CUP). Popularly referred to as the Young Turks, the CUP soon evolved into a broad umbrella organization that sought to coordinate a number of secular groups that opposed the sultan's rule. Those bodies varied widely in terms of membership and goals. They included organizations created by officials angry that political power had shifted from the bureaucracy to the court, groups formed by non-Turkish Muslims such as Albanians who feared that their territory might be taken by Christian European states, and secret societies established by military cadets who doubted the sultan's ability to maintain the territorial integrity of the empire in the face of

⁷ Findley, 165–84.

the challenge that Russia and the newly independent Balkan states posed. It even included representatives of Armenian nationalist groups.⁸

The CUP faced two significant hurdles that prevented it from constituting an effective opposition during its early years. First, harassed by Abdülhamit II's secret police, the Young Turks could not operate within the empire and instead confronted the difficult task of trying to shape public opinion from exile. Second, and more importantly, the groups that made up the CUP were unified only by their fervent opposition to the sultan and their shared desire to restore the constitution that he had suspended in 1878. In other regards—particularly in terms of the objectives that they sought and the methods by which they hoped to attain them—they differed substantially.⁹

While the constituent groups in the CUP were able to look past their differences during the 1890s, they began to split over them shortly after the turn of the century. The final break occurred in 1902 over a proposal that called for the CUP to overthrow Abdülhamit II with British help after which the organization would transform the empire into a decentralized state that afforded ethnic and religious minorities substantial autonomy. While this plan found support among liberal former bureaucrats and many non-Turkish elements, it was an absolute non-starter for the Turkish organizations in the CUP. Fearing that foreign intervention and decentralization were merely waystations on the road to the empire's eventual partition and liquidation, those groups split from the CUP in protest and formed a new umbrella organization called, confusingly, the Committee of Progress and Union (CPU). Rejecting foreign intervention, it called for the restoration of the constitution and for the transformation of the empire into a highly centralized state.¹⁰

The proponents of decentralization appeared to be the better positioned faction immediately following the Congress of Ottoman Liberals, but it was the CPU that ultimately emerged as the dominant opposition organization. Its success was a function of its clearer understanding of the structure of political power in the empire. Concluding that no revolutionary movement could hope to overthrow the sultan without the support of the military, the group began to make contact with sympathetic officers in the Ottoman army. Their efforts bore fruit in 1907 when the CPU absorbed the secret Ottoman Freedom Society, a revolutionary organization made up of Turkish officers in the Third Army stationed in Macedonia. Now enjoying military backing, the CPU was finally in a position to move against the government. Accordingly, over the next year, it

⁸ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 504–5; Findley, *Turkey, Islam, Nationalism, and Modernity*, 161–62.

⁹ Erik J. Zürcher, *Turkey: A Modern History*, Third Edition (London: I.B. Tauris, 2004), 87–89.r

¹⁰ M. Şükrü Hanioglu, *Preparation for a Revolution: The Young Turks, 1902-1908* (New York: Oxford University Press, 2001), 3–5.

began laying plans for a revolution in which the Third Army would first seize control of Macedonia and then force Abdülhamit II to restore the system of constitutional rule.¹¹

The Revolution of 1908

Two events impelled the CPU to launch its revolution prematurely. First, it worried that a meeting between Russian Tsar Nicholas II (r. 1894-1917) and British King Edward VII (r. 1901-1910) presaged the impending partition of the empire. Second, the CPU received word that Abdülhamit II's secret police were readying an imminent preemptive action against it in Macedonia. Fearing disaster if they failed to act, the organization's leaders responded by ordering sympathetic elements of the Third Army to rise up against the Hamidian regime in July 1908. Quickly securing control of Macedonia, the revolutionaries proclaimed the reestablishment of the constitution on July 23 and began marching on Istanbul. The end came swiftly. Bowing to the inevitable, the sultan declared the next day that he was restoring the constitution and recalling parliament. His announcement met with an enthusiastic reception among his subjects. As elated as they were stunned, they poured into the streets where they organized a series of huge, spontaneous mass celebrations throughout the empire. A new day appeared to have dawned.¹²

The popular response to the restoration of democratic rule created a critical misperception about the events of July 1908. Bewitched by the euphoric crowds cheering for liberty, fraternity, justice, and equality, Western observers and Ottoman subjects alike assumed that a mass democratic movement had sparked the revolution and concluded that the Ottoman Empire was en route to becoming a pluralist, liberal state. Though widespread, this interpretation was far off the mark. As the fact that the celebratory crowds emerged only *after* the sultan had announced the reinstatement of the constitution makes clear, the revolution was not a popular movement effected by a broad, democratic coalition. Instead, it was a narrow military insurrection organized by a secretive, conspiratorial society. More significantly, the CPU—which, confusingly, had reappropriated the CUP name—did not revolt in order to reform the empire along liberal lines or to promote democracy as many people had assumed. Instead, it had seized power in pursuit of a fundamentally conservative objective: the maintenance of the Ottoman Empire's territorial integrity in the face of both foreign aggression and the challenge of internal nationalist movements. In other words, despite appearances, the revolution was not a popular uprising, but instead a `military putsch—albeit one clothed in the verbiage of liberal constitutionalism.¹³

¹¹ M. Şükrü Hanioglu, *A Brief History of the Late Ottoman Empire* (Princeton: Princeton University Press, 2008), 147–48.

¹² McMeekin, *The Ottoman Endgame*, 45–47.

¹³ Hanioglu, *A Brief History of the Late Ottoman Empire*, 147–49.

The Struggle for Power

In the meantime, unaware of the CUP's reactionary ideology, middle-class liberals and others threw their support behind the party. Indeed, the restoration of constitutional rule had created a honeymoon period during which the Young Turks enjoyed overwhelming popular backing—sufficient to give them a majority of seats in parliamentary elections held in December 1908. By that point, however, the loss of additional Ottoman territory was already beginning to eat into the party's support. Seizing the opportunity that the revolution presented, a group of predatory states had moved against it in a series of coordinated actions in October 1908. Austria-Hungary annexed Bosnia-Herzegovina, Greece took formal possession of Crete, and Bulgaria declared its full independence. It is true that these were paper losses: all of those territories had been under foreign rule for some time and had been connected to the empire by only the thinnest of legal fictions. Nonetheless, this new round of contraction was demoralizing and seemed to demonstrate that the Young Turks were no more capable than Abdülhamit II had been at preserving the Ottoman state's territory. Reeling from the domestic fallout of these diplomatic defeats, the CUP thereafter struggled to maintain its dominant political position in the face of competing challenges from a diverse array of groups including conservative religious scholars, liberal proponents of civil liberties, and ethnic minorities who, with growing volume, demanded immediate autonomy or outright independence.¹⁴

The Young Turks responded by gradually abandoning their already lukewarm commitment to constitutionalism and by moving to solidify their grip on the government. They secured their right flank in April 1909 by successfully putting down a reactionary “counter coup” organized by a combination of ‘ulama’, madrasa students, and non-commissioned military officers and by using the failed revolt to compel Abdülhamit II to abdicate in favor of his more pliant brother, Mehmet V (r. 1909-1918). With the sultan and the religious conservatives contained, they next moved to curb the threat that the liberals posed. Using the counter coup as a pretext, they enforced limits on public assembly, banned strikes, and curbed press freedoms in order to weaken advocates of reform; they also made abundant use of violence and electoral fraud to score a resounding victory over the Entente Libérale Party in the 1912 parliamentary elections.¹⁵

In so doing, however, the Young Turks badly overplayed their hand. While their efforts to limit popular protests, to ban strikes, and to censor the press had been successful, their use of vote rigging and intimidation in what came to be known appropriately as the “big stick election” was so glaring that it produced a powerful political counterreaction. Dismayed by the CUP's disregard for democratic norms, constitutionalists and disaffected army officers came together immediately after the vote

¹⁴ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 513–15.

¹⁵ McMeekin, *The Ottoman Endgame*, 197–98.

to force the dissolution of parliament and to compel the formation of a new cabinet composed of liberals. As a result, the reign of the Young Turks appeared to be over.¹⁶

The Italo-Turkish War and the Balkan Wars, 1911-1913

Thanks to a series of wars fought between 1911 and 1913, however, the CUP managed to achieve a remarkable reversal of fortune. The first of those conflicts, the Italo-Turkish War, began in 1911 when Italy invaded the last Ottoman possessions in Africa, the provinces of Tripoli and Cyrenaica in modern-day Libya, in a bid to enlarge its empire. The ensuing war did not go well for the Ottomans. Despite the organization of a successful guerrilla resistance movement by army officers, the Italian navy's control of the supply lines to Libya compelled the liberal government of Kamil Pasha (1833-1913) to grudgingly agree to cede Tripoli and Cyrenaica to Rome in October 1912.¹⁷

Italy was far from the only vulture circling the empire. Immediately on the heels of its defeat in Libya, the Sublime Porte faced a new and more-dangerous threat much closer to home when Bulgaria, Serbia, Montenegro, and Greece invaded the empire in an effort to strip it of its remaining European territories. The combination of so many enemies overwhelmed the Ottoman military. By the time a ceasefire brought the First Balkan War to a halt in December 1912, the empire had lost Macedonia and most of Thrace and was in imminent danger of surrendering the besieged city of Edirne—the empire's symbolically important former capital.¹⁸

These military defeats prodded the Young Turks into action. Fearing that Kamil Pasha was going to yield Edirne in peace negotiations, a group of CUP officers led by Enver Pasha (1881-1922) and Talat Pasha (1874-1921) launched a bloody coup called the Raid on the Sublime Porte in January 1913 that deposed the liberal cabinet and replaced it with a CUP-controlled government committed to pushing on with the war. The decision to continue the conflict was popular at first and resulted in a surge in support for the CUP. However, the Young Turks proved no more capable of defending the empire than had the liberals they had overthrown. Indeed, rather than producing an Ottoman victory, the decision to abandon diplomacy in favor of a continuation of the war resulted in the surrender of Edirne in March 1913 and left the empire vulnerable to further territorial losses.¹⁹

Reeling from the defeat, the CUP had no choice but to return to the bargaining table. The resulting agreement reflected the Ottoman Empire's feeble position. Signed in May 1913, the Treaty of London cost the empire the city of Edirne, four-million subjects, and 155,000 square kilometers of territory including nearly all of its remaining

¹⁶ McMeekin, 51–54; Findley, *Turkey, Islam, Nationalism, and Modernity*, 197–98.

¹⁷ Eugene Rogan, *The Fall of the Ottomans: The Great War in the Middle East*, Reprint edition (Basic Books, 2015), 14–18.

¹⁸ Richard J. Evans, *The Pursuit of Power: Europe 1815-1914* (New York, New York: Penguin Books, 2016), 693–95.

¹⁹ Rogan, *The Fall of the Ottomans*, 18–21.

possessions in the Balkans. Given these harsh terms, the agreement was a crushing political blow for the CUP. Saddled with responsibility for the treaty, it quickly lost all of the support that its decision to continue the war had gained it in the winter. Once again, as a result, it appeared doomed to political irrelevance.²⁰

Surprisingly, however, the Young Turks managed to stage a remarkable comeback in the summer of 1913. Their revival was a product of both the party's domestic initiatives and the vicissitudes of international relations. Internally, the Young Turks broke the power of the liberal opposition by using the assassination of popular Grand Vezir Mehmet Sevket Pasha (1856-1913) as a pretext to arrest a number of Entente Libérale Party leaders and to force most of the rest into exile. Internationally, meanwhile, they scored a rare international victory when they managed to retake some of the territory that the empire had lost during the First Balkan War. They were able to do so not because of Ottoman military prowess but instead because the victors of that conflict had begun falling out over the spoils almost before the ink had dried on the Treaty of London. The result was the outbreak in June 1913 of another conflict: the Second Balkan War. Pitting an overmatched Bulgaria against a coalition that included Montenegro, Romania, Serbia, and Greece, the new conflict created an opportunity that the Young Turks were quick to seize. Exploiting Bulgaria's isolation, they ordered the army to reoccupy much of the territory in Thrace that they had recently ceded in the Treaty of London. Facing no meaningful opposition, Ottoman troops easily retook control of the region including, most importantly, the symbolically important city of Edirne. It was a substantial—and welcome—change of fortune. Long conditioned to expect military and diplomatic defeat, the Ottoman public was ecstatic upon receiving news of the victory that the Young Turks had secured. As a consequence, support for the CUP soared to heights not seen since the heady days of the revolution.²¹

With its position temporarily solidified, the CUP followed by mounting a successful effort to institutionalize a system of one-party rule. It did so by making advancement in the bureaucracy contingent on loyalty to the party's shadowy central committee and, more importantly, by creating an extraconstitutional, CUP-controlled apparatus that paralleled the government's official administrative structure. Giving the Young Turks unprecedented control over the workings of the state, this mechanism proved invaluable to Enver, Talat, and Djemal (1872-1922), the triumvirate of men who took control of the cabinet in a shakeup in 1913. Known as the Three Pashas, these members of the CUP inner circle thereafter used it to rule through a series of nominally temporary decrees that permitted them to sideline parliament.²²

Meanwhile, the Young Turks also moved to strengthen the empire militarily and diplomatically in an effort to prevent any further territorial loss. Here, they had at best mixed success. On the one hand, the order for two modern, dreadnought battleships from Britain in 1913, the *Sultan Osman I* and the *Reshadieh*, and the establishment of a

²⁰ Rogan, 18–21.

²¹ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 524.

²² Hanioglu, *A Brief History of the Late Ottoman Empire*, 159–63, 174–75.

German military mission under General Otto Liman von Sanders (1855-1929) in 1913 promised to go a long way toward improving the empire's military capabilities. On the other, despite their vigorous efforts, the Three Pashas were unable to secure a defensive pact with one of the Great Powers that could guarantee the empire's territorial integrity—hardly a favorable sign for a state facing the threat of further dismemberment. Even Germany, which had enjoyed close diplomatic relations with the Hamidian regime, was unwilling to discuss a binding alliance with an empire that many believed was on its last legs.²³

Nationalism Comes to the Middle East

Since the early nineteenth century, the ideology of nationalism that had produced the Balkan Wars had constituted the driving force behind the Ottoman Empire's loss of European territory. Whether used by indigenous elites or manipulated by Great Powers such as tsarist Russia, it provided a powerful new conception of identity that had fueled independence movements throughout the empire's Balkan territories. Indeed, the impact of nationalism on the region had been so enormous that, by 1914, the sultanate's European possessions had been reduced to little more than Istanbul and its immediate hinterlands.

In contrast, nationalism had a much more difficult time establishing roots in the Middle East. Indeed, it only began to percolate into the region during the Hamidian Age and the Second Constitutional Period, and, even then, commanded far less popular support there than it did in the Balkans. What explains this discrepancy? Why was it not as popular in the empire's Middle Eastern provinces than it was in its European territories? To a substantial degree, nationalism's more-limited penetration of the Middle East stemmed from the enduring appeal in the region of supranational forms of identity such as Ottomanism and, especially, Pan-Islamism. With deep roots among the empire's Muslim subjects, Pan-Islamist notions of identity resonated far-more deeply with the people of the region than did ethnically and linguistically based appeals to national identity. As a result, on the eve of the First World War, nationalism in most of the Middle East had no mass following and remained confined largely to small groups of intellectuals and cultural leaders—particularly among minority groups that faced discrimination or the threat of organized violence.

Armenian Nationalism

The Armenians were one such group. Experiencing stepped-up violence during the late-nineteenth century, they were the first people in the Middle East to develop a full-blown nationalist movement. It began among expatriate communities in Europe where the gravitational influence of Western ideas of nationalism was strongest. There, Armenian activists developed two rival organizations: the Dashnak Party in Russia and the Hunchak Party in Switzerland. The two groups differed largely in terms goals. Aware that the Armenians did not constitute a majority in any part of Anatolia, the more

²³ McMeekin, *The Ottoman Endgame*, 84–86.

moderate Dashnaks accepted continued Ottoman rule as inevitable and pressed for greater autonomy within the empire. In contrast, the more-radical Hunchaks demanded nothing less than outright independence.²⁴

Both the Dashnaks and Hunchaks used a combination of approaches to advance their goals. Many of their methods were peaceful. For example, both groups organized demonstrations calling for autonomy and pressed for Armenians to have the right to possess arms so that they could defend themselves from the depredations of local Muslims—particularly from the nomadic Kurds who had long extorted protection money from them. At the same time, however, the two organizations were not afraid to use violence to advance their objectives. Indeed, both took actions designed to provoke Ottoman responses of such overwhelming violence that they would compel the international community to intervene on the Armenians' behalf. In this effort, they were partially successful. Dashnak and Hunchak terrorist operations sparked massive reprisals by mobs of Muslims in eastern Anatolia and Istanbul between 1894 and 1896 that resulted in the deaths of tens of thousands of Armenians. While they were able to arouse violence against the empire's Armenian population, however, these actions failed to secure the sustained Western intervention that the nationalists sought. Worse, at times—such as when the Dashnaks seized a Western-owned bank in Istanbul in 1896—the violence even alienated the European states whose help the Armenian nationalists hoped to win.²⁵

Shortly after the turn of the century, as a consequence, both the Hunchaks and Dashnaks felt compelled to abandon violence. In its place, they began to cooperate with the more liberal elements in the CUP in hopes of achieving greater rights for their people if and when the Young Turks managed to depose Abdülhamit II. Their moderate strategy appeared to pay off with the revolution of 1908. Not only did the Armenians gain greater liberties and privileges as a result of the restoration of the constitution, but they even managed to elect fourteen representatives to the new parliament. For a time, as a result, it seemed as though the Armenians would finally be able to secure real autonomy within the empire.²⁶

Violence soon destroyed that sense of hopefulness, however. Most notably, Turkish mob attacks in the city of Adana in the aftermath of the counter coup in 1909 resulted in the massacre of as many as 30,000 Armenians. It was a sobering moment for the Armenian nationalist groups. The sheer scale of the violence threw cold water on the idea that the revolution marked the start of a new era in intercommunal relations.

²⁴ Findley, *Turkey, Islam, Nationalism, and Modernity*, 142–44.

²⁵ Rogan, *The Fall of the Ottomans*, 9–14; Finkel, *Osman's Dream: The History of the Ottoman Empire*, 502–3.

²⁶ Rogan, *The Fall of the Ottomans*, 13–14.

More importantly, it dashed once and for all any hope that the Armenians would be able to achieve genuine autonomy within the structure of the Ottoman Empire.²⁷

From Ottomanism to Turkism

In contrast to Armenian nationalism, Turkish nationalism, or Turkism, emerged only very gradually in the years leading up to the First World War. That it was slow to develop was largely a function of the fact that the governing elite had not endorsed it. Instead, seeking to preserve the empire, successive regimes had promoted broader forms of identity. For example, the Tanzimat reformers had championed Ottomanism as a way to bind all of the empire's ethnic and confessional groups together through shared loyalty to the dynasty. Abdülhamit II's Pan-Islamism was narrower in scope, meanwhile, but similarly sought to provide a conception of identity sufficiently broad that it could subsume particularist forms of national identification and thus preserve the empire.²⁸

Once in power, the Young Turks picked up where their Tanzimat forebears had left off and officially espoused the ideology of Ottomanism as a unifying principle. Privately, it is true, growing numbers of Young Turks rejected Ottomanism in favor of Turkism. Influenced by the intellectual Yusuf Akcura's (1876-1935) contention that Ottomanism was a dead end in a world increasingly defined by national identity, they believed that the Turks needed to abandon the effort to maintain a multi-ethnic empire in favor of a focus on establishing a unified and homogeneous Turkish state. During the first few years of the Second Constitutional Period, however, they felt compelled to keep such views to themselves. Instead, ruling an empire in which Turks did not constitute a majority, the adherents to Turkism within the CUP saw utility in downplaying their nationalist sympathies and thus continued to publicly endorse Ottomanism.²⁹

The disastrous Balkan Wars radically altered the logic that underlay this approach. Rampant desertion by ethnic Greek and Bulgarian troops in the Ottoman army combined with the victors' brutal campaigns of ethnic cleansing in Macedonian and Thrace—the winners expelled over 300,000 Muslims immediately following the conflict—inflamed the empire's Turkish population and produced a surge of Turkish-nationalist sentiment. At the same time, the loss of nearly all of the empire's European territory meant that Turks constituted a majority of Ottoman subjects for the first time since the early-fifteenth century. As a result, while the CUP continued to officially espouse Ottomanism, it increasingly conflated that idea with Turkism. State policy reflected this change. Thereafter, the government promoted Turkish nationalist policies such as requiring that education and court proceedings be conducted in the Turkish language and adhered to a new construction of Ottomanism wherein ethnic Turks would

²⁷ Rouben P. Adalian, "The Armenian Genocide," in *Centuries of Genocide: Essays and Eyewitness Accounts*, ed. Samuel Totten and William Spencer Parsons, Fourth Edition (New York: Routledge, 2013), 131–32.

²⁸ Hanioğlu, *A Brief History of the Late Ottoman Empire*, 147.

²⁹ Hanioğlu, 147.

hold a dominant cultural and political position in the empire similar to that occupied by Germans in the Austro-Hungarian Empire. The Young Turks also began to cast Anatolia as not merely an important part of the empire, but as its Turkish heart—a mental remapping that bode ill for the millions of Armenians who lived in that region.³⁰

Zionism

Zionism, the movement dedicated to the establishment of a Jewish nation-state in the biblical land of Canaan, arose in Europe at the same time that Turkism took root in the Ottoman Empire. Named for a hill in Jerusalem that had come to symbolize the promised land, Zionism emerged in response to the shifting circumstances that the Ashkenazi, or Central and Eastern European Jewish people, had experienced in the 1800s. During the early part of the century, their position had improved dramatically. Thanks to the egalitarian ideas of the French Revolution, the Ashkenazi had enjoyed rapid improvements in their social, economic, and political circumstances. Napoleon Bonaparte (r. 1804-1814, 1815) was central to this progress. Dominating or outright ruling much of Europe during his reign, he had compelled most of the Continent's states to follow France's lead in emancipating the Jews. Thereafter, Jewish people no longer had to wear clothing that identified their religion, could live outside the ghettos to which urban Jews had long been confined, and were free to enter the professions. It was a heady time for the Ashkenazi people. Enjoying new freedoms and the economic benefits that the Industrial Revolution had brought, many in the growing Jewish middle-class concluded that the bad days of outright discrimination and pogroms—state sponsored riots—were behind them.³¹

A series of high-profile anti-Semitic incidents at the turn of the century soon disabused them of that notion, however. The Dreyfus Affair—the wrongful conviction of a Jewish French army officer in 1894—and the election of the openly anti-Semitic Karl Lueger (1844-1910) as mayor of Vienna in 1897 made clear that a new and more insidious form of anti-Semitism based not on religion but instead on race had emerged as nationalism's handmaiden. Everywhere in Europe, it seemed, nationalists were constructing the Jewish people as the Other—those who did not belong within the uniform “imagined community” of the nation-state.³²

As bad as it was in Western and Central Europe, the situation was far bleaker in the Russian Empire. Home to more than five million of Europe's nine-million Jews, Russia had not followed the other European states in emancipating the Jewish people. Instead, it continued to confine them to the Pale of Settlement: the region of far-western Russia where they resided in urban ghettos, or, more commonly, in poor agricultural

³⁰ Mustafa Aksakal, “The Ottoman Empire,” in *Empires at War: 1911-1923*, ed. Robert Gerwarth and Erez Manela (New York: Oxford University Press, 2014), 21; Rogan, *The Fall of the Ottomans*, 23–24; Hanioglu, *A Brief History of the Late Ottoman Empire*, 166–87; Findley, *Turkey, Islam, Nationalism, and Modernity*, 202–3.

³¹ Evans, *The Pursuit of Power*, 475–76.

³² Evans, 475–76.

communities called *shtetls*. Barred from most professions, facing strict quotas for admission to universities, and lacking steady employment, they lived in abject poverty. Astonishingly, as bad as life had been for them in mid-nineteenth-century Russia, it grew dramatically worse following the assassination of Tsar Alexander II (r. 1855-1881) in 1881. Blaming the Jews for his death, Russian mobs took part in a series of huge pogroms that resulted in the murder of thousands of Jewish people and in the large-scale destruction of Jewish-owned property.³³

In response, a nascent Zionist movement arose in Russia. Inspired by the Russian physician Leo Pinsker's (1821-1891) Zionist pamphlet *Auto Emancipation*, a trickle of Jewish people began to leave Russia for Palestine following Alexander II's assassination in a wave of migration called the First *Aliyah*, a term derived from the Hebrew word for "going up . . . to Jerusalem." This early Zionist effort was not very successful, however. It involved only small numbers of people, and it lacked the financial reserves needed to adequately support those who emigrated. As a result, while several thousand Russian Jews managed to relocate to Palestine in the 1880s and 1890s, a large proportion of them could not overcome the lack of resources or the harsh environment and soon left. Zionism appeared to be stillborn.³⁴

The movement gained a new life thanks to the writings of Theodore Herzl (1860-1904). Widely considered the father of Zionism, Herzl was a middle-class Viennese journalist who initially championed integration and assimilation in Europe. His views underwent a rapid change in the 1890s, however, as a result of his experience covering the Dreyfus Affair. Shocked by the anti-Semitism that the episode had revealed, he came to the conclusion that the Jewish people would never gain genuine acceptance and safety in Europe, and he began to argue that they could only ensure their security by establishing their own nation-state. But where should they construct it? To Herzl, the answer was clear: they should do so in the Jewish people's historical homeland of Palestine. Hoping to persuade European Jews to support Zionism, he put these ideas into words in 1896 in the influential book, *Der Judenstaat*, or *The Jewish State*. Pleased with the interest it generated, he followed by helping to organize the First Zionist Congress in Basel, Switzerland the next year. There, the assembled delegates established the Zionist Organization, which, under Herzl's direction, lobbied European governments for support and purchased land in Palestine on which pioneers could settle using funds provided by wealthy Jews.³⁵

At the same time that Herzl and the Zionist Organization were creating the necessary infrastructure to sustain the movement, a sudden upsurge of anti-Semitism in Russia persuaded a large number of Jewish people to leave the tsarist empire for

³³ Walter Laqueur, *A History of Zionism: From the French Revolution to the Establishment of the State of Israel* (New York: Schocken, 2003), 56–60.

³⁴ Michael Stanislawski, *Zionism: A Very Short Introduction* (New York: Oxford University Press, 2017), 17–20.

³⁵ Stanislawski, 22–31 Note, though, that scholars have recently cast doubt on Herzl's claim that the Dreyfus Affair sparked his change of heart regarding assimilation.

Palestine. Once more, political instability in Russia lay at the heart of the rise in anti-Jewish sentiment. Defeat in the Russo-Japanese War in 1904 and the outbreak of a revolution in 1905 had sparked a series of brutal pogroms throughout the empire. In response, thousands of the tsar's Jewish subjects decided to make a better life for themselves by moving to the Holy Land in a wave of immigration known as the Second *Aliyah*. They were far-more successful in establishing a permanent Zionist presence in Palestine than their predecessors who had come a generation earlier. Now able to draw on the substantial financial resources of the Zionist Organization, the immigrants of the Second *Aliyah* were able to establish the sinews of a Jewish nation there. They founded settlements such as the city of Tel Aviv, started communal farms called *kibbutzim*, and revived Hebrew as the language of Jewish people living in Palestine. Above all, they turned the idea of a Jewish community in Palestine from a dream into a reality; indeed, without the Second *Aliyah*, Zionism may well have died out.³⁶

Despite this success, however, Zionism very much remained a niche movement. Most Central European Jewish people were indifferent to it and expressed no interest in leaving their comfortable lives for the impoverished and undeveloped Holy Land. Others outright opposed Zionism out of fear that it would undermine the rights and status that Jewish people had gained in Europe over the course of the nineteenth century. More importantly, Palestine simply compared unfavorably to the other, more-attractive options that beckoned to those who sought to leave Europe. Indeed, of the staggering 2.5 million Jewish people who emigrated from the Russian Empire between 1904 and 1913, only 60,000 were willing to endure the difficult conditions in Palestine. The overwhelming majority instead opted to move to the *Golden Medene*—the Golden Land of the United States.³⁷

There was another problem as well. Palestine was not an unoccupied land available for colonization, but instead a region that had a substantial population of Christian and Muslim Arabs. Herzl and his allies were well aware of this issue. Indeed, there is a famous tale of two rabbis who, sent by the Zionist Organization to Palestine to assess its suitability as a homeland for the Jewish people, reported back after a few weeks that “[t]he bride is beautiful, but she is married to another man.” The story is of course fictive, but, like many apocryphal accounts, it also conveys an important truth. That is, it accurately captures the fact that the early Zionists knew that the land on which they sought to establish a Jewish state was already occupied. In the face of surging anti-Semitism in Europe, however, they chose to brush aside the idea that the hundreds of thousands of Arabs who lived in Palestine would pose a serious obstacle to their

³⁶ Charles D. Smith, *Palestine and the Arab-Israeli Conflict: A History with Documents*, Sixth Edition (Boston: Bedford/St. Martin's, 2007), 38–40.

³⁷ Kirsten E. Schulze, *The Arab-Israeli Conflict* (New York: Pearson, 2008), 4; Stanislawski, *Zionism*, 24–25.

goals. Instead, airily assuming that the issue would resolve itself in due course, they pushed forward with their effort to establish a Jewish state in the territory.³⁸

Arab Nationalism

Like Zionism, Arab nationalism also emerged during the final decades of the Ottoman Empire. It first originated in Egypt, where it gained a small following among the educated classes thanks to the tireless efforts of two newspaper editors, Ahmad Lufti al-Sayyid (1872-1963) and Mustafa Kamil (1874-1908). Drawing inspiration from the 'Urabi revolt, they promoted the idea that Egyptians were a distinct people who, as such, should rule themselves in their own independent nation-state. The two men made some inroads with this idea at the turn of the century among educated Egyptians; however, the relatively light hand of the British colonial regime limited the appeal of their message. That situation changed suddenly in 1906 as a result of the Dinshaway Incident: a confrontation between overbearing colonial officials hunting pigeons and Egyptian peasants in the Nile River Delta that resulted in the death of a British officer. The episode itself was comparatively unremarkable. However, Britain's decision to make an example of the incident by sentencing two of the peasants involved to hard labor for life and by condemning four others to death shocked the Egyptian people. The result was a sudden, brief upsurge of nationalist sentiment.³⁹

Syria, too, had an embryonic nationalist movement prior to World War I. Influenced by European ideas, educated Arabs in Damascus took umbrage in the years immediately preceding the conflict with the Ottoman government's new requirement that education and court proceedings in the Arab provinces had to take place in Turkish—a decree that was itself, as we have seen, a function of rising nationalist sentiment. In response, a small group of academics, bureaucrats, and military leaders in Damascus began to develop a distinctive Arab sense of identity and to emphasize the uniqueness and glory of the Arabs' history; in other words, they began to create a distinctive, Arab form of nationalism.⁴⁰

Despite these early stirrings, however, Arab nationalism failed to gain traction in the years before the First World War. Two factors accounted for its relatively slow development. First, it faced stiff competition from the existing ideology of Pan-Islamism. Well established, Pan-Islamism provided a popular and well-established form of identity that muted the appeal of nationalism among the Arab people. Second, the Arab nationalists were themselves divided between those who favored Pan-Arabism—the idea that all Arabs shared an identity and should live together in a unified state—and those who promoted narrower conceptions such as Egyptian or Syrian nationalism.

³⁸ Avi Shlaim, *The Iron Wall: Israel and the Arab World, Revised and Expanded* (New York: W. W. Norton, 2014), 3–4.

³⁹ Eugene Rogan, *The Arabs: A History – Revised and Updated Edition* (New York: Basic Books, 2012), 144–46.

⁴⁰ Albert Hourani, *A History of the Arab Peoples* (Cambridge, Mass: Belknap Press of Harvard University Press, 2002), 308–9.

Facing these headwinds, Arab nationalism simply could not make meaningful progress prior to 1914. As a result, it would remain a fringe movement until European dominance of the Middle East in the interwar period and the defeat of the Ottoman Empire in World War I established conditions more conducive to its growth.⁴¹

War Comes to the Middle East

Breaking out in July 1914, the First World War seemed destined to effect huge changes to the Middle East. Most obviously, as nearly all observers at the time assumed, the conflict appeared almost certain to end in the rapid defeat of Ottoman armies and in the subsequent dismemberment of the empire. They were half right. The Ottoman Empire lost the war and soon thereafter faced partition at the hands of the victors. However, it did not go down to defeat quickly or easily. On the contrary, in the short term, the sultan's armies acquitted themselves very well. Indeed, they fought so effectively in 1915 and 1916 that they compelled France and, especially, Britain to make a series of fateful decisions that would powerfully shape the future of the Middle East.

The Outbreak of the First World War

As many had feared, the First World War originated in the Ottoman Empire's former Balkan territories. On June 28, 1914, a member of a Serbian terrorist organization called the Black Hand assassinated the Austro-Hungarian Archduke Franz Ferdinand (1863-1914) in Sarajevo, the capital of the former Ottoman province of Bosnia-Herzegovina, as part of a terror campaign aimed at compelling Vienna to cede the territory to Serbia. The killing produced the "July Crisis," a gradually worsening diplomatic impasse that gripped the capitals of Europe for the next five weeks. Despite hopes for a peaceful resolution, the situation slid irrevocably toward war when Austria-Hungary, after first securing German support, issued a harsh ultimatum to Serbia in late July that, by design, proved unacceptable to the Serbian government. From that point, the crisis quickly spun out of control. Beginning with Tsar Nicholas II's decision to order a general mobilization of the Russian army on July 30, the three Triple Entente powers, France, Britain, and Russia, and the leading states of the Triple Alliance, Germany and Austria-Hungary, quickly entered the conflict. Italy, the third member of the Triple Alliance, did not honor its treaty commitment in 1914 and would, as we shall see, instead join the war on the Allied side in 1915.⁴²

Many in Europe cheered on the impending conflict, believing that it would be brief, glorious, and triumphal for their side. Indeed, vast crowds gathered in cities across the continent to express their unbridled enthusiasm for the coming war. The conflict that

⁴¹ Hourani, 308–11.

⁴² Ian Kershaw, *To Hell and Back: Europe 1914-1949* (New York: Penguin Books, 2015), 9–36.

ensued did not live up to their euphoric predictions, however. Neither brief nor glorious, it was, at least until its third year, anything but triumphal for its participants.⁴³

The war also defied Berlin's strategic assumptions. The commander of the German military, Helmuth von Moltke (1848-1916), had hoped to deal with the dilemma of fighting a two-front war by initially going on the defensive in the east so that he could concentrate his forces for an offensive designed to quickly knock France out of the conflict. Germany would then shift its armies to the Russian front where they would defeat the tsar's huge military and bring the conflict to a swift and victorious conclusion. Despite a vast amount of prewar preparation and planning, however, the offensive did not go off as designed. Instead, while Germany scored a surprising victory over Russia in the fall of 1914, it proved unable to achieve the decisive triumph over France on which von Moltke's entire strategy hinged. Now facing the prospect of fighting a two-front war against a coalition of enemies that enjoyed vast superiority in terms of resources and economic productivity, Berlin was forced to shift to the defensive on the Western Front.

Thereafter, neither side could achieve a meaningful victory. Despite enjoying a huge material advantage thanks to their vast empires and control of the seas, The French and British repeatedly failed in their efforts to crack their way through the elaborate trench fortifications that the Germans had constructed in northern France. Denied access to critical raw materials as a result of a highly effective Allied blockade, meanwhile, Germany faced the daunting prospect of fighting a long war while enduring severe material shortages and declining economic productivity. The result was a brutal, slowly escalating stalemate in which literally millions died in futile frontal attacks on heavily defended positions.

Germany Gains an Ally, October 1914

In the meantime, the outbreak of the conflict put the Ottoman government in a difficult spot. As we have seen, many of the European states had designs on the empire's territory but had heretofore been prevented from acting on those desires by the other Great Powers. Now, thanks to the exigencies of war, they no longer faced those diplomatic constraints and were free to seize Ottoman territory. Thus, the outbreak of the war left the sultanate facing almost-certain destruction.⁴⁴

In response, the Three Pashas frantically stepped up their search for a binding defensive alliance with one of the European states in hopes that such an agreement would dissuade the other Great Powers from helping themselves to Ottoman land. They differed, however, over which state they should approach. Arguing that geography prevented Germany from providing meaningful aid to the empire, Djemal insisted that they needed to strike a deal with Britain or France. Accordingly, he made a series of approaches to them in July 1914. Unsurprisingly, his overtures went nowhere. Well

⁴³ Kershaw, 9–36.

⁴⁴ Rogan, *The Arabs*, 37–41.

aware that the tsar's government would insist on acquiring the straits as the price of its participation in the war, Paris and London felt that they had no choice but to reject his advances. Making matters worse, the British government followed by requisitioning the nearly completed dreadnoughts *Sultan Osman I* and *Reshadieh* on August 1 for service in the war against Germany. Meanwhile, just as Britain and France were rebuffing Djemal's entreaties, Talat and Enver approached Berlin about an alliance. Unlike Djemal's pitch to Paris and London, their proposal met with success. Signed on August 2, the alliance—which remained secret for the time being—called for Germany to guarantee the Ottoman Empire's borders and to provide it with military aid; in exchange, the Three Pashas promised that they would bring their state into the war on the side of the German-led Central Powers.⁴⁵

Why did Djemal, Enver, and Talat take the fateful step of entering into this agreement? After all, Djemal was correct in pointing out that Berlin could provide only limited assistance to the Ottoman Empire. Indeed, as events would prove, the alliance with Berlin failed spectacularly to provide the security that the triumvirate sought even as it obligated the Ottoman Empire to join a war that it could not win. As such, historians have long concluded that the decision to ally with Germany and thus to join the war had been a shortsighted one rooted in desperation and wishful thinking—a clumsy mistake that doomed the empire to defeat and dismemberment.⁴⁶

This interpretation has dominated historical understandings of Turkey's decision to enter the conflict since the war's conclusion, and has done so, in part, because it contains more than a few grains of truth. After all, the decision to enter the war *did* put the empire in a conflict that it could not win. Likewise, the alliance with Germany *did* give sanction to the Allies' decision to apportion Ottoman territory among themselves at the conclusion of the war.⁴⁷

In recent years, however, a new generation of historians has challenged this view. They argue that it ahistorically locates Ottoman decision making in the context of the war that occurred, rather than in the environment and expectations that existed in 1914. As the historian Mustafa Aksakal notes, most educated observers did not expect a long, bloody conflict that would conclude with one side achieving a total victory but instead believed that any general European war would be brief and would end in a negotiated settlement. Seen in this light, the treaty makes a great deal more sense. In the short run, it offered the security that the triumvirate believed the empire required to survive the war intact; in the long term, it promised to provide the permanent defensive alliance needed to guarantee the empire's borders.⁴⁸

⁴⁵ Rogan, *The Fall of the Ottomans*, 37–41.

⁴⁶ Mustafa Aksakal, *The Ottoman Road to War in 1914: The Ottoman Empire and the First World War* (Cambridge: Cambridge University Press, 2010), 12–13, 190–91.

⁴⁷ Aksakal, 12–13, 190–91.

⁴⁸ Aksakal, 12–13, 190–91.

What did Germany stand to gain from the alliance? Kaiser Wilhelm II (r. 1888-1918) and his cabinet perceived three benefits in the agreement. First, they believed that by closing the straits that were so vital to Russian trade—37 percent of its exports passed through them—the alliance with the Ottoman Empire would greatly weaken the economy on which the tsar's massive military depended. Second, they were confident that Ottoman armies could divert troops away from other fronts by invading Russian territory in the Caucasus and by attacking the Suez Canal. Finally, and most importantly, Berlin was convinced that the Ottoman Empire possessed a unique and potentially devastating ideological weapon: the authority that the sultan held in his capacity as caliph to issue a fatwa calling for all Sunnis to wage jihad against the Allies. German officials placed great stock in such a declaration—believing that it would spark a series of debilitating revolts in British, Russian, and French colonial territories.⁴⁹

Having finally secured the alliance that they had been desperately seeking since the end of the Balkan Wars, the Ottomans promptly repaid the favor by deliberately extorting and stonewalling their new ally. They did so in two ways. First, they exploited the perilous situation of two German warships then facing imminent destruction at the hands of Britain's Mediterranean fleet, the powerful dreadnought S.M.S. *Goeben* and the light cruiser S.M.S. *Breslau*, in order to squeeze new concessions from Berlin. To rescue the vessels, they proposed a fictive sale of the ships to the Ottoman government—a paper exchange that would save the *Goeben* and *Breslau* from destruction and partially offset the loss of the dreadnoughts that London had requisitioned. It was a far-more one-sided proposal than it appeared at first blush, however. In exchange, Germany would have to agree to a series of concessions including promises to help the Sublime Porte regain several strategic islands in the Aegean Sea from Greece and recover territory in the Caucasus it had ceded to Russia in the nineteenth century. Most importantly, the proposal required Germany to abolish the capitulatory privileges its subjects enjoyed in the Ottoman Empire. The kaiser's ministers spluttered with rage at what they saw as blackmail on the part of their putative ally. Aware that their ships faced sure destruction, however, they gritted their teeth and agreed to the Sublime Porte's terms.⁵⁰

Second, citing the need for further military and financial assistance, the Three Pashas repeatedly rebuffed German demands that they honor the terms of the alliance and declare war on the Allies. Indeed, it was only after the Sublime Porte had squeezed out one last concession from Berlin that the two governments finally came to terms regarding Ottoman entry into the conflict. According to the arrangement they worked out, the Sublime Porte agreed to permit the now-nominally Ottoman *Goeben* and *Breslau* to attack Russian ports. In exchange, Berlin promised to provide the Ottoman government with a loan of five-million Turkish pounds in gold.⁵¹

⁴⁹ Hew Strachan, *The Outbreak of the First World War* (New York: Oxford University Press, 2004), 62; Rogan, *The Fall of the Ottomans*, 46–47.

⁵⁰ Aksakal, *The Ottoman Road to War in 1914*, 110–18, 137–52, 156–63.

⁵¹ McMeekin, *The Ottoman Endgame*, 126–28.

To this point, the Three Pashas had played a weak hand very well. They had exploited the July Crisis and Germany's growing desperation for help against the Allies with great skill and effectiveness. They had gained valuable concessions, acquired urgently needed financial aid, and secured vital war materiel. They had even won Germany's agreement to abolish the hated capitulatory agreements. The decision to authorize the bombardment of Russian territory marked the point of no return, however. Once the *Goeben* and *Breslau* began shelling the Russian naval base at Sevastopol on October 29, the Ottoman Empire could no longer turn back from the brink and was thereafter irrevocably bound to its ally's fate. It was, in sum, a momentous decision—one that was pregnant with significance for the empire and its people. With it, the conflict that would so profoundly shape the Middle East's future had finally come to the region.⁵²

The Early War

Ottoman Offensives

Having finally brought the empire into the war, the Three Pashas wasted little time in joining the fight. They began by readying large offensives against Russian positions in the Caucasus and against the Suez Canal. More importantly, they moved aggressively to play the religious card on which Berlin had placed such high hopes. Almost immediately after entering the conflict, the Three Pashas had twenty-nine legal scholars draft five fatwas justifying holy war; they followed by arranging for sultan Mehmet V (r. 1909-1918)—acting in his capacity as caliph—to formally call a jihad against the Allies on the basis of those legal rulings. The Germans were elated. After a rocky start, their alliance with Istanbul appeared finally to be bearing fruit.⁵³

The results turned out to be bitterly disappointing. To German dismay, neither of the military offensives proved successful in the least. The first, the advance on the Suez Canal, was merely ineffectual. Under Djemal's direction, the attackers did manage to get as far as the waterway's east bank in December 1914; facing entrenched Indian troops that the British had deployed on the opposite side, however, they were forced to retreat back to Palestine with substantial losses.⁵⁴ Meanwhile, the second offensive, Enver Pasha's effort to unhinge the Russian position in the Caucasus in late 1914, was an outright debacle. Seeking a quick victory, he sent inadequately provisioned soldiers—many even lacked shoes—on a winter offensive in the mountains against entrenched Russia formations. Despite the poorly contrived and supported attack, the Ottoman troops initially made good progress. The arrival of Russian reinforcements and, worse, the onset of blizzard conditions that brought temperatures as low as -36°C dramatically altered the situation, however. Exposed, suffering enormous casualties,

⁵² McMeekin, 124–33.

⁵³ Rogan, *The Fall of the Ottomans*, 47–52.

⁵⁴ More than one-million Indian soldiers fought for the British Empire during the First World War.

and out of supplies, Enver's army soon broke and retreated in disarray. By the time it returned to its starting position, it had lost 60,000 of its 95,000 soldiers.⁵⁵

Still, the Germans were buoyed by their expectation that a declaration of jihad would turn the war in favor of the Central Powers. Like many Europeans, they put great stock in the power of Muslim holy war. They believed that the caliph's formal call to jihad would produce massive uprisings of zealous Muslims rendered fearless by their belief that Heaven awaited those who died fighting to defend Islam. Facing the rebellion of millions of fanatical Muslim subjects, the Allies would be unable to prosecute the war against the Central Powers and would thus be compelled to sue for peace.⁵⁶

Ultimately, the call to holy war failed to live up to the German government's inflated expectations. Despite a vigorous and well-funded propaganda effort, the caliph's declaration of jihad produced no meaningful uprisings in the Allied empires. In part, its failure was the product of an effective public relations campaign by the Allies. Central to that effort was King George V's (r. 1910-1936) public assurance to the British Empire's 100 million Muslim subjects that his government would refrain from attacking or damaging Mecca, Medina, or the important religious centers in Iraq. Reassured in this way, prominent Muslim leaders including the Aga Khan, the nawab of Dhaka, and the bey of French-ruled Tunis instructed their subjects to continue to abide by the authority of the colonial powers. Thus, to the disappointment of the German government, the caliph's declaration of jihad ultimately had no impact on the war.⁵⁷

Britain's Response

Still, the failure of the caliph's declaration aside, the Ottoman Empire's entry into the war seriously complicated Britain's position in the Middle East. Two vulnerabilities stood out. First, Ottoman forces located in modern-day Iraq menaced the British-owned oil wells and refinery facilities in western Iran that lay just across the Shatt al-Arab Waterway. This threat was a serious one. Iranian petroleum was not merely an economically valuable commodity but one of great strategic significance thanks to a key policy change made just before the war. In 1912, First Lord of the Admiralty Winston Churchill (1874-1965) had ordered the conversion of the British fleet from coal powerplants to more powerful and flexible oil-fired ones. This move gave the Royal Navy vastly greater mobility but came at the cost of making the fleet—the first line of imperial defense—dependent on imported petroleum from Britain's oilfields in Iran. Second, Ottoman control of Palestine posed a constant danger to the Suez Canal—the vital artery through which massive quantities of resources, goods and personnel vital to the war effort flowed. This threat was nothing short of existential. As British officials

⁵⁵ Kristian Coates Ulrichsen, *The First World War in the Middle East* (London: C. Hurst and Co., 2014), 103; Michael A. Reynolds, *Shattering Empires: The Clash and Collapse of the Ottoman and Russian Empires, 1908–1918* (New York: Cambridge University Press, 2011), 124–25.

⁵⁶ Rogan, *The Fall of the Ottomans*, 47–52.

⁵⁷ Rogan, 70–72.

keenly grasped, without the canal, Britain would no longer be able to prosecute the war with Germany.⁵⁸

London moved aggressively to safeguard these vital assets. First, on the basis of contingency plans drafted shortly after hostilities had broken out, Indian troops landed in Ottoman territory at the head of the Persian Gulf the day after Britain declared war on the sultanate. Quickly taking control of Basra, they ensured that the oil that had become so essential to British security continued to flow unimpeded.⁵⁹

Second, London moved to secure the Suez Canal by assuming complete control of the province of Egypt. Doing so was more complicated than it appeared at first blush. While a de facto British colony, Egypt was still technically part of the Ottoman Empire; London would thus have to tread carefully lest its actions breathe life into any latent loyalty to the sultan. To address this challenge, it pursued a carrot-and-stick approach. On the one hand, it imposed censorship, maintained a substantial number of troops in the Canal Zone, implemented martial law, and, most importantly, decreed that Egypt was henceforth a British protectorate rather than a part of the Ottoman Empire. On the other, it softened these harsh policies by promising that the Egyptians would not have to contribute in any way to the war effort and by hinting that formal inclusion in the British Empire could lead to eventual self-government.⁶⁰

This approach was a savvy one and it might have kept Egypt quiescent during the war if the British government had been able to prevent the global conflict from adversely affecting its people. London proved unable to do so, however. Instead, in two critical ways, the conflict soon imposed substantial costs on Egypt. First, the insatiable, war-induced demand for textiles led Egyptian farmers to switch from growing cereal crops to raising cotton—a change that produced surging food prices, sustained inflation, and hunger. Second, despite Britain's promise that the conflict would not affect Egypt, military exigencies quickly led London to seek to harness its economy to the empire's broader war effort. Within a year of the conflict's start, colonial officials began conscripting men for service in labor battalions and confiscating transport animals. As a result, the Egyptian people grew increasingly frustrated and restive under British rule and came more and more to expect that they would, as compensation, enjoy some significant measure of self-government immediately following the conclusion of the conflict.⁶¹

⁵⁸ McMeekin, *The Ottoman Endgame*, 135–45.

⁵⁹ Rogan, *The Fall of the Ottomans*, 79–86.

⁶⁰ Rogan, 68–70.

⁶¹ M W Daly, "The British Occupation, 1882-1922," in *The Cambridge History of Egypt. Volume 2, Modern Egypt, from 1517 to the End of the Twentieth Century*, ed. M. W. Daly (Cambridge: Cambridge University Press, 1998), 245–48.

The Gallipoli Campaign, March 1915 to January 1916

Meanwhile, British officials appraising the broader significance of Ottoman entry into the war were gradually coming to conclude that the empire's participation on the German side presented them with both a crisis and an opportunity. As they were well aware, its belligerency prevented Britain and France from sending Russia the munitions and resources needed to prosecute fully the conflict on the Eastern Front and thus imperiled the broader Allied war effort. At the same time, however, Istanbul's vulnerable coastal location and the abysmal performance of the Ottoman army in the early fighting suggested to some British officials that the Allies could knock the sultanate out of the conflict by seizing its capital in a surprise naval attack. If successful, such a move would pay manifold dividends. It would enhance British prestige among its Muslim subjects, potentially sway neutral Balkan states such as Greece and Bulgaria to join the Allied cause, and, most importantly, permit France and Britain to supply Russia with war materiel. Given the mounting frustration and staggering casualty count on the Western Front, it was an appealing idea. Accordingly, Secretary of State for War Horatio Kitchener (1850-1916) ordered Churchill and his staff to draw up plans for a sudden attack on Istanbul through the narrow Dardanelles Strait.⁶²

The proposal that Churchill presented to the cabinet was a bold one. Arguing that the Ottoman gun emplacements along the Dardanelles were obsolete and poorly supplied, he proposed having a squadron of Anglo-French battleships force its way through the narrow strait in a sudden attack. Once past the Turkish defenses, the ships could then sail unmolested across the Sea of Marmara to Istanbul where they could train their guns on the city and compel the Ottoman government to surrender.⁶³

Unfortunately for the Allies, the attack turned out to be an unmitigated disaster. Churchill's plan had much to recommend it and might well have succeeded had they proceeded with it in December 1914. By the time the Allies launched the attack in March 1915, however, the Ottomans and their German partners had dramatically reinforced its defenses; as a result, the attempt to force the strait failed with the loss of three battleships. Realizing that they could not push their way through without first silencing the Ottomans' shore defenses, the Allies followed by landing British, French, Australian, and New Zealand troops on the Gallipoli Peninsula that commanded the western side of the strait. This effort, too, ended in disaster. Under the direction of Liman von Sanders, the well-led, deeply entrenched, and highly motivated Ottoman Fifth Army defending the peninsula prevented the invading troops from seizing the high ground and instead pinned them on their vulnerable beachheads. Eight grueling months of brutal trench warfare followed before the Allies finally admitted defeat and withdrew. The campaign thus ended in triumph for the Ottoman army. Though it had suffered 300,000 casualties

⁶² Rogan, *The Fall of the Ottomans*, 130–32.

⁶³ McMeekin, *The Ottoman Endgame*, 169–79.

to the Allies' 230,000, it had protected the capital and had achieved a decisive victory over the powerful British Empire.⁶⁴

For the Allies, on the other hand, the Gallipoli campaign had been an unmitigated disaster that had set the war effort back on multiple levels. First, the successful Ottoman defense of the peninsula persuaded Bulgaria to throw in with Germany in September 1915; in other words, rather than leading key Balkan states to join the fight against the Central Powers, the campaign had instead worsened the Allied position in Southeastern Europe. Second, the British defeat was a huge morale boost for the Ottoman Empire that breathed new life into its heretofore faltering war effort. Finally, and of greatest concern to London, the abject failure of the Gallipoli campaign greatly eroded British military prestige in the Middle East and South Asia and raised renewed fears that, emboldened by the apparent weakness of imperial military forces, the empire's Muslim subjects might rise in revolt in response to the caliph's summons to jihad.⁶⁵

Kut al-Amara, December 1915-April 1916

But what could they do to forestall such an uprising? To British officials, the answer seemed obvious: the best way to put to bed any lingering questions about the efficacy of the British army was to achieve a swift, symbolic triumph over the recently victorious Ottoman military. Conveniently, Mesopotamia seemed to offer a golden opportunity to do so. The region contained a city, Baghdad, of enormous historical and symbolic importance to Muslims, and the Ottoman army defending it was small, lacked competent leadership, and suffered from poor morale. Accordingly, seeing substantial benefits and little cost in acting, British officials decided to mount what they believed would be a quick and inexpensive offensive in Mesopotamia that would end with imperial troops taking Baghdad and, in the process, restoring Britain's battered military reputation.⁶⁶

Begun with high hopes, the campaign ended in disaster. In September 1915, Major General Charles Townshend (1861-1924) departed from Basra with a division of Indian troops. The operation started auspiciously. After brushing aside early resistance, his soldiers headed north confident that they would take Baghdad by Christmas. At that point, however, the campaign began to go off the rails. Under a new commander, the reinforced Ottoman Sixth Army stopped Townshend's force near the old Persian capital of Ctesiphon in November and then forced it to retreat. Overwhelmed by the speed of the enemy pursuit, Townshend's beleaguered troops were compelled to take refuge in the town of Kut al-Amara on the Tigris River where they waited for a relief army from Basra to rescue them. It never came. Instead, while their supplies rapidly dwindled, successive efforts to break through to Townshend's troops faltered in the face of determined Ottoman resistance. Finally, with no prospect of relief in sight and with his

⁶⁴ Ulrichsen, *The First World War in the Middle East*, 79–92.

⁶⁵ Rogan, *The Fall of the Ottomans*, 207–15.

⁶⁶ Ulrichsen, *The First World War in the Middle East*, 127–33.

stores exhausted, Townshend surrendered his force of thirteen-thousand men in April 1916.⁶⁷

The loss of Townshend's army created a panic among British officials. Coming so soon after the disaster at Gallipoli, this second defeat at the hands of the Ottoman military not only emboldened the Central Powers, but, more importantly, seemed to call into question the fundamental pillar on which British colonial authority had long rested: the belief that its military was invincible against non-Europeans. Doubts about the efficacy of the British armed forces were enormously worrying for London. Coupled with the sultan's call to jihad, they had the potential to produce massive uprisings among his majesty's Muslim subjects—rebellions that could threaten not merely the war effort but even the empire itself. Thus, coming as it did on the heels of the Gallipoli disaster, the defeat at Kut created a sense of urgency among British officials; they had to do something to ensure the loyalty of the empire's millions of Muslim subjects, and they had to do it fast.⁶⁸

The Historical Debate: The Armenian Genocide

In the meantime, the Allied attacks in the Dardanelles and eastern Anatolia helped to set the stage for the worst incident of ethnic cleansing to occur during the war: the Armenian Genocide. The mass murder of Armenians that occurred under the Young Turk government was at root the product of the increasingly reactionary and militant Turkish nationalist thinking that had become more-and-more prevalent in the empire during the prior two decades. Its immediate trigger, however, was a series of scattered Armenian uprisings in eastern Anatolia that raised fears among Muslim Ottoman subjects that the Armenians constituted what one leading CUP member called an “enemy within”—a disloyal fifth column that was ready to help the Allies defeat and dismember the empire.⁶⁹

The large-scale murder of Armenians began in earnest in the spring of 1915. In April, the government ordered the arrest and deportation of 240 prominent Armenians in Istanbul in order to prevent the community they led from cooperating with the Allied forces that were about to land at Gallipoli. The campaign picked up steam dramatically the following month when Interior Minister Talat issued the infamous Deportation Law that authorized provincial governors to remove Armenians from the critical border provinces of eastern Anatolia and to resettle them in Syria. The Kurds who composed the irregular formations that oversaw that effort did so in brutal fashion. In each city or town targeted for deportation, the militias murdered all male Armenians aged twelve and up before force marching the poorly provisioned women and children into the desert.

⁶⁷ McMeekin, *The Ottoman Endgame*, 263–70, 291–93.

⁶⁸ Ulrichsen, *The First World War in the Middle East*, 132–37.

⁶⁹ Rogan, *The Arabs*, 163–65.

While en route, they killed any stragglers and permitted armed bands to periodically attack the columns. Unsurprisingly, precious few Armenians made it to Syria alive.⁷⁰

For over a century, a highly politicized debate has raged over the massacre of the Armenians. It has centered on two related points: the intentions of the Ottoman government and whether the attacks constituted a deliberate act of genocide. Drawing on evidence provided by contemporary observers such as American Ambassador Henry Morgenthau (1891-1967), Armenian nationalists today maintain that the CUP murdered between 1 million and 1.5 million Armenians and argue that the violence was a centrally directed, deliberate act of genocide designed to purge Asia Minor of Armenian Christians. In contrast, the Turkish government claims that the Young Turks did not plan a campaign of systematic murder and that they merely ordered the deportations in response to Armenian collaboration with Russian forces and attacks on Turks. A substantial number of Armenians were killed in intercommunal violence, it concedes, but those people died as a result of a civil war between Muslim and Armenian militias in eastern Anatolia and not due to deliberate state policy. As evidence, the Turkish government cites the fact that Talat's deportation orders explicitly instructed government forces to treat the Armenians well. Thus, in the official Turkish view, the mass murder of Armenians during World War I was neither a case of genocide nor the product of deliberate, state policy but instead a function of deep-seated intercommunal animosity.⁷¹

While some scholars accept the broad outlines of the Turkish government's interpretation, others instead agree with the historian Taner Akçam that it badly and deliberately mischaracterizes what was, in fact, a clear case of genocide. Drawing on heretofore-unexamined documents from the Turkish archives, he rejects the view that the massacres were an unfortunate but also unintended consequence of the war and the contention that the people who perpetrated them did so without official sanction. Instead, he argues that the mass slaughter of Armenians stemmed from a policy that the CUP developed immediately after the Balkan Wars that was aimed, in his words, at "demographically restructuring Anatolia" to ensure that no minority group was sufficiently concentrated in a given area that it could, with Great Power help, secure its independence.⁷²

It was the First World War, Akçam continues, that provided the ruling party with the opportunity to put this program into action. Taking advantage of the cover that the conflict afforded them, the Young Turks moved to homogenize Asia Minor through the relocation of Armenians and other ethno-religious minorities such as Assyrian Christians so that those groups constituted no more than 10 percent of the population of any

⁷⁰ Finkel, *Osman's Dream: The History of the Ottoman Empire*, 533–35.

⁷¹ Taner Akçam, *The Young Turks' Crime against Humanity: The Armenian Genocide and Ethnic Cleansing in the Ottoman Empire* (Princeton: Princeton University Press, 2012), ix–xxxvi, quote from page xvii.

⁷² Akçam, *The Young Turks' Crime against Humanity: The Armenian Genocide and Ethnic Cleansing in the Ottoman Empire*, ix–xxxvi, quote from page xvii.

province. Since the large size of the empire's Armenian community rendered impossible the achievement of that goal through relocation alone, the CUP instructed local officials to murder hundreds of thousands of Armenians during the deportations. The Young Turks did not order those deaths openly, however. Instead, taking advantage of the extraconstitutional, party-controlled apparatus that the CUP had set up to control the bureaucracy following the Balkan Wars, leaders such as Talat employed a "dual-track mechanism" wherein they sent one set of orders through official channels instructing local authorities to take great pains to treat the Armenians well during the deportations and another set through the unofficial, CUP-controlled administrative structure ordering its agents to engage in mass murder. Thus, Akçam concludes, the massacre of the Armenians was neither accidental nor a function of wartime exigencies; instead, it was a carefully camouflaged but ultimately deliberate act of genocide on the part of the CUP government designed to ensure permanent Turkish dominance of all of Anatolia.⁷³

Conclusion

Meanwhile, even as the genocide of the Armenian people worsened, the situation on the Ottoman front remained shockingly dire for the Allies in early 1916. Not only had the British suffered costly reversals at Gallipoli and Kut at the hands of Ottoman forces that they had believed to be incapable of standing up to Western troops, but, thanks in part to those defeats, Bulgaria had joined the Central Powers while countries favorably disposed to the Allies such as Greece continued to remain on the fence. Meanwhile, the Ottoman Empires' stunning victories over British forces threatened to breathe new life into the caliph's call for jihad and thus raised the specter of armed revolts breaking out among the hundreds of millions of Muslim subjects in the French, Russian, and, especially, British Empires.

In response to this danger, London would pursue a series of actions that would prove enormously consequential for the future of the Middle East. Most notably, the reversals at Gallipoli and Kut and the dangers inherent in the sultan's call for jihad impelled British officials to seek new allies by dangling before them the possibility of acquiring parts of the Ottoman Empire if they joined the fight. Critically, the desire to offset its military setbacks in the Middle East also led London to promise territory to non-state actors that had the potential to undermine the Central Powers or to mute the caliph's call to holy war. Britain's diplomatic outreach would ultimately bear fruit and would play a small role in the Allied sides' eventual victory. In the process, however, it would also result in the negotiation of a series of clashing diplomatic accords that would, in heavily modified form, come to define the contours and enduring conflicts of the modern Middle East. It is to the formulation of those agreements and to the region they created that we shall now turn.

⁷³ Akçam, ix–xxxvi, quote from page xviii.

Chapter Ten: “A Peace to End All Peace”

Despite its early success in the war, the Ottoman Empire could not hold out against superior Allied forces indefinitely and was compelled to sue for peace in October 1918. Its defeat created a vast power vacuum in the Middle East—one that the main European victors in the war, France and Britain, intended to fill. They were not the only parties staking claims to parts of the region, however. Their wartime allies Italy and Greece also sought territory in the Middle East—the former to justify its massive wartime losses and the latter to advance a dubious campaign to reconstruct the Byzantine Empire. To further complicate the situation, non-state actors including Arab princes, nationalist movements, and Zionists advanced competing and seemingly irreconcilable claims that were backed, in several cases, by a series of contradictory wartime agreements.

As a consequence, the Middle East underwent a complex and drawn-out process of negotiation, diplomatic maneuvering, resistance, and accommodation in the immediate postwar era. Eventually, the various parties arrived at an arrangement—what the historian David Fromkin calls “the settlement of 1922”—that, reflecting the relative power of the participants, determined the political systems, rulers, and frontiers of the post-Ottoman Middle East. While many of the people who helped to produce the new map of the region were pleased with the results, other, more far-sighted observers grasped that the settlement was fundamentally flawed and would likely doom the Middle East to future instability. Perhaps the most far sighted among them was future British Field Marshal Archibald Wavell (1883-1950). Parodying President Woodrow Wilson’s (r. 1913-1921) famous declaration that World War I was a “war to end all war,” he declared prophetically in 1922 that the diplomats, bureaucrats, and regional experts who had contrived the modern Middle East between 1918 and 1922 had created a “peace to end all peace”—one, he predicted, that would ensure instability in the region for decades to come.¹

Diplomatic Maneuvering

Peace was little more than a dream in 1916. Instead, the conflict that had produced so much death and misery in 1914 had, remarkably, grown even worse over the next two years. The deadliest fighting took place on the Western Front. Though they deployed new weapons including tanks and poison gas, the French and British armies fighting there proved abjectly incapable of breaking through the increasingly complex and formidable German trenchworks that stretched from the English Channel to the Swiss border and struggled to make even minor gains in a war that quickly degenerated into a mass slaughter on an industrial scale. No campaign demonstrated this reality more clearly than the Franco-British push along the Somme River in the summer and fall of 1916. Despite months of furious attacks and the loss of a staggering 600,000

¹ David Fromkin, *A Peace to End All Peace: The Fall of the Ottoman Empire and the Creation of the Modern Middle East* (New York: Avon Books, 1989), 5, 18.

Allied soldiers, the offensive only managed to advance the front line a mere seven miles.²

Casualties on this scale were unsustainable and impelled Britain, France, and Russia to seek new allies that could share the burden. Commencing not long after the war began, their efforts began to bear fruit in 1915. That spring, they persuaded Germany's former Triple Alliance partner, Italy, to join the conflict on the Allied side. Later, in June 1917, they induced Greece to enter the war. The participation of these countries in the fighting certainly boosted the Allied war effort, but it came at a cost: in exchange for joining, the new belligerents demanded that the Allies grant them significant territorial concessions at the expense of the defeated states. For example, Italy declared war only after the Allies had first promised that it would receive the South Tyrol and territory along the Adriatic from the Habsburg Empire. Similarly, Greece entered the conflict with the understanding that Britain, France, and Russia would help it achieve the *megali idea*, or great idea: the Greek nationalist dream of reconstituting the Byzantine Empire.³

As the death toll mounted, however, the territorial promises secured early in the war seemed increasingly inadequate. Accordingly, the belligerents—new and old alike—began to demand greater and greater compensation to justify the slaughter. But where could London, Paris, and St. Petersburg—renamed Petrograd in 1914—find sufficient territory to satisfy both their own aspirations and the expectation of their new allies? After all, Britain and France had already staked claims to Germany's overseas colonies and were disinclined to give up their gains. The answer, increasingly, was the Ottoman Empire. In a process they referred to as the "Great Loot," Entente diplomats began to promise their new European allies territory in Asia Minor. Italy would receive Antalya along Anatolia's southern coast, while Greece would acquire the predominantly Greek city of Smyrna and its hinterlands.⁴

The McMahon-Husayn Correspondence, July 1915-June 1916

British officials did not focus exclusively on gaining additional military help, however. Instead, they also sought ideological assistance in the war. Most notably, they became increasingly interested in 1915 and 1916 in securing assistance in neutralizing the Ottoman declaration of jihad that they feared was gaining renewed traction among the British Empire's Muslim subjects thanks to the Ottoman Empire's victories at Gallipoli and Kut. Fortune appeared to smile on the British in this effort. In the summer of 1915, the Ottoman governor of the Hijaz, the Emir Husayn (r. 1916-1924), approached the ranking British official in Egypt, High Commissioner Henry McMahon (1862-1949), in an effort to secure formal recognition of Arab independence under his

² John Keegan, *The First World War* (New York: Vintage, 1998), 298–99.

³ Margaret MacMillan, *Paris 1919: Six Months That Changed the World* (New York: Random House, 2007), 283, 348.

⁴ Scott Anderson, *Lawrence in Arabia: War, Deceit, Imperial Folly and the Making of the Modern Middle East* (Anchor, 2013), 152–53.

family's rule. There followed a lengthy exchange of letters between McMahon and Husayn that came, subsequently, to be known as the McMahon-Husayn Correspondence.⁵

At first, McMahon and the British did not believe that they had much to gain from the negotiations. The head of the Hashimite clan and thus sharifian, or descended from Muhammad, Husayn had initially asked London to help him and his sons establish a series of independent, Hashimite-ruled Arab kingdoms encompassing Syria, Mesopotamia, and Arabia; in exchange, he pledged to lead a revolt against the Ottoman Empire and promised that huge numbers of Arab soldiers would desert Djemal Pasha's (1872-1922) army and join their rebellion. McMahon found the emir's expansive demands for territory to be wildly out of scale with his likely ability to affect the war; more importantly, he was aware that London had its own designs on the Arab lands that the emir sought. At first, accordingly, he carefully refrained from discussing territorial claims with Husayn.⁶

The abject failure of a major push to break the deadlock at Gallipoli in the late summer of 1915 put the emir's request in a different light, however. Thereafter, British officials became more enthusiastic about the assistance that Husayn could provide. In part, they warmed to a deal because they had come to believe that the Hashimites might make a modest but valuable military contribution. Specifically, they concluded that Arab guerrillas operating out of Arabia might draw off Ottoman soldiers and disrupt lines of communication and thus aid an upcoming offensive into Palestine. Ultimately, however, it was the Hashimite's soft power rather than their military potential that rekindled British interest in concluding a deal. Fearful that the recent defeat at Gallipoli might reinvigorate the Ottoman sultan's call to jihad, they concluded that having the Sharifian ruler of Mecca and Medina on the Allied side could help to maintain the continued loyalty of the millions of Muslim subjects of the French, British, and Russian empires.⁷

McMahon consequently shifted his tone and began to indicate in his letters that London was willing to accommodate Husayn's territorial aspirations. He insisted on two exceptions, however. First, aware that London had designs on the *vilayets*, or provinces, of Basra and Baghdad, he made clear that they had to remain separate from the Arab territories that Husayn sought. Second, he demanded that "portions of Syria lying to the west of the districts of Damascus, Homs, Hama and Aleppo," remain apart from the proposed Hashimite states, ostensibly because "they cannot be said to be purely Arab" but really because France had made clear that it wanted to add them to its empire—a point we shall return to in a moment. Husayn was disappointed with these conditions. Nonetheless, with the British dangling promises of weapons, recognition,

⁵ Fromkin, *A Peace to End All Peace*, 173–84.

⁶ Eugene Rogan, *The Fall of the Ottomans: The Great War in the Middle East*, Reprint edition (Basic Books, 2015), 281–83.

⁷ Rogan, 281–83.

and gold in front of him, he decided to accept McMahon's conditions and pledged to begin his rebellion against the Ottoman Empire in mid 1916.⁸

Though each side had its reservations, both Husayn and the British were pleased with the outcome of the negotiations. That they were mutually satisfied owed substantially to the fact that neither party fully understood that the other had been bargaining in bad faith. Husayn, for his part, had in effect signed checks that could not be cashed. That is, despite his assertions to the contrary, he exercised no meaningful influence over Arabs living outside the Hijaz; accordingly, his claim that a revolt under his banner would lead huge numbers of Arabs to switch sides was an empty one. If Husayn had overpromised, McMahon had been downright duplicitous. A subtle negotiator, he had made liberal use of evasive language and careful qualifiers to ensure that while Britain *appeared* to have agreed to help the emir secure Arab independence under his family's rule, it in fact had made no concrete commitments whatsoever. Thus, while Husayn came away from the exchange believing that London had pledged to help him and his sons establish themselves as the sovereign monarchs of independent states, it had in fact agreed only to make them the titular rulers of indirect British colonies. Indeed, as Foreign Secretary Edward Grey (1862-1933) noted privately at the time, London was promising little more than "a castle in the air which would never materialize."⁹

Sykes-Picot Agreement, May 1916

As earlier alluded to, McMahon had good reason to dissemble in his discussions with Husayn. At the same time that he was negotiating with the emir, another British official Mark Sykes (1879-1919) was engaged in secret talks with the French diplomat François Georges-Picot (1870-1951) over the partition of the Arab Middle East; as a result, the high commissioner needed to avoid making promises to Husayn that might conflict with France's territorial designs. Sykes, a rank amateur, and Picot, a hardened French nationalist, met several times between October 1915 and March 1916 to determine how the Allies would divide the region. Picot drove a hard bargain. He insisted that France gain direct control of territory in southeast Asia Minor and along the coast of the Levant, and he demanded that it indirectly rule the Syrian interior as far as the city of Mosul. These were expansive demands, particularly given the widespread belief that the *vilayet* of Mosul contained substantial oil deposits. Nevertheless, London perceived advantage in yielding to Picot's request. With one eye fixed firmly on the future, it concluded that French possession of Syria would give Britain both a buffer and a natural ally should Russia resume the Great Game following the war.¹⁰

⁸ "Letters between Hussein Ibn Ali and Sir Henry McMahon," accessed May 26, 2020, https://wwi.lib.byu.edu/index.php/Letters_between_Hussein_Ibn_Ali_and_Sir_Henry_McMahon; Anderson, *Lawrence in Arabia*, 160–63.

⁹ Fromkin, *A Peace to End All Peace*, 184–87.

¹⁰ Rogan, *The Fall of the Ottomans*, 285–87.

In any event, control of Syria was one of London’s goals in the Middle East. Instead, its focus in the region was on taking possession of Mesopotamia—a region that it correctly believed contained substantial oil deposits—and on controlling a belt of territory connecting Egypt in the west to modern-day Iraq in the east that could safeguard the empire’s vital communications with India. Sykes consequently proposed splitting the Middle East along a line stretching from Acre in Palestine to Kirkuk in Mesopotamia, with France gaining control of territory north of the line and Britain securing the region south of it. Picot agreed to the broad outlines of Sykes’s division of the region; however, he refused to sign off on the proposal to assign Palestine to Britain. Insisting that the territory had to retain its historical connection to Syria, the French diplomat demanded that France take control of it following the war.¹¹

By March 1916, the two men had finally hammered out an agreement. According to its terms, Britain would directly rule Basra and Baghdad and would exert indirect control of other Arab territories south of the line Sykes had drawn, while France would directly administer southeastern Anatolia and Lebanon and would indirectly oversee the Syrian interior as far as Mosul. They compromised over Palestine. In recognition of the importance of its religious sites, they agreed that it would come under international control following the war. Pleased with the deal, they presented it to Russian Foreign Minister Sergey Sasanov (1860-1927). After reviewing the agreement, he gave it his formal blessing. However, he conditioned Petrograd’s acceptance on Paris and London’s willingness to grant Russia new territorial concessions in Asia Minor beyond Istanbul and the straits, which the Allies had already agreed would come under the tsar’s control at the conclusion of the conflict. Picot and Sykes accepted this qualification. Accordingly, with an exchange of letters in May 1916, the three powers formally signed off on what has been known ever since as the Sykes-Picot Agreement.¹²

The Historical Debate: The Sykes-Picot Agreement

Over the past few decades, the agreement has come to assume an outsized role in popular understandings of the modern Middle East’s genesis and problems. According to the conventional narrative, Sykes and Picot divided the region for nakedly imperialist reasons rather than to meet the aspirations of the people living there. As a result, the agreement imposed artificial and problematic borders that embedded within the Middle East state system a series of intractable conflicts. Accordingly, to the educated public, this externally imposed agreement—the modern Middle East’s

¹¹ James Barr, *A Line in the Sand: The Anglo-French Struggle for the Middle East, 1914-1948* (New York: W. W. Norton & Company, 2011), 15–26; Fromkin, *A Peace to End All Peace*, 188–93.

¹² Rogan, *The Fall of the Ottomans*, 285–87.

diplomatic original sin—rather than internal conflict has been the ultimate source of the many problems that the region has experienced over the past century.¹³

While popular, this view is not universally shared by scholars. On the contrary, a number of historians argue that it distorts and oversimplifies what was a much more complex process. The historian Sean McMeekin, for one, argues that while the conventional view is accurate in its characterization of British and French motives, it also substantially overstates in two critical ways the agreement’s role in defining the region’s future development. First, the frontiers and divisions that Sykes and Picot hammered out in 1915 and 1916 bore only a superficial resemblance to the Middle East that finally emerged from the shadow of war in the early 1920s; therefore, it cannot be held responsible for the problems that those frontiers may have created. More importantly, he continues, the idea that the agreement was the source of the region’s modern-day tribulations denies agency to the many states, peoples, and groups that contested the shape and structure of the Middle East in the years immediately following World War I. Indeed, it is McMeekin’s view that the creation of the modern Middle East occurred not in 1916 with the Sykes-Picot Agreement but instead in the immediate postwar years in a complex process that demarcated the borders and established the states that continue to define the region to the present day.¹⁴

The Balfour Declaration, November 1917

Further complicating understandings of the Sykes-Picot Agreement was the fact that the British were not yet done apportioning the Ottoman Empire’s territory in the Arab Middle East. Most notably, at the urging of the influential chemist and Zionist, Chaim Weizmann (1874-1952), British Foreign Secretary Arthur Balfour (1848-1930) prepared a third document dealing with the region’s postwar territorial arrangements, the famous Balfour Declaration, well after Sykes and Picot had concluded their agreement. Issued on November 2, 1917, the statement proclaimed that “His Majesty’s Government view with favour the establishment in Palestine of a national home for the Jewish people and will use their best endeavours to facilitate the achievement of that object.” At just 128-words, it was a remarkably brief statement. What it lacked in length it more than made up for in impact, however, for—in issuing it—Britain had committed itself to backing the Zionist effort to establish a homeland for the Jewish people in Palestine.¹⁵

Why did the British government further complicate the already difficult question of the postwar Middle East’s disposition? It did so in part because several key members of the cabinet were personally supportive of Zionism. Balfour and Prime Minister David Lloyd George (r. 1916-1922), in particular, were genuinely sympathetic to the situation

¹³ Sean McMeekin, *The Ottoman Endgame: War, Revolution, and the Making of the Modern Middle East, 1908-1923* (New York: Penguin Books, 2015), xix–xx.

¹⁴ McMeekin, xx–xxii.

¹⁵ “The Balfour Declaration, November 2, 1917,” accessed May 29, 2020, https://avalon.law.yale.edu/20th_century/balfour.asp.

of the Jewish people. As a rule, however, the British government did not make foreign policy on the basis of sympathy. Far more important in leading London to back Zionism were two views that had taken root among British officials over the course of 1916: first, the idea that possession of Palestine—which controlled the eastern approaches to the Suez Canal—was critical to the long-term security of the empire, and, second, the belief that Zionism could serve as a stalking horse for British efforts to assume control of that territory.¹⁶

This last point was critical. Much though British officials wanted simply to annex Palestine at the end of the war, they understood that France and, especially, the United States, would, for different reasons, block any attempt on their part to do so. French disapproval was straightforward and reflected traditional power politics: having just signed the Sykes-Picot Agreement, Paris was highly unlikely to agree to alter the deal to suit Britain's imperial ambitions without receiving unacceptably costly concessions in return. American opposition, in contrast, was more ideological in nature. Its moralistic president, Woodrow Wilson, had made clear in his "Peace Without Victory Speech" in January 1917 that he opposed formal colonialism and that he would work to stop any postwar transfer of territory that did not enjoy the consent of those affected. Thus, with its two most important allies' firm in their opposition, London's effort to secure Palestine appeared to be stillborn in early 1917.¹⁷

So how could they get around this resistance? Increasingly, British officials came to see Zionism as the answer. That is, they came to believe that support for the movement could, if properly presented, neutralize both American and French objections and thus permit London to add Palestine to the empire.¹⁸

They first used Zionism to disarm French opposition. The key was the situation in Russia. In March 1917, frustration with the war effort and the management of the economy had resulted in a revolution that overthrew Tsar Nicholas II. Initially, revolutionary Russia remained committed to the war; as the economic and military situation continued to worsen over the course of 1917, however, its enthusiasm began to waver. This state of affairs deeply troubled London, and, especially, Paris. If Russia concluded a separate peace, Germany could redeploy its eastern armies to the Western Front and thus shift the balance of forces in France—enough, possibly, to tip the balance in the war. British statesmen were quick to take advantage of its ally's concern. Beginning in 1917, they began to argue to their French counterparts that Entente support for a Zionist homeland in a British-controlled Palestine would lead Russia's substantial Jewish population to press the revolutionary government to remain in the conflict and would thus ensure against the possibility that Germany could concentrate its forces in the west. Rattled by the prospect of a renewed offensive by a reinforced

¹⁶ Benny Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920* (New York: Vintage Books, 2001), 71–73.

¹⁷ Barr, *A Line in the Sand: The Anglo-French Struggle for the Middle East, 1914-1948*, 28–29.

¹⁸ Barr, 28–29.

enemy, Paris found this argument persuasive and softened its previously firm opposition to British control of Palestine.¹⁹

Support for Zionism also neutralized Woodrow Wilson's objections to London's acquisition of the territory. British policymakers had deduced that they could avoid American opposition to their territorial aspirations only if they could credibly present those acquisitions as selfless acts aimed at preparing disenfranchised peoples for autonomy or independence. Zionism permitted them to do just that in Palestine. By backing the movement, Britain could cloak what was an act of plain-and-simple imperialism as a selfless effort to help the long-suffering Jewish people at last acquire a home. It was an effective ploy. The American president not only refrained from opposing the British action but even formally endorsed the Balfour Declaration.²⁰

Thus, while the Balfour Declaration appeared on the surface to focus on achieving the Zionists' goals, its real purpose was to provide the ideological cover Britain needed to advance its imperial interests. This is not to suggest that the Balfour Declaration was anything less than a major victory for the Zionists; in fact, as we shall see in chapter thirteen, the statement gave their movement a degree of international legitimacy that it heretofore lacked and thus marked a critical turning point in its evolution. Rather, it is only to stress that Britain's motives were rooted fundamentally in promoting its self-interest rather than in supporting an idealistic cause. The verbiage of the Balfour Declaration made this reality clear. London was not calling for a sovereign Jewish state in Palestine but was instead proposing the establishment of a mere "national home" for the Jewish people. This choice of language was very deliberate: Balfour and the cabinet drafted the declaration to ensure that the Zionist project would be incorporated within a British colony rather than emerge as an independent state. London's support for Zionism was thus, at root, a cynical strategy for colonial aggrandizement. As such, it succeeded smartly.²¹

Unsurprisingly, London's myriad efforts to secure new allies in the Middle East landed Britain in an awkward position immediately following the war. Simply put, it had entered into three wartime agreements regarding the Ottoman Empire's Arab lands that promised the same territory to different parties. Paris and London expected to rule certain areas directly—the northern coast of the Levant in France's case; Basra, Baghdad, and Palestine in Britain's—and to divide the interior into a series of Arab client states. For their part, the Zionists saw the Balfour Declaration as merely a first step that would lead quickly and inevitably to the creation of a Jewish state in Palestine. Finally, while Husayn had grudgingly agreed to accept French control of Lebanon and coastal Syria, he had also justifiably come away from his negotiations with McMahon believing that Britain had pledged to recognize independent Hashimite rule in territory that included Palestine and inland Syria. In other words, the promises that London had

¹⁹ Fromkin, *A Peace to End All Peace*, 286–89.

²⁰ MacMillan, *Paris 1919*, 422–23.

²¹ Barr, *A Line in the Sand: The Anglo-French Struggle for the Middle East, 1914-1948*, 28; Quote from "The Balfour Declaration, November 2, 1917."

made were irreconcilable: in zero-sum fashion, each party could obtain its territorial aspirations only at the expense of the others. The result would be bitterness, feelings of betrayal, and outright conflict between the various parties following the war—particularly in Palestine, or, as some were coming to call it, the “thrice-promised land”—that continue to reverberate to the present day.²²

The Allies Victorious

The problems that the wartime agreements created were ones of the future, however; in the meantime, the Allies first had to win the war. Doing so seemed as far away as ever at the start of 1917. The Western Front remained a brutal stalemate in which the British and French sacrificed ever-larger numbers of young men in a futile effort to break Germany’s seemingly impenetrable trench lines. The situation was little better in the Middle East. Checked at Kut and Gallipoli by the sultanate’s surprisingly tenacious armies, the British had been unable to take any Ottoman territory beyond the small amount they had secured in 1914. It is true that the tsar’s armies had made substantial inroads in eastern Anatolia over the course of 1915 and 1916. Thanks to Russia’s gradual, post-revolutionary disengagement from the war, however, the Ottoman military had even succeeded in stabilizing that front.²³

To London’s relief, the military situation in the Middle East began to change in 1917. First, in Mesopotamia, a reinforced army of Indian troops erased the lingering stain of Kut by taking Baghdad in March 1917. More dramatically, General Edmund Allenby’s (1861-1936) Egyptian Expeditionary Force mounted a successful offensive aimed at taking possession of Palestine late in the year. It was, by the standards of the First World War, a lightning campaign. Breaking through the Ottoman lines along the frontier with Egypt in November 1917, Allenby’s troops advanced rapidly northward into the Holy Land. The drive was a major success—one made all the sweeter by the capture of the symbolically important city of Jerusalem on December 9.²⁴

The Arab Revolt that the Emir Husayn had launched in June 1916 played a small supporting role in this campaign. Under the command of Husayn’s sons, Faysal (1883-1933) and Abdallah (1882-1951), Arab irregulars attacked Ottoman troops defending the Hijaz Railway east of the Jordan River in a series of operations designed to disrupt Ottoman supply lines. While the revolt’s actual contribution to the war was comparatively minor, it came to assume an outsized role in popular understanding of Allenby’s Palestine campaign thanks to the work of Faysal’s British liaison officer, T. E. Lawrence (1888-1935), better known as Lawrence of Arabia. An effective self-promoter and a man of great imagination, Lawrence created the impression that the Arab Revolt had been pivotal in winning the campaign in Palestine in his exaggerated account of his

²² Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 75.

²³ Rogan, *The Fall of the Ottomans*, 355–57.

²⁴ Kristian Coates Ulrichsen, *The First World War in the Middle East* (London: C. Hurst and Co., 2014), 111–12; McMeekin, *The Ottoman Endgame*, 400–402.

service in the Middle East, the book *Seven Pillars of Wisdom* (1926), on which the award-winning film *Lawrence of Arabia* (1962) was based.²⁵

The conclusion of the war in the Middle East came with startling suddenness in October 1918. The endgame began in September when Entente armies launched simultaneous attacks against Ottoman forces on multiple fronts. Their first big breakthrough came in Palestine. After a long pause to bring forward supplies, the Egyptian Expeditionary Force punched through the Ottoman defenses north of Jerusalem on September 19 and raced toward Damascus. To the jubilation of people on the home front, Australian troops occupied the city just twelve days later. That account was not the story that London chose to present to the world, however. Instead, it concocted an elaborate ruse in which it pretended that Faysal's troops rather than the Australians had been responsible for the city's capture.²⁶

What accounted for this move? Why did London seek to elevate the Hashimite contribution to the war effort in Syria? It did so in service of the new, more expansive objectives that the British government had come to seek in the Middle East. Prime Minister David Lloyd George was the primary proponent of the change in London's goals in the region. By early 1918, he had come to conclude that the British Empire's future security required it to assume a position of hegemonic dominance in the Middle East following the war—a situation that would only work if it prevented France, a potential future enemy, from having a substantial presence in the region. Accordingly, he directed his subordinates to claim that Faysal's forces had been instrumental in taking Damascus—an achievement that, fake or not, would entitle the Hashimites to keep Syria after the war and thus ensure that the territory came under the control of a pliable British proxy rather than the rival French Empire.²⁷

Meanwhile, Allied armies in the Balkans were scoring an even more decisive victory. In late September 1918, a multi-national army under French command shattered the Central Powers' defenses in southern Bulgaria and began racing toward the undefended city of Istanbul. With no reserves remaining, the Ottoman government was out of options and had little choice but to sue for peace. It was not the Three Pashas who surrendered the empire, however. Aware that they would be compelled to answer for the deaths of hundreds of thousands of Armenians, Enver (1881-1922), Talat (1874-1921), and Djemal hurriedly resigned and slipped out of Istanbul on a German torpedo boat. Thus, it was a new government that on October 30, 1918, signed the draconian Armistice of Mudros that formally ended the First World War in the Middle East.²⁸

²⁵ McMeekin, *The Ottoman Endgame*, 342–46.

²⁶ Rogan, *The Fall of the Ottomans*, 374–79.

²⁷ Ulrichsen, *The First World War in the Middle East*, 94–95, 171.

²⁸ Ulrichsen, 94–95.

Dividing the Spoils

Sorting out the Middle East following the fighting was not going to be easy. Even with Russia's claims voided as a result of its exit from the war, the various agreements that London and Paris had made to secure new allies and to achieve their own territorial aspirations in the region promised to render efforts to establish a durable settlement in the Middle East difficult. Adding to the challenge was the fact that the victors had begun to fall out over the spoils. As we just saw, the imperial rivalry between France and Britain had rekindled during the waning days of the conflict as London moved to keep the French out of Syria in service of Lloyd George's goal of assuming hegemony over the Middle East. Likewise, jockeying for control of southwestern Anatolia, Italy and Greece were eyeing each other with growing suspicion. Ultimately, however, it was a pair of epochal events that occurred well outside the Middle East—the Bolshevik Revolution and the emergence of Woodrow Wilson as the globe's dominant political figure—that did the most to complicate the victors' efforts to arrive at a stable settlement in the region.

The Bolsheviks

From the very start, the Bolshevik seizure of power in November 1917 greatly confounded French and, especially, British efforts to secure dominion in the Middle East. It did so in two ways. First, Commissar for Foreign Affairs Leon Trotsky (1879-1940) revealed the hypocrisy of the French and British governments' claims that they fought to spread democracy and to help advance the interest of small and weak nations by publishing the secret treaties that the Allied states had concluded with one another—including the Sykes-Picot Agreement. Coming on the immediate heels of the Balfour Declaration, the revelation of this nakedly imperialist treaty embarrassed London and Paris and raised doubts in the Hashimite camp about whether they could trust the British. Second, Bolshevik leader Vladimir Lenin (r. 1917-1924), who had long called for self-determination for the many millions of people across the globe living under foreign rule, had stepped up his criticism of imperialism after his party seized power in Russia in November 1917. In doing so, Lenin was being disingenuous. His opposition to imperialism sprang far less from any interest he may have had in helping people living under the yoke of colonialism secure their independence than from his desire to weaken the capitalist states. That he was being duplicitous was ultimately unimportant in the immediate postwar era, however. What mattered was that his words helped create the expectation among people ruled by a foreign power or facing the prospect of being incorporated into a colonial empire that they could and should enjoy self-determination in the aftermath of the war.²⁹

²⁹ Erez Manela, *The Wilsonian Moment: Self-Determination and the International Origins of Anticolonial Nationalism* (New York: Oxford University Press, 2007), 37–38.

The “Wilsonian Moment”

The Bolsheviks thus posed a serious ideological challenge to the democratic-capitalist states—one that Woodrow Wilson, who had brought his country into the war in April 1917, felt he needed to counter. The result was his most important address: The Fourteen Points Speech that he gave before Congress on January 8, 1918. In it, he outlined four general principles—open diplomacy, freedom of the seas, free trade, and disarmament—that he argued should govern international relations in the postwar world. He also proposed the establishment of a new, supranational institution, the League of Nations, that would be empowered to enforce peace and prevent the outbreak of another general war, and he outlined nine specific territorial adjustments designed to achieve self-determination and the equality of nations. His twelfth point dealt specifically with the Middle East. It called for the creation of a Turkish state, but also stipulated that the other nations in the Ottoman Empire should have “an absolutely unmolested opportunity of autonomous development.”³⁰

It was a bravura performance. Wilson’s call for a new system of international relations based on the Fourteen Points captured the imagination of people across the globe. Coming from the leader of the world’s most powerful nation, it dramatically overshadowed Lenin’s message and led people to believe that an Allied victory would remake the world along liberal-democratic lines. It thus set the stage for what the historian Erez Manela calls the “Wilsonian Moment”: the period between November 1918 and March 1919 when colonized people, whose expectations had risen dramatically as a result of Wilson’s promotion of self-determination, shifted from their earlier calls for limited reforms or autonomy to demands for outright independence.³¹

The Paris Peace Conference, January-June 1919

In two critical ways, Wilson’s Fourteen Points and the Wilsonian Moment dramatically altered the global diplomatic landscape. First, they inspired nationalists to take their case to the Paris Peace Conference at which the postwar order was to be determined. Believing that they had the support of the American president, dozens of unaccredited groups, individuals, and officials arrived in France to press for independence in the winter and spring of 1919. From the Middle East alone, delegations came from Armenia, Syria, Lebanon, and Persia seeking to enlist the support of their champion in their newfound efforts to secure self-government.³²

Second, as we saw earlier with regard to the Balfour Declaration, the American president strongly opposed formal colonialism and was determined to prevent the peace settlement from turning into a land grab. There were, to be clear, significant limits to his influence. Most notably, the need to secure the Allies’ approval of the League of

³⁰ “President Woodrow Wilson’s Fourteen Points,” accessed June 4, 2020, https://avalon.law.yale.edu/20th_century/wilson14.asp.

³¹ Manela, *The Wilsonian Moment*, 4–6.

³² Manela, 59–60.

Nations—the centerpiece of his global vision—compelled the president to repeatedly agree to vast transfers of territory that violated the spirit of the Fourteen Points. Still, as the leader of the world's most powerful state—a country to which the Allies owed almost-incalculably large sums of money—he was more than strong enough to compel them to justify the acquisition of any new territory and to pay at least lip service to the idea that they were acting to help the affected people achieve self-determination.³³

Britain was far-better positioned to do so in the Middle East than were its imperial rivals. Indeed, the McMahon-Husayn Correspondence and the Balfour Declaration allowed London to portray its territorial ambition in the region not as a land grab but instead as a selfless effort on its part to help Zionists and Arabs alike achieve self-determination. Britain was thus able to win the president's approval for its territorial acquisitions in the Middle East and was even able to secure his assistance in its efforts to deny Syria to the French. Paris and Rome, in contrast, had no similar agreements that they could use to disguise their expansionist goals in the region and faced an uphill fight in persuading Wilson to give his blessings to their imperial ambitions. Thus, as the peace talks heated up in the spring of 1919, Britain appeared to have smartly outmaneuvered its erstwhile allies—casting them as naked imperialists at the very moment, ironically, that it was massively expanding its own empire in the Middle East and elsewhere.³⁴

Wilson further complicated the disposition of the Ottoman Empire by preventing the victors from simply adding newly gained territory into their colonial empires. Great Power pressure and his own deeply held racist beliefs may have compelled Wilson to accept the Allies' acquisition of territory after the war, but he remained implacably opposed to traditional imperialism. Indeed, he forthrightly insisted that any territorial gains had to be temporary and that the affected places had to be administered in the best interest of their inhabitants. The result was a compromise solution known as the Mandate System. Overseen by the League of Nations, it designated the former German colonies and large sections of the former Ottoman Empire as mandates: territories that were not-yet ready for self-rule and thus needed tutelage under an established power. It then assigned each to a so-called mandatory power and tasked those states with preparing the peoples that they administered for eventual independence. The Mandate System thus reconciled Wilson's opposition to formal colonialism with the Allies' insistence that they acquire territory at the end of the war. If it was a compromise, however, it was one that leaned heavily toward traditional imperialism. That is, despite the rhetoric about preparing the people of the mandates for self-governance, it amounted in practical terms to little more than a perpetuation of the existing colonial system under a new name.³⁵

³³ Manela, 59–60.

³⁴ Fromkin, *A Peace to End All Peace*, 376–78.

³⁵ Zara Steiner, *The Lights That Failed: European International History 1919-1933* (Oxford: Oxford University Press, 2005), 105.

Given this background, the negotiations at the Paris Peace Conference in the winter and spring of 1919 proved to be disheartening for the Arabs. Wilson's acceptance of the mandate system for the Arab part of the Ottoman Empire at the start of the conference had already taken genuine independence off the table. Meanwhile, the president's refusal to meet with the unaccredited delegations from the Middle East denied the Arabs any opportunity to plead their case for self-determination.³⁶

Even in the rare instances in which the Arabs enjoyed diplomatic successes, their achievements came with significant caveats. With London's help, for example, Faysal had succeeded in assuming the title of king of Syria and in establishing a government in Damascus. Exploiting his dependence, however, British officials significantly undermined those achievements by compelling him to sign the Weizmann-Faysal Agreement in which he formally accepted the terms of the Balfour Declaration and effectively ceded Palestine to the Zionists. Likewise, the King-Crane Commission, a fact-finding mission sent to the region to determine the wishes of its residents, accurately reported both Syrian hostility to the prospects of France becoming the mandatory power over their country and Palestinian-Arab opposition to the establishment of a Jewish homeland in the mandate and thus raised hopes that Wilson might rein in Paris and London. Thanks to the president's continued need for French and British support in his efforts secure the League of Nations, however, the commission's report had no impact on the disposition of Arab territory.³⁷

Meanwhile, even as the Arabs were reduced to bystanders, the talks over a final settlement in the Middle East dragged on. The primary holdup was Syria. With France and Britain at loggerheads over which would take possession of the *vilayet*, precious little progress could be made regarding the final disposition of Ottoman territory. In the end, it took two significant changes to the global diplomatic situation to break the deadlock. First, with Wilson incapacitated by a severe stroke, the US Senate rejected membership in the League of Nations in November 1919. As a result, the Europeans were now free to structure the peace among themselves without having to adhere fully to the president's demands. Second, with costly crises in Ireland, India, and Egypt straining the treasury, the badly overextended British government determined that it needed to trim its commitments. As part of that retrenchment, Lloyd George and his colleagues decided to abandon Faysal and permit the French to occupy Syria.³⁸

The Treaty of Sèvres, August 1920

With London having cut the Gordian knot, representatives of Britain, France, and Italy were finally able to arrive at an agreement regarding the Middle East. Meeting in the Italian resort town of San Remo in April 1920, they finally hashed out the terms of the Treaty of Sèvres, the peace settlement that would formally end the war between the Allies and the Ottoman Empire. It stipulated the division of the Arab portion of the

³⁶ Manela, *The Wilsonian Moment*, 59–60.

³⁷ MacMillan, *Paris 1919*, 390–94, 406.

³⁸ Steiner, *The Lights That Failed*, 107.

defeated empire along what amounted to a modified version of the Sykes-Picot Agreement. France would assume separate mandates over Syria and Lebanon but agreed to cede Mosul to London in exchange for the rights to a one-quarter share of the province's oil. Paris also indicated its willingness to have Faysal continue to serve as king of Syria, albeit in a strictly titular capacity. For its part, Britain would acquire a mandate over a new country, Iraq, that encompassed the *vilayets* of Mosul, Baghdad, and Basra; it would also become the mandatory power for Palestine where it was formally tasked with implementing the Balfour Declaration. Further north, the treaty partitioned the south coast of Anatolia between France, Italy, and Greece and established an autonomous Kurdish region and an independent Armenia. Both the city of Istanbul and the straits would fall under international administration. Finally, what little remained of the Ottoman Empire would be compelled to accept the reinstatement of the capitulations.³⁹

The treaty was thus a draconian one that effectively dismembered the empire—a bitter pill for a dynasty and state that had once held substantial territory on three continents. With the capital literally under the guns of the British navy, however, it was also an agreement that the defeated Ottomans were in no position to refuse. Accordingly, desperate to hang on to power, the final Ottoman sultan, Mehmed VI (r. 1918-1922) signed the treaty in August 1920.⁴⁰

London was, without question, the primary beneficiary of the Treaty of Sèvres. With it, British officials had obtained the arc of territory stretching from the Sinai Peninsula in the west to the Persian border in the east needed to prevent a rival from threatening the vital communications that connected India and Britain. They had also gained direct or indirect control of all of the Middle East's known oil reserves—a commodity that had demonstrated its enormous strategic value in the recent war. As such, London had achieved the position of hegemonic dominance in the region that Lloyd George had sought since becoming prime minister.⁴¹

What is more, the British did not anticipate facing significant challenges to their rule in the region moving forward. Instead, the cabinet was quite confident that the Arabs would quickly accommodate themselves to trading the corrupt and inefficient rule of the Ottoman Empire for the stability, progress, and order that they believed London's mastery of the region would bring. Indeed, at the time that the prime minister signed the Treaty of Sèvres, he and his colleagues expected that the people of the region would not merely welcome Britain's enlightened administration but would also enthusiastically cooperate with their new imperial rulers.⁴²

³⁹ McMeekin, *The Ottoman Endgame*, 439–40.

⁴⁰ McMeekin, 439–40.

⁴¹ MacMillan, *Paris 1919*, 395; Kourosh Ahmadi, *Islands and International Politics in the Persian Gulf: The Abu Musa and Tunbs in Strategic Context* (New York: Routledge, 2008), 66–67.

⁴² Fromkin, *A Peace to End All Peace*, 420.

Resistance

These assumptions proved to be almost comically off the mark. The people of the Middle East sought not tutelage or exploitation, but instead independence and the opportunity to freely develop their economies and political institutions. As a result, rather than producing the gratitude and the positive, cooperative relationship between the people of the region and their new masters that British officials had anticipated, the efforts to secure hegemony over the Middle East birthed a series of vigorous resistance movements in nearly every part of the region. Popular demands for rapid postwar demobilization in Britain and elite insistence that the cabinet trim expenses so as to return the government to a balanced budget greatly complicated London's ability to respond to these challenges. As a result, while these resistance movements lacked the power to expel the British and its allies from the Middle East, neither could London impose its will on the region.⁴³

Egypt—the Wafd Party

That the resistance to the Treaty caught the British unaware is surprising in light of the fact that it first emerged more than a year before the formal signing of the Treaty of Sèvres. It did so in Egypt, and it occurred as a result of the hopes that Wilson's call for self-determination had raised. Until nearly the end of the conflict, Egyptian nationalists had sought only modest goals such as greater autonomy and a larger Egyptian role in the administration of the protectorate. As was the case in other parts of the world, however, Wilson's Fourteen Points Speech had dramatically altered expectations in Egypt. Thereafter, nationalists no longer hoped for mere autonomy, but instead anticipated that their country would, with the president's help, win complete independence following the war.⁴⁴

Accordingly, in November 1918, a group of nationalists led by the former Minister of Justice, Saad Zaghul (1859-1927), requested that Britain permit a delegation, or wafd, to travel to Paris to present Egypt's case for independence. Their demand alarmed British officials. Implacably opposed to granting independence to Egypt and thus ceding control of the Suez Canal, they flatly refused to permit the wafd to leave Egypt for France. The nationalists were not easily dissuaded, however. Anticipating that Wilson would help them once he learned of their plight, they responded by organizing rallies, collecting petitions, and forming a new political organization, the Wafd Party, to mobilize the people in support of independence. Still, the British refused to budge.⁴⁵

The standoff turned to open revolt in March 1919. Fearing that the nationalists' campaign was making inroads with the Egyptian masses, High Commissioner Reginald Wingate (1861-1953) arrested Zaghul and several other Wafd Party leaders and deported them to the island of Malta. It was a bold move—and a counterproductive one.

⁴³ Fromkin, 386–88, 420.

⁴⁴ Manela, *The Wilsonian Moment*, 63–75.

⁴⁵ Manela, 63–75.

Rather than quelling the growing opposition to British rule, the arrests outraged people throughout the country and sparked a province-wide mass uprising called the Revolution of 1919. Over the next few months, Egyptians rioted, participated in protests, tore up railways, cut telegraph lines, and went on strike to make clear their opposition to Britain's detention of the Wafd leaders and its refusal to discuss independence. The demonstrations were enormously popular and enjoyed support from all elements of Egyptian society. They brought Muslims and Christians together and, in a break with longstanding custom, even involved substantial numbers of women. The protests and, especially, the British response, were also violent. By the time the colonial authorities managed to restore some measure of order in April, eight-hundred Egyptians and sixty Britons had died in the unrest.⁴⁶

London responded by replacing Wingate with the more diplomatic Allenby in late March. Aware that the arrest of Zaghul and his colleagues had amounted to heaping fuel on a raging fire, the field marshal announced in short order that he was releasing the Wafd leaders. The move worked. To London's relief, even while the Egyptians continued to seethe at the British, stability returned to the province.⁴⁷

Meanwhile, now free to plead Egypt's case before the peace conference, Zaghul and his colleagues traveled directly from Malta to Paris. They got there too late. Just before the delegation finally arrived, the US formally declared that it was recognizing Britain's protectorate over Egypt. It was a sobering moment for the nationalists. Not only had Egypt failed to gain its independence, but Wilson, its purported champion, had repudiated their call for self-determination. Importantly, however, the president's apparent retreat from the principles that he had outlined in the Fourteen Points Address did not lead the nationalists to revert to the more-limited demands that they had held before the Wilsonian Moment. Instead, they pressed on with their call for nothing less than full independence.⁴⁸

Syria and Transjordan

Resistance to the Treaty of Sèvres next appeared in Syria. It did not, at least initially, come from the top. With the treaty having assigned Syria to France, the British had instructed Faysal to work out an agreement with Paris. Faysal was furious with London for abandoning him, but he was also a pragmatist. As such, he was open to accepting Paris's proposal that he serve as king of Syria in a titular capacity under French protection. Syrian nationalists were far less flexible than the king, however. On the contrary, adamant that Syria become a genuinely independent state rather than a nominally autonomous part of the French Empire, they pressured Faysal to adopt a hard line. They also began firing on French troops along the border between Syria and French-occupied Lebanon. Paris was not amused. Determined to assert control of Syria, it ordered High Commissioner Henri Gouraud (1867-1946) to invade Syria from

⁴⁶ Manela, 141–43.

⁴⁷ Fromkin, *A Peace to End All Peace*, 419–20.

⁴⁸ Manela, *The Wilsonian Moment*, 143–57.

Lebanon in July 1920. Disciplined and well-armed, his powerful army easily brushed aside a much smaller Arab force on July 24 and established control of Syria. A strident nationalist, Gouraud promptly set the tone for French rule by ostentatiously going to the Mausoleum of Saladin and declaring "arise Saladin, we have returned, and my presence here consecrates the victory of the Cross over the Crescent."⁴⁹

The Hashimites felt justifiably betrayed. London had first browbeaten Faysal into accepting the separation of Palestine from Syria and had then abandoned him to the French despite its earlier promises of support—a move that had resulted in Faysal fleeing, humiliated, into exile. Indeed, with his expulsion from Syria, the only provision in the McMahon-Husayn Correspondence that London was still honoring was the one calling for it to formally recognize Husayn's Kingdom of the Hijaz—a territory, the Hashimites pointed out, that they had controlled long before they had begun negotiating with Britain in 1915. The family was not entirely without recourse, however. In November 1920, Faysal's brother Abdallah, the future king of Jordan, and a small force of irregulars moved to contest Paris and London's overbearing actions by occupying Transjordan: the anarchic, lightly populated no man's land that lay on the eastern side of the Jordan River. While there, he announced that he was recruiting volunteers to join him in a campaign aimed at expelling the French from Syria and at restoring his brother to the throne. It was a bold action—one that greatly alarmed the British, who feared that the French would use his attacks as a pretext to add Transjordan to its Syrian Mandate.⁵⁰

Palestine

In the meantime, opposition to British rule and the Treaty of Sèvres had spread to Palestine. The Balfour Declaration's formal expression of support for a Jewish homeland was the primary source of this resistance. Though the statement had forthrightly declared that "nothing shall be done which may prejudice the civil and religious rights of existing non-Jewish communities in Palestine," the Arabs feared that London's pledge of support for Zionism would leave them a subordinate people in an alien and hostile state. A surge in Zionist immigration and an uptick in Jewish land purchases following Britain's assumption of control of the territory confirmed these concerns and rapidly raised tensions in the mandate. The mounting frustration finally boiled over with the outbreak of massive anti-Zionist riots in Jerusalem in April 1920 and Jaffa in May 1921. The intercommunal violence shocked the British. Resulting in the deaths of sixty-four Arabs and ninety-five Jewish people, it made clear to them for the

⁴⁹ MacMillan, *Paris 1919*, 406–7; Quote from Michael Provence, *The Last Ottoman Generation and the Making of the Modern Middle East* (Cambridge: Cambridge University Press, 2017), 3.

⁵⁰ Eugene Rogan, *The Arabs: A History – Revised and Updated Edition* (New York: Basic Books, 2012), 183–84.

first time the depths of Palestinian hostility to both Zionism and Britain and led at least some officials to begin to question the wisdom of the Balfour Declaration.⁵¹

The Iraq Revolt, 1920

The most serious resistance to Britain's play for dominance in the Arab world occurred in Mesopotamia in the summer of 1920. The territory, which British officials had taken to calling by its old Arab name, Iraq, was something of a Frankenstein creation. London had fashioned it by joining the heavily Kurdish *vilayet* of Mosul to the predominantly Arab *vilayets* of Basra and Baghdad. While the latter pair had strong historical links, ethnically distinct Mosul had only limited economic connections with them—a fact that suggested to some that welding the three *vilayets* into a coherent state would be a serious challenge. Indeed, one American missionary cautioned that merging them together was tantamount to “flying in the face of four millenniums of history.” Such warnings would prove prescient. As we shall see, the union of these *vilayets* would prove problematic. However, their merger was not the reason that Iraq erupted in the summer of 1920.⁵²

Instead, the revolt was a product of London's heavy-handed approach to the administration of the territory. Britain's establishment of a mandate without any consultation with the region's inhabitants quickly aroused intense nationalist sentiments among urban Arabs. More importantly, the imposition of levies designed to cover the costs of the mandate's administration infuriated the tribal Arabs of the rural areas who had not previously had to pay taxes. As a result, in early July, a *thawra*, or revolt, broke out along the Lower Euphrates River and quickly spread north to the region around Baghdad. The British were almost completely unprepared for it. Failing to grasp the degree to which Iraqis opposed the mandate, London had sought to maintain order in the territory with a garrison that was inadequate even in times of peace. Thoroughly overmatched by a rebellion supported by an estimated 131,000 insurgents, it could do nothing to prevent the guerillas from seizing huge swaths of territory in central Iraq.⁵³

London acted quickly to regain control of Iraq. Deploying reinforcements from India, it used a combination of overwhelming force and divide-and-conquer tactics to roll back the revolt. The British campaign was as brutal as it was effective. With the assistance of aircraft that dropped both conventional and mustard-gas bombs on villages believed to be supporting the insurgents, British and Indian troops were able to restore control over the main population centers and lines of communication by

⁵¹ Quote from “The Balfour Declaration, November 2, 1917”; Rogan, *The Arabs*, 197–98.

⁵² Fromkin, *A Peace to End All Peace*, 450–51.

⁵³ Toby Dodge, *Inventing Iraq: The Failure of Nation-Building and a History Denied* (New York: Columbia University Press, 2003), x, 8–9, 134; Quote from Fromkin, *A Peace to End All Peace*, 451.

October. Additional operations further suppressed the rebellion to the point that by February 1921, the British had succeeded in fully extinguishing the *thawra*.⁵⁴

Putting the revolt down had proven very costly, however. The fighting had resulted in the deaths of 500 Indian and British troops and an estimated 6,000 Iraqis. Worse, in the eyes of the budget-minded cabinet, moreover, it had set the exchequer back a whopping £40 million—a sum larger than the amount Britain had spent on the entire Arab Revolt against the Ottomans in 1917 and 1918.⁵⁵

More broadly, the revolt revealed the tension that existed between London's expansive goals in the region and the more-limited resources it was willing to expend to attain those ends. That gap was substantial. On the one hand, popular demands for demobilization and elite insistence on a balanced budget resulted in the government cutting military expenditures in half every year from 1919 to 1923. On the other, Lloyd George's cabinet was determined to secure the fruits of Britain's victory in the Middle East by establishing control of the territories that its armies had taken possession of during and just after the war—a task that appeared to require a great deal of money and troops. The revolt thus brought the contradiction between Britain's limited means and its expansive objectives into clear focus. Simply put, hegemony in the Arab Middle East along the lines outlined in the Treaty of Sèvres demanded money and troops that Britain did not have.⁵⁶

Turkey

Meanwhile, the most successful and consequential opposition to Britain's effort to achieve hegemony in the Middle East occurred in Turkey. The resistance there was so effective, in fact, that the country was able to renegotiate the peace that the victors had imposed on it—something none of the other members of the Central Powers would manage to achieve. That it was able to do so was astonishing in light of the feeble position that the Ottoman Empire occupied immediately following World War I. Brutally one sided, the armistice of Mudros had compelled Istanbul to demobilize its military forces and gave the Allied powers carte blanche to occupy any territory they deemed strategic—thus allowing the victors to dominate completely the Ottoman Empire and its government. The armistice agreement also included provisions permitting the Allied powers to take control of key territories and to neutralize the straits. By early 1919, as a result, the Turks had hit rock bottom. Crushed militarily and diplomatically and lacking a genuinely autonomous government, they appeared to be a thoroughly beaten people, powerless to resist further territorial demands.⁵⁷

⁵⁴ Christopher Catherwood, *Churchill's Folly: How Winston Churchill Created Modern Iraq* (New York: Basic Books, 2005), 85; Fromkin, *A Peace to End All Peace*, 452–53.

⁵⁵ Charles Tripp, *A History of Iraq* (Cambridge: Cambridge University Press, 2002), 8–9.

⁵⁶ Tripp, 43–44; Dodge, *Inventing Iraq*, 8–9, 134–35.

⁵⁷ McMeekin, *The Ottoman Endgame*, 406–11.

Ironically, their revival was the unintended consequence of a dispute among the Allies in early 1919 over the disposition of Turkish territory. At issue was the predominantly Greek city of Smyrna that both Italy and Greece coveted. Annoyed with Rome’s increasingly expansive territorial aspirations, Wilson and Lloyd George had sought to preempt an Italian move on the city by authorizing the Greek government to deploy troops to occupy it in May 1919. The decision delighted Greece’s nationalistic prime minister, Eleftherios Venizelos (r. 1917-1920). Committed to achieving the *megali idea*, he wasted no time in sending a substantial army to Smyrna. Upon its arrival, it received a hero’s welcome from a Greek community that was ecstatic that the city was going to become part of Greece. Unsurprisingly, the Turks did not share their enthusiasm. On the contrary, the deployment of Greek soldiers to Asia Minor aroused widespread anger and produced a surge of nationalist sentiment among Anatolia’s Turkish majority. In fact, it was the arrival of Greek troops that catalyzed the emergence of Turkish resistance to the peace plan. As future President Kemal Mustafa (r. 1923-1938), later wrote, “[if] the enemy had not stupidly come here, the whole country might have slept on heedlessly.”⁵⁸

Determined to prevent the carve up of Anatolia, Kemal Mustafa assumed leadership of an embryonic Turkish nationalist movement in May 1919. A veteran of the Gallipoli campaign, he quickly established a new political infrastructure and a reformed military force in Anatolia. He also oversaw the drafting of the National Pact: a declaration of principles that renounced the Arab portion of the Ottoman Empire while simultaneously calling for the establishment of a Turkish state free of foreign interference in those areas that had a Turkish majority—including Armenia and Kurdistan.⁵⁹

Even as Kemal Mustafa steadily gained the support of a growing number of Turks in 1919, the British remained oblivious to the rapidly maturing resistance he led. In fact, they only awoke to the danger that his movement posed to their designs for the Middle East when the newly elected Ottoman Chamber of Deputies formally adopted the National Pact in February 1920. Finally grasping the threat that the nationalists posed, Lloyd George moved quickly to crush the movement by ordering the occupation of Istanbul and the dissolution of the Ottoman legislature in March 1920. This attempt to hobble the nationalists backfired badly, however. The occupation destroyed what little credibility the sultan still possessed even as Kemal Mustafa’s movement gained legitimacy thanks to its resolute opposition to the Allies—particularly after he made Ankara the new capital by convening a new legislative body there in April 1920 called the Grand National Assembly. Collectively, these events dramatically raised the nationalists’ standing. From that point forward, the sultan’s regime may have continued

⁵⁸ Erik Goldstein, *The First World War Peace Settlements, 1919-1925: From Versailles to Locarno, 1919-25* (Routledge, 2013), 59.

⁵⁹ Erik J. Zürcher, *Turkey: A Modern History*, Third Edition (London: I.B. Tauris, 2004), 149–52.

to retain international recognition, but it was Kemal Mustafa's government that enjoyed popular support—and thus genuine sovereignty—among the Turks.⁶⁰

With his seizure of Istanbul having failed to arrest the nationalist movement, Lloyd George found himself caught in a steadily worsening situation. By June, Kemal Mustafa's increasingly aggressive soldiers were menacing British troops guarding the neutral zone that the Allies had set up around the straits. Compounding the situation, the British government's self-imposed fiscal restraints and prior military commitments meant that the prime minister had no reserves with which to reinforce them. Nonetheless, Lloyd George refused to scale back his goals in Asia Minor. Instead, seeking to uphold the terms of the Treaty of Sèvres, he authorized Greece—effectively Britain's catspaw in the region—to march its army out from Smyrna and secure the coastal plain. Venizelos happily complied. Eager to pursue his country's irredentist claims in Anatolia, he ordered the Greek army to occupy the lowlands of western Asia Minor—a move that compelled Kemal Mustafa's forces to retreat to the central plateau. It was an impressive victory on paper, one that both secured control of substantial territory and protected the straits. It had no substantive effect on the strategic balance, however. So long as the nationalists maintained an army in central Anatolia, the peace settlement on which Britain's grand designs for the Middle East rested could not be guaranteed and was thus incomplete.⁶¹

Despite the mounting risks, Lloyd George and the Greeks continued to bet on a military solution. It was a gamble with ever-worsening odds thanks to the Turks' dramatically improved position in 1920 and early 1921. Two factors explain the nationalists' newfound strength. First, seeking to curb Britain's influence in the region, Moscow had agreed in late 1920 to supply them with arms—a deal that greatly redressed the imbalance between Turkish forces and the Greek army. Second, nationalist diplomatic efforts to peel off London's allies bore fruit. Having concluded that Kemal Mustafa's movement could not be beaten at an acceptable cost, the French and Italian governments had agreed to evacuate their troops from Asia Minor in 1921 in exchange for economic concessions. As a result, the nationalists were able to free up further troops for use against the Greeks.⁶²

Unwilling to cut their losses, Athens and London pressed on. With Lloyd George's blessing, the Greek army launched an offensive toward Ankara in June 1921 aimed at destroying the nationalist movement. It was a long and brutal campaign in which the Greeks systematically burned Turkish villages in order to secure the tenuous supply lines needed to sustain their army so far from its bases. Finally, at the end of the summer, the invaders compelled Kemal Mustafa's nationalist military to stand and fight along the Sakarya River just fifty miles from Ankara. While the Greeks prevailed in the first few days of the battle, a powerful Turkish counterattack broke their morale and

⁶⁰ Caroline Finkel, *Osman's Dream: The History of the Ottoman Empire* (New York: Basic Books, 2005), 541–42.

⁶¹ Fromkin, *A Peace to End All Peace*, 430–32.

⁶² McMeekin, *The Ottoman Endgame*, 443–52.

compelled them to retreat to a new line on the edge of the Anatolian plateau. It was a huge triumph for the nationalists—one that further solidified the movement's political standing among the Turks.⁶³

The nationalist campaign of resistance ended in victory the following year. In August 1922, Kemal Mustafa defeated the Greek army so completely that it first retreated in disorder to the coast and then abandoned Anatolia altogether. Retribution soon followed. Most notoriously, after Venizelos's forces withdrew, a massive fire destroyed the Greek quarter of Smyrna and rendered hundreds-of-thousands of its residents homeless. Facing the prospect of further violence, nearly all of them fled Anatolia for Greece as refugees. More would follow. Finally, in October, the arrival of Turkish troops opposite British positions guarding the straits at Chanak put Britain and the nationalists on the brink of war. Some of the hardliners in the British government wanted to fight, but, this time, more moderate voices prevailed. As a result, London and the nationalists agreed to a ceasefire that called for the Turkish army to stand down in exchange for a British promise to renegotiate the Treaty of Sèvres. Lloyd George would not be involved in those talks, however. Coming on the heels of the other problems that his aggressive policies in the Middle East had created, the Chanak Crisis brought down his government and ended his stint as prime minister.⁶⁴

The Settlement of 1922

The extent and intensity of the resistance to British hegemony that culminated in the Chanak Crisis compelled London to rethink its approach to the Middle East. Not only had Lloyd George's plan for regional dominance failed, but it had imposed costs that the British government was unwilling to shoulder. London did not completely abandon the former prime minister's drive for regional hegemony, however. Instead, it articulated in piecemeal fashion the Settlement of 1922—a new approach toward the region designed to replace the Treaty of Sèvres on which Lloyd George's vision of British dominance had rested with a new arrangement designed to sustain London's control of the Middle East at a tolerable cost by meeting the minimal demands of the region's peoples. It was a complex settlement—one that included a renegotiated peace settlement with the Turkish nationalists, nominal Egyptian independence, and, crucially, an effort by British officials to reconcile the McMahon-Husayn Correspondence, the Balfour Declaration, and the Sykes-Picot Agreement. The settlement was successful in the short run insofar as it helped London achieve its goal of retaining a dominant position in the Middle East during the interwar period at a cost that London found acceptable. It did not set the stage for the region's long-term stability or development, however. On the contrary, it bred resentments, undermined prospects for the institutionalization of representative governments, and created states whose legitimacy—and even their very right to exist—remained widely questioned.⁶⁵

⁶³ McMeekin, 454–58.

⁶⁴ MacMillan, *Paris 1919*, 450–52.

⁶⁵ Fromkin, *A Peace to End All Peace*, 559–63.

The Treaty of Lausanne, July 1923

The effort to resolve the Chanak Crisis ultimately produced the most enduring part of the settlement: the Treaty of Lausanne. Its negotiation was nearly as hard fought as the combat that had preceded it. Determined to drive a tough bargain, Britain set the tone by inviting the sultan’s government to send a separate delegation in a bid to weaken the nationalists’ position; no stranger to hardball tactics, Kemal Mustafa deftly countered that move by abolishing the sultanate in November 1922. The tough negotiations continued after the conference finally got underway. Most notably, the Turkish representative, the former general, Ismet Inonu (1884-1973), habitually signaled his opposition to London’s proposals by demonstratively removing his hearing aid when the British negotiator, Lord Curzon (1859-1925) was speaking and only replacing it when he had finished.⁶⁶

The resulting treaty was a major victory for the nationalists. To be sure, Kemal Mustafa and his allies did have to make some concessions. For example, Inonu had no choice but to agree to permit the League of Nations to determine the disposition of the *vilayet* of Mosul—a region that Kemal Mustafa had insisted should revert from Iraqi to Turkish control. The balance of the treaty more than offset such provisions, however. It absolved Turkey of the burden of paying reparations, ended the much-resented capitulations, and—though it called for their neutralization—returned the straits to Turkish sovereignty. More importantly, it assigned to Turkey all of the parts of the former Ottoman Empire in which Turkish people composed a majority of the population and thus secured international recognition of the National Pact. Relatedly, the agreement helped the nationalists achieve their goal of making their country more ethnically homogenous. Most notably, it provided for “a ‘compulsory exchange’ of populations” between Greece and Turkey that resulted in the eviction of 1.3 million Greek people from Turkey and the expulsion of 800,000 Muslims from Greece in what was the first instance of internationally sanctioned ethnic cleansing following World War I. Sadly, it would not be the last.⁶⁷

The Cairo Conference

While Curzon directed the negotiations with Turkey, Winston Churchill (1874-1965), in his new capacity as secretary of state for the colonies, oversaw the restructuring of Britain’s empire in Palestine and Iraq. As Churchill knew all too well, 1920 had been a rough year for Britain’s imperial pretensions in the region. The rioting in Jerusalem suggested that it had stumbled into a costly quagmire in Palestine, while Abdallah’s threat to invade French-controlled Syria raised the possibility that the Hashimites could drag Britain into a crisis with its French ally-cum-rival. Worst of all, the revolt in Mesopotamia had turned Iraq into a money pit at a time when the government

⁶⁶ MacMillan, *Paris 1919*, 452; James William Spain, *In Those Days: A Diplomat Remembers* (Kent, OH: Kent State University Press, 1998), 162.

⁶⁷ Steiner, *The Lights That Failed*, 120–23; Quote from McMeekin, *The Ottoman Endgame*, 489.

was committed to aggressively cutting the budget. Coming in rapid succession, these problems were dispiriting and seemed insoluble to many observers. Indeed, the situation appeared so intractably bleak in early 1921 that some officials were beginning to quietly wonder whether Britain should abandon the Middle Eastern mandates.⁶⁸

Churchill did not share that view. A committed imperialist, he moved with characteristic energy to find a solution to the empire's problems in the Arab world. He had three explicit goals. First, believing that Britain's Middle Eastern mandates were valuable additions to the empire, he sought to ensure that London retained control of them. Second, aware that Britain needed to economize in order to get back on the gold standard, he wanted to dramatically reduce the cost that they imposed on the exchequer. Finally, he hoped to undo the damage that London had done to its international reputation by failing to live up to the terms of its wartime agreements regarding the Middle East. That is, he wanted to find a way to reconcile the seemingly incompatible terms of the Sykes-Picot Agreement, McMahon-Husayn Correspondence, and Balfour Declaration.⁶⁹

These were daunting goals, even for a person with Churchill's diplomatic savvy. Retaining possession of the territory it had secured through the Treaty of Sèvres required the maintenance of substantial civil and military apparatuses in the region—structures that, by their very existence, complicated efforts to reduce the budget. Likewise, harmonizing diplomatic agreements that had promised the same territory to different parties would be no simple task.⁷⁰

To achieve these challenging ends, Churchill convened the Cairo Conference at the Semiramis Hotel in Cairo in March 1921. The list of British officials and experts who took part was impressive. The attendees included high-ranking bureaucrats and military officers stationed in the region as well as Middle East experts such as T. E. Lawrence and Gertrude Bell (1868-1926), the writer and traveler who had assumed a position as a key administrator in Iraq. Tasked with resolving the problems that Britain had experienced in its Middle Eastern mandates, the attendees would make a series of important decisions that would have momentous long-term consequences for the region.⁷¹

The conference first dealt with the situation in Iraq. Wracked by a costly insurgency in the countryside and by seething opposition in the cities, it was the most troubled part of Britain's Middle Eastern empire in 1921. After vigorous debate, the attendees determined that they could best ensure the mandate's long-term stability by shifting to a system of indirect rule—one that they believed would be more acceptable to the territory's Arab population. Accordingly, in August 1921, they arranged for the now-

⁶⁸ Catherwood, *Churchill's Folly*, 94–103.

⁶⁹ Fromkin, *A Peace to End All Peace*, 493–99, 528–29.

⁷⁰ Fromkin, 493–99, 528–29.

⁷¹ Barr, *A Line in the Sand: The Anglo-French Struggle for the Middle East, 1914-1948*, 112.

unemployed and unhappy Faysal to become the king of Iraq. In Churchill's eyes, it was a savvy solution that neatly killed two birds with one stone: it both promised to reduce unrest among urban Iraqis and made good London's failure to adhere to the McMahon-Husayn Correspondence—thus helping to restore Britain's battered diplomatic reputation.⁷²

What it did not do, however, was address the urgent problem posed by the need to contain costs. How, the attendees debated, could London maintain control of Iraq's vast rural areas in a way that was consistent with its determination to drastically pare back government expenditures? For Churchill, technology and brute force provided a way out of this conundrum. That is, Britain would uphold order and internal stability in Iraq by using airpower to coerce obedience. Cheap and powerful, airplanes could, through the threat of bombing, compel the seemingly ungovernable rural tribes to pay the taxes that would make mandatory Iraq financially self-sustaining—thus relieving Britain of the cost of administering it.⁷³

Having arrived at a solution for Iraq, the conferees next addressed the problems of Palestine. Here, London faced a series of peculiarly challenging tasks. It needed to end communal violence between Palestinian Arabs and Jewish immigrants, reconcile the conflicting promises it had made in the Balfour Declaration and the McMahon-Husayn Correspondence, prevent Abdallah's presence in Transjordan from embroiling Britain in a conflict with France, and, lastly, satisfactorily dispose of the territory on the east bank of the Jordan River that Faysal had ruled until July 1920. After lengthy debates, the attendees concluded that they could neatly resolve all of these issues through a combination of Solomonic division and the appeasement of the Hashimites. Implemented in 1922, this new approach involved two steps. First, it split the mandate of Palestine along the Jordan River and declared that Transjordan, the territory on the east bank, would not be subject to the Balfour Declaration. Second, using the same indirect system of colonial government that the conference had adopted for Iraq, it formally installed Abdallah as Transjordan's nominal ruler.⁷⁴

Churchill was delighted with this solution. In his view, it deftly resolved all of the problems that had emerged in Palestine in 1920. By creating an explicitly Arab state through the division of the mandate, it would ease Arab opposition to Zionism and would thus end the resentment that had fueled the riot in Jerusalem in 1920. By making Abdallah the ruler of Transjordan, it would dissuade the emir from dragging Britain into a potentially costly conflict with France. Above all, by dividing the mandate of Palestine, his plan would, in his eyes, reconcile the commitment Britain had made in the McMahon-Husayn Correspondence to foster the creation of Hashimite states with the promise it had expressed in the Balfour Declaration to oversee the establishment of a

⁷² Fromkin, *A Peace to End All Peace*, 503–4.

⁷³ Catherwood, *Churchill's Folly*, 129–38.

⁷⁴ Eugene L. Rogan, "The Emergence of the Middle East in the Modern State System," in *International Relations of the Middle East*, 4th ed. (New York: Oxford University Press, 2016), 50–51.

Jewish homeland in Palestine and would thus allow Britain to meet all of the obligations it had made during the war. In hindsight, of course, it achieved only the second and third of these aims. That is, while the conference's plan did ease Hashimite frustration with London and did permit Britain to claim that it had stood by its word, it proved utterly incapable of resolving either the competition between the Zionists and the Arab population of Palestine or, more broadly, the contradiction inherent in the Balfour Declaration's pledge to create a Jewish homeland while simultaneously respecting the rights of non-Jewish people in the mandate. Churchill was blind to these shortcomings, however, and genuinely believed that he had resolved in Cairo the issues that had bedeviled mandatory Palestine.⁷⁵

Thus, at the conclusion of the conference, Churchill felt that he had put Britain's policy toward the Middle East mandates back on track. Indeed, he was quite pleased with his work. In conjunction with the renegotiation of the Sykes-Picot Agreement that had earlier taken place at the San Remo Conference, he believed that his reorganization of the terms of the mandates in Palestine and Iraq had finally reconciled the contradictory promises that Britain had made in its wartime agreements regarding the Middle East and had thus restored its tarnished diplomatic reputation. All the relevant parties had to make compromises, he acknowledged, but each also received essentially what Britain had promised them. Meanwhile, by shifting to a system of indirect, Hashimite rule in Transjordan and Iraq—the so-called "Sherifian Solution"—and by relying on the threat of bombing villages in Iraq to uphold order, he was confident that he had devised a way to ease the expense of empire to the point that Britain would be able to retain possession of its new and strategically valuable Arab territories at an acceptable cost. Unaware of the conference's genuine shortcomings, in sum, Churchill left Cairo confident that he had resolved the serious problems that had emerged in the mandates.⁷⁶

Egypt: The Four Reserve Points

Separately, meanwhile, Britain finally managed to clarify its relationship with Egypt in 1922. It had been no simple task. From nearly the start of his tenure as high commissioner, Allenby had sought to end the unrest in the protectorate by negotiating a treaty of independence with representatives from the Wafd Party. Despite the fact that both sides were in agreement that the talks should end with Egypt becoming independent, however, the discussions quickly deadlocked over the status of the Suez Canal. London remained adamant that the final agreement permit it to maintain a permanent military presence in the canal zone; Zaghul and the Wafd Party in contrast were equally determined to end the occupation altogether. As a result, the two sides

⁷⁵ Fromkin, *A Peace to End All Peace*, 510–29.

⁷⁶ Quote from Rogan, *The Arabs*, 183; Fromkin, *A Peace to End All Peace*, 528–29.

could not make meaningful progress toward an agreement and the unrest that had started in 1919 continued to grind on.⁷⁷

Allenby finally broke the impasse in February 1922. Unable to stabilize the situation in Egypt through negotiations, he resolved it by issuing the Unilateral Declaration of Egyptian Independence. It promised the quick termination of martial law and formally ended the protectorate—thereby making Egypt an independent state. However, the independence that Egypt acquired was incomplete thanks to a series of colonialist qualifications that Allenby included in his declaration. Known as the four “Reserved Points,” these conditions retained for Britain control of the Suez Canal, responsibility for Egypt’s defense, the authority to protect the interests of foreigners and minorities in the country, and, finally, the administration of Sudan—which Egypt and Britain had held in condominium since 1882.⁷⁸

Unsurprisingly, Allenby’s declaration met with a mixed reception among Egypt’s increasingly nationalist-minded people. They were thrilled that their country had achieved independence, of course, but they were equally bitter in their resentment of the Reserved Points—provisions that, in their view, had merely transformed Egypt into a subtler kind of protectorate. This criticism had merit. While Allenby’s statement certainly marked the passing of an important milestone for Egypt on its march to full independence, it also ensured that Britain would continue to dominate their country for another generation.⁷⁹

For London, in contrast, the declaration was a major success. By meeting Egypt’s minimal demands, Allenby had succeeded in taking the air out of the protest movement that had threatened Britain’s position in the country. As such, while the Egyptians continued to seethe with resentment, they no longer engaged in the civil disobedience and mass protest that had threatened to undermine Britain’s position in Egypt and, more critically, its continued control of the Suez Canal.⁸⁰

Conclusion

With the situation in Egypt having finally reached a resolution, the Settlement of 1922 was complete, and, with it, the broad outlines of the modern Middle East had taken form. Politically, it was a vastly different region than it had been at the start of World War I. The Allies had crushed the Ottoman Empire in the conflict and, in so doing, appeared to have finally resolved the “Eastern Question.” Meanwhile, the process of imperial aggrandizement, resistance, negotiation, and accommodation that occurred in

⁷⁷ M W Daly, “The British Occupation, 1882-1922,” in *The Cambridge History of Egypt. Volume 2, Modern Egypt, from 1517 to the End of the Twentieth Century*, ed. M. W. Daly (Cambridge: Cambridge University Press, 1998), 249–51.

⁷⁸ Daly, 249–51.

⁷⁹ Jayne Gifford, *Britain in Egypt: Egyptian Nationalism and Imperial Strategy, 1919-1931* (London: I. B. Tauris, 2020), 180.

⁸⁰ Gifford, 180.

the years immediately after the war had resulted in the division of the region into a series of nation-states—albeit ones that were, in nearly all cases, *de facto* colonies of France or Britain. Those states still exist today; likewise, with some notable exceptions, the frontiers that the victors drew in the early 1920s continue to divide the modern Middle East.

This new state system would not prove to be a stable one, however. The French and British had chosen leaders, drawn up borders, and created political structures based on a mix of self-interest and ignorance; as a consequence, most of the governments that emerged in the region either lacked legitimacy with their people or controlled territory to which, in whole or in part, other states staked claims. Indeed, directly or indirectly, the contested peace settlement that followed the First World War gave birth to nearly all of the modern Middle East's major geopolitical issues including, most notably, the Israeli-Palestinian conflict, Kurdish demands for an independent country, instability in Lebanon, and the weakness of the Iraqi state. This is not to deny agency to the people of the Middle East; after all, as we have seen, regional actors played a part in shaping the outcome of the settlement of 1922. Rather, it is to make clear the degree to which the negotiated peace reflected the imbalance in power between the indigenous parties and the imperialist states in the early 1920s. Indeed, even those Middle Eastern countries that had managed to remain independent after the settlement of 1922 found themselves compelled to walk a careful line *vis-à-vis* the Great Powers during the interwar era. It is to their development during that period and their efforts to remain free from Western dominance that we shall now turn.

Chapter Eleven: The Independent States, 1918-1948

As we saw in chapter ten, most of the Middle East came under direct European rule following the Settlement of 1922. For different reasons, however, three of the region's states—Saudi Arabia, Iran, and Turkey—managed to retain their independence. Among the small number of non-Western countries to avoid direct imperial control during the interwar era, these states were able to autonomously develop their political institutions, economies, and societies.

Still, all three found the period between the end of World War I and the conclusion of World War II to be a challenging time during which they had only mixed success in maneuvering through the treacherous waters of the Western-dominated, interwar Middle East. Of the three, Iran was by far the least successful. It failed to develop either a strong, modern economy or a political system that enjoyed close connections to the people; worse, it ended up coming under Allied military occupation during World War II. Saudi Arabia fared better thanks to its military power and its possession of substantial oil reserves but still chafed at British dominance of the region. Turkey was a different story. In contrast to Saudi Arabia and Iran, it appeared to enjoy genuine independence and, for a time, seemed to have blazed a path that other states could follow to achieve autonomy and industrialization. Yet even Turkey had to walk a very careful line to avoid foreign domination during the Second World War.

Turkey: Independence and Cultural Revolution

During the interwar period, Turkey was widely considered to be one of the world's great success stories. Its leaders had skillfully navigated the collapse of the Ottoman Empire, secured independence, and modernized both the economy and the political system. More dramatically, they had launched a far-reaching, top-down cultural revolution that had succeeded in effecting a radical restructuring of Turkish society and identity. Indeed, under its powerful president, Kemal Mustafa (r. 1923-1938), Turkey appeared to have successfully undergone a nearly unprecedented nation-building campaign that had fundamentally remade the predominantly Turkish part of the Ottoman Empire into a new nation-state replete with a new Turkish identity to replace the old Ottoman one. The regime's reform program was so successful, in fact, that it came to constitute an influential model that other independent, non-European states would embrace during the interwar period and that many newly decolonized African, Asian, and Middle Eastern states would adopt following World War II.

Turkey in 1923

Having been involved in war almost continuously since 1911, Turkey was a deeply troubled state following the signing of the Treaty of Lausanne in July 1923. Demographic and economic evidence make plain the degree to which a decade of intense conflict had affected the country. The fighting, ethnic cleansing, war-induced famine, and genocide had combined to shrink the population of Turkey from an

estimated 16.3 million people in 1914 to 13.6 million at the time of its first census in 1927—a staggering 17 percent drop. The population was not only much smaller, moreover, but also far-more ethnically homogenous. Thanks to the departure of more than one-million Greek refugees between 1912 and 1923 and the genocide of between 800,000 and 1,000,000 Armenians between 1915 and 1918, Turkey emerged from the wreckage of the Ottoman Empire an overwhelmingly Turkish-Muslim state that had only one substantial minority, the Sunni Kurds, who accounted for about 20 percent of the population. Economically, meanwhile, the country's situation following the Treaty of Lausanne was similarly dismal. The strain of more than a decade of continuous warfare, the loss of the comparatively entrepreneurial Greeks and Armenians who had dominated the commercial economy, and the establishment of new borders that disrupted longstanding trade relations left the country with a severely depressed economy. Thus, while the Turkish people greeted the signing of the Treaty of Lausanne with euphoria, their joy was more-than tempered by the grim situation in which their country found itself in 1923.¹

Consolidating Power, 1923-1927

Perhaps unsurprisingly given Turkey's difficult circumstances, the leadership of the resistance movement that had secured independence began to fracture immediately following the ratification of the Treaty of Lausanne. The primary fault line was not, at least initially, ideological or economic in nature; instead, it centered on the amount of power that President Kemal Mustafa would wield in the new government. In this struggle, the Turkish president would win a substantial victory. As savvy a political operator as he was a military commander, he deftly handled the challenge to his authority that emerged following Turkey's diplomatic victory in Lausanne by leveraging his status as the main hero of the war for independence to outmaneuver and swiftly neuter the opposition. In the process, he succeeded in transforming Turkey into an autocratic, one-party state under his tight control.

Latent opposition to Kemal Mustafa had long existed in the national independence movement, but it only took concrete form following the nationalists' victory. It did so in response to two acts that he successfully forced through parliament in late 1923 and early 1924. First, to foreclose a possible restoration of the sultanate, he exploited his control of the newly created Republican People's Party (RPP) to secure passage of a law in October 1923 declaring Turkey a republic. He followed by persuading the legislature to formally abolish the caliphate in March 1924 and to adopt a new, republican constitution the following month—moves that ensured that Turkey would be a secular rather than religious state.²

Many in the independence movement found these actions deeply troubling. Hoping that Turkey would retain a Sunni Muslim identity, conservatives were furious

¹ Douglas A. Howard, *The History of Turkey* (Westport, Conn: Greenwood Press, 2001), 99.

² Ahmad Feroz, *The Making of Modern Turkey* (New York: Routledge, 1993), 53–56.

that he had orchestrated the abolition of the caliphate. For their part, many politically moderate veterans of the independence effort shared Kemal Mustafa's belief that Turkey had to become a republic and abolish the caliphate if it wished to emerge as a modern nation-state, but they found his heavy-handed parliamentary tactics troubling and feared that he was acting less to advance Turkey's interests than to consolidate authority in his hands. In response, they moved to form a new political party, the Progressive Republican Party (PRP), in November 1924. A typical, European-style secular-nationalist party, the PRP promoted a moderate, decentralized, and democratic political vision as an alternative to the RPP's autocratic and more revolutionary approach. It quickly won a substantial following among the public.³

Bridling at the growing popularity of the PRP, Kemal Mustafa and his supporters—known as Kemalists—schemed to undermine the moderates' challenge. A Kurdish rebellion that broke out in southeastern Turkey gave them just the pretext they were seeking. The revolt was the product of Kurdish anger about the recent changes the RPP had effected. Concerned that Kemal Mustafa's promotion of Turkish nationalism and secularism would render them second-class citizens, the Kurds rose up in revolt in February 1925 under the leadership of the charismatic dervish leader Shaykh Sait (1865-1925). The rebels enjoyed a few early successes but ultimately could not stand up to Kemal Mustafa's hardened troops. By late spring, as a result, the government had suppressed the rebellion.⁴

The Turkish president and his supporters wasted little time in using the rebellion to move against their political enemies in the PRP. In March 1925, the Grand National Assembly passed the Law on the Maintenance of Order, which gave the government sweeping powers to muzzle the press and to ban organizations that threatened to undermine law and order. Though they had justified it as necessary for the suppression of the Shaykh Sait revolt, the Kemalists quickly exploited the law to crush their political rivals. They permanently shuttered all independent national newspapers, abolished the PRP, arrested 7,500 opponents of the RPP's secularizing program, and had 660 of them executed.⁵

A failed assassination attempt against the president in June 1926 gave the Kemalists the chance they needed to finish off their political opponents. Once again relying on the Law on the Maintenance of Order, they arrested all prominent members of the PRP. They followed by holding a series of show trials aimed at discrediting those PRP leaders who had been prominent figures in the independence movement and who could thus credibly challenge Mustafa Kemal. Though discontent in the army compelled the Kemalists to eventually release those who had served prominently in the national resistance struggle, the tribunals had the desired effect: they ended all organized

³ Erik J. Zürcher, *Turkey: A Modern History*, Third Edition (London: I.B. Tauris, 2004), 166–69.

⁴ Carter V. Findley, *Turkey, Islam, Nationalism, and Modernity* (New Haven: Yale University Press, 2010), 251.

⁵ Howard, *The History of Turkey*, 95.

political opposition to Kemal Mustafa and his platform. Thereafter, the Turkish president ruled without restraint.⁶

Kemal Mustafa capped his consolidation of power with a multi-day address to the RPP's Second Congress in October 1927. Known as the *Nutuk*, or the Speech, its purported aim was to outline the history of the national resistance movement and the early republic. In fact, its goal was to use the past to justify the recent purges and to rationalize continued Kemalist dominance of Turkey. Casting most other leaders of the independence movement as either feckless or traitorous, the Speech presented what came to constitute, in the historian Erik Zürcher's words, a Kemal Mustafa-centered, establishment "master narrative" of the national resistance struggle and the early years of the republic—one that diminished the role that others had played in the national liberation movement even as it solidified the Turkish president's status as the unassailable leader of the republic.⁷

The *Nutuk* proved enormously successful. Coming from a figure of such prominence, it became the starting point for the nationalist history curriculum subsequently taught in Turkish schools. More importantly, it created the collective historical memory of the Turkish War for Independence that would dominate understandings of the nationalist movement for the next two generations—one that transformed Kemal Mustafa from the movement's leading figure into the very personification of the struggle itself. Thereafter, accounts of the war for independence centered almost exclusively on the role that he played in leading the nationalist effort and downplayed or outright erased the contributions that other leaders made. The *Nutuk* thus eviscerated the organized political opposition, led, as it was, by prominent veterans of the independence movement, and ensured that Kemal Mustafa and the RPP retained total dominance of the Turkish government. Indeed, save for a brief period in 1930, Turkey remained an autocratic, one-party state through the end of World War II—a situation that permitted the Kemalists to undertake an increasingly intense drive to radically restructure and modernize Turkish society.⁸

The Reforms of the 1920s

Startlingly rapid in pace, the reform program began with a bang in 1924. That year, the government asserted state control of religion by shuttering the madrasas, taking control of education, closing the shari'a courts, abolishing the office of sheikh ül-Islam, and establishing a Directorate of Religious Affairs. It followed the next year by shutting down the popular Sufi dervish convents, by strongly discouraging women from veiling, and by banning the fez—the hat, ironically, that Sultan Mahmud II (r. 1808-1839) had introduced in the 1820s as part of an earlier reform effort. The pace did not let up in 1926. That year, the government switched Turkey to the twenty-four-hour international

⁶ Findley, *Turkey, Islam, Nationalism, and Modernity*, 251.

⁷ Erik J. Zürcher, *The Young Turk Legacy and Nation Building: From the Ottoman Empire to Atatürk's Turkey* (New York: I.B. Tauris, 2010), 6–16.

⁸ Findley, *Turkey, Islam, Nationalism, and Modernity*, 257.

clock and required people to abandon the Islamic calendar in favor of the Western Gregorian one. It also adopted a new legal system, the Swiss Civil Code, that outlawed polygamy and formally ended 'ulama' control of family law. The state continued the secularization and modernization drive in 1928 by making the use of European (Arab) numerals mandatory and by removing the clause from the constitution that declared that Islam was the religion of the republic.⁹

That same year, Kemal announced what was perhaps the most far-reaching reform yet implemented. On August 9, 1928, he declared that Turkey would replace the traditional Ottoman Arabic-Persian script with a modified version of the Latin alphabet. Proponents of the new, twenty-nine-letter writing system argued that it was more streamlined and better suited to spoken Turkish than the written Arabic of the Ottoman era. They also maintained that it was easier to learn and would thus permit a rapid increase in Turkey's comparatively anemic literacy rate. These were persuasive arguments for the reform-minded Kemalists. Accordingly, on November 1, 1928, the RPP-dominated Turkish parliament mandated that the use of the new script would be compulsory for all publications effective January 1, 1929. Thus, the regime gave literate Turks the unenviable task of learning to read and write a new, and very-foreign alphabet in just two months!¹⁰

What did Kemal Mustafa and his supporters seek to achieve by decreeing so many sweeping reforms so quickly? After all, a more deliberate pace would have given people time to acclimate to the changes. Two objectives impelled them to reject a moderate timeline in favor of an aggressive one. First, they moved quickly because they wanted to rapidly weaken the authority of the religious scholars—the group that the historian Douglas Howard argues was “[t]he single most important potential challenger to the legitimacy of the republican regime.” Indeed, many of the Kemalists' early reforms—abolishing the sheikh ül-Islam, closing the shari'a courts, ending religious control of education, and adopting the Swiss Civil Code—were designed in large part to weaken the 'ulama'. Second, Kemal Mustafa and his supporters promoted the rapid transformation of Turkish society in order to shock the Turkish people into a complete reconstruction of their identity. By compelling Turks to live according to the Western calendar and the international clock, making them adopt European modes of dress, and, especially, forcing them to communicate using the Latin script, the Kemalists aimed to reorient their country away from its Islamic and Ottoman past and toward the dynamic, modern, and powerful societies of Europe. Doing so, they hoped, would lead the Turkish people to effect the fundamental mental remapping and thoroughgoing change in identity that the Kemalists believed had to occur if Turkey were to assume its rightful place among the forward-looking, modern states of the world.¹¹

⁹ M. Şükrü Hanioglu, *Atatürk: An Intellectual Biography* (Princeton: Princeton University Press, 2011), 155–59, 209, 214–18.

¹⁰ Feroz, *The Making of Modern Turkey*, 80–81.

¹¹ Howard, *The History of Turkey*, 98–99.

Unsurprisingly, given these aims, support for the secularization program was far from universal. While the reforms were popular with town and city residents—especially those who belonged to the professional, business, and military classes—they generated substantial resistance among the more conservative urban craftspeople and, especially, among the peasantry of rural Anatolia. Perhaps surprisingly, those groups raised few objections to the abolition of formal religious institutions such as the caliphate, the sheikh ül-Islam, or the madrasas. Instead, it was the extension of the secularization program to popular religious and cultural practices such as wearing the fez or undertaking pilgrimages to saints' tombs that fueled their resentment. As we shall see in chapter sixteen, their discontent would gradually deepen until, to the shock of the Kemalists, it spurred the emergence of a populist Islamic opposition movement after World War II.¹²

Kemalism

Having solidified their control of the state and implemented the first phase of their reform program, Kemal and his supporters next moved to formally articulate their governing philosophy. They did so in the RPP party platform of 1931, which laid out the so-called Six Arrows—republicanism, secularism, nationalism, populism, revolutionism, and etatism—that were the guiding principles of Kemalism. The first, republicanism, called for sovereignty to emanate from the people rather than from a monarch—though the fact that the Kemalist state was, at root, a fundamentally undemocratic one suggests that the RPP's embrace of representative government was more theoretical than real. The second arrow, secularism, was, as we have seen, the guiding principle of the reforms of the 1920s. The third, nationalism and the effort to construct a new Turkish identity would in turn dominate the reforms of the 1930s. The fourth, populism, went hand in hand with nationalism. It stressed collective solidarity, called for putting the nation ahead of group or individual interests, and denied both the existence of class conflict in Turkey and, by extension, the need for class-based political parties.¹³

While the first four arrows enjoyed overwhelming support within the RPP, the remaining two, Etatism and *Inkılâpçılık*, or revolutionism, were more controversial. The principle of revolutionism, which called for the embrace of ongoing change in support of Kemalism, troubled the capitalists and landlords who dominated the party's conservative wing. Worried that it might result in Turkey entering a disruptive state of permanent revolution, they responded by choosing to interpret *Inkılâpçılık* as a call not for revolution but instead for reform. Etatism, the statist economic doctrine that the party adopted at its Third Conference, proved similarly unpopular with the RPP's conservative factions. Their opposition to it was grounded in three fears: that large-scale state intervention in the economy would be inefficient, that it would almost inevitably depress economic growth, and, predictably, that it would threaten their financial interests.¹⁴

¹² Zürcher, *Turkey*, 191–92.

¹³ Zürcher, 181–82.

¹⁴ Feroz, *The Making of Modern Turkey*, 63–64.

Kemalist Economic Policies

Advocates of etatism were able to overcome these objections thanks to the failings of the more conventional economic policies that Turkey had pursued during the 1920s. With the nation's economy having fallen into ruin in the wake of the war for independence and with the Treaty of Lausanne having barred the government from raising tariff rates until 1929, the Kemalists had been compelled in the early 1920s to adopt a fairly conventional economic-development policy that relied primarily on private actors and the market to drive growth.¹⁵

The stunning, Great Depression-induced decline in international trade that began in 1929 compelled Turkey to rethink its economic-development strategy. The crisis came on with great suddenness. Following the crash of the American stock market in October 1929, demand for primary products—particularly for wheat, the mainstay of the country's exports—evaporated. The result was something of a perfect storm for Turkey. Not only did the staggering, two-thirds drop in wheat prices immiserate the peasant class that was dependent on the export market, but it also left the country unable to earn the foreign currency that its industrialists and merchants needed to purchase imported capital equipment.¹⁶

The crisis produced a sharp debate within governing circles in 1931. Arguing that a free-market system would generate higher growth rates, some remained committed to a liberal-capitalist system such as existed in Britain or France. Others, including Prime minister and future president, İsmet İnönü (r. 1938-1950), countered that a market-based system could not resolve the economic crisis and maintained that Turkey could best deal with the Great Depression by adopting a centralized, state-directed, etatist economic system. The debate did not last long. The combination of the democratic-capitalist world's inability to achieve a speedy economic recovery and the Soviet Union's tremendous growth in heavy industries during the First Five Year Plan (1928-1932) tipped the balance in favor of those who championed a state-led, economic order. As a result, beginning in 1931, the Kemalists formally launched an etatist program centered on three components: the promotion of Import Substitution Industrialization (ISI), direct state involvement in the economy, and central planning.¹⁷

The adoption of Import Substitution Industrialization—a policy wherein a developing state seeks to foster domestic industries by insulating them from foreign competition—was to a substantial degree a case of making a virtue of necessity. With the prices of the commodities that Turkey produced having fallen far more than the cost of the finished goods that it imported, the government could only meet the demand for many heretofore imported products by ensuring that domestic enterprises manufactured them. It did so by raising tariffs precipitously and by subsidizing domestic producers—

¹⁵ Findley, *Turkey, Islam, Nationalism, and Modernity*, 271; Zürcher, *Turkey*, 195–97.

¹⁶ Zürcher, *Turkey*, 196–97.

¹⁷ Zürcher, 197.

the latter beginning with companies that manufactured the so-called “three whites,” clothing, sugar, and flour.¹⁸

Meanwhile, at the same time that it was nurturing domestic producers, the Turkish government was also shifting to a system of direct, state involvement in the economy and laying the groundwork for central planning. The expansion of state economic activity came first. Focusing on heavy industries, the government established two holding companies to oversee its investments: the Sumerbank, which managed developments in the manufacturing sector, and the Etibank, which promoted mining. Well capitalized, they produced a substantial increase in the state’s share of economic activity over the course of the 1930s. Central direction of the economy quickly followed. With Soviet technical assistance, the government launched Turkey’s first Five Year Plan in 1933 to coordinate state and private investment—a move that further strengthened the government’s control of economic activity.¹⁹

Despite these efforts, the Kemalist program produced, at best, mixed results. To be sure, Turkey’s statist approach to the economy did enjoy genuine success in the 1930s. Etatism resulted in Turkey’s GDP rising by an average of 7 percent in the late 1930s, and it certainly furthered the Kemalists’ efforts to cultivate a modest industrial sector. Its ability to ameliorate the worst aspects of the Great Depression also likely helped the country avoid the descent into regime change and outright dictatorship that the economic crisis had produced in much of Latin America and Europe.²⁰

At the same time, however, the RPP’s statist policies failed to live up to the favorable press they received. Turkey’s economic performance during the early years of etatism was strong, but, coming at a time of broad, global economic recovery, it was hardly unusual. Meanwhile, the ISI model of development would prove incapable of sustaining economic growth in the long term. The problem was one of incentives. By protecting domestic industries, the ISI development model created opportunities for patronage, corruption, and waste and discouraged efforts to control costs. As a result, after World War II, Turkey would find itself falling further-and-further behind states such as Japan that had instead opted to follow an Export Oriented Industrialization (EOI) strategy that focused on exploiting their comparative advantages to produce goods that were competitive on global markets.²¹

The Reforms of the 1930s

Concurrent with the shift to etatism in the early 1930s, the Kemalists launched the second phase of their cultural revolution. It began where the reform program of the 1920s had left off: with an effort to sever the link between Turkish and the other languages of the Middle East. Building on the adoption of the Latin alphabet in 1929,

¹⁸ Findley, *Turkey, Islam, Nationalism, and Modernity*, 273–74.

¹⁹ Zürcher, *Turkey*, 197–98.

²⁰ Findley, *Turkey, Islam, Nationalism, and Modernity*, 274–76.

²¹ Findley, 274–76.

Kemal and his supporters moved in the early 1930s to purge the language of its many Arab and Persian loanwords in order to emphasize the break with Turkey's Muslim and Ottoman past. To do so, they founded the Turkish Language Society in 1932 and tasked it with replacing foreign expressions with neologisms, archaic terms, or words derived from Turkish dialects. This campaign met with resistance initially but ultimately proved successful. Indeed, thanks to the efforts of the Turkish Language Society, modern Turkish is so different from that spoken during the late Ottoman Empire or even the early republic that the *Nutuk* that Kemal gave in 1927 is nearly as indecipherable to Turkish people today as Chaucer is to twenty-first century Americans.²²

Kemal Mustafa's interest in the Turkish language went much further than merely seeking to purge it of Arab and Persian words, however. He also sought to enlist modern linguistic studies in support of a larger effort to imbue the Turks with a triumphal narrative that both further distanced them from their Ottoman and Islamic past and rendered them the cultural equals of Western peoples. The Turkish president struck gold in this effort in 1935. That year, he became acquainted with the Sun Language Theory, a pseudoscientific linguistic hypothesis developed by a little-known Austrian academic named Hermann Kvergic. It held that the proto-Turkish of the early Turks of Central Asia was the first spoken language and that all subsequent languages evolved from it. In other words, Kvergic was arguing that Turkish was the mother tongue of all peoples and that it was the Turks who had invented spoken language.²³

Kemal Mustafa was quick to grasp the benefits that the Sun Language Theory offered in support of his effort to ideologically reconstruct Turkish society. Not only did it give the Turks a glorious past that was wholly separate from and far older than their time as Ottoman subjects or even their history as Muslims, but, in arguing that the Turkish people invented something as foundational as the spoken word, it also gave them a cultural and intellectual cachet that even the Europeans could not match. Unsurprisingly, the Kemalists enthusiastically threw their support behind the Sun Language Theory. They required university faculty to embed it in their teaching and research, made it a core component of the curriculum in primary education, and lavished funds on efforts by the Turkish Language Society to promote it.²⁴

Along with linguistics, the Kemalists also recruited the field of history in their effort to create a new Turkish national identity. Keenly grasping history's ability to shape contemporary perceptions of identity, Kemal Mustafa established the Turkish Historical Society in 1931 and tasked it with creating a suitably glorious nationalist past for the republic. This initiative began to bear fruit the following year with the development of the Turkish Historical Thesis. A work of nationalist pseudohistory, it asserted that the Turkish people were a branch of the Caucasian race and that their ancestors were the first to develop a sophisticated civilization. Facing periodic drought in their Central Asian homeland, they migrated outward in a series of waves that brought civilization to the

²² Howard, *The History of Turkey*, 190.

²³ Hanioglu, *Atatürk*, 173–80.

²⁴ Hanioglu, 173–80.

peoples of China, Greece, India, and Egypt. In other words, the Turkish Historical Thesis asserted that that the Turks were not just the originators of civilization, but also the founders of what were regarded as the world's greatest and most influential cultures.²⁵

Like the Sun Language Theory, the Turkish Historical Thesis proved central to the Kemalists' effort to construct a new, modern Turkish national identity. It did so in two ways. First, by establishing that the Turkish people were Caucasian at a time when Eugenics, Social Darwinism, and racism enjoyed broad acceptance, it linked Turkey to the modern West and accorded it a high place in both formal and informal racial hierarchies. Second, by giving the Turks an heroic past that stretched far back in time, it furthered the goal of creating an identity that was distinct from the older Ottoman and Muslim ones to which many still clung. Placing great stock in these ideological benefits, the state aggressively promoted the Turkish Historical Thesis and made it the backbone of the history curriculum in both universities and primary schools. They also celebrated it through official propaganda. For example, the development banks that the Kemalists had set up—the Etibank and Sumerbank—were named for two of the groups that the Turkish Historical Thesis claimed were among the first to leave Central Asia: the Hittites and Sumerians.²⁶

Meanwhile, the Kemalists also moved in the 1930s to bring Turkey in alignment with the Western European states by legally and socially emancipating women. Scholars refer to this top-down program of reform as “state feminism.” It involved the promotion of female education, the revision of laws that had restricted women's employment, and the encouragement of activities such as ballroom dancing and beauty contests that broke with traditional Turkish gender norms. It also granted women the right to vote in municipal elections in 1930 and in national ones in 1934, meaning that Turkish women had the franchise before their counterparts in purportedly more-modern Western nations such as France and Switzerland. The Kemalist program of state feminism even abolished laws that had banned women from running for parliament—as seventeen women successfully did in 1935—and opened the door for many urban middle-class women to go into professions such as medicine, teaching, architecture, and the civil service. These were important achievements, and they initiated a genuine process of change in Turkey.²⁷

Unsurprisingly, however, the gains that Turkish women secured in the 1930s were largely cosmetic. Like many of his supporters, Kemal Mustafa was personally quite conservative regarding gender issues and promoted women's empowerment less to improve their position in Turkish society than to support propaganda efforts designed to reinforce Turkey's new image as a modern, Western state. Meanwhile, in rural areas, men and women alike clung tenaciously to traditional gender norms and resisted the

²⁵ Zürcher, *Turkey*, 190–91.

²⁶ Hanioglu, *Atatürk*, 163–68; Zürcher, *Turkey*, 191.

²⁷ Nikki R. Keddie, *Women in the Middle East: Past and Present* (Princeton: Princeton University Press, 2012), 77–82.

new laws and social mores. As a result, while the position of women in Turkish society changed, it did so very slowly.²⁸

Finally, the reform program of the 1930s changed how people referred to one another through the Surname Law of 1934. It had two provisions. First, it banned the use of venerable, but-still-popular Turkish and Ottoman noble titles such as *effendi* and *bey* that the Kemalists viewed as old fashioned and outmoded. Second, and more far reaching, it mandated that all Turks had to adopt a Western-style family name and thus broke with the traditional Muslim custom of having only a first name. The law was a key part of the Kemalist modernization effort. It not only further updated the nation's image and strengthened its symbolic links with the West but, by requiring people to take a Turkish rather than an Arabic or Persian name, further solidified the emerging Turkish identity that the Kemalists had championed at the expense of older Ottoman and Muslim ones. Kemal Mustafa set the tone by choosing the name Ataturk, or "father of the Turks."²⁹

The Surname Law capped what had been a sweeping, decade-long top-down social transformation of Turkey. The changes had been enormous. The adoption of family names and the embrace of state feminism had eroded traditional, Ottoman era social practices and had given Turkey the more-modern image that the Kemalists sought. Meanwhile, the Sun Language Theory and the Turkish Historical Thesis had helped both to institutionalize the new, Turkish identity that they continued to champion and to undermine the older, Ottoman and Muslim identities that they wished to erase. Finally, the secularization reforms of the 1920s had broken the power of the only group with the credibility and influence to oppose the ruling party: the religious scholars. Thus, at the end of the 1930s, the Kemalists could look back with satisfaction on their efforts to remake Turkey and to secure broad acceptance of the new modern, Turkish identity that they viewed as essential to the modernization of their country.

The driving force behind these changes was no longer there to celebrate those achievements, however. Instead, afflicted with cirrhosis of the liver, Ataturk passed away in 1938 not long after the reform program of the 1930s came to an end. He may have died, but he was most certainly not forgotten. All-but deified as the RPP's "eternal party chairman," he became in death the embodiment of the Turkish republic and a powerful and lasting symbol of secularism and Turkish nationalism. Indeed, Ataturk is so revered in Turkey that, to the present day, it remains a crime to defame him.³⁰

²⁸ Keddie, 77, 81–82; Howard, *The History of Turkey*, 105; Findley, *Turkey, Islam, Nationalism, and Modernity*, 278–79.

²⁹ Howard, *The History of Turkey*, 105–6; Findley, *Turkey, Islam, Nationalism, and Modernity*, 253.

³⁰ Feroz, *The Making of Modern Turkey*, 69–70.

Ismet Inonu, the Montreux Convention, and World War II

He was succeeded by his former prime minister, Ismet Inonu, who would serve as president until 1950. That Inonu rose to the presidency was surprising in light of the fact that he had become something of a rival to Ataturk during the last few years of the latter's life. Indeed, he was a threat of sufficient scale that Kemal Mustafa had forced him out of his post as prime minister in 1937. Despite his fall from the inner circle, however, Inonu's continued commitment to one-party rule and his determination to see the domestic reforms of the Six Arrows carried through to their conclusion favorably impressed the RPP. Accordingly, the party selected him to succeed Ataturk shortly after the latter's death. Ironically, in light of their importance in vaulting him into the presidency, domestic affairs took a back seat during the early part of his tenure. Instead, faced with the existential threat that the rise of fascism and the outbreak of World War II posed to the republic, Inonu was compelled to focus nearly all of his attention on foreign affairs during the early part of his tenure.³¹

Even before Inonu became president, Turkey had taken steps to secure the country's territorial integrity and its recently won independence. Concerns about possible aggression from Benito Mussolini's (r. 1922-1945) fascist Italy had precipitated that effort. Worried that the Italian dictator was casting a covetous eye at Turkish territory, Ataturk had requested in the mid 1930s a revision of the regulations governing the straits that would give Turkey more control over the vital waterway connecting the Black and the Mediterranean Seas and thus a greater ability to defend itself. The resulting agreement, the Montreux Convention of 1936, was a significant diplomatic victory for Ataturk. It restored to Ankara the right both to militarize the straits and to close them to foreign warships when it believed that Turkey confronted the "imminent danger of war."³²

Inonu followed Ataturk's success in regaining control of the straits by pursuing an adroit diplomatic line aimed at keeping Turkey from becoming drawn into the Second World War. Fearing the effects of German and Italian aggression if Turkey were isolated, his government initially leaned towards Paris and London. Shocked by Adolph Hitler's (r. 1933-1945) defeat of France in 1940 and his conquest of the Balkans in 1941, however, he subsequently abandoned his close ties with the Allies in favor of a tacitly pro-German neutrality. Once the military balance had shifted irrevocably back towards the Allies in 1943, Inonu once again reversed course. Even then, he moved carefully and shrewdly. Most notably, he waited until February 1945 to declare war on Germany—early enough to ensure that Turkey would become a charter member of the United Nations but late enough to guarantee that its participation in the war was *pro forma* and that no Turkish troops would take part in the fighting.³³

³¹ Feroz, 69–70.

³² Howard, *The History of Turkey*, 110–13.

³³ Howard, 110–13.

Inonu ultimately earned mixed marks from the Turkish people regarding his handling of the Second World War. On the one hand, with memories of the disastrous consequences of the Ottoman Empire's participation in World War I still fresh in their minds, many applauded his ability to keep Turkey from becoming embroiled in the conflict. On the other, war-induced inflation, severe material shortages, and the brutally confiscatory Capital Tax Law that the government levied disproportionately on non-Muslims caused many to lose faith in the RPP. As the war neared its conclusion, as a result, support for the Kemalist program slumped badly amid stepped-up demands for political liberalization.³⁴

The Turkish Straits Crisis

Inonu's government faced another serious foreign-policy challenge just after the conflict ended—one that bore more than a superficial resemblance to some of the crises the Ottoman Empire had experienced in the nineteenth century. Angry that Turkey had permitted German warships to traverse the straits during World War II, Joseph Stalin (r. 1924-1953) sought to strongarm Turkey into revising the Montreux Convention to give the USSR joint control of the straits and the right to maintain military bases along them. He also demanded territory along the Black Sea in northeast Anatolia. Given the Soviet Union's enormous military capabilities, Stalin's threat was a serious one for Inonu's government. Turkey did not have to face this renewal of aggression from the north alone, however. Fearing that Stalin sought control of the straits as a springboard for further aggrandizement, US President Harry Truman (r. 1945-1953) deployed the battleship USS *Missouri* and the aircraft carrier USS *Franklin D. Roosevelt* to the region as a show of force in 1946. Faced with the superior might of the United States, Stalin had little choice but to back down; as a result, the situation quickly deescalated. Still, despite its quick ending, the straits crisis had a lasting effect on Turkey's diplomatic alignment. Indeed, it marked not merely the start of close ties with the US but also the key event that drove Ankara to join the North Atlantic Treaty Organization (NATO) in 1952.³⁵

The Kemalist Legacy

As the republic neared its twenty-fifth anniversary in the late 1940s, the Kemalists could take pride in the changes they had made to Turkey. In less than two decades, they had implemented an innovative and apparently successful economic response to the Great Depression and had launched a successful cultural revolution that had effected a clean break with the immediate Ottoman past. They had also erected a cult of personality around the head of state that turned him in both life and death into a powerful symbol of national unity and had skillfully avoided becoming embroiled in World War II until the fighting was all-but over. Given the extent of these achievements, Turkey and the Kemalist program unsurprisingly came to serve as a

³⁴ Feroz, *The Making of Modern Turkey*, 69–70.

³⁵ George C. Herring, *From Colony to Superpower: U.S. Foreign Relations since 1776* (New York: Oxford University Press, 2008), 609.

highly influential model of development that many states in Latin America, Asia, Africa, and, notably, the Middle East would adopt in the decades immediately after the Second World War.

At the same time, however, keen-eyed observers could detect more than a few cracks in the Kemalist façade. By the end of World War II, the ruling party faced growing discontent among both urban liberals, who chafed at limits to political expression, and rural conservatives, who opposed the RPP's radical secularization program. Ironically, as a result, at the very moment that its international reputation and influence were peaking, Kemalism was beginning to confront mounting protests to its domestic monopoly on political power—challenges that would, as we shall see, succeed in compelling the RPP to relinquish the iron grip it had held on political power since 1924.

The Historical Debate: Modernization Theory

From the founding of the Turkish Republic through the 1960s, a framework called modernization theory dominated interpretations of Turkey's birth and early development. Articulated most clearly by the historian Bernard Lewis in his influential 1961 book, *The Emergence of Modern Turkey*, it held that modernization in the country was an inevitable process that began after World War I when the Turkish people rejected their preindustrial and undemocratic past in the Islamic and Ottoman worlds in favor of modern, Western-style industrialism and republicanism. Lewis acknowledged that the process was not entirely smooth. He noted, for example, that a minority of reactionaries such as rural peasants and, especially, religious leaders resisted these changes. However, he concluded, these traditionalists constituted only a shrinking minority of dead enders who gradually melted away as the benefits of modernity spread.³⁶

Lewis's study proved enormously influential, and it dominated understandings of Turkey's emergence as a modern nation-state for decades. Beginning in the 1970s, however, a growing number of scholars began to challenge both his interpretation and the broader modernization paradigm. They focus their critiques on two core aspects of his argument. First, they reject his contention that the creation of Turkey was the inevitable product of a newly awakened people seeking to free themselves of their decadent Ottoman and Islamic past in favor of modernity; instead, they argue that its birth was a function of historical contingency. That is, they point out that nearly everyone who took part in the national resistance did so to perpetuate the Ottoman Empire and the caliphate rather than to create a republican nation-state. The move away from the sultanate only became inevitable, they argue, when Atatürk and a small band of co-conspirators used their dominant position in the RPP to force through a law in 1923 making Turkey a republic. Second, they dispute modernization theory's binary categorization of Turkish society as either traditional or modern and instead maintain that people then and now embrace elements of both. As the historian Erik Zürcher notes, for example, the Islamist Turkish Welfare Party of the 1980s and 1990s was not

³⁶ Bernard Lewis, *The Emergence of Modern Turkey* (Oxford: Oxford University Press, 1961), Passim.

simply a reactionary, fundamentalist movement, but instead a political organization trying to address—in an Islamic context—the problems typical of a modern industrial society.³⁷

Saudi Arabia and the Gulf: The Emergence of the Oil Economy

Though it emerged at the same time as the Turkish republic, Saudi Arabia followed a decidedly different path in its development as a state. Like the original Saudi emirate that Muhammad ibn Saud (r. 1726-1765) had founded in the eighteenth century, the modern kingdom developed out of an alliance between the Saudi family, which hoped to secure political control of Arabia, and the Wahhabi 'ulama', who sought to impose their Salafist interpretation of Islam on the people of the region. Given its modest population and forbidding climate, the Saudi state did not appear likely to be anything more than an impoverished and isolated theocracy at the time of its revival in the early twentieth century. Its prospects changed dramatically in 1938, however, thanks to the discovery of seemingly limitless oil beneath its heretofore valueless desert sands. The result was a paradoxical state—one that would assume an increasingly central role in the modern global economy even as it enforced adherence to an austere religious movement that focused narrowly on the past.

Ibn Saud

That the Saudis emerged as economic kingpins is surprising in light of the family's weak position at the turn of the twentieth century. In 1891, the rival Rashidi family had defeated the Saudi emir, Abdul Rahman bin Faisal (r. 1889-1891), and had seized control of Riyadh, the capital of the Emirate of Najd that the Saudis had established in 1824. The setback was a dispiriting one for Abdul Rahman, and it compelled him and his family to flee to Kuwait where the exiled Saudi leader plotted the recovery of his state. Those plans came to naught. Defeated once again by the Rashidis in 1900, the dispirited Saudi leader abandoned all hope of retaking Riyadh—leaving him and his family fated to disappear into political irrelevancy.³⁸

Instead, it achieved a stunning recovery under Abdul Rahman's son, Abdulaziz Ibn Saud (r. 1902-1953). Broad chested and standing six feet three inches in height, the charismatic young Saudi leader was as determined to recover his patrimony as he was physically intimidating. A war between the Kuwaitis and the Rashidis that began in 1901 gave him his chance. Seeking to distract the Rashidis, the Emir of Kuwait sent ibn Saud and a small group of his most dedicated followers to attack Riyadh in early 1902. Their first assault on the walled city failed badly; however, their second effort, a bold night assault that concluded when ibn Saud killed the Rashidi governor, succeeded. It was a

³⁷ Zürcher, *The Young Turk Legacy and Nation Building*, 41–52.

³⁸ James Wynbrandt, *A Brief History of Saudi Arabia*, Second Edition (New York: Checkmark Books, 2004), 162–66.

great victory—one that both restored the Saudis to power and cemented Ibn Saud's reputation as a daring military commander.³⁹

The Conquest of the Najd, 1902-1918

Ibn Saud was not content merely to recover his inheritance or to serve as the emir of a small state, however. A man of great ambition, he instead wanted to dominate the Arabian peninsula. Accordingly, he followed the recapture of Riyadh by waging a campaign of conquest against the nearby bedouin tribes that ended with him adding much of the Najd to his expanding state. He followed with another campaign that secured for him control of the large, heavily Shi'i Eastern Province that lay along the Arabian coast of the Persian Gulf. Thus, by the eve of the First World War, the Saudi state had taken control of a substantial amount of territory and had assumed a commanding position in central Arabia.⁴⁰

How did Ibn Saud achieve such success so quickly? How could the scion of an exiled family lacking in resources and wealth secure legitimacy and build an army strong enough to assume control of such a large territory in only a little more than a decade? Ibn Saud's leadership skills and charisma certainly played a significant part in the emirate's expansion; at root, however, his success then and later stemmed from his unique ability to satisfy the aspirations of two very different groups: the Wahhabi 'ulama' within Arabia and the British who increasingly dominated the peninsula's periphery.⁴¹

A shrewd man with a keen grasp of politics, he understood that the Saudi family's traditional relationship with the Wahhabis was central to its power. Accordingly, from the start of his reign he renewed the family's close alliance with the *mutawwaa*, a class of Wahhabi religious scholars unique at that time to the Najd who used violence—they typically caned people—to enforce observance of Islamic rituals like communal prayer and to prohibit banned activities such as drinking. To secure their support, Ibn Saud made the *mutawwaa* salaried employees of the state and sanctioned their authority to enforce behavior; in particular, he permitted them to impose Wahhabi doctrine on the bedouin tribes. In exchange, they used the newly created Committee for the Propagation of Virtue and Prohibition of Vice to compel submission both to shari'a law and to Ibn Saud's secular authority. The emir followed by having the *mutawwaa* recruit the most zealous bedouin warriors into a new military organization, the al-Ikhwān, or brotherhood, that supplemented his traditional forces. These moves paid multiple dividends. They ensured the continued support of the influential *mutawwaa* clerics, ended the bedouins' disruptive tribal feuds, and rendered rebellion against him all-but

³⁹ Daniel Yergin, *The Prize: The Epic Quest for Oil, Money & Power* (New York: Touchstone Books, 1991), 284–85.

⁴⁰ Rachel Bronson, *Thicker Than Oil: America's Uneasy Partnership with Saudi Arabia* (Oxford: Oxford University Press, 2006), 28–30.

⁴¹ Madawi al-Rasheed, *A History of Saudi Arabia*, Second Edition (Cambridge: Cambridge University Press, 2010), 40–41.

impossible. Above all, they gave him a powerful cadre of disciplined and religiously motivated troops that he could use to further expand his territory.⁴²

Having deftly used the Saudi family's traditional connection with the Wahhabis to establish dominance in the Najd, he followed by pursuing a relationship with the British. London was receptive to his entreaties. Seeking to ensure its control of the Persian Gulf during World War I, it opened talks with him in 1915 aimed at enlisting the Saudi emirate in the fight against the Ottoman Empire and its local proxies. The result was the Anglo-Saudi Treaty. By its terms, Ibn Saud pledged to refrain from raiding Britain's Persian Gulf protectorates and promised to attack their common enemy in the Najd, the Ottomans' Rashidi allies. In exchange, London agreed to provide him with arms, £20,000 in gold up front, and a further subsidy of £5,000 in gold per month. The deal was a very good one for Ibn Saud. At the modest cost of going to war against his family's main rival in the Najd and promising to respect the territory of Britain's protectorates—something he was wise enough to have done even if he had not signed a formal treaty with London—he was able to equip his army with modern weaponry and to secure sufficient gold to ensure the absolute loyalty of key tribal leaders. Coupled with the alliance he had struck with the *mutawwaa* and his creation of the powerful al-Ikhwan, he was well positioned to continue his campaign of territorial expansion in the postwar era.⁴³

Consolidation, 1919-1932

Acting carefully so as to avoid alarming London, Ibn Saud did so with great skill. He started by finally finishing off the Saudi family's old enemies, the Rashidis, in 1921. That victory gave him control of the northern Nadj and a common border with Hashimite-controlled Transjordan and Iraq; more broadly, it established that his state was the dominant force in Arabia. Even then, however, Ibn Saud was careful to remain on good terms with London. For example, he agreed to ease British concerns about disruptive cross-border raids into Hashimite territory by signing the Uqayr Protocol in 1922, which demarcated the frontier between his state and the Mandates of Iraq and Transjordan. He also scrupulously refrained from seizing Mecca and Medina during the early 1920s out of deference to London's close ties with King Husayn (r. 1916-1924).⁴⁴

Thanks to the rapidly worsening state of relations between Britain and Husayn, however, he did not have to restrain himself for long. Ties between London and the Hashimite king deteriorated rapidly in the wake of World War I. Believing that the British had reneged on the terms of the Husayn-McMahon Correspondence, Husayn had steadfastly refused to sign a treaty with Britain recognizing the new postwar order in the region—a choice that British diplomats found unacceptable. Grasping that London had reached the end of its rope and intuiting that they were giving serious consideration to

⁴² al-Rasheed, 47–59.

⁴³ Wynbrandt, *A Brief History of Saudi Arabia*, 175.

⁴⁴ al-Rasheed, *A History of Saudi Arabia*, 41–42; Wynbrandt, *A Brief History of Saudi Arabia*, 180–82, 188–89.

abandoning Husayn, Ibn Saud realized that he could now invade the Hijaz without running afoul of London—he just needed a reasonable pretext to do so. Fortunately for him, Husayn soon provided the necessary rationale by declaring himself caliph shortly after Kemal Mustafa had abolished the office in early 1924. Announcing that the Hashimite king was unfit to hold such an exalted position, Ibn Saud promptly invaded the Hijaz. The ensuing war was entirely one sided. Without British backing, Husayn could do nothing to resist Ibn Saud's powerful al-Ikhwan warriors and soon fled into exile.⁴⁵

He may have defeated his last substantial external rival with his conquest of the Hijaz, but he soon faced a far-more dangerous internal challenge in the form of the Ikhwan and their *mutawwaa* supporters. By the late 1920s, the brotherhood and the *mutawwaa* had become thoroughly disillusioned with their relationship with the Saudi state. They resented Ibn Saud's attempts to curb their enforcement of Wahhabi beliefs in the Hijaz, bridled at his refusal to permit them to destroy tombs and shrines as their Wahhabi beliefs dictated, and fumed at his realpolitik willingness to extend limited religious toleration to the Shi'i of the Eastern Province. They also opposed his introduction into Arabia of modern innovations such as the telephone, telegraph, airplane, and automobile on grounds that those devices did not exist at the time of the early *umma* and thus had no place in a Muslim society. Above all, they were angry with Ibn Saud because his pragmatic decision to agree to respect the borders that divided his state from the British protectorates that surrounded it had put an end to their jihad to extend Wahhabism.⁴⁶

The ikhwan and *mutawwaa* were not alone in their dissatisfaction with the relationship between the brotherhood and the Saudi state, moreover. By the late 1920s, Ibn Saud, too, had grown disenchanted with the arrangement. Already weary of the discontent that their strict enforcement of the shari'a code had caused in the Hijaz, he was becoming more-and-more concerned that their loud, religiously based antagonism toward Britain might drag him into a war he could not win.⁴⁷

The conflict between Ibn Saud and the al-Ikhwan finally came to a climax in 1929. The proximate cause was a brotherhood attack on an Iraqi police outpost near the border that set off an escalating cycle of reprisals. Intent on sending a strong signal of resolve so as to prevent further attacks, Britain retaliated by bombing one of the brotherhood's encampments; upping the ante, the al-Ikhwan responded in turn by launching a series of raids into both Kuwait and Iraq.⁴⁸

The worsening situation deeply troubled Ibn Saud. Fearing that further reprisals could escalate into a full-scale war with Britain, he pressed the brotherhood to end the

⁴⁵ Eugene Rogan, *The Arabs: A History – Revised and Updated Edition* (New York: Basic Books, 2012), 182.

⁴⁶ al-Rasheed, *A History of Saudi Arabia*, 64–67.

⁴⁷ al-Rasheed, 64–67.

⁴⁸ Wynbrandt, *A Brief History of Saudi Arabia*, 185.

conflict. They refused. Instead, to assert their autonomy, they raised the banner of revolt in March 1929 and went to war with the emir. Ibn Saud responded to this challenge with characteristic speed and skill. Determined to ensure his monopoly over the levers of power in his state, he pressured the Wahhabi 'ulama' in Riyadh to issue a fatwa justifying his use of force against the al-Ikhwan. He followed by swiftly destroying the brotherhood in battle.⁴⁹

Substantially strengthening his domestic standing, the triumph over the al-Ikhwan paid Ibn Saud two immediate political dividends. First, in 1932, it permitted him to announce the unification of the Najd and the Hijaz into the new Kingdom of Saudi Arabia. Second, the brotherhood's defeat meant that he could now deal with his partners-cum-rivals, the Wahhabi religious scholars, from a position of strength. He did so by compelling them to assume a subordinate position in which their authority would not extend to political issues—as many of the Wahhabi religious scholars wished—but would instead be confined to education, religion, and social law. Thus, his defeat of the brotherhood marked a pivotal turning point in the emergence of Saudi Arabia. It not only permitted him to establish the kingdom itself, but, by transforming his family's traditional Wahhabi allies into junior partners, it seemed to have ensured that he and his successors would rule without serious internal challenge or constraint.⁵⁰

Appearances in this case were deceiving. Ibn Saud may have won in the short term, but over the long haul, the 'ulama' were able to use their monopoly over education and religious law to transform themselves into a major domestic force in Saudi Arabia—one to which even the ruling dynasty had to defer. Later, as we shall see, they would leverage their influence and the tremendous wealth the kingdom derived from the sale of petroleum to export their fundamentalist views widely. Thus, as the historian Rashid Khalidi argues, the arrangement Ibn Saud worked out with the 'ulama' after the defeat of the al-Ikhwan turned out to be a “devil's bargain”—one that would bring extremist religious ideas and vast oil money together in a way that permitted the Wahhabi to spread their austere teachings throughout the Sunni world.⁵¹

Fiscal Issues

In the meantime, Ibn Saud had a new crisis to deal with. While his domestic political position had improved as a result of his victory over the al-Ikhwan, the fiscal health of his state was deteriorating rapidly. Part of the issue was a steady increase in state spending. The war against the brotherhood, the need to increase subsidies to tribal chiefs to ensure their continued loyalty, the start of a substantial development program, and the cost of Ibn Saud's lavish personal lifestyle had combined to dramatically increase government spending in the late 1920s and early 1930s. Ibn Saud's government would likely have been able to deal with these rising expenditures

⁴⁹ Bronson, *Thicker Than Oil*, 32–33.

⁵⁰ al-Rasheed, *A History of Saudi Arabia*, 66–71.

⁵¹ Quote from Rashid Khalidi, *Resurrecting Empire: Western Footprints and America's Perilous Path in the Middle East*, Reprint edition (Boston: Beacon Press, 2005), 108.

had the Saudi economy been growing, but the added expenses coincided with the onset of the Great Depression. The global economic contraction bit hard into the kingdom's primary source of revenue, the *hajj*. Thanks to the worldwide collapse in economic activity, the number of pilgrims who visited Mecca fell from 129,000 in 1926 to the *hajj*'s modern, pre-Covid-19 low of just 20,000 in 1933. The consequences for the Saudi treasury were dire. Whipsawed between rising expenditures and declining revenues, the country quickly amassed a £300,000 deficit—a figure far beyond its capacity to service.⁵²

An astute man, ibn Saud grasped that the fiscal crisis made for a combustible situation. Without revenue, he could neither pay the army that maintained internal order nor continue to offer the subsidies to prominent tribal leaders that had ensured the allegiance of the bedouin. Absent the loyalty of the military and the tribes, in turn, dynastic authority would weaken and the state that he had so painstakingly constructed would collapse. In other words, ibn Saud had to find some way to get his fiscal house in order and he had to do so fast if he hoped to prevent his budgetary problems from evolving into a dangerous political crisis.⁵³

The Oil Boom Begins, 1933-1948

One possible solution was to sell mineral rights to Western firms—particularly the rights to any oil located beneath the kingdom's seemingly limitless expanse of sand. There was an established way to do so. Before World War II, states such as Saudi Arabia typically auctioned rights to mineral deposits such as petroleum to foreign companies through a mechanism called a concession. An oil concession gave a foreign firm the exclusive right to explore for petroleum deposits and to exploit any commercially viable fields that it discovered in an agreed upon territory for a limited—but usually lengthy—period of time. In exchange, the state granting the concession received a lump sum of money up front as well as a comparatively modest annual rent payment on the land on which the company was drilling and—most consequentially—royalties on the sale of any oil.⁵⁴

Concerned about the impact of inviting non-Muslims into his kingdom, Ibn Saud had previously expressed doubts about the wisdom of entering into an oil concession. His dire fiscal circumstances produced a quick change of heart, however. Now desperate for funds, he set aside his misgivings and made clear that he would enthusiastically welcome the opportunity to grant a concession to one of the Western oil companies. As he told one of his advisors, “if anyone would offer me a million pounds, I would give him all the concessions he wanted.”⁵⁵

⁵² Wynbrandt, *A Brief History of Saudi Arabia*, 187.

⁵³ Bronson, *Thicker Than Oil*, 14; Wynbrandt, *A Brief History of Saudi Arabia*, 187.

⁵⁴ Rogan, *The Arabs*, 356.

⁵⁵ Yergin, *The Prize*, 281–88.

In any event, an oil concession seemed to Ibn Saud like a low-risk opportunity to earn some revenue. Doubting that his kingdom possessed any petroleum whatsoever, he saw a concession as little more than a way to secure an infusion of quick and desperately needed cash that came without any long-term costs or consequences. He was not alone in his skepticism about Saudi Arabia's prospects as an oil producer. John Cadman, the chair of the Anglo-Persian Oil Company (APOC) maintained that geological surveys yielded "little room for optimism" that the kingdom might contain oil. One of his subordinates was even blunter, declaring that Saudi Arabia was "devoid of all prospects" of possessing petroleum.⁵⁶

The discovery of petroleum on the island of Bahrain in 1932 altered perceptions of Saudi Arabia's potential as an oil producer, however. Given Bahrain's similar geology and its proximity—it lies just thirty-two kilometers off the kingdom's Eastern Province in the Persian Gulf—the presence of oil there suggested strongly that Saudi Arabia might contain substantial deposits. Even then, however, few firms sought the concession. Recent discoveries and a depression-induced collapse in demand for petroleum in the industrialized states had resulted in a persistent, global oil glut and a broad retrenchment in the industry; coming in such circumstances, any new production from Arabia would merely serve to further depress the already sagging price of oil. As a result, only one comparatively small company, Standard Oil of California (Socal), pursued a concession in Saudi Arabia.⁵⁷

Socal's bid more than satisfied Ibn Saud. By its terms, the company agreed to provide the king with two interest free loans in gold—a £30,000 one distributed immediately and a second £20,000 one issued eighteen months later—that Saudi Arabia would repay only out of any future royalties. If Socal failed to find any oil, in other words, the loans would amount to an outright gift to the kingdom. Socal also promised to pay £5,000 annually in rent as well as royalties on the sale of any oil that it extracted. In exchange, it won the exclusive right to explore for petroleum in a 930,000 square-kilometer swath of the kingdom and to exploit any oil that it found there. The concession's terms were to last for sixty years.⁵⁸

Socal's efforts to turn the concession into a profitable venture got off to a rocky start thanks to two significant problems. First, as a relatively small company, it lacked the downstream markets in which to sell any petroleum that its newly created subsidiary, the California-Arabian Oil Company (Casoc) might discover in the kingdom. It solved this problem in 1936 by signing a deal with the Texas Oil Company—later known as Texaco—that brought it and its valuable distribution network into Casoc as an equal partner. That move may have secured the needed markets, but it did not solve the second, more fundamental problem: despite multiple efforts, the company had repeatedly come up empty in its effort to discover petroleum in Saudi Arabia. Indeed, its exploratory drilling in the kingdom's most geologically promising location near the town

⁵⁶ Yergin, 281–88.

⁵⁷ Yergin, 289–91.

⁵⁸ Yergin, 291.

of Dammam in the Eastern Province had repeatedly failed to produce commercially viable wells, though it did yield sufficient trace amounts of oil to justify continued exploration. Only in March 1938, a full five years into the concession, did one of Casoc's wells, Dammam-7, finally strike oil. Producing 1,500 barrels per day at a time when the daily output of the typical US well was a paltry 100 barrels, it alone made the concession a success. Further wells in the area proved equally productive and, to the delight of its shareholders, confirmed that Casoc had discovered a major field.⁵⁹

Still, it took some time for the oil industry in Saudi Arabia to really take off. Thanks to wartime disruptions and a lack of equipment, Casoc's daily output only rose from 11,000 barrels per day in 1939 to 21,000 barrels per day in 1944. Thus, at the end of World War II, the kingdom remained a bit player in the global oil market.⁶⁰

Saudi production crossed an important watershed immediately after the war, however. Between 1944, when Casoc renamed itself the Arabian-American Oil Company (Aramco), and 1950, the kingdom's output surged by more than 2500 percent and reached the then-staggering figure of 548,000 barrels per day. What accounted for this change? How did Aramco achieve such a startling increase in oil production following the war? Two events explain the massive surge. First, in 1948, Aramco geologists discovered the Ghawar Field southwest of Dhahran. An elephant field of almost unimaginable dimensions, it remains the largest oilfield in the world and alone accounts for roughly half of Saudi Arabia's current daily production. Second, Aramco brought on Socony-Vacuum and Standard Oil of New Jersey as partners. Large companies that controlled substantial downstream markets, they gave Aramco the capital and distribution networks the consortium needed to fully exploit its massive new find and to make Saudi Arabia a major player in global energy production.⁶¹

Birth of the US-Saudi Alliance

Relations between the US and Saudi governments warmed in parallel with the growing involvement of American petroleum companies in the kingdom. For Washington, the desire for closer ties stemmed from concerns about declining domestic supply. America had long been the world's leading producer of petroleum; however, policymakers in the administration of Franklin Roosevelt (r. 1932-1945) had grown increasingly worried during the war that a combination of declining domestic production and rising demand could hamstring the American economy following the conflict and might even induce a second depression. Secretary of the Interior Harold Ickes (1873-1952) provided perhaps the clearest exposition of this view in an article published in *American Magazine* in December 1943 entitled, appropriately, "We're Running Out of Oil!" With such concerns in mind, the US government took a series of steps during the latter stages of the war to establish closer links with the kingdom. Most notably, it opened a diplomatic legation in Jeddah in 1942 and began to provide Lend Lease aid to

⁵⁹ Wynbrandt, *A Brief History of Saudi Arabia*, 191–92.

⁶⁰ Bronson, *Thicker Than Oil*, 19–20.

⁶¹ Bronson, 19–20.

Saudi Arabia in 1943. Ibn Saud was receptive to Washington's entreaties. Seeking both a patron to protect his kingdom and a counterweight to Britain's dominant position in the Middle East, he, too, hoped to strengthen bilateral ties.⁶²

In 1945, accordingly, the two nations established close links that would, with occasional hiccups, endure to the present day. The crucial event that solidified the budding relationship was the meeting between Ibn Saud and Roosevelt that occurred aboard the cruiser USS *Quincy* on the Suez Canal's Great Bitter Lake on February 14, 1945. It was a momentous occasion for both countries. While Roosevelt and Ibn Saud disagreed over the question of Palestine's future and while the meeting produced no formal agreement, their brief encounter nonetheless marked the moment at which Saudi Arabia entered the US sphere of interest. Indeed, it was not long after that meeting that the two states entered into an agreement that gave the US military access to bases in the kingdom including one at Dhahran from which B-29 strategic bombers could strike targets in the USSR. Later, with Europe's postwar recovery increasingly dependent on Arabian oil, President Harry S. Truman (r. 1945-1953) established an implicit security guarantee of the kingdom by declaring that Washington was "interested in the preservation of the independence and territorial integrity of Saudi Arabia." Thus, by the end of the decade, the American special relationship with the kingdom—one in which Saudi Arabia provided religious sanction of Washington's conduct of the Cold War and a stable supply of oil in exchange for military protection—had been firmly established.⁶³

The Gulf States

While Saudi Arabia came to lie within the US sphere immediately following World War II, the principalities of the Persian Gulf instead remained firmly in Britain's orbit. London's domination of the region stretched back all the way to 1820 when it compelled the Gulf emirates to adhere to an agreement banning piracy called the General Treaty. It had followed by entering into a series of protectorate agreements with the Trucial States—today's United Arab Emirates (UAE)—in 1835 and by signing similar agreements with Bahrain in 1861, Kuwait in 1899, and Qatar in 1916. Designed to deny hostile European states a foothold in the Persian Gulf from which they might menace Britain's vital commerce with India, those treaties granted London control of the emirates' foreign relations in exchange for a promise to provide them with military protection. London supplemented those bilateral agreements with a general statement shortly after the turn of the century. Announced by Secretary of State for Foreign Affairs Lord Lansdowne (1845-1927) in 1903, it amounted to a British Monroe Doctrine for the Persian Gulf.⁶⁴

⁶² Wynbrandt, *A Brief History of Saudi Arabia*, 196; Bronson, *Thicker Than Oil*, 40; "U.S. Consulate General Jeddah | U.S. Embassy & Consulates in Saudi Arabia," accessed August 26, 2020, <https://sa.usembassy.gov/embassy-consulates/jeddah/>.

⁶³ Bronson, *Thicker Than Oil*, 36–40, 46–49.

⁶⁴ Kristian Coates Ulrichsen, *Insecure Gulf: The End of Certainty and the Transition to the Post-Oil Era* (Oxford: Oxford University Press, 2015), 15–20.

Meanwhile, like Saudi Arabia, the Gulf States enjoyed a sudden windfall with the discovery of oil in the region beginning in the interwar period. Bahrain was the first of the emirates to become a producer state when Socal struck oil on the island in 1932. Kuwait was next. Here, the timing of the find was peculiarly serendipitous. Only a few years earlier, the pearl-diving industry that had long constituted the mainstay of its economy had collapsed after a Japanese noodle seller invented an inexpensive process for developing cultured pearls. The result was a period of severe economic contraction for the people and government of Kuwait. The bad times did not last, however. Instead, in 1938, the emirate's prospects took a dramatic turn for the better when the Kuwait Oil Company, a joint venture of Gulf Oil and APOC, discovered a major field in the country. Qatar followed quickly with the discovery of the Dukhan Field in 1939, while the Trucial States and Oman joined the petroleum producers' club in 1958 and 1964 respectively.⁶⁵

Rumblings of Discontent

Initially, the leaders of the oil-producing states of the Persian Gulf and Arabian peninsula were happy with the concession system under which the petroleum industry operated. From their perspective, the arrangement was a highly favorable one that had bequeathed upon them fantastical amounts of money at little cost. For example, Ibn Saud's first royalty check was for a staggering \$1.5 million. To be sure, the system gave companies like Aramco the lion's share of the profits; however, it also saddled them with all of the costs and uncertainty inherent in the extraction and sale of oil. Those expenses and risks were substantial. The petroleum industry was and is a highly capital-intensive one that required firms to spend huge amounts of money on tankers, refineries, pipelines, downstream marketing and distribution operations, and storage tanks—sums that were wildly beyond what countries like Saudi Arabia could raise at that time. It was also a capricious business in which many concessions failed to result in commercially viable operations. As the historian Eugene Rogan points out, "companies [often] drilled for years without so much as an oily rag to show for their efforts." During the 1930s and early 1940, as a result, Arab rulers such as ibn Saud were perfectly content with an arrangement that provided them with heretofore unimaginable streams of income while leaving all of the headaches, costs, and risks to the oil companies.⁶⁶

By the late 1940s, however, attitudes were beginning to change. Whatever risks the oil companies had once shouldered were long gone. Instead, having paid off the infrastructure needed to produce and distribute petroleum products, they were reaping what increasingly seemed like windfall profits. By 1949, for example, Aramco's earnings had grown so much that the company was earning three times the amount that Saudi Arabia was receiving for the sale of its own petroleum. Even more galling, the company paid \$43 million in taxes to the United States in 1949 but only \$39 million in royalties to

⁶⁵ Yergin, *The Prize*, 292–300.

⁶⁶ Rogan, *The Arabs*, 356; Bronson, *Thicker Than Oil*, 19.

Ibn Saud's government—meaning, in other words, that the US government took in \$4 million more on the sale of the kingdom's oil than did the Saudi state!⁶⁷

It was not just the distribution of profits that the petrostates were upset about, moreover. More fundamentally, their governments were increasingly concerned about the question of *control*. The concession system had rendered the producing states little more than passive participants in the petroleum industry. While they collected regular royalty checks, it was the oil companies that made all of the decisions pertaining to production—meaning how much petroleum is drilled—and marketing—meaning the price at which the firms sold the oil. Worse, the companies frequently made pricing and production decisions that were clearly contrary to the interests of the host countries. By 1950, as a result, pressure was building for a wholesale restructuring of the petroleum industry—one that would, at a minimum, dramatically increase the share of profits that went to the producing states.⁶⁸

Iran: Contested Independence

Surprisingly, in light of its apparent similarities with Saudi Arabia and Turkey, Iran had a far-more difficult experience during the interwar period than did those states. On the surface, its commonalities with them were strong. Like Kemalist Turkey, Iran undertook an ambitious modernization program centered on secularization and Import Substitution Industrialization aimed at transforming the country into a Western-style state; like Ibn Saud's kingdom, meanwhile, it was a major oil producer. These similarities ultimately proved superficial, however, and Iran was unable to match their success in navigating the difficult waters of the interwar period. Instead, thanks to a combination of its internal weakness, its strategic location, and the pro-German sentiments of its ruler, Reza Khan (r. 1925-1941), it came under Allied occupation during World War II and emerged from the conflict a shaky and divided country—one that was ill-prepared for the challenges of the mid-twentieth century.

Qajar Iran

To understand Iran's development during the interwar period, we must first briefly trace its history from the end of the Safavid Dynasty in 1722 through the early-twentieth century. For Iran, the seven-decade period that followed the demise of the Safavid Dynasty was a peculiarly fractious and challenging one. Lacking a powerful central government, it devolved into an anarchic place in which tribal chiefs, high-ranking Shi'i clerics, and regional warlords divided the country among themselves. Unsurprisingly, it was a period of destructive civil wars, costly invasions, and steep economic decline.⁶⁹

⁶⁷ Yergin, *The Prize*, 446–47.

⁶⁸ Rogan, *The Arabs*, 356–57.

⁶⁹ Ira M. Lapidus, *A History of Islamic Societies*, 2nd ed. (Cambridge: Cambridge University Press, 2002), 469.

Agha Mohammad Khan's (r. 1796-1797) establishment of the Qajar Dynasty in 1796 finally restored a degree of order to Iran. Of Turkic origin, the Qajar appeared on the surface to be powerful rulers cut from the same cloth as the great shahs of the earlier Safavid Dynasty. This image was certainly one that the dynasty sought to project. For instance, Fath Ali Shah (r. 1797-1834) professed to rule Iran as an unchallenged and absolute monarch and went so far as to claim that as shah he owned all of the country's land.⁷⁰

In reality, Qajar authority was quite limited. Fath Ali Shah may have exercised absolute authority in the territory immediately surrounding the capital of Tehran, but he lacked the military strength and bureaucratic capacity to exert direct control over territory that lay outside that area. Instead, he and his successors could only indirectly influence events in the rest of the country by entering into strategic political marriages with key aristocrats, by exploiting—and at times stoking—conflict among tribes and ethnic groups, and by trading the legitimacy that the dynasty could bestow on local notables in exchange for their nominal subordination to the central government.⁷¹

One consequence of the dynasty's political weakness was its tendency to pursue policies that were harmful to the country's economy. Fiscal issues played a key role here. Perpetually in budgetary arrears, the Qajar were compelled to supplement the taxes they raised directly in the region surrounding Tehran by resorting to the sale of offices—particularly governorships and positions as tax collectors. Holding their posts for terms of only one year, the winning bidders had little incentive to invest in the territories they acquired and instead squeezed the taxpayers as hard as they could so as to extract maximum profits before their terms expired. As a result, while the sale of offices secured quick cash for the dynasty, it came at the cost of long-term economic growth. Making matters worse, the Qajar devoted precious little of the revenue raised through this system to the construction and maintenance of infrastructure such as roads, irrigation works, or educational institutions that might have contributed to economic development. On the contrary, they devoted the vast bulk of the tax revenue they raised to the military, to the small central bureaucracy and, most notably, to the extravagant lifestyles of the ruling family and its hangers-on.⁷²

The Western Onslaught

Qajar weakness would have proven challenging enough for Iran during ordinary times, but it was particularly problematic coinciding as it did with surging Western aggression. As was the case for the Ottoman Empire, the greatest threat came from the north. Eager to acquire warm-water ports, Russia set its sights on Iran beginning shortly after the turn of the eighteenth century. In a pair of conflicts fought during and shortly

⁷⁰ Ervand Abrahamian, *A History of Modern Iran* (Cambridge: Cambridge University Press, 2008), 8–22.

⁷¹ Abrahamian, 8–22.

⁷² Nikki Keddie, *Modern Iran: Roots and Results of Revolution* (New Haven: Yale University Press, 2006), 24–27.

after the Napoleonic Wars, Tsar Alexander I's (r. 1801-1825) modern army handily defeated Fath Ali Shah's outmatched force of undisciplined tribal levies and irregular troops. Unsurprisingly, given the one-sided nature of the wars, the agreements that ended the fighting imposed stiff terms on the Qajars. They required Iran to cede substantial territory in the Caucasus, to cap Iranian tariffs on imported Russian goods at 5 percent, to exempt the tsars' subjects from having to pay internal customs, and to accord them extraterritorial legal rights.⁷³

Russia was not alone in seeking strategic and commercial advantage in Iran. Eager both to secure new outlets for its industrial goods and to protect its vital Indian colony by preventing St. Petersburg from establishing a presence on the Persian Gulf, Britain, too, sought influence in Iran and access to its markets. London got its chance thanks to a succession dispute among rival Qajar princes following Fath Ali Shah's death in 1834. In exchange for backing Shah Mohammad Mirza's candidacy (r. 1834-1848), the British government secured treaties in 1836 and 1841 that accorded its merchants the same commercial and legal rights that the tsarist government had won for its subjects.⁷⁴

Ironically, it was London and St. Petersburg's competition for dominance that allowed Iran to retain its independence during the age of imperialism. Determined to prevent each other from assuming a commanding position in a country that each deemed essential to its security, the two powers tacitly agreed in the mid-nineteenth century to turn Iran into a neutral buffer state in the ongoing Great Game. This situation did not mean that Tehran enjoyed genuine, independence, however. As Mohammad Mirza's succession made clear, British and Russian diplomats regularly involved themselves in Iran's internal affairs—a practice that severely eroded the dynasty's standing and legitimacy among the country's increasingly nationalistic subjects.⁷⁵

The rise in Iranian nationalist sentiment in the late-nineteenth century was, itself, a direct result of the growing Western economic penetration that the commercial treaties with Russia and Britain had made possible. The impact of those agreements was enormous. With Iranian tariff rates limited to a mere 5 percent, inexpensive, industrially produced Western finished goods such as guns, sugar, kerosene, tea, and, most consequentially, textiles, flooded into the country. The result was a sudden, wholesale restructuring of Iran's economy. Unable to compete with the West's cheaper, machine-made goods, many traditional businesses rapidly folded; in their place, new, less-profitable extractive industries emerged to supply European factories with the raw materials that they demanded. Thus, while onetime mainstays of the urban economy such as handicrafts and textiles withered, the production of primary products for export such as raw cotton, rice, and opium thrived. Concurrently, a combination of the global

⁷³ Gene R. Garthwaite, *The Persians* (Malden, MA: Blackwell Publishing, 2005), 192–94.

⁷⁴ Keddie, *Modern Iran*, 44.

⁷⁵ Keddie, 34–35.

collapse of the price of silver—the metallic basis of Iran’s currency—and the rapid rise in the country’s current account deficit produced biting inflation.⁷⁶

The ongoing economic transformation that resulted from the West’s sudden commercial penetration proved deeply unsettling. To be sure, some benefited. Most notably, landlords who switched to the production of cash crops for export and a few far-sighted merchants who were able to successfully navigate the country’s incorporation into the global market did quite well. Most Iranians fared poorly, however. Workers in the traditional craft and textile industries lost their livelihood when factory-produced goods began to flood into Iran, for example, while many of the politically and symbolically important bazaar merchants faced bankruptcy in the face of the huge influx of inexpensive Western products. Peasants suffered the most. Already enduring a decline in wages owing to a general rise in the population and a parallel increase in the labor supply, they suffered long periods of immiseration over the nineteenth century punctuated by bouts of famine. It was a combustible mix—one that produced surging Iranian nationalism, rising anger at the Western powers, and growing frustration with the Qajar shahs.⁷⁷

Concessions

Most observers then and now agree that the Iranian people’s dissatisfaction with the ruling dynasty was justified. Over the course of the nineteenth century, the Qajar regime provided little effective leadership in response to the threat that the West posed. Most notably, in sharp contrast to the substantial defensive modernization programs that the Egyptian and Ottoman states pursued in the nineteenth century, the Iranian government made no more than halfhearted efforts to reform or strengthen the country.⁷⁸

Owing to the length of his reign, Nasser al-Din Shah (r. 1848-1896) bears the bulk of the responsibility for this failing. While he did authorize a few modernization initiatives such as the establishment in 1879 of a small, Western-style military unit—the Russian-officered Cossack Brigade—he also ended a promising effort to initiate a Tanzimat-style modernization program by having its leading proponent, his chief minister, Amir Kabir (1807-1852), killed in 1852 at the behest of courtiers whose salaries and power his restructuring proposal had threatened. Thereafter, the shah staunchly opposed most reforms, and instead focused his efforts on selling concessions to Westerners in order to generate the funds needed to sustain his government and his lavish lifestyle.⁷⁹

⁷⁶ Abrahamian, *A History of Modern Iran*, 37.

⁷⁷ Abbas Amanat, *Iran: A Modern History* (New Haven: Yale University Press, 2017), 240–42.

⁷⁸ Garthwaite, *The Persians*, 206.

⁷⁹ Michael Axworthy, *A History of Iran: Empire of the Mind*, Second Edition (New York: Basic Books, 2016), 192–99.

At first, the concessions were a boon for the Qajar shah. Though the income raised was modest relative to the privileges that their Western purchasers were gaining, they nonetheless provided the Iranian government with desperately needed funds. The concessions seemed to come at little political cost, moreover. Early agreements such as one that granted Britain the right to construct a telegraph line across the country did not contribute to the wrenching economic changes that Iran was experiencing or directly impact ordinary people. As such, they aroused neither significant popular attention nor, more importantly, meaningful opposition.⁸⁰

Beginning in the late-nineteenth century, however, concessions became a lightning-rod issue in Iran. The key event that transformed public perceptions of them was the signing of a new contract between the government and the news-agency pioneer Baron Julius de Reuter (1816-1899) in 1872. The concession he secured was as huge as it was one sided. It gave him the exclusive right to construct and operate railroads, streetcars, new irrigation works, mines, and a national bank in Iran; in exchange, Nasser al-Din Shah's government would receive a modest share of the profits. The concession was to last for seventy years. The leak of the deal's terms produced an immediate backlash in Iran. Furious at the shah for entering into an obviously detrimental agreement that would directly and adversely affect the lives of many Iranians, his subjects launched spontaneous, nationwide protests demanding its immediate revocation. Faced with this intense and broad-based opposition movement, the shah had little choice but to comply. Accordingly, at great expense, he broke the agreement with de Reuter.⁸¹

A subsequent concession that granted the British merchant Major G. F. Talbot a monopoly on the sale and production of tobacco in 1891 generated an even fiercer nationalist response. The cultural and economic centrality of tobacco products in Iran explains the intensity of the opposition. With a huge share of the population either consuming or producing tobacco, the concession promised to directly and viscerally affect the lives of millions of Iranians. As a result, news of the agreement spurred the immediate emergence of a new opposition coalition. Broad in composition, it included groups as diverse as secular Western reformers, bazaar merchants, and urban workers.⁸²

Significantly, it also involved Shi'i clerics. Their participation in the protest movement marked a significant change. Heretofore, the 'ulama' had refrained from intruding directly in politics. Now, they were not only participating in the protests against the tobacco concession but were leading them. Indeed, the core element of the protest effort—a boycott on the production and consumption of tobacco—originated in 1892

⁸⁰ Keddie, *Modern Iran*, 54–55.

⁸¹ Keddie, 54–55.

⁸² Amanat, *Iran*, 304–14.

with a fatwa issued by the *mujtahid* Mirza Hasan Shirazi (1815-1895), at the behest of the Islamic modernist Jamal al-Din al-Afghani (1838-1897).⁸³

Iranians of all classes enthusiastically took up the ban. In short order, tobacco production halted, and Iranian smokers went cold turkey. Even Nasser al-Din Shah's wives honored the boycott. Facing such intense and unified opposition, the shah had little option but to cancel the concession. Doing so proved to be expensive. Breaking the deal with Talbot came at the staggering cost of £500,000, an amount so high that the shah had to borrow the funds on international money markets—thereby incurring Iran's first foreign debt.⁸⁴

Iran's Constitutional Revolution, 1905-1911

The tobacco protest set the stage for further and more far-reaching popular political action under Nasser al-Din Shah's successor, Mozaffar ad-Din Shah (r. 1896-1907). The new shah's reign generated substantial opposition from the very start. With his hands tied by his state's feeble administrative capacity and by Western-imposed tariff limits, he could neither arrest skyrocketing inflation nor address a serious cholera epidemic then ravaging Iran. Accordingly, any reservoir of support that the new shah may have enjoyed quickly evaporated. Meanwhile, his profligacy soon put the country yet further under the West's thumb. Most notably, he took three trips to Europe of such extravagance that they required his government to take two costly loans from Russian banks and to grant a British subject, William Knox D'Arcy (1849-1917), a new, sixty-year concession for the exploration and production of petroleum in southern Iran. The oil concession and the loans infuriated the now-politically conscious Iranian people. Indeed, news of them not only spurred the quick revival of the tobacco protest coalition shortly after the turn of the century but also left the country standing on the brink of a revolution.⁸⁵

It finally came in 1905. As revolutions so often have, it began with women protesting food shortages and rising prices in the cities. Mozaffar ad-Din Shah's government could do little to arrest the revolt's progress once it got going. Indeed, its transparent effort to scapegoat sugar merchants for the rise in prices not only failed to deflect blame but further infuriated the public. Even the use of force failed to slow the growing size of the demonstrations that rocked the country's cities. Finally, in 1906, the protests reached a crescendo. Early in the year, three leading *mujtahids* took *bast*, or sanctuary, at a shrine near Tehran and demanded the creation of an *Adalat Khaneh*, or house of justice, by which—though they were vague—they meant some form of consultative or representative body. Shortly thereafter, 14,000 bazaar merchants upped the ante by taking *bast* in the gardens of the British legation, where, guided by secular modernizers, they called for the promulgation of a constitution and the establishment of a majlis, or parliament. These events ended any hope that Mozaffar ad-Din Shah's

⁸³ Amanat, 304–14.

⁸⁴ Garthwaite, *The Persians*, 208; Keddie, *Modern Iran*, 60–61.

⁸⁵ Keddie, *Modern Iran*, 62–65, 72.

government might have entertained that it could rein in the protests. Accordingly, the shah capitulated in August and reluctantly agreed to hold elections for a representative assembly.⁸⁶

Convened in October 1906, the National Assembly began by drafting a constitution for Iran. It came in two parts. The first, the Fundamental Law, established Iran as a constitutional monarchy in which parliament was to be the dominant branch of government. The second, the Supplementary Fundamental Law, guaranteed Western-style civil rights such as freedom of the press and assembly and granted all citizens—including religious minorities—equal rights. With great reluctance, Mozaffar ad-Din Shah agreed to abide by the Fundamental Law in December 1906. His successor, Mohammad Ali Shah (r. 1907-1909), was even less enthusiastic about ceding power to the people; nonetheless, he committed the dynasty to observe the Supplementary Fundamental Law by signing it in October 1908. It was a heady time in Iran. Only the second Asian country to adopt a written constitution, it appeared to have joined the ranks of the world's democratic states.⁸⁷

In reality, democracy continued to face substantial challenges from both inside and outside Iran. Within the country, the resistance to constitutional government came primarily from an alliance between the shah and the clerical establishment, which had become alarmed when it came to realize that the reformers' version of democracy was very different from the system of traditional Muslim elite consultation that they had in mind when they had called for *Adalat Khaneh*. The monarchy and the clerical establishment were powerful institutions in Iran, and, acting in unison, they were able to mount a strong resistance to constitutional government. Indeed, at one point in 1908, the conservatives were able to effect a counter coup that temporarily ended democratic rule in Iran.⁸⁸

Ultimately, democratic forces in Iran had the strength to contain the challenges that the monarchy and the 'ulama' posed to representative government. They quickly reined in the shah and restored the majlis to its position of authority in Iran. They were not strong enough to overcome the combined hostility of London and St. Petersburg to Iranian democracy, however. Their shared opposition to the political changes in Tehran was a byproduct of a recent reconciliation between them. Fierce competitors for much of the nineteenth century, they had signed an agreement in August 1907 called the Anglo-Russian Convention that was designed to resolve their differences so that they could jointly address the common danger posed by Germany. Focused largely on ending their Great Game rivalry in Asia, the agreement dealt with their competing interests in Iran by dividing the country into a northern, Russian-dominated region, a

⁸⁶ Axworthy, *A History of Iran*, 2nd page, slide to revolution section.

⁸⁷ Amanat, *Iran*, 337–41.

⁸⁸ Keddie, *Modern Iran*, 69–71.

British sphere of interest in the south, and a neutral, buffer zone in the middle of the country.⁸⁹

The agreement did not end Iran's independence. Indeed, the two powers continued to recognize Iranian sovereignty and, for a time at least, accepted its democratic government and nationalist aspirations. What they were not willing to tolerate, however, were efforts by the Iranian government to challenge their dominant positions in their respective spheres of interest. Accordingly, when the majlis threatened the tripartite division of the country by seeking to enforce tax collection in the Russian zone, the tsarist government—with London's tacit support—compelled the cabinet to dissolve the majlis. This intervention in Iranian affairs proved pivotal. With it, Iran's experiment with constitutional government and representative democracy came to an ignominious end.⁹⁰

The Discovery of Oil

Meanwhile, even during the constitutional period, foreign economic dominance of Iran continued to increase. Most notably, after several years of fruitless searching and worsening finances, D'Arcy and his partner, the Burmah Oil Company, struck oil in the province of Khuzestan in southwest Iran in 1908. This was a major discovery—one that was well beyond D'Arcy and Burmah's ability to exploit. Consequently, the following year, they formed a new, publicly held corporation, the Anglo-Persian Oil Company, to raise the capital needed to extract and market the concession's petroleum. APOC's production totals were modest at first; it pumped just 1,600 barrels per day in 1912. However, the fact that it possessed the only proven, British-controlled source of oil in the world made the company and its concession strategically significant—especially after the admiralty opted in 1912 to convert the fleet from coal-fired powerplants to more flexible and efficient oil-powered engines. Indeed, the navy's shift from coal to oil rendered the D'Arcy concession so essential to imperial defense that the British government itself assumed a majority stake in APOC in 1914. Thus, from that point forward, the British state directly owned the oil that lay under the sand of southwest Iran—a situation that indignant Iranian nationalists could do nothing about.⁹¹

World War I and Britain's Play for Dominance

The outbreak of the First World War served to further catalyze foreign domination of Iran. In the north, Russian forces not only fought with Ottoman troops along both sides of Iran's western border but also regularly intervened in its domestic politics. In the south, meanwhile, London expanded its sphere of interest and took steps to protect APOC assets in Khuzestan Province from Arab saboteurs incited to action by Wilhelm Wassmus (1880-1931)—Germany's "Lawrence of Arabia." It did so by organizing a new military formation composed of Iranian troops under the command of British officers

⁸⁹ Abrahamian, *A History of Modern Iran*, 49–50.

⁹⁰ Keddie, *Modern Iran*, 69–71.

⁹¹ Yergin, *The Prize*, 150–64.

called the South Persia Rifles. Numbering 8,000 men at its peak, it proved essential both in safeguarding the oil fields and in ensuring that Khuzestan remained firmly within the British sphere of interest.⁹²

Despite being officially neutral, Iran suffered enormously during and just after the First World War. Some of its problems stemmed directly from the fighting that took place on its soil. Most obviously, combat operations in the north between Ottoman and Russian forces resulted in the deaths of substantial numbers of Iranians and left many more refugees. Far more consequential for Iran, however, was the indirect, war-induced scarcity of food. Stemming from a combination of labor impressment, fighting, and requisitioning, which redirected scarce food supplies to foreign armies, famine killed untold thousands across Iran. Worse, it left millions malnourished and thus susceptible to epidemic diseases like cholera, typhus, and, especially, the Spanish Flu. Demographic data makes clear just how devastating the war was for Iran. Collectively, famine, disease, and combat operations resulted in the deaths of two million Iranians—about 10 percent of the prewar population—between 1917 and 1921.⁹³

Britain sought to exploit Iran's weakness and Russia's revolution-induced paralysis immediately following the armistice to advance its imperial interests in the region. Indeed, Iran was a central part of Prime Minister Lloyd George's (r. 1916-1922) postwar plan to establish Britain as the hegemonic power in the Middle East. In contrast to its approach in other parts of the region, however, London did not seek formal control in Iran. Instead, it moved to establish a de facto protectorate over the country through a new, bilateral agreement it drafted in 1919 called the Anglo-Persian Treaty. By its terms, London would provide Tehran with a £2-million loan to pay for a powerful army and for transportation infrastructure that would help the central government restore its authority over the country following the war. In exchange, Britain would assume formal control over Iran's finances and customs collection and would become the country's sole provider of military equipment and advisors.⁹⁴

Ultimately, London's gambit failed. Its attempt to implement key parts of the treaty before the Iranian government had formally endorsed it ignited a furious opposition campaign that promised violence against any politician who favored the agreement. That threat worked. Successive prime ministers—even ones supportive of the deal—refused to submit it to the majlis for approval for fear of their lives. As a result, the treaty did not go into effect. In response, London decided to abandon its bid for dominance in Iran in 1921 and to instead consider alternative ways of ensuring its interests in the country. It quickly settled on the idea of sponsoring a strongman who could maintain order and uphold key commercial agreements. Accordingly, British officials in Tehran dropped their efforts to secure passage of the treaty and began

⁹² Garthwaite, *The Persians*, 214–15.

⁹³ Abrahamian, *A History of Modern Iran*, 59–60; Keddie, *Modern Iran*, 73–75.

⁹⁴ Keddie, *Modern Iran*, 75–80.

searching for a suitably pliable military figure who could impose order on a country that had slid into chaos.⁹⁵

Reza Khan Takes Control, 1921-1925

They found their man in a dynamic, if autocratic, officer in the Cossack Brigade: the modernizing nationalist Reza Khan. With tacit British encouragement, he rode into Tehran in February 1921 at the head of a force of 3,000 troops and deposed the existing government. At first, it appeared that Reza Khan had acted merely to restore order to Iran. Indeed, he did not seize the reins of government himself but instead took the subordinate posts of minister of war and head of the army and helped the journalist and modernizer Sayyid Zia al-Din Tabatabai (r. 1921) become prime minister. At British insistence, he also retained Ahmad Shah Qajar (r. 1909-1925) as head of state. It soon became clear, however, that these actions were little more than window dressing designed to obscure what had in fact been a coup d'état in which Reza Khan had assumed control of the Iranian state.⁹⁶

He devoted the first few years of his rule to consolidating his authority and engaging in state building. These were tall orders. Bereft of an effective central government since the outbreak of World War I, Iran had devolved into a decentralized patchwork of de facto statelets ruled by tribal khans and regional warlords who expressed little more than token fealty to Tehran. Thanks to steadily rising oil royalties, however, Reza Khan was well positioned to overcome this resistance. Moving aggressively, he devoted the lion's share of Iran's oil revenue to the army. He then used his beefed-up military to expand the proportion of the country over which the central government exercised real authority, and, importantly, to compel landlords who fell under his control to pay both current *and* back taxes. In short order, as a result, he had set in motion a virtuous cycle wherein his growing army secured control of territory and new taxes that in turn permitted a further expansion of the military and, with it, the acquisition of additional land and revenue.⁹⁷

By 1925, the process was essentially complete. Thereafter, the central government's writ extended to all parts of Iran and Reza Khan's rule stood without serious challenge internally or externally. Confident in his position, he chose that time to depose Ahmad Shah Qajar and to arrange for the majlis, which he had reduced to little more than a debating society, to proclaim him the new monarch: Reza Shah Pahlavi. He chose his dynastic name with great deliberation. Dating to the Sassanid Dynasty

⁹⁵ David Fromkin, *A Peace to End All Peace: The Fall of the Ottoman Empire and the Creation of the Modern Middle East* (New York: Avon Books, 1989), 455–60.

⁹⁶ Abrahamian, *A History of Modern Iran*, 63–65.

⁹⁷ Abrahamian, 66–72.

(224-651), the term Pahlavi linked his family to one of Iran's most powerful pre-Islamic ruling houses and thus helped to legitimate his reign.⁹⁸

“Modernization from Above,” 1925-1941

Reza Khan moved aggressively after his coronation to complete the campaign of nation building that he had begun in 1921. An autocratic ruler who was committed to remaking his country along modernist, Western lines, he forcefully pursued a rapid, top-down restructuring program modeled on Mustafa Kemal's widely celebrated effort to restructure Turkey. Appropriately, he referred to his reform effort as “modernization from above.”⁹⁹

Like Ataturk Reza Khan made secularization a central component of his restructuring program. Looking to break the power and independence of the Shi'i clerics, he pursued secularization along three tracks. First, he stripped the religious scholars of their control over education by establishing a secular public school system patterned on the French *lycée* model and by gradually folding existing religious schools into it. Second, he ended the clerics' dominance of the law in Iran by replacing the shari'a code with new civil, commercial, and criminal legal systems and by requiring judges to possess degrees from secular law schools—a stipulation that effectively barred the religiously educated 'ulama' from serving on the bench. Finally, he undermined the clerics' financial independence by first promulgating a law in 1932 that forbade them from engaging in the profitable practice of registering legal documents and then, later that decade, by seizing the *awqaf* that sustained them. By 1941, as a result, the once mighty religious scholars—who, ironically, had been among Reza Khan's earliest supporters—had lost much of their independence and power.¹⁰⁰

“Modernization from above” also involved an array of Kemalist-inspired changes designed to turn Iran into a modern country unified by a common national identity. Those reforms ran the gamut. To bring Iran into accord with modern European practices, Reza Khan introduced the metric system, replaced the traditional Muslim lunar calendar with a new, solar one—albeit still dated from the hijra—and required Iranians to adhere to a common time zone. Meanwhile, to unify the people and to bind them to the state, he promoted a series of initiatives aimed at replacing particularist identities with a common, nationalist one. Issued in 1928, for example, the Law of Uniformity of Dress sought to weaken competing conceptions of identity by banning people from wearing clothing that indicated their tribe, class, ethnicity, or religion and by requiring all men, save for clerics, to wear Western-style, brimmed hats. Later, and more controversially, he outlawed the practice of veiling. Finally, he implemented a

⁹⁸ Ali Ansari, *Modern Iran Since 1921: The Pahlavis and After* (New York: Longman, 2003), 36.

⁹⁹ Keddie, *Modern Iran*, 87.

¹⁰⁰ Lapidus, *A History of Islamic Societies*, 477–78; Abrahamian, *A History of Modern Iran*, 84–85.

system of conscription in 1925 designed as much to create a sense of shared, national identity as to ensure an adequate supply of recruits for the army.¹⁰¹

Like Ataturk, Reza Khan also recruited the social sciences to the task of nation building. To create a strong sense of common identity, he established the Cultural Academy and tasked it with replacing Armenian, Turkish, and Arabic loanwords with archaic Persian terms or neologisms in an effort to purge the Persian language of foreign influence. A more radical attempt to follow the Kemalist example by replacing the Arabic alphabet drew interest for a time, but ultimately foundered because the task was deemed too difficult.¹⁰²

History and archaeology figured prominently in Reza Khan's nation-building program. Looking to weaken Iran's connection with its immediate Muslim past, he insisted on a public-school curriculum that emphasized the country's rich, pre-Islamic history. Likewise, he sponsored archaeological digs designed to connect Iran to earlier dynasties such as the Achaemenids and Parthians, and he insisted that newly constructed public buildings reference pre-Islamic architectural styles. In 1935, meanwhile, he moved to change Iran's international image by requesting that Western states cease referring to his country as Persia and instead begin calling it Iran—a term derived from the word Aryan. Doing so would achieve two ends. First, it would deepen the country's connection to its ancient past and thus help to foster a distinct national identity. Second, by casting Iranians as members of the Aryan race, it would elevate his people's relative standing in the Western racial hierarchies that loomed so large in the interwar period.¹⁰³

"Modernization from above" cleaved most closely to the Kemalist program in its pursuit of economic development. Like Ataturk, Reza Khan promoted a state-led, ISI-development model in which domestic factories sheltered by protective tariffs produced goods for import substitution. He laid the basis for this program in 1927 by unilaterally abrogating the nineteenth-century era unequal treaties that had accorded extraterritorial privileges to Westerners and had limited Iran's customs duties to 5 percent. He followed in 1933 by creating two holding companies, the Agricultural and Industrial Banks, to determine how state-raised capital could best be allocated for economic development. Between their establishment and 1941, they underwrote the construction of more than three-hundred factories, mills, and refineries that produced a mix of durable materials like cement and lumber and consumer goods such as cigarettes, sugar, processed foods, textiles, glass, soap, and matches. Concurrently, the state laid thousands of kilometers of roads to link Iran into a single, unified market. Finally, it imposed taxes on tea and sugar to finance the construction of the north-south, Trans-Iranian Railway.¹⁰⁴

¹⁰¹ Keddie, *Modern Iran*, 91–100; Abrahamian, *A History of Modern Iran*, 77–83.

¹⁰² Abrahamian, *A History of Modern Iran*, 86–87.

¹⁰³ Garthwaite, *The Persians*, 229; Abrahamian, *A History of Modern Iran*, 86–88.

¹⁰⁴ Quote from Keddie, *Modern Iran*, 95–102; Abrahamian, *A History of Modern Iran*, 77.

Though Reza Khan trumpeted his “modernization from above” program as a great achievement, it was, in reality, a mixed success at best. On the positive side of the ledger, his campaign had fostered the development of a bourgeoisie, secured limited emancipation for women, improved transportation infrastructure, and provided almost all of the country’s cities with at-least limited electrical power. More importantly, it had succeeded in unifying Iran and at replacing tribal and ethnic identities with an Iranian nationalist one. At the same time, however, Reza Khan’s reform program had a number of substantial shortcomings. Politically, it had imposed a despotic government on Iran and had cleaved the country into “two cultures”: a small, wealthy, Westernized class of bureaucrats, military officers, professionals, and businesspeople and a much larger and poorer traditional one of bazaar merchants, clerics, and peasants. Economically, meanwhile, his development program had failed utterly to live up to his promises. Marred by corruption, Iran’s nascent industrial sector was inefficient, uncompetitive, and exploitive of labor; it also produced goods that were—as is typical in cases of ISI development—both shoddier and substantially more expensive than the imports that they replaced.¹⁰⁵

Reza Khan’s biggest failure, however, was his inability to achieve a favorable revision of the terms of the APOC oil concession. Like other Iranian nationalists, the shah was dissatisfied with nearly every part of the agreement. He was unhappy that it did not supply Iran with a minimum royalty payment, upset with the company’s discriminatory labor practices, and, above all, angry that the British state earned substantially more from the sale of his nation’s oil than did his government. Accordingly, in the early 1930s, he insisted on opening negotiations with APOC aimed at altering the terms of the concession in Iran’s favor. The discussions went nowhere. While the company was willing to talk, it was unwilling to make more than cosmetic changes to the terms of the concession and was adamant that the profit-sharing arrangement could not be altered. Frustrated, Reza Khan broke off the discussions in 1932 and announced that Iran would unilaterally annul the D’Arcy concession and assume control of the oil fields.¹⁰⁶

This threat finally got APOC and the British government’s attention. At their request, talks recommenced in 1933. They soon produced a revised agreement that called for Britain to grant a larger share of the profits to Tehran and to train more Iranians for skilled positions in the oil fields—terms that Reza Khan trumpeted as a great victory for his country. In reality, as many Iranians grasped, the renegotiated arrangement was in fact a poor agreement for his state that offered little more than minor alterations to the terms of the original concession. Most notably, it only raised Iran’s share of the profits from an anemic 16 percent to a marginally better 20 percent—an increase that APOC’s creative accounting succeeded in substantially reducing. Likewise, it offered no more than symbolic changes to the company’s discriminatory employment practices. Perhaps worst of all, Reza Khan was only able to win those

¹⁰⁵ Garthwaite, *The Persians*, 228; Keddie, *Modern Iran*, 95–103.

¹⁰⁶ Ervand Abrahamian, *The Coup: 1953, the CIA, and the Roots of Modern U.S.-Iranian Relations* (The New Press, 2013), 28–29.

modification to the original deal by agreeing to extend the concession's term by thirty-two years, thus giving Britain control of Iran's oil through 1993.¹⁰⁷

Reza Khan's inability to secure a meaningful revision of the APOC concession left him angry and led him to gradually drift toward Britain's rival, Germany. The relationship blossomed quickly—particularly in the area of trade. Indeed, commerce between the two countries grew so rapidly over the next few years that Germany displaced Britain as Iran's largest trading partner by the end of the decade. Tehran's ties with Berlin did not remain limited to the economic sphere for long, moreover. Seeing Nazi Germany as a potential counterweight to Britain, Reza Shah moved to strengthen political relations during the late 1930s. He permitted German intelligence operatives free rein in Iran, pursued a series of diplomatic policies supportive of Berlin, and even appropriated Nazi racial and nationalist views.¹⁰⁸

World War II and Reza Shah's Abdication

Reza Khan's dalliance with Hitler's Germany proved to be his downfall. During the early part of World War II, Britain had been willing to tolerate his passive animosity and tacitly pro-German position. Hitler's invasion of the USSR in June 1941 drastically altered Iran's standing in the Allies' strategic calculus, however. Not only did his country possess vital oil reserves, but it also occupied the best route for the shipment of desperately needed war materiel and resources to the Soviet Union. In other words, it was too vital of a place to leave in the hands of someone as unreliable as Reza Shah. Accordingly, the USSR and Britain launched a bloodless invasion in August 1941 in which they assumed control of Iran and deposed the unreliable shah. Joined by the United States after it entered the war, the occupying powers followed by using the Trans-Iranian Railroad to ship war materiel to the USSR. Moscow and London did not, however, end the Pahlavi Dynasty. Instead, they arranged for Reza Khan's twenty-one-year-old son, Muhammad Reza (r. 1941-1979), to replace his father on the throne.¹⁰⁹

Young and inexperienced, the new shah lacked the force of will or the political skills to take control of his country; as a result, a power vacuum emerged in Tehran. Taking advantage, a variety groups and institutions were able to assert themselves culturally and politically. Some of them were old ones that Reza Khan had suppressed. Most notably, the clerics exploited the shah's deposition to win back control of both the *awqaf* that financially sustained them and the religious schools that the government had nationalized; in so doing, they were able to restore much of the autonomy that they had lost in the "Modernization from Above" program. Other groups that gained prominence at this time were new. One of the most prominent of them was the Soviet-supported, far-left Tudeh Party. Thanks to its support for unionization and call for aggressive land reform, it quickly gained the backing of most workers, many intellectuals, and some

¹⁰⁷ Abrahamian, 28–29.

¹⁰⁸ Keddie, *Modern Iran*, 98–101.

¹⁰⁹ Amanat, *Iran*, 493–99.

peasants. It even secured the support of a significant portion of the Westernized middle class.¹¹⁰

Postwar Crises

The instability continued in the years immediately following World War II as Iran became embroiled in an early dispute in the emerging Cold War. The crisis was a product of Moscow's territorial and economic aspirations. As per the agreements that governed the occupation, British and American soldiers withdrew from Iran within six months of the conflict's conclusion. The USSR did not follow suit, however. Instead, leaving its troops in place as leverage, it set up pro-Soviet separatist Azerbaijani and Kurdish republics in the northeast and demanded that Tehran grant it a highly favorable concession to exploit oil deposits in Iran's northern provinces. Given the power imbalance between it and the USSR, Iran appeared to have little choice but to capitulate to these demands.¹¹¹

Instead, its savvy prime minister, Ahmad Qavam (1874-1955), was able to deftly maneuver his country through the crisis. He did so in stages. First, he negotiated an agreement that seemed to meet most of Moscow's demands. In exchange for the withdrawal of Soviet troops, he agreed to respect the autonomy of the Kurdish and Azerbaijani republics and to submit a bill to the majlis granting the USSR an oil concession. It looked like a good deal for Moscow, which responded by promptly pulling the Red Army from Iran. Qavam followed the Soviet withdrawal by implementing the second phase of his plan, which involved crushing the Kurdish and Azerbaijani republics and their leftist governments. Moscow was unhappy with this move but was willing to tolerate Tehran's betrayal so long as it got its oil concession. Here, too, however, Qavam double crossed the Soviets. Aware all along that strident nationalist sentiment rendered the oil-concession bill a dead letter in the majlis, he effected the final stage of his plan in October 1947 when he formally submitted the legislation only to see it quickly and predictably go down to defeat—thus permitting Iran to meet the obligations of its agreement with Moscow while simultaneously freeing it from having to make any real concessions.¹¹²

The crisis significantly altered the political balance in Iran in two ways. First, it greatly weakened the heretofore rapidly growing Tudeh Party. While the party had achieved new levels of popularity as a result of a successful strike it had organized against APOC in the spring of 1946, the party's willingness to defend the USSR's actions in Iran severely damaged its nationalist credentials and cost it much of its support. Second, the crisis solidified the emerging view in Washington that Iran was a critical component of its broader Cold War effort to contain the USSR and thus spurred the US to step up its engagement with the country. Notably, Washington agreed in 1947

¹¹⁰ Garthwaite, *The Persians*, 235–37.

¹¹¹ Barry Rubin, *Paved with Good Intentions: The American Experience and Iran* (New York: Oxford University Press, 1980), 33–35.

¹¹² Rubin, 33–35.

to extend the terms of an existing military advisory mission under the command of Colonel Norman Schwarzkopf (1895-1958) and to provide Iran with \$60 million in military aid.¹¹³

Conclusion

As Iran's experience makes clear, the interwar period was a challenging time for the sovereign states of the Middle East. Operating in the shadow of the West, the countries Iran, Saudi Arabia, and Turkey struggled to transform themselves into stable, fully independent nation-states and to develop their economies. These were difficult enough goals to achieve on their own, but they were rendered even more challenging by the huge power imbalance that existed between those states and the Western powers. Indeed, Iran proved unable to maintain its independence in the face of Allied pressure during World War II while Saudi Arabia was compelled to settle for a smaller share of the profits generated by its oil wealth than did either Aramco or the US government. Even Turkey had to walk a very careful line to retain its independence during the Second World War. Still, even with these limits, the citizens and subjects of these states enjoyed vastly greater autonomy than did the majority of the people of the Middle East, who lived not under independent governments but instead under de facto European colonial rule. It is to their very different situation and their efforts to change it that we shall next turn.

¹¹³ Keddie, *Modern Iran*, 113–14.

Chapter Twelve: The Middle East Under European Control, 1922-1948

When the dust cleared from the final peace settlement in the early 1920s, France and Britain had emerged as the masters of nearly the entire Arab world. Paris directly governed Lebanon and Syria, while London, in keeping with its indirect colonial approach, controlled Egypt and the mandates of Transjordan and Iraq through local rulers.¹ The imperial powers did not formally rule these territories as permanent colonies, however. Instead, on the basis of the mandatory agreements into which they had entered—or, in the case of Egypt, the public statements of policy they had made—France and Britain administered those places on a temporary basis. Moreover, at President Woodrow Wilson's (r. 1913-1921) insistence, Paris and London had formally committed themselves to preparing the peoples they ruled in the Middle East for eventual self-government.

They did not honor that obligation. Instead, seeking to turn their new possessions in the region into permanent—if nominally independent—neo-colonies, the imperial states hindered the emergence of effective governments and eschewed making any real effort to help the Arab states develop the close connection between citizens and government on which genuine legitimacy ultimately rests. Worse, in several cases, they bequeathed to the Arab governments substantial coercive capacities designed to permit those states to maintain order.

Paris and London possessed the most power and thus had the biggest impact on political developments in those territories, but they were not alone in inhibiting the development of effective governments in the region. Indigenous groups including greedy elites, populist extremist organizations, and, especially, autocratic military officers also pursued policies that interfered with the development of popular, stable governments. Coupled with the self-interested approach of the imperial powers, the actions of these groups ensured that Egypt and the Arab mandates emerged from World War II not as secure, self-governing countries, but instead as unstable and undemocratic states—ones that were simultaneously powerful due to their possession of substantial coercive capacities and weak owing to their lack of legitimacy and absence of close connections with the people they ruled.

Iraq: Quasi Independence

Iraq's experience between the end of the First World War and the conclusion of World War II was emblematic of this story. On the surface, the country appeared to be an interwar success. As the mandatory power, Britain had provided it with modern political institutions and had even granted it its independence in 1932. A charter

¹ Britain was also the mandatory power for Palestine. We will examine that territory separately in the next chapter.

member of the United Nations, Iraq seemed to have become a stable member of the international order by the end of the Second World War.

In fact, Iraq emerged from that conflict a deeply troubled country. That it did so was largely a function of London's self-interested administration. During its time as the mandatory power, Britain had imposed a variety of political, economic, military, and social policies on Iraq that were designed to serve its imperial interests rather than the needs of the Iraqi people. Collectively, they ensured that this multi-ethnic, multi-confessional country failed to develop the strong civic-political culture—meaning a sense of identity with the state and an acceptance of its legitimate power—that modern polities need to function effectively. Britain certainly bore primary responsibility for Iraq's warped development, but it was not solely culpable for its challenging political situation. Drawing on the values they had acquired in Ottoman military academies, the Sunni elite that dominated Iraq in the 1930s also contributed to the country's fractured political development by promoting an authoritarian and militarized conception of the state, and by affording the numerically larger Kurdish and Shi'i Arab people with little power or even sense of belonging in the national community.

Mandatory Iraq, 1922-1932

The peculiar contours of the Iraqi state first took shape under indirect imperial rule during the mandatory period. They emerged as a result of Britain's pursuit of two broad objectives for the country. First, it sought to prepare Iraq for self-government as required by the League of Nations' mandate. Second, and far more important, it also acted to ensure that Iraq would remain an informal British colony after it had secured its independence. That these two goals were clearly in conflict was not obvious to British officials at the time. On the contrary, they perceived them to be entirely complementary in nature.

On the surface, Britain achieved a great deal for the country during its brief tenure as the mandatory power. Drawing on its long history of representative government, it helped nurture the key institutions needed for Iraq to function as a democratic state. Under its tutelage, Iraq acquired a constitution, an elected parliament, a military, a constitutional monarch in the form of the Hashimite leader of the Arab Revolt, Faysal I (r. 1921-1933), and a bureaucracy. Britain also helped to ensure the development of Iraq's economy by persuading the League of Nations to assign the disputed *vilayet* of Mosul and its substantial oil reserves to Iraq instead of Turkey in 1925, and it established a new way of assuring internal stability through the use of cost-effective air power rather than through the maintenance of an expensive conscription army.²

² Phebe Marr, *The Modern History of Iraq*, Second Edition (Boulder: Routledge, 2004), 21, 24–26, 30–31; Toby Dodge, *Inventing Iraq: The Failure of Nation-Building and a History Denied* (New York: Columbia University Press, 2003), 132–36.

Drawing on their extensive historical and anthropological knowledge, British officials also made sweeping changes to the political, legal, and economic structure of rural Iraq designed to secure the interests of its tribal people. Noting the similarity between rural Iraqis and the reputedly fierce-but-honorable Pathan people of the northwest frontier region of Britain's India colony, British officials devised a romantic construction of the Iraqi tribespeople that held that they were, like their Pathan counterparts, "noble savages" who were uncorrupted by civilization, and therefore vulnerable. On the one hand, their closeness to nature left them admirably honest and innately good; on the other, it also rendered them too naive and credulous to avoid being exploited by the town Arabs whom racist colonial officers perceived as being wily and cynical. This situation deeply troubled British officials. How, they asked themselves, could they safeguard the noble-but-simple tribespeople from their unscrupulous, urban neighbors?³

While anthropological knowledge of the subcontinent raised the need to protect the tribespeople, familiarity with their own history led the British to a solution to this problem. Looking back on the role that the rural gentry had played in guarding peasants from an absolutist government under the Tudor monarchs, they reasoned that a comparable class of landlords could shield the mandate's benighted tribal people by similarly interceding between them and the state. Fortunately for the British, rural Iraq had an existing class of leaders that they believed could, if properly strengthened, satisfy that need: the shaykhs who served as tribal leaders. Accordingly, British officials gave those leaders the tools needed to serve as a mediating force. They granted the shaykhs judicial power over the people of their tribes, made them responsible for collecting taxes and maintaining order, and, most critically, assigned them legal title to land that the tribes had heretofore held on a communal basis.⁴

Independence, 1932

Iraqi independence came quickly. Committed to a balanced budget, Britain was eager to shed itself of expenses—even the comparatively minimal costs it incurred as the mandatory power in Iraq. Accordingly, London opened talks with Faysal I's government in 1929 aimed at establishing Iraq as an independent state while simultaneously ensuring Britain's interests. Signed in 1932, the resulting accord seemed beneficial for both parties. By its terms, Iraq gained its independence and assumed a seat at the League of Nations; in exchange, Britain secured control of strategic airbases at Habbaniya near Baghdad and at Shuaiba in the south, and the right to move troops through Iraq during time of war. It also won the exclusive right to supply Iraq with military advisors and equipment. The terms of the treaty were to last for twenty-five years.⁵ With the signing of the agreement, Britain appeared to have brought the mandate to a successful conclusion for all parties. Baghdad had won its

³ William R. Polk, *Understanding Iraq* (New York: Harper Perennial, 2005), 21, 24–26, 30–31.

⁴ Dodge, *Inventing Iraq*, 46, 63.

⁵ Marr, *The Modern History of Iraq*, 34–36.

independence—the first of the mandates to do so—and seemed to possess the institutional infrastructure to function as a sovereign state. For its part, meanwhile, London had secured its key strategic goals in Iraq for at least the next few decades.

Problems

Reality most certainly did not match appearances, however. Britain had structured Iraq based on a set of assumptions and preconceived notions that proved wildly inaccurate, and it had imposed policies that worked to ensure its short-term interests at the expense of Iraq's long-term needs. The result was the emergence of a deeply flawed and troubled country riven by class and ethno-religious conflict. These outcomes were the product of a number of internal and external forces. However, four British actions during the mandatory period were particularly important in producing them: its handling of the negotiations over Mosul, its reliance on airpower to ensure order, its elevation of the tribal shaykhs to a position of authority, and its insistence on an independence treaty that left Iraq in a state of neocolonial dependence.⁶

The disposition of Mosul reflected most clearly how London pursued its interests at the expense of Baghdad, and how, in the process, it weakened and delegitimated the Iraqi state. Britain had administered the former Ottoman *vilayet* of Mosul through the Iraq mandate since the end of World War I. As we saw in chapter ten, however, it had agreed at Turkish insistence to defer the issue of its final disposition to the League of Nations. Thus, whether oil-rich Mosul would—or would not—be part of Iraq remained very much an open question during the mid 1920s.⁷

In the meantime, an important event was taking place in the oil industry that would, to the detriment of Iraq, intersect with the question of Mosul's disposition. In a complex series of negotiations, a consortium of oil companies had assumed control of the German and Ottoman shares of the Turkish Petroleum Company (TPC)—renamed the Iraq Petroleum Company (IPC) in 1929—and had thus acquired an exclusive concession to find and exploit oil in the territory of the former Ottoman Empire. The consortium was an international one, consisting of the French *Compagnie Française des Pétroles*, Britain's Anglo-Persian Oil Company, British-controlled Royal Dutch Shell, and a group of five American firms headed by Standard Oil of New Jersey. Each party received a 23.75 percent share of the TPC, with the remainder going to the Armenian businessperson Calouste “Mr. Five Percent” Gulbenkian (1889-1955), who had put the original TPC consortium together in 1912. Critically, the constituent companies in the TPC agreed to adhere to a self-denial clause known as the Red Line Agreement that bound them to pursue concessions in the territory of the former Ottoman Empire on a joint basis through the TPC. This provision applied to Mosul and had the effect of

⁶ Dodge, *Inventing Iraq*, 31.

⁷ Zara Steiner, *The Lights That Failed: European International History 1919-1933* (Oxford: Oxford University Press, 2005), 120–22.

limiting competition for the right to exploit oil and thus helped to keep the price of the concession low.⁸

Still, the consortium faced the undesirable prospect of having to share a substantial portion of the profits on any oil found in Mosul with the Iraqi government. This situation stemmed from an agreement negotiated in 1920 that had allocated to Baghdad not merely a portion of the profits but also a twenty percent ownership share of any oil produced in Iraq by the TPC. London had accepted that arrangement in the early 1920s, but it had, for two reasons, become unalterably opposed to it by mid-decade. First, an ownership position would result in Baghdad receiving a substantially larger share of the profits than was the case in similar states. Second, and more importantly, an ownership stake would give Iraq a voice in setting production rates and prices and would thus establish a dangerous precedent that other producing countries might seek to follow.⁹

As British officials were well aware, however, they possessed a powerful card that they could—and did—play to block Baghdad from participating in the management of the concession. It involved the negotiations with Turkey. Exploiting its role as the mandatory power, London threatened to inform the League of Nations that it was willing to cede the *vilayet* of Mosul to Turkey if the Iraqi government did not abandon its ownership stake. This move was heavy handed and imperialistic, but it was also successful. Unwilling to risk the loss of Mosul and its oil, Faysal I reluctantly agreed to relinquish Iraq's 20 percent ownership position in the TPC in favor of a modest increase in royalty payments and a British promise to work aggressively to keep the *vilayet* in Iraq.¹⁰

Britain's hardball approach to Mosul adversely affected Iraq's political and economic development in two critical ways. First, it resulted in the country receiving a substantially smaller share of the TPC's profits following the discovery of oil near Kirkuk in 1927 than it would have received had it retained its ownership position. As a result, interwar Iraq was perpetually starved of the funds needed to promote economic development. Second, Britain's ability to compel Faysal I's government to yield its ownership position made abundantly clear the degree to which the monarchy was subordinate to London and thus undermined the credibility of the Iraqi state in the eyes of its people.¹¹

Britain's strategy of using airplanes to enforce internal order similarly warped Iraq and delegitimated its state. As we saw in chapter ten, the use of aircraft in the mandate

⁸ William Stivers, "A Note on the Red Line Agreement," *Diplomatic History* 7, no. 1 (January 1983): 23–24; Quote from Daniel Yergin, *The Prize: The Epic Quest for Oil, Money & Power* (New York: Touchstone Books, 1991), 187.

⁹ Charles Tripp, *A History of Iraq* (Cambridge: Cambridge University Press, 2002), 58–60.

¹⁰ Tripp, 58–60.

¹¹ Marr, *The Modern History of Iraq*, 30–31.

had its roots in the financial constraints that London faced in the immediate aftermath of the First World War. Determined to return to a balanced budget, the British government had sought at that time to aggressively cut the cost of its Middle Eastern empire. Iraq had been especially problematic in this regard. Not only was it a large country and thus expensive to police, but it had already been the scene of a major rebellion—one that had required an expensive deployment of troops to bring under control. Indeed, Iraq had become so costly that some officials had given serious thought to abandoning it altogether.¹²

As we saw in chapter ten, Secretary of State for the Colonies Winston Churchill (1874-1965) was not among them. Determined to keep the mandate, he and his advisors had determined in 1921 that Britain could retain control of it at an acceptable cost through the use of airplanes. Inexpensive yet powerful, aircraft and the threat of bombing would both ensure order and compel the heretofore autonomous rural tribes to pay the taxes that the mandatory government imposed; their use would thus make Britain's presence in Iraq financially self-sustaining.¹³

The approach worked. Indeed, from London's perspective, the employment of airpower to coerce rural Iraqis appeared to be one of the great successes of its Middle East policy in the early 1920s. It not only compelled obedience from the tribespeople—just as Churchill had argued that it would—but it also all but eliminated the cost of maintaining order in the mandate. As a result, Britain was able to retain possession of Iraq in a way that was consistent with its commitment to a balanced budget.¹⁴

The reliance on air power may have served Britain's financial interests, but it ran diametrically counter to Baghdad's political goals. Intent on nation building—that is, on creating a shared sense of national identity—Faysal I's government had been pushing for the development of a large conscription-based army that would bring multiple benefits. First, such a force would provide the state with the means both to defend itself from external attack and to maintain internal order. Equally important, it would also help to inculcate a shared sense of identity with the Hashimite monarchy and the Iraqi state among the people, tying Faysal I's heterogeneous country together into a coherent whole. It was a sound, if expensive, proposal. As noted, however, it was also incompatible with Britain's desire to control costs. Accordingly, London rejected it outright and instead continued to depend on airpower for security through the early 1930s.¹⁵

This decision had profound implications. The absence of a conscription army meant that the Iraqi state lacked an important institution for forging a nation out of its divergent ethnic and confessional groups—one that could have helped it to develop the

¹² Christopher Catherwood, *Churchill's Folly: How Winston Churchill Created Modern Iraq* (New York: Basic Books, 2005), 94–103.

¹³ Dodge, *Inventing Iraq*, 131.

¹⁴ Polk, *Understanding Iraq*, 83–84.

¹⁵ Dodge, *Inventing Iraq*, 136–41.

“imagined political community” that is so essential to a shared sense of identity. Meanwhile, the use of airplanes to enforce internal order and compel the payment of taxes meant that the state’s connection with its subjects rested, at root, on violence and intimidation rather than on positive interactions that might bind the people to the government and help to create a positive civic-political culture. In other words, Britain’s decision about the nature of the Iraqi military meant that Faysal I’s government would hold substantial coercive capacities but would lack the positive connections with the people on which genuine legitimacy ultimately rests. This reliance on what the scholar Toby Dodge calls “despotic state power” at the expense of a broad popular mandate would cast long shadows. Thereafter, nearly all subsequent Iraqi governments eschewed any meaningful effort to win the consent of the governed and instead ruled by means of brute force.¹⁶

Meanwhile, the effort to protect the rural tribespeople from the presumably rapacious town Arabs also created lasting issues that undermined Iraq’s political development. The problem here stemmed from the fact that the historical and anthropological analogies that British officials had drawn on in approaching the mandate were almost comically off the mark and had led them to implement policies that radically—and detrimentally—reshaped the structure of Iraqi society. The shaykhs were not, as the British had assumed, powerful, charismatic figures who represented the will of the tribes; rather, they were first-among-equals whose authority rarely extended beyond the power of persuasion. London’s decision to vest judicial power in these men and, more importantly, to grant them title to tribal land thus did not reinforce existing practices; rather, it imposed a top-down social revolution on rural Iraq that turned these weak figures of tenuous influence into a powerful class of landlords whose authority ultimately rested on the coercive power of the state. The consequences were substantial. In short order, rural Iraqis fell into thrall to the landlord-shaykhs and their financial backers, the town Arabs—the very class, ironically, from whom the British had sought to protect the tribespeople. The passage of the Law Governing the Rights and Duties of Cultivators in 1933 confirmed these momentous changes. Imposing draconian restrictions on the ability of indebted peasants to move and binding them to the land, this sweeping law allowed the shaykhs to complete the enserfment of the heretofore free tribespeople that British policy had inadvertently started.¹⁷

Unsurprisingly, the attempt by London to restructure Iraqi society based on flawed historical and anthropological assumptions adversely affected the mandate’s development. It did so in two ways. First, it warped the country’s already unstable political system. The transformation of the shaykhs into a politically and economically powerful class helped fuel the emergence of an oligarchic system of government in which competing factions composed of well-connected political figures, newly empowered landlord-shaykhs, and affluent town Arabs competed for access to the

¹⁶ First quote from Benedict Anderson, *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, Revised edition (London: Verso, 2006), 6; Second quote from Dodge, *Inventing Iraq*, 145.

¹⁷ Tripp, *A History of Iraq*, 51–52, 84–85.

levers of power on which the acquisition of new wealth depended. Thanks to British actions, in other words, Iraq ended up with a government of and for the elite rather than one attuned to the needs of ordinary Iraqis. Second, Britain's efforts to restructure Iraqi society dramatically widened economic inequality in the country and, as a consequence, helped to create deep-seated feelings of bitterness and resentment—particularly among the tribespeople who increasingly fled the shaykhs' exploitation by moving to the shantytowns then popping up on the outskirts of the country's cities.¹⁸

Just like the social revolution that Britain had inadvertently effected among the tribespeople, the treaty that established Iraq as an independent state created broad resentment in key groups that would come to undermine the state's legitimacy. The agreement was not without its supporters, of course. Some, including Faysal I, the court faction headed by the important politician Nuri Said (1888-1958), and many of the landlord-shaykhs, continued to perceive value in the connection with Britain and saw the treaty beneficial for their country. In contrast, many other politically conscious Iraqis viewed it as an embarrassingly unequal agreement that rendered their country little more than an informal British colony. Educated Sunnis, especially those in the armed forces, were particularly furious about it and the privileges it accorded to Britain. They bitterly resented London's continued role advising the government and armed forces, railed about its dominance of Iraq's economy, loudly complained about its retention of airbases, and fiercely assailed its right to move troops through the country during time of war. Unsurprisingly, there emerged among them a stridently nationalist and anti-imperialist sentiment and a deep skepticism about a government that they perceived to be a collaborationist one.¹⁹

Thus, thanks to a mix of ignorance and its efforts to uphold British commercial and strategic interests, London's brief tenure as the mandatory power left a destructive legacy for Iraq. It is true that British officials had provided the country with the critical institutions needed to function as a modern state and that they moved quickly to grant it independence. At the same time, however, they had pursued self-interested policies that left Iraq with an exploitive elite and an oppressive state that lacked positive connections to the people. The result was a country hobbled by tremendous inequality, class conflict, seething resentment toward the British, widespread alienation toward a political system that relied on its coercive powers to compel obedience, and the absence of a shared sense of national identity—conditions that were hardly conducive to stable political development.²⁰

The Army in Politics

Britain continued to exert a great deal of influence over Iraq following its independence, but it no longer enjoyed mastery over the country as it had before 1932. Instead, with the tacit support of the youthful new king, Ghazi (r. 1933-1939), the Sunni-

¹⁸ Tripp, 75–78, 84–85.

¹⁹ Marr, *The Modern History of Iraq*, 34–36, 40.

²⁰ Tripp, *A History of Iraq*, 74–76.

dominated officer corps quickly displaced both London and the civilian factions as the primary force in Iraqi politics. The army's political ascendancy ultimately proved to be short-lived, but its enthusiastic embrace of authoritarianism and its aggressive promotion of a Sunni-dominated, secular, Arab-nationalist state—what the scholar Vali Nasr calls “the secularization of Sunni political identity”—would indelibly and adversely shape Iraq through the end of the twentieth century.²¹

The Sunni Arab officers and bureaucrats who dominated Iraq in the 1930s held a common vision for their country's future rooted in their shared Ottoman background. Educated in the military academies of the late empire, these veterans of the Committee for Union and Progress's (CUP) autocratic government had little use for liberal, British-supported political concepts like representative government or civil rights—particularly if those practices enfranchised Kurds and Shi'a—and instead hoped to modernize Iraq along authoritarian and militarized lines. Accordingly, they instituted a stridently pan-Arab nationalist curriculum in the schools, promoted centralization, dismissed democracy as a distraction that would weaken the state, and established a conscription-based military. While they rejected Britain's liberal political ideas, however, they enthusiastically embraced the coercive powers that London had given the Iraqi state and freely used them to enforce both the central government's authority and Sunni social and political dominance. For example, the Iraqi military brutally suppressed Kurdish and Assyrian Christian insurgencies in the mid 1930s and made liberal use of aerial bombing to bring revolting Shi'i tribes to heel.²²

Emboldened by its success suppressing these uprisings, the military dramatically intensified its involvement in politics in the mid 1930s. This process started in 1936, when General Bakr Sidqi (1890-1937), a Kurd, used the army to support a coup that made the reformer Hikmat Sulayman (1889-1964) prime minister and installed members of a radical faction called the Ahali group in the cabinet. Seeking to dramatically restructure the country, Sulayman's government immediately pressed for a series of sweeping changes that were designed to end Iraq's severe economic and political inequality, such as the legalization of unions and the repeal of the Law Governing the Rights and Duties of Cultivators. These proposals went nowhere thanks to the opposition of the more conservative Sidqi. However, the general did enthusiastically throw his support behind Sulayman's promotion of a new, broader construction of Iraqi nationalism that was inclusive of Shi'a and Kurds, as well as his “Iraq first” foreign policy that downplayed pan-Arabism in favor of closer ties to Turkey and Iran.²³

²¹ Tripp, 77–82; Vali Nasr, *The Shia Revival: How Conflicts within Islam Will Shape the Future* (New York: W. W. Norton, 2006), 92.

²² Michael Provence, “Stateless Revolutionaries and the Aftermath of the Ottoman Great War,” *Journal of Modern European History* 13, no. 3 (2015): 404–11; Marr, *The Modern History of Iraq*, 39–44.

²³ Quote from Tripp, *A History of Iraq*, 88–99; D. K. Fieldhouse, *Western Imperialism in the Middle East 1914-1958* (Oxford: Oxford University Press, 2006), 104.

Those last initiatives proved to be the regime's downfall. Angry that Sidqi had abandoned pan-Arabism in favor of a more inclusive conception of national identity, a group of Sunni army officers assassinated the general in 1937. Thereafter, the increasingly emboldened officer corps made regular use of violence and coups—six occurred between 1936 and 1941—to compel successive civilian cabinets to adhere to their authoritarian, Sunni, pan-Arabist vision for Iraq.²⁴

The Coup of 1941

Regional and global events soon put the Sunni officers and the British on a collision course. Opposition to London's continued dominance had surged to new heights among Sunni nationalists in the late 1930s in response to Britain's violent suppression of an Arab uprising in Palestine and its subsequent proposal to partition that territory into Jewish and Palestinian states. Keenly aware of the vast power imbalance that existed between Baghdad and London, however, the officers understood that any effort to try to force the British out would be a fool's errand. Thus, though they fumed, they made no meaningful attempt to challenge London's position in Iraq during the late 1930s. Germany's spectacular success during the early stages of World War II appeared to change the political calculus, however. Not only did the Third Reich's victories seem to indicate that London's power was in eclipse, but they also suggested that the nationalists could secure help from Berlin in their efforts to expel Britain. In response, a clique of officers known as the Golden Square began to lay plans for a move against London's position in Iraq.²⁵

Even as the outbreak of the war seemed to create opportunities for the Sunni nationalists, however, it also powerfully reinforced Britain's determination to retain a dominant position in Iraq. Indeed, the realities of modern warfare had dramatically heightened the country's importance in London's strategic calculations. Not only did Iraqi oil fuel Britain's critically important Mediterranean fleet, but its territory also served as a vital waystation for the transfer of troops from South Asia to the Middle East and Europe. As such, London was deeply concerned about the threat that the Sunni officers posed to its position in a territory it deemed vital to the war effort.²⁶

To the relief of the British, a leadership change in Iraq shortly before the outbreak of World War II not only eased those fears but also seemed to create an opportunity to curb the growing power of the Sunni nationalists. The shift was a product of Ghazi's untimely death in a car accident in 1939 and the subsequent accession of his son, Faysal II (r. 1939-1958) to the throne. Still an infant, the new king could not yet rule; accordingly, real power ending up in the hands of the regent, his uncle, Abd il-Alah (1913-1958). Staunchly pro-British, Abd il-Alah made clear from the start that he

²⁴ Marr, *The Modern History of Iraq*, 49, 52–53.

²⁵ Marr, 51–52.

²⁶ Tripp, *A History of Iraq*, 99–105.

intended to use his considerable constitutional authority—which included the power to approve new prime ministers—to check the pan-Arabists.²⁷

Bolstered by Abd il-Alah's presence, the British moved to achieve a favorable resolution of the political situation in the spring of 1941. Aware that the Golden Square was conspiring with Germany to push it out of Iraq, London pressed both the regent and Prime Minister Taha al-Hashimi (r. 1941) to adopt a more openly pro-British foreign policy. Al-Hashimi was amenable but was aware that he would need to first weaken the Golden Square if he hoped to align Baghdad more closely with London. Accordingly, he quietly began to reassign its leaders to less-important commands located far from the capital and to replace them with more reliable men. Though well laid out, the plan backfired badly. Catching wind of his scheme before al-Hashimi could put it into effect, the officers quickly organized a countermove aimed at overthrowing the prime minister and at replacing him with the pro-German politician Rashid Ali (r. 1941).²⁸

The operation was an almost total success. It removed al-Hashimi from office and secured the Golden Square's complete control of the government. Critically, however, thanks to the intervention of American ambassador Paul Knabenshue (1883-1942), the plotters failed to arrest il-Alah, who slipped out of the palace by hiding under a blanket in the backseat of the ambassador's official car. As a result, Ali and the Golden Square were unable to force the regent to issue a royal decree legitimating the change in government. Still, this issue aside, the pan-Arabists were pleased with their work. Solidly in control of the country and anticipating German military assistance, they were confident that they would soon be in a position to eject the British from Iraq.²⁹

The Anglo-Iraqi War

Coming at a time in which German fortunes in the Eastern Mediterranean and North Africa were rapidly improving, the coup deeply troubled British officials and impelled them to move to force the Iraqi government to make clear whether it intended to abide by the terms of the 1932 treaty. They did so in the spring of 1941 by landing Indian troops in Basra—thereby compelling the Iraqi government either to back down or to openly break with Britain. Furious, the Golden Square rose to the bait. Emboldened by the prospect of German support, it countered London's provocative action by ordering Iraqi troops to surround the British airfield at Habbaniya in late April and by threatening to shoot down any planes that attempted to take off from the base. Having maneuvered the pan-Arabists into formally violating the terms of the 1932 treaty, London followed by ordering British forces defending Habbaniya to first push Iraqi soldiers away from the base and to then advance on Baghdad. Despite the deployment of German aircraft to northern Iraq, the initial British move was successful—as was a

²⁷ Daniel Silverfarb, *Britain's Informal Empire in the Middle East: A Case Study of Iraq 1929-1941* (New York: Oxford University Press, 1986), 103, 120.

²⁸ Silverfarb, 122–24.

²⁹ Peter L. Hahn, *Missions Accomplished?: The United States and Iraq Since World War I* (New York: Oxford University Press, 2012), 7–8.

second attack on Iraqi positions outside Fallujah shortly thereafter. At that point, the Golden Square's position gave way. With morale collapsing, the troops defending the capital abandoned their posts en masse, leaving Rashid Ali and the leaders of the Golden Square with little recourse but to flee for safety to Iran. Thus, by the beginning of June, Britain had not only ended the threat to its position in Iraq but had once again taken complete control of the country.³⁰

The officers may have gone down to defeat in their showdown with Britain, but their brief time in power nonetheless had an enormous impact on Iraq's emerging political culture. During their period of dominance in the 1930s and early 1940s, they had succeeded in institutionalizing the vision of authoritarian nationalism that they had acquired during their time studying in the Ottoman Empire's military academies. Just as important, they had also succeeded in ensuring that both their conception of Iraq as a country dominated by Sunni Arabs and their embrace of pan-Arabism continued to command broad acceptance in the Iraqi government and—especially—in the military. The impact of these views was as baleful as it was lasting for Iraq. Coming on the heels of Britain's failure to foster either the development of a strong, inclusive sense of national identity or a government that enjoyed broad legitimacy, the growing prevalence of the officers' pan-Arabist beliefs prevented the state from constructing a notion of identity that was inclusive of all Iraqis or from establishing a governing system that enjoyed the genuine consent of the people.³¹

The Second British Occupation

Lasting from 1941 to 1947, Britain's second occupation gave the civilian political elite a new chance both to purge Iraq of the officers' influence and to finally establish a strong connection between government and society. Initially, the pro-British politicians who dominated wartime Iraq appeared to move in that direction. They vigorously persecuted the leaders of the coup in 1941 and aggressively purged pan-Arab extremists from the government. They also broadened support for the state by bringing a number of prominent Kurds and Shi'a into the ruling circle.³²

Ultimately, however, Nuri Said, il-Alah, and the other political figures who dominated Iraq during the second occupation did not take advantage of the opportunities that wartime exigencies had presented them. They failed to build strong connections between citizens and state, refused to give Iraqis genuine representation, and made only halfhearted efforts to create a robust, shared identity that would be inclusive of the Kurds and Shi'i. Instead, the landlord-shaykhs and wealthy town Arabs who dominated the government largely left unchanged what was, at root, a thoroughly unrepresentative state. Self-interested and with little commitment to genuine democracy, they focused on engaging in intra-elite factional struggles and on using the state's powers of coercion and patronage to their advantage—just as they had in the

³⁰ Fieldhouse, *Western Imperialism in the Middle East 1914-1958*, 105–6.

³¹ Tripp, *A History of Iraq*, 77–82, 94–96.

³² Marr, *The Modern History of Iraq*, 56–58.

1930s. Largely Sunni Arabs themselves, meanwhile, the members of the ruling factions saw little to gain from ending the Sunni dominance of Iraq that had existed from the start of the mandate. Thus, despite the substantial coercive powers it continued to hold, the Iraqi state remained a weak one that enjoyed little more civic-political culture or legitimacy in 1945 than it had when it had become independent in 1932.³³

Interwar Egypt: The Liberal Age

A British protectorate rather than a mandate, Egypt seemed substantially further along the road to independence than did Iraq at the start of the interwar era. It had a well-established monarchy, exercised substantial autonomy, and was in the process of drafting a constitution. Importantly, it also had a popular, well-established political organization that could connect the people to the state: Saad Zaghlul's (1859-1927) Wafd Party. As such, despite Britain's retention of a powerful position in the country thanks to the Four Reserved Points, Egypt appeared poised to emerge as a stable, independent, and democratic state in the early 1920s.

The imminence of self-rule in Egypt was, in fact, an illusion. In reality, neither genuine independence nor the institutionalization of a robust system of representative government were at hand. Instead, thanks to a combination of British manipulation, royal obstruction, and Wafd Party elitism, Egypt failed to establish the civic-political culture needed to legitimate the state and to bind its people to the liberal political order. As a result, after the Second World War, the country would experience the spread of political violence, the rising power of undemocratic and extremist non-parliamentary movements, and, most importantly, the complete delegitimation of the liberal system—developments that would, collectively, doom the constitutional order to a quick demise.³⁴

Constitution and Democracy

In contrast to the gloom that Egypt would experience following World War II, hopes ran high in 1922 that full independence was imminent. Indeed, the Egyptian people at that time were enthusiastically working to establish the contours of the new government that they believed would lead them to independence. They focused first on drafting a constitution. Taking effect in 1923, it reflected divergent impulses. On the one hand, it guaranteed basic rights and ensured through the adoption of universal male suffrage that the Egyptian people had a strong voice in the government. On the other, it vested an unusual degree of authority in the king—giving him the power to prorogue

³³ Marr, 58–61.

³⁴ Steven A. Cook, *The Struggle for Egypt: From Nasser to Tahrir Square*, Reprint edition (Oxford University Press, 2011), 22–23.

parliament, to dismiss cabinet ministers, to veto legislation, and to name two-fifths of the members of parliament's upper house.³⁵

Egypt followed in 1924 with its first parliamentary elections under the new system. To the surprise of no one, the Wafd Party won a decisive victory and Zaghlul took office as the country's first prime minister. It was a heady time. With the Wafd Party promising to negotiate an end to the Four Reserved Points, the country appeared ready at last to become a fully independent, democratic state.³⁶

The Trials of Representative Government

It failed to do so, however. In fact, Egypt would have a peculiarly difficult time securing full independence under the direction of the Wafd Party—then or later. What explains this failure? Why in spite of having both broad popular support and a large majority in parliament, was the party unable to achieve its primary goal? The answer lay in the peculiar structure of Egyptian politics in the interwar era. During the 1920s and 1930s, three power centers—Britain, the king, and the Wafd Party—defined the contours of the Egyptian political system. Each had a different goal. While the secular Wafd sought genuine independence and representative government, King Fuad (r. 1922-1936) instead strove to establish a royal dictatorship. In the background, meanwhile, London worked to retain the informal control over Egypt and the Suez Canal that the Four Reserved Points had accorded it. The result was a tripartite struggle that saw the parties form tacit, shifting alliances to achieve their policy objectives. At times, Britain and the Wafd collaborated to contain the monarch; as their interests were generally more closely aligned, however, it was more typically the British and Fuad who cooperated to frustrate the Wafd's approach. As a result, the Wafd Party's efforts to renegotiate the terms of Egypt's relationship with Britain or to shift power from the monarch to the legislature met with repeated failures.³⁷

Indeed, a cyclical pattern soon took shape. Each phase would begin with the Wafd Party winning a decisive electoral victory that gave it a mandate to seek full independence from London. Taking advantage of the vast authority that the constitution vested in him, Fuad would then use a conflict—typically an impasse in the negotiations—to assert his power. He would dismiss the cabinet, dissolve parliament, and install a pliable government that was beholden to him; the palace would then rule the country autocratically for several years. Eventually, a new political conflict would

³⁵ Selma Botman, "The Liberal Age," in *The Cambridge History of Egypt. Volume 2, Modern Egypt, from 1517 to the End of the Twentieth Century* (Cambridge: Cambridge University Press, 1998), 285–91.

³⁶ Botman, 285–91.

³⁷ Arthur Goldschmidt Jr., *Modern Egypt: The Formation of a Nation State* (New York: Routledge, 2004), 74–76.

emerge that would compel the king to call new elections. The Wafd would once again win a substantial majority, after which the cycle would once more repeat.³⁸

The pattern first took shape immediately following the Wafd's huge victory in elections held in January 1924. Having secured a large parliamentary majority on the basis of Zaghul's promise to achieve full independence, the party immediately opened talks with London aimed at attaining complete sovereignty. For a time, hopes ran high. To the consternation of the Wafd, however, the negotiations deadlocked over the question of whether London or Cairo would control Sudan. The diplomatic impasse—and the subsequent assassination in Cairo of the British governor-general of Anglo-Egyptian Sudan—then created a crisis that compelled Zaghul to resign. Using the powers that the constitution had vested in him, Fuad followed by dissolving parliament and by naming a new cabinet consisting of royalist favorites who were disinclined to continue the push for full independence.³⁹

To Zaghul's dismay, as a result, the talks he had opened with Britain came to an end. It was a bitter pill for a man who had fought so hard and for so long for Egyptian independence. It was also his last hurrah: dying just three years later, he never again headed the government.⁴⁰

The cycle repeated several more times before culminating in the temporary end of democratic rule in the early 1930s. Ironically, the retreat from democratic government came in response to the Wafd's growing success. Running once again on a platform calling for negotiations aimed at achieving full independence, the party won a resounding electoral victory in 1929 that made its new leader, Mustafa al-Nahhas (1879-1965), the prime minister. Genuine independence once more seemed imminent. Unfortunately for the Wafd, however, al-Nahhas was no more successful than Zaghul had been at translating electoral victory into diplomatic progress, and the talks again broke down over the question of whether London or Cairo would rule the Sudan. Worse, the deadlock once more presented Fuad with a golden opportunity to assert his power. Taking advantage of the crisis, he secured al-Nahhas's resignation and appointed a new prime minister—the autocratic royalist, Ismail Sidqi (1875-1950)—and tasked him with implementing a new constitution that vested nearly all power in the monarchy. Once it was enacted, Fuad himself ruled Egypt as royal dictator while the Wafd could only look on, impotently, from the sidelines.⁴¹

As these events make clear, Fuad and the British bore most of the responsibility for both the Wafd's difficulties and for Egypt's failure to develop a robust democratic system in the interwar period. Indeed, both parties worked assiduously to foil the will of the Egyptian people. At the same time, however, the Wafd itself was far from

³⁸ Botman, "The Liberal Age," 290.

³⁹ Jayne Gifford, *Britain in Egypt: Egyptian Nationalism and Imperial Strategy, 1919-1931* (London: I. B. Tauris, 2020), 87–116.

⁴⁰ Botman, "The Liberal Age," 291–93.

⁴¹ Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 77.

blameless—either for the predicament it found itself in during the early 1930s or, more broadly, for the failings of Egyptian democracy. Its shortcomings were largely the product of the values and outlook of its leadership. While the Wafd was by far the largest political organization in Egypt and stood as the self-described champion of the liberal order, it was, ironically, not itself a democratically structured organization. Instead, dominated by educated, affluent members of the bourgeoisie, the party was an elitist institution that greeted bottom-up political activities with great suspicion and that saw little utility in institutionalizing democratic practices. As such, it made no meaningful effort to engage with or organize the Egyptian masses.⁴²

The result was good neither for the party nor for Egyptian democracy writ large. Having failed either to mobilize the masses or to bind them to the state, the Wafd ended up with electoral support that was broad but shallow and thus unreliable. More importantly, the party's lack of interest in connecting the people to the democratic state or in building the robust civic-political culture on which stable representative government ultimately rests meant that Egyptian democracy never developed deep roots—an issue that would emerge as a serious problem in the postwar era.⁴³

The Anglo-Egyptian Treaty, 1936

The cyclical political pattern finally broke in the mid 1930s thanks to a change in British policy. To that point, London had tacitly supported royal government in Egypt because neither Sidqi nor Fuad challenged its imperial position in Egypt. By the middle of the decade, however, evolving international circumstances compelled it to reexamine that approach. Against the backdrop of the broader threat that the rise of fascism in Europe posed to British security, Italy's invasion of Ethiopia in 1935 dramatically raised the stakes in Egypt. Were disgruntled Egyptian nationalists to throw in with the fascist states during time of war, defending the Suez Canal—the vital artery that linked Britain to its colonies in Asia and Oceania—would be immeasurably more challenging. London thus concluded that it needed to quickly secure a compromise agreement with Cairo that would satisfy Egyptian demands for independence while simultaneously permitting Britain to retain its critical base complex along the canal. There was one hitch, however. British officials believed that only a treaty negotiated by a popularly elected government would be legitimate in the eyes of the Egyptian people; in other words, Egypt would need to revert to democratic rule before productive talks could occur.⁴⁴

Accordingly, London joined the Wafd in demanding that the king restore the constitution of 1923 and that he hold new elections. Facing intense domestic and international pressure, Fuad had no choice but to schedule a vote for the spring of 1936. To no one's surprise, the Wafd won another decisive victory in the election, and al-Nahas once again became prime minister. Coupled with the accession of the popular new king, Faruq (r. 1936-1952), following the death of his father in early 1936,

⁴² Botman, "The Liberal Age," 288.

⁴³ Botman, 288.

⁴⁴ Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 77–79.

the subsequent opening of talks created a surge of optimism. Egypt once more appeared to be on the cusp of real independence.⁴⁵

With both sides looking to conclude a deal, the negotiations quickly produced a settlement. Structured as a twenty-year defensive agreement, the Treaty of Preferential Alliance signed in August 1936 satisfied each party's primary objectives. London kept control of Sudan, retained the right to station 10,000 troops in the Canal Zone during peacetime and an unlimited number during time of war, and secured a provision calling for it to come to Egypt's assistance should that country be threatened with invasion. In exchange, it ceded responsibility for the protection of minorities and foreigners to the government in Cairo, assented to the abolition of the hated capitulations, and, most importantly, agreed to fully recognize Egypt's sovereignty—thus paving the way for it to join the League of Nations and to oversee its own foreign policy. It was a momentous occasion. To the jubilation of most Egyptians, their state had finally won its formal independence.⁴⁶

Challenges to the Liberal Order

Not everyone shared in the enthusiasm about the agreement, however. On the contrary, bitterness regarding its limits fed a growing dissatisfaction with both the Wafd and Egypt's broader democratic order. Criticism of the party came from all sides. Secular rightists such as the small-but-growing Young Egypt Party attacked the Wafd by arguing that the treaty al-Nahas had negotiated had not ended Britain's colonial position in Egypt but had instead merely obscured it in the guise of a defensive alliance. Though their reasons differed, feminists were equally vociferous in their condemnation of the Wafd and Egypt's constitutional order. Decrying the party and the government for their culpability in the ongoing subjugation of women, they sought a series of legal and social reforms. They demanded female suffrage, called for the creation of jobs for women, pressed for the legalization of birth control and abortion, and appealed for sweeping cultural changes such as permitting wives to have an identity outside of marriage. On the left, meanwhile, mounting disenchantment with the Wafd Party's failure to address the Great Depression's impact on the lower classes helped fuel the growing popularity of a communist party among the skilled workers and intelligentsia of Cairo. While it remained small, the party would prove to be unusually influential. As we shall see, many of the ideas that it popularized—including, most importantly, the concept of central economic planning—would powerfully influence the policies of Gamal Abdel Nasser's (r. 1952-1970) Arab-nationalist government beginning in the 1950s.⁴⁷

Despite the lasting influence of their economic policies, however, the communists proved unable to capitalize on growing disillusionment with the liberal order to win mass appeal. Instead, it was a relatively new Islamist organization, the Muslim Brotherhood, that was able to take advantage of the secular Wafd Party's missteps to emerge as

⁴⁵ Goldschmidt, 77–79.

⁴⁶ Botman, "The Liberal Age," 295–196.

⁴⁷ Botman, 296–97, 303–4.

Egypt's most important non-parliamentary opposition movement in the interwar era. Founded in 1928 by a twenty-two-year-old schoolteacher named Hassan al-Banna (1906-1949), the organization was initially a typical Muslim charitable society that provided education and jobs to the poor. It did not retain this orientation for long, however. Instead, influenced by the ideas of Islamic modernists such as Muhammad Abduh (1849-1905), the brotherhood soon adopted an explicitly political approach that called for Egypt to abandon liberal, Western-style secularism in favor of a government based on Islamic principles.⁴⁸

It is important to note that the Muslim Brotherhood was not, like the Wahhabi, backward looking. That is, it neither called for a return to the economic arrangements of the past nor condemned Western science and technology. What the brothers did reject, however, was the Wafd Party's view that modern economic practices and technologies should be embedded in a Western-style social system. Arguing that such an approach had merely served to exacerbate inequality and to weaken the governments of those Muslim states that had pursued it, the brothers instead maintained that Western technologies and systems could only benefit Muslims if those practices were incorporated into a social arrangement that was based firmly on Islamic social values. It proved to be a highly appealing message. Blending the benefits of modernity with familiar touchstones, it resonated strongly with many lower- and middle-class Egyptian Muslims. Indeed, in conjunction with the brotherhood's strong anticolonialism and implementation of successful economic development projects, it resulted in the organization emerging after World War II as the most important popular opposition movement in Egypt.⁴⁹

Thus, in spite of the treaty with Britain, the Wafd party and the broader parliamentary system that it represented had begun to see their standing decline in the late 1930s. The liberal order's failure to achieve complete independence—meaning one in which there was *no* British presence in Egypt whatsoever—had subjected it to a barrage of criticism from opposition groups spanning the political spectrum. Likewise, the Wafd's inability to ameliorate the economic impact of the Great Depression or to satisfy the demands of Egyptian feminists had gradually eroded its base of support. Still, it is important to keep the party's situation during the late 1930s in perspective. Popular feelings toward it may have been more tepid than in the past, but the events of the interwar period including the treaty signed in 1936 had hardly ended its appeal or cost it the backing of most Egyptians.⁵⁰

World War II and the Abdin Palace Incident

World War II would prove to be a different story, however. Indeed, it was the Wafd's inability to navigate the tricky events of the conflict that ultimately discredited

⁴⁸ Cook, *The Struggle for Egypt*, 27–30.

⁴⁹ Cook, 27–30.

⁵⁰ Eugene Rogan, *The Arabs: A History – Revised and Updated Edition* (New York: Basic Books, 2012), 196.

both the party and the broader constitutional order. From the very start of the war, Egyptians chafed at London's not-insubstantial interference in their country's internal affairs. As per the terms of the Anglo-Egyptian Treaty of 1936, London compelled Cairo to impose martial law and censorship, to permit hundreds of thousands of Commonwealth forces to occupy bases in the country, and to break relations with Germany. It also succeeded in forcing the dismissal of a pro-Axis prime minister and a prominent general. Coming on top of near famine conditions and the biting inflation that wartime economic disruptions had caused, these events embittered Egyptians of all social classes and left many disillusioned with the Wafd and the liberal order for their failure to deliver on the promise of real independence.⁵¹

Ultimately, however, it was not the sum of the war's impact on Egypt's economic health or the restrictions that the treaty had placed on its independence that turned people against the liberal order. On the contrary, it was a single event, The Abdin Palace Incident of February 1942, that fatally wounded the credibility of Egypt's parliamentary system. The episode grew out of a change in Egyptian leadership earlier that year. Angry that Prime Minister Husayn Sirri (1894-1960) had broken diplomatic relations with a German ally, Vichy France, without first seeking royal input, Faruq had secured the prime minister's resignation in early 1942. In less-fraught times, the British might have been willing to let this action slide. With Germany's powerful Afrika Korps mounting an offensive in Libya toward the Egyptian frontier, however, London felt that it needed to respond aggressively to Faruq's move in order to ensure its position in the country. Accordingly, at a meeting with the king at his residence, the Abdin Palace, on the evening of February 4, 1942, Ambassador Miles Lampson (1880-1964) issued a blunt ultimatum: either appoint Wafd leader al-Nahhas as prime minister or abdicate. With little room to maneuver thanks to the fact that Lampson had preemptively arranged for British tanks to surround the palace, the king quickly acceded to the ambassador's demands. Believing that he could secure additional resources for Egypt if he assumed control of the government, al-Nahhas also agreed to go along with Lampson's demands. As a result, the crisis came to a speedy end.⁵²

Despite its quick conclusion, however, the Abdin Palace Incident irrevocably damaged the position of all three parties. Faruq's failure to abdicate or even to stand up to the British utterly destroyed his standing with the Egyptian people. Meanwhile, though Lampson's naked interference in Egypt's internal affairs had secured London's immediate ends, the move also produced a huge upsurge in anti-British sentiment that imperiled Britain's long-term presence in the country. If London and Faruq had weakened their respective positions, however, the Wafd had—through its willingness to assume power behind British bayonets—all-but destroyed its popular standing and nationalist credentials. Thereafter, as a result, it hemorrhaged support to movements

⁵¹ Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 85–86.

⁵² Botman, "The Liberal Age," 298–300.

such as the Muslim Brotherhood that had remained uncompromising in their opposition to the British.⁵³

Egypt After World War II

With the liberal order discredited and with resentment toward the British surging to new heights, Egypt began to descend into political chaos following World War II. The prime movers of the country's political collapse were the non-parliamentary groups that had been steadily gaining strength since the signing of the 1936 Anglo-Egyptian Treaty. Believing that their time was at hand, the most important of those organizations, the Muslim Brotherhood took steps to complete the destruction of the liberal system and to force a complete break with London following the war. It organized demonstrations to protest the failings of the liberal government, ordered a wave of assassinations including the shocking murder of Prime Minister Ahmad Mahir (1888-1945) in parliament in February 1945, and organized increasingly violent riots against the continued British presence in Egypt. The worst clash occurred in Cairo in February 1946. Faced with a huge crowd of students and workers attempting to storm their barracks, British troops opened fire with machine guns; when the smoke cleared, twenty-three Egyptians lay dead and more than one hundred were injured.⁵⁴

Despite the chaos, the liberal order nearly succeeded in redeeming itself late in 1946. Reappointed prime minister that year, Sidqi initiated negotiations with London aimed at finally securing the withdrawal of British troops from Egypt. After several rounds of difficult talks, he and Foreign Minister Ernest Bevin (1881-1951) succeeded in hashing out a tentative agreement that called for London to remove its remaining forces from the Canal Zone. More importantly, they came up with a compromise arrangement for the administration of Sudan that papered over the two sides' differences regarding whether Egypt or Britain would retain control of that territory. For a brief and exciting moment, as a result, complete independence at last appeared to be at hand. Once again, however, the status of Sudan prevented the two sides from reaching a deal. Bevin and Sidqi's effort to finesse the issue of which party would control that territory ultimately failed to appease either the Wafd or British imperialists; as a result, the talks soon came to a bitter end amid a flurry of finger pointing and bickering over who was to blame for their collapse.⁵⁵

Thereafter, the liberal order steadily lost what little support it had managed to retain through the war's conclusion. Over the rest of the decade, it faced constant criticism from the non-parliamentary political movements and took the blame for Egypt's disastrous involvement in the 1948 Arab-Israeli War. It also continued to lose legitimacy as a result of a surge in political violence that included the Muslim Brotherhood's

⁵³ Rogan, *The Arabs*, 210.

⁵⁴ Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 90.

⁵⁵ Botman, "The Liberal Age," 304–5.

assassination of Prime Minister Nuqrashi (1888-1948) in December 1948 and the government's retaliatory killing of al-Banna in February 1949.⁵⁶

The constitutional system might have survived even these serious challenges had the Wafd Party previously built close links between the Egyptian people and their government during the interwar era and thus established the strong civic-political foundation on which the stability and success of representative government depends. Having failed to do so, however, the party found itself utterly incapable of arresting the collapse in support for the constitutional order that occurred in the late 1940s. Thus, by the end of the decade, Egypt's liberal age—begun with such high hopes just a few decades earlier—was nearly at an end.⁵⁷

Syria

Syria's experience under the mandate system was every bit as frustrating as Egypt's or Iraq's. Despite Paris's claims that it was bringing representative government and modernity to the country, it made little effort to establish a durable democratic political system in Syria or to create strong links between the country's people and their government. On the contrary, in an effort to perpetuate its continued control, it imposed a system of direct rule on the mandate designed to deprive Syrians of the opportunity to gain experience in self-government, and it divided the country along confessional lines in order to prevent the emergence of a unified nationalist resistance movement. As a result, when Syria finally won its independence following World War II it lacked the experienced administrators, durable political institutions, and established civic-political culture needed for its government to function effectively or even to maintain political stability.

The Division of Syria

From the start, the French approach to Syria reflected its transparently imperialist goals. With longstanding economic and cultural interests in the region, France had intended since early in the First World War to make Syria a permanent addition to its empire and had viewed the mandate system as little more than a device to appease critics like US President Woodrow Wilson (r. 1913-1921). Unsurprisingly, in light of these goals, Paris moved to satisfy its mandatory obligation to prepare Syria for self-government in a manner that was largely performative. Thus, while High Commissioner General Henri Gouraud (1867-1946) may have scrupulously established representative political institutions and made a point of bringing educated Syrians into the mandatory administration, he also took steps to ensure that those governing structures were powerless and that French advisors rather than Syrian officials made all meaningful decisions. More broadly, Gouraud not only failed to foster any meaningful positive

⁵⁶ Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 91–100.

⁵⁷ Goldschmidt, 91–100.

connection between the Syrian people and its government, but actively worked against the establishment of such links⁵⁸

France also pursued a cynical strategy of divide-and-conquer in Syria in which it split the mandate into a series of statelets. Gouraud's Secretary General and successor as High Commissioner, Robert de Caix (1869-1970) oversaw this effort. He began to put it into effect in 1920 by creating the predominantly Christian state of Lebanon. Later, he established microstates for two comparatively small religious communities: the Alawites, a Shi'i sect that predominated along Syria's northern Mediterranean coast, and the Druze, a syncretic ethno-religious group that occupied the mountainous south. Assuming that these heretofore persecuted religious minorities would be thankful for the autonomy they had received and would therefore be supportive of France's presence in Syria, De Caix opted to rule them indirectly through existing tribal leaders according to a model pioneered by French colonial authorities in Morocco. Finally, he also created two Sunni Arab microstates, Aleppo and Damascus, that France would rule directly.⁵⁹

The division of the mandate promised to pay substantial dividends for the French. Direct rule of Aleppo and Damascus, which colonial officials combined into a single state in 1925 as part of a cost-cutting bid, would permit them to maintain close control of the educated, urban Sunni notables that they viewed as the biggest threat to continued French rule—particularly the former Ottoman bureaucrats and military leaders whom they correctly suspected of attempting to organize an anti-colonialist nationalist movement. Meanwhile, the division of Syria into multiple states would also limit the ability of the urban elite to establish ties with the rural population and would thus forestall any effort to create a mandate-wide anticolonial movement. Finally, as previously noted, the establishment of religiously based microstates promised to create constituencies that were dependent on France for their continued autonomy and would thus assist the French in retaining control of the mandate. In sum, it was a nakedly imperialistic strategy but also a shrewd and well-thought-out one that seemed likely to ensure that Syria remained a French possession for a long time.⁶⁰

The Great Syrian Revolt, 1925-1927

Unfortunately for de Caix's successor, Maurice Sarrail, (1856-1929), just as British historical and anthropological knowledge had proven to be a poor guide to Iraq, so, in two ways, did French assumptions turn out be disastrously off the mark in Syria. First, France's divide-and-conquer approach failed to keep the different confessional groups from organizing together against them. Second, the religious minorities that Paris had assumed would back their rule not only failed to help the French but also took an active part in resisting them. Ironically, in fact, it was one of those groups, the Druze,

⁵⁸ Michael Provence, *The Last Ottoman Generation and the Making of the Modern Middle East* (Cambridge: Cambridge University Press, 2017), 84–87.

⁵⁹ Rogan, *The Arabs*, 225–27.

⁶⁰ Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, 90–91.

that played the lead part in precipitating the most substantial challenge to their control of the mandate: the Great Syrian Revolt of 1927.⁶¹

France's unwillingness to respect Druze autonomy sparked the rebellion. Overseen by an arrogant and domineering French officer, the Druze chafed from the start of the mandate at interference in their internal affairs and at efforts to impose unpopular reforms. Sarrail compounded the situation by refusing to meet with a delegation to hear their grievances in the spring of 1925—a snub that proved to be the final straw for many Druze leaders including the powerful Sultan Pasha al-Atrash (1888-1982). Concluding that his people would never enjoy real autonomy so long as the French remained in Syria, he began organizing a massive revolt in 1924. He understood that his people faced an uphill fight against the French. He thus made contact with the Arab nationalist Dr. Abd al-Rahman Shahbandar (1879-1940), Faysal I's former foreign minister and the leader of the independence-minded People's Party, to see if his group would like to launch a coordinated uprising against the French. Believing that an alliance of Druze and Arabs could defeat France, Shahbandar enthusiastically agreed. Accordingly, drawing on Kemal Mustafa's (r. 1923-1938) successful independence campaign in Turkey as a blueprint, the two men began planning a synchronized rebellion aimed at achieving self-rule and at ending the partition of Syria.⁶²

Known as the Great Syrian Revolt, the uprising constituted a serious threat to France's position in Syria. It began with a series of spectacular Druze victories in the south. In July 1925, Sultan al-Atrash's troops seized control of the regional capital after decimating a French column of three-hundred men. In a triumph that electrified the people of Syria, his forces followed by destroying a huge punitive expedition sent to avenge France's earlier defeat. The revolt soon spread. Waiting until Sarrail had deployed the bulk of his troops against the Druze, Fawzi al-Qawuqji (1890-1977), a one-time Ottoman military officer, opened a new front two months after al-Atrash's first victory by leading an Arab uprising in Hama. Soon forced out of the city by superior French forces, al-Qawuqji and his men relocated to the almost-impenetrably thick orchard groves outside Damascus. From well-concealed bases among the fruit trees, they followed with a series of effective guerrilla attacks against French troops in the capital.⁶³

To the horror of the French, the revolt quickly captured the imagination of the Syrian public. Inspired by al-Atrash and al-Qawuqji's success, people from all of the mandate's confessional groups joined the rebellion by the thousands. By the end of the year, as a result, the uprising had become precisely the kind of threat that the French

⁶¹ John McHugo, *Syria: A History of the Last Hundred Years* (New York: The New Press, 2014), 82–83.

⁶² Rogan, *The Arabs*, 224–30.

⁶³ Michael Provence, *The Great Syrian Revolt and the Rise of Arab Nationalism* (Austin: University of Texas Press, 2005), *passim*.

had so carefully designed their divide-and-conquer policy to prevent: a multi-religious, multi-ethnic, mandate-wide movement dedicated to achieving full independence.⁶⁴

Despite the severity of this challenge, French colonial officials were soon able to restore order to the mandate. Old hands at the game of imperialism, they were neither easily rattled nor shy about using brutal methods to restore order. This they proceeded to do. They began by ordering French artillery to subject Damascus to a punishing, twenty-four-hour bombardment that killed 1,400 people and flattened the old commercial district. When terror attacks such as this one failed to end the uprising, they followed by flooding the mandate with reinforcements in a bid to overwhelm the rebels. The French first used those troops in a successful effort to root al-Qawuqji's guerrillas out of the dense orchards that surrounded Damascus. Having secured the capital, they followed by sending them south to defeat Sultan al-Atrash and restore control of the Druze microstate. This campaign finally broke the rebellion. French mopping-up operations continued until early 1927, but the revolt was effectively over by the end of 1926.⁶⁵

Stalemate

The experience of the Great Syrian Revolt left both French colonialists and Syrian nationalists willing to make concessions. For France, the motivating issue was fiscal in nature. The campaign to end the rebellion had proven so unsustainably expensive that it compelled Paris to abandon its commitment to direct rule and to instead seek a less-costly arrangement patterned after the empire-by-treaty system that Britain had established in Iraq. The mandate authorities consequently issued a series of amnesties and permitted the establishment of a moderate, secular political party of conservative landowners and notables called the National Bloc. They also signaled their willingness to enter negotiations aimed at permitting a greater degree of autonomy in the mandate in exchange for formal recognition of France's strategic and economic interests.⁶⁶

For the elites who dominated the National Bloc, in contrast, the primary factor driving moderation was the vast imbalance in military power that existed in the mandate. Making clear France's overwhelming military advantage, the suppression of the rebellion had persuaded Syrian nationalists to abandon the effort to achieve independence through the use of force in favor of the more-limited objective of greater autonomy through negotiations. Accordingly, the National Bloc indicated that it was willing to take part in talks aimed at achieving a system of indirect rule that it referred to as "honorable cooperation." Thus, with the two sides now seeking similar goals, the

⁶⁴ McHugo, *Syria*, 85–86.

⁶⁵ James Barr, *A Line in the Sand: The Anglo-French Struggle for the Middle East, 1914-1948* (New York: W. W. Norton & Company, 2011), 129–38.

⁶⁶ Fieldhouse, *Western Imperialism in the Middle East 1914-1958*, 290–97.

stage was set for discussions aimed at formalizing French interests in Syria in exchange for nominal independence.⁶⁷

Begun with high hopes, the subsequent negotiations proved frustratingly contentious and failed to produce an agreement. The primary stumbling block was the question of Syrian unity. The French were insistent that the mandate remain divided into multiple statelets; the National Bloc was equally adamant that it be reunified. With the two sides unable to find a way to finesse their differences on this point, the talks conducted in the early 1930s collapsed in failure. Subsequent efforts to restart the negotiations likewise foundered over the parties' continued inability to find common ground.⁶⁸

By the mid 1930s, frustration was once more boiling over in Syria. In response, the National Bloc shifted to a more aggressive approach designed to target the French in the pocketbook. In January 1936, it initiated a non-violent pressure campaign that came to be known as the Syrian General Strike. Enjoying near universal support in the mandate, it aimed to shut down the entire Syrian economy in order to stop the flow of tax revenue that underwrote French control of the mandate. This new strategy worked. Confronted with the unpalatable prospect of having to shoulder the cost of the mandate during a time of economic contraction, Paris capitulated and agreed to hold serious negotiations in 1936.⁶⁹

Prospects looked good for an agreement. With war clouds gathering in Europe, the government of leftist Prime Minister Leon Blum (1872-1950) was eager to ensure stability in the mandate; accordingly, it signaled its willingness to compromise on the question of Syrian unity. With the main stumbling block removed, the ensuing negotiations were productive and soon produced a draft agreement that satisfied both sides' minimum conditions. Patterned loosely after the Anglo-Iraqi Treaty, it granted Paris the right to use two strategic airbases for twenty-five years and required Damascus to back France during time of war; in exchange, the French government agreed to reunify Syria and to recognize its independence. Delighted to have finally secured an agreement, the Syrian parliament enthusiastically approved it in December 1936. Independence at last appeared to be at hand.⁷⁰

To the frustration of the National Bloc, however, even the limited sovereignty envisaged in the treaty remained tantalizingly out of reach. Determined to maintain control of the mandate, France's influential colonial lobby succeeded in preventing ratification of the treaty in the French senate. Making matters worse, in 1938, the French government agreed to cede the coastal region of Alexandretta in northwest Syria to Turkey in a bid to keep Ankara from allying with Nazi Germany. It was a bitter pill for Syrian nationalists, and it seemed to make emphatically clear that the National

⁶⁷ Rogan, *The Arabs*, 240.

⁶⁸ Provence, *Great Syrian Revolt*, 228–30.

⁶⁹ Provence, 228–30.

⁷⁰ Fieldhouse, *Western Imperialism in the Middle East 1914-1958*, 265–72.

Bloc's strategy of moderation and compromise had failed to produce results. Indeed, the truth was that Syria was no closer to independence and unity in 1938 than it had been at the end of the Great Syrian Revolt a decade earlier.⁷¹

World War II and Independence

In the end, it was neither further negotiations nor additional protests that permitted Syria to gain its independence. Instead, it was France's defeat and occupation in World War II that finally brought the mandate to a close. Even then, independence did not come quickly or easily. Indeed, during the first few years of the war, Syria's position was essentially unchanged. Controlled by a mandatory government that was loyal to the collaborationist Vichy regime, it remained a de facto colony of France just as it had been prior to the war's outbreak.⁷²

Only in the spring of 1941 did the mandate's circumstances finally begin to change. They did so as a result of the shifting military balance. Determined to secure the Middle East at a time when German and Italian forces operating out of Libya were threatening the Suez Canal, British soldiers and Free French troops led by future French President Charles de Gaulle (r. 1959-1970) invaded the mandate and quickly overwhelmed its Vichy defenders. It was an important victory for the Allies. Along with the near-simultaneous military operation against the Golden Square in Iraq, it solidified their uncontested control of the Middle East and thus ensured that the region's oil and strategic bases remained in their hands. If the Allies were thrilled with the outcome of the operation, however, the Syrians were considerably more ambivalent. While they were pleased to be out from under the thumb of the Vichy government, they were unhappy that they were now under Free French rule—particularly after de Gaulle declared that his government intended to retain control of Syria following the war.⁷³

At that point, however, France and Britain's longstanding imperial competition resurfaced to the benefit of Syrian nationalists. Eager to push its colonial rival out of the Middle East once and for all, London began to apply pressure to de Gaulle to compel him to relinquish the mandate. With France still under German occupation, the frustrated French leader was in no position to argue and quickly—if begrudgingly—agreed to London's demands. Once again, as a result, genuine self-government seemed to be imminent.⁷⁴

Syria managed to secure its independence this time, but only after French colonialists mounted a desperate—and violent—last-minute bid to hang on to the territory. That effort began in the war's waning days. Determined to retain a dominant position in Syria, Paris refused to withdraw its remaining troops from the country in the

⁷¹ McHugo, *Syria*, 97–99.

⁷² Rogan, *The Arabs*, 240–41.

⁷³ Rogan, 241.

⁷⁴ Barr, *A Line in the Sand: The Anglo-French Struggle for the Middle East, 1914-1948*, 202–21.

spring of 1945 in a transparent effort to blackmail the nationalists into signing a new treaty recognizing France's long-term interests in the mandate. Enraged by this eleventh-hour attempt to compromise their country's independence, the Syrian people responded with a series of massive demonstrations against France's continued presence. Just as it had in 1920 and 1925, Paris replied with violence. Bombarding Damascus with artillery in May 1945, French troops killed over four-hundred Syrians and leveled large parts of the city. This time, however, the use of force proved to be France's swan song in Syria. Looking to bolster its reputation among the Arabs, London forced Paris to end its terror campaign in 1945 and followed by compelling it to remove its last troops in 1946. As a result, after twenty-six-years of French rule, Syria had finally become an independent state.⁷⁵

France's rule of Syria may have been comparatively brief, but it nonetheless powerfully—and adversely—shaped the country's development. To a large degree, the mandate's legacy was a function not of what the French had done, but of what they had failed to do. Even more so than Britain in Iraq, Paris had made no more than cosmetic attempts to meet its obligation as Syria's mandatory power to help prepare the country for independence. As a result, at the time it became independent, Syria had neither the experienced administrators nor the durable political institutions that modern states require; more importantly, it lacked the civic-political culture that is so essential for the effective functioning of a representative polity.⁷⁶

France did not merely fail to foster the development of robust political institutions or strong connections between people and state, moreover. Instead, in its efforts to render its control permanent, it also actively saddled the mandate with a destructive political inheritance that all-but ensured continued volatility and despotic rule. This legacy manifested itself in three ways. First, Paris's refusal to ratify the 1936 treaty that the National Bloc had negotiated thoroughly discredited both the political moderates and the broader democratic system that they represented. As a result, Syria's weak parliamentary government would enter independence burdened by a lack of legitimacy—a condition, as we shall see, that would pave the way for its replacement by a succession of autocratic governments. Second, French reliance on martial law and the naked use of force to ensure its control during the mandate period provided those undemocratic regimes with a brutal, ready-made template that they would make liberal use of to crush any opposition. Finally, the creation of a separate Lebanese state following World War I embittered nationalists who believed that it was rightly a part of Greater Syria and would later lead Damascus to intervene in Lebanon to the detriment of both countries.⁷⁷

⁷⁵ Barr, 283–195.

⁷⁶ Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, 271–72.

⁷⁷ Provence, 271–72.

Lebanon

Lebanon would avoid the intense unrest that plagued Syria during the interwar era. Dominated by the Christian Maronites, it instead enjoyed two decades of relative stability under French administration. To many observers at the time, its cohesion seemed to be a direct function of the mandatory government's decision to give it a peculiar political structure—the confessional system—in which the different religious groups divided power in proportion to their share of the population. Seemingly successful, that arrangement would become sufficiently entrenched during the mandate period that it would come to form the basis of the system of government that the Lebanese adopted upon gaining independence in 1943.

Mandatory Lebanon

That Lebanon emerged as a separate state was a result of both French colonial ambitions and the aspirations of the Maronite Christian community that dominated the landlocked region of Mount Lebanon. Unlike Syrian elites who universally opposed French rule, the Maronite Christians welcomed Paris's involvement in the region. They had two reasons for doing so. First, they had a history of close cultural, religious, and economic ties with France stretching back to the nineteenth century. Second, and more importantly, Maronite leaders had territorial and political aspirations of their own. Eager to break free of Muslim rule, they hoped to take advantage of the fluid postwar situation to establish an independent, Maronite-dominated "Greater Lebanon" expanded to its "natural boundaries" through the inclusion of the coastal strip to the west and the fertile Bekaa Valley to the east.⁷⁸

As they were aware, however, the enlarged state that they sought would not be universally popular among the people of the region. Uninterested in living in a Maronite-dominated country, other confessional groups—not to mention Syrian nationalists—were certain to oppose its creation and to work aggressively to prevent its establishment. How could the Maronites overcome that resistance? How could they secure both their independence and the additional territory they would need if they were to have a functional state? The answer, they believed, lay in Paris. By making clear to the French government that they would be happy to accept its suzerainty, they could secure its help in achieving their dream of an enlarged Maronite state.⁷⁹

Paris was more than happy to comply. It did so not for altruistic reasons, however, but instead in furtherance of France's imperial goals. Experienced colonialists, the French had long used the Maronites as a stalking horse for their interests in the region and saw the establishment of a separate, Christian-dominated Lebanese state both as a way of dividing any potential indigenous opposition to its rule and as an insurance policy that could safeguard its presence in the Levant in the event that Syria became independent. Accordingly, High Commissioner Gouraud enthusiastically

⁷⁸ Rogan, *The Arabs*, 212–16.

⁷⁹ Rogan, 212–16.

announced in September 1920, the establishment of Greater Lebanon and the inclusion within it of the additional territory that the Maronites sought.⁸⁰

The creation of the new Lebanese state posed a difficult demographic challenge for the Maronites, however. While they amounted to 76 percent of the population of Mount Lebanon, they constituted only a negligible minority in the surrounding areas. As a result, while the territory they hoped to add to Mount Lebanon was essential for the establishment of a functional independent state, its inclusion also threatened to dilute the Maronite share of the population to the point that they would no longer be either numerically or politically dominant.⁸¹

Determined to push forward with a separate Lebanese state, France responded to this challenge in two ways. First, French officials drew the new state's borders with great care in order to ensure both that the Maronites continued to constitute a majority of Lebanon's population and that the non-Maronites remained sufficiently divided between Sunni, Shi'a, Druze, and other groups such that no single religious community had adequate numbers to contest their dominant position. In this effort they were quite successful: while substantially larger than the Maronites' traditional power base of Mount Lebanon, Greater Lebanon nonetheless remained 58 percent Maronite with the balance so divided that none of the other confessional groups was strong enough to challenge their preeminent political position. Second, Paris moved to divide the Lebanese through the institutionalization of the confessional system: a political arrangement with antecedents in the old Ottoman *millet* system that apportioned representation among the country's confessional groups in proportion to their size. Accordingly, rather than having the seventeen seats on the elected Representative Council open to all people, the French instead divided the positions among the major religious communities in proportion to each groups' share of the population: ten for the Maronites, four for the Sunnis, two for the Shi'a, and one for the Druze. This policy, too, proved enormously beneficial to the French. By keeping the Lebanese divided, it succeeded in preventing the emergence of a mandate-wide, anti-French nationalist movement. As a result, in sharp contrast to its experience in neighboring Syria, Paris faced almost no organized opposition to its rule in Lebanon during the mandate era⁸²

What these policies could not do, however, was to offset the precipitous drop in French power that had occurred during World War II—a decline that would permit Lebanon to gain its independence. As in Syria, the beginning of the end for France's control of the mandate came when British and Free French troops seized the country from its Vichy rulers in 1941. Just as it did with regard to Syria, London followed the campaign by compelling de Gaulle to promise to grant the mandate its independence. In no position to refuse thanks to his dependence on Britain, the Free French leader

⁸⁰ Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, 89–90.

⁸¹ Rogan, *The Arabs*, 212–16.

⁸² Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, 121–22, 193.

quickly complied. He was being disingenuous in doing so, however. Despite public assurances to the contrary, he and his colonialist supporters remained steadfastly committed to keeping Lebanon within Paris's sphere and schemed to restore full French control. Given this goal, they grew alarmed in late 1943 when the Lebanese government announced that it intended to strike from its constitution all provisions granting France special powers and privileges in the country. Fearing that those revisions would end any chance of restoring French rule, de Gaulle responded by ordering his soldiers to arrest President Beshara al-Khoury (r. 1943-1952).⁸³

The crackdown proved to be France's last hurrah in Lebanon. Massive demonstrations in Beirut and stiff pressure from Great Britain quickly compelled Paris to capitulate and to free al-Khoury. France managed to preserve limited influence in Lebanon for a few more years, but al-Khoury's release from custody effectively ended any chance that it might reassert control. Thus, by the end of 1943, Lebanon had secured its full independence.⁸⁴

The National Pact

Having achieved independence, the Lebanese now faced the task of devising a new system of government to replace the mandatory one. Doing so would prove to be tricky in light of the fact that the confessional system through which the French had ruled had created powerful vested interests that were disinclined to cede power to a central state. In response, the leaders of the different sectarian groups negotiated an informal, extraconstitutional power-sharing arrangement called the National Pact that was, in essence, a modified version of the confessional system through which the French had ruled. Designed to ensure that each of the major religious groups enjoyed representation in proportion to its share of the population in 1932—the year of Lebanon's most recent census—it allocated seats in parliament according to a formula that guaranteed six positions for the numerically dominant Maronites to every five held by representatives of the other confessional groups. In similar fashion, the National Pact divided key positions in the government along religious lines. It stipulated that the president of Lebanon would always be a Maronite, the prime minister a Sunni, and the speaker of parliament a Shi'i; other cabinet positions were reserved for smaller confessional groups such as the Druze and Greek Orthodox Christians.⁸⁵

At first, the National Pact seemed to be a durable and effective agreement. Under it, Lebanon enjoyed substantial economic growth and a degree of stability that stood in sharp contrast to the experience of Syria and many other Arab states. Indeed, through the mid 1950s, Lebanon enjoyed a reputation as a stable and pragmatic state with a capital so cultured and cosmopolitan it came to be known as “the Paris of the Middle East.” Over the long term, however, the National Pact revealed itself to be

⁸³ Rogan, *The Arabs*, 242–44.

⁸⁴ Barr, *A Line in the Sand: The Anglo-French Struggle for the Middle East, 1914-1948*, 244–152.

⁸⁵ Rogan, *The Arabs*, 240–42.

enormously problematic for the country. Reinforcing narrow, particularist religious identities, it inhibited the development of a strong civic-political culture that could bind the different confessional groups together into a cohesive whole; in so doing, it institutionalized the division of the country along confessional lines and sowed the seeds of future confessional conflict that would result in political strife in the 1950s and the outbreak of a long and debilitating civil war in the 1970s. Thus, the French decision to apportion political power between confessional groups—an action that was designed, at root, to achieve Paris's colonial ends—saddled Lebanon with a legacy of conflict and division that has long outlasted its brief time as a French-controlled mandate.⁸⁶

Transjordan

The most stable of the Middle Eastern mandates was the Emirate of Transjordan, so named because it lies on the east side of the Jordan River. About the size of Indiana, the territory came under the rule of Faysal I's brother, the Emir Abdallah (r. 1921-1951), in 1921 before formally becoming a British mandate in 1922. Thanks to Abdallah's skilled leadership, this nearly landlocked territory would overcome its peculiar origins and a number of daunting early challenges to develop into a stable, if undemocratic, state.

Transjordan's relative success is surprising given its accidental—and complicated—emergence as a state. From the end of the First World War until the French put him to flight in July 1920, Faysal I had administered it as part of his Arab Kingdom of Syria. As we have seen, Paris had followed his ejection by establishing a mandatory regime in Syria; however, it could not incorporate Transjordan into it since the territory lay within the region that the Sykes-Picot Agreement had assigned to Britain. As a consequence, after July 1920, Transjordan became a no-man's land that lacked a recognized government. It remained so until the spring of 1921 when, as part of the "Sherifian Solution," the Cairo Conference acknowledged Abdallah, who had occupied the territory in the fall of 1920, as the emir of Transjordan. To give legal sanction to this move, Britain appended the territory to the Mandate for Palestine; however, it never intended for the two regions to be unified or for Transjordan to be part of a Jewish homeland. London made this clear in 1922 when it formally announced that the land east of the Jordan River would neither be subject to the Balfour Declaration nor politically connected to Palestine.⁸⁷

Transjordan's eventual emergence as a relatively stable state is also remarkable in light of the significant issues that Abdallah faced in the early 1920s. Some of those problems were long-term ones. With a population of only 250,000 people, Transjordan was uneasily divided between two groups—the townspeople and the bedouin—and was

⁸⁶ Quote from Jeremy Salt, *The Unmaking of the Middle East: A History of Western Disorder in Arab Lands* (University of California Press, 2008), 186; Rogan, *The Arabs*, 241–42.

⁸⁷ Fieldhouse, *Western Imperialism in the Middle East 1914-1958*, 223–25; quote from Rogan, *The Arabs*, 183.

desperately short of raw materials including ones as basic as water. As a consequence, it remained in dire economic circumstances even in comparison to its impoverished neighbors. Transjordan also faced a number of more transitory but still serious problems including rural tax revolts in 1921 and 1922, raids by Ibn Saud's (r. 1902-1953) Ikhwan warriors, and efforts by Zionists and pro-Zionist British officials to depose Abdallah and annex his emirate to Palestine.⁸⁸

On top of those challenges, Abdallah also confronted the fact that native Transjordanians and British officials alike were unhappy about the influence in his government of the *Istiqbalis*: former officials from Faysal I's regime in Syria who constituted Abdallah's closest confidants. Urban Transjordanians opposed the *Istiqbalis* because they viewed them as an alien ruling class that monopolized the best positions in the bureaucracy; London distrusted them because it perceived them to be troublemakers who might embroil Britain in a crisis with France by raiding Syria. London's worries about the *Istiqbalis* rose steadily in the early 1920s as their influence over Abdallah seemed to grow. Indeed, by early 1924, fears that the *Istiqbalis* were going to drag Britain into a costly conflict with France had become so acute that British officials began to give serious consideration to the idea of replacing Abdallah as emir.⁸⁹

The emir was a more effective leader than the British had initially assumed, however, and he ultimately proved that he could overcome the issues that the mandate faced. Central to his success was his ability to secure good relations with the nomadic tribes. Cultivating close personal connections with the tribal shaykhs, he was able to bind the militarily powerful bedouins to his government. In turn, his links with the tribal leaders led key British officials to come to grasp his value in maintaining order in the mandate. As a result, rather than deposing him, they gave Abdallah an ultimatum in 1924 requiring him to agree to abide by a series of conditions including, most importantly, one demanding that he expel the *Istiqbalis* from Jordan within five days. Intent on maintaining his grip on the throne, Abdallah agreed—a decision that may have cost him the backing of his old supporters but one that also improved his standing with both native Transjordanians and, more importantly, the British.⁹⁰

Abdallah's subsequent success in maintaining order in Transjordan and his continued compliance with British wishes solidified his position. His standing rose so high, in fact, that London granted his principality official recognition in 1928 through a treaty that formalized the bilateral relationship. It was a good deal for Abdallah, if perhaps not for his subjects. At the cost of ceding control of Transjordan's foreign relations to London, the agreement committed Britain to defend the country from

⁸⁸ Maan Abu Nowar, *The Development of Trans-Jordan 1929-1939: A History of the Hashemite Kingdom of Jordan* (Reading, UK: Ithaca Press, 2006), 4–5; Philip Robins, *A History of Jordan*, Second Edition (Cambridge: Cambridge University Press, 2019), 23–31.

⁸⁹ Robins, *A History of Jordan*, 23–31.

⁹⁰ Yoav Alon, *The Making of Jordan: Tribes, Colonialism and the Modern State* (New York: I. B. Tauris, 2007), 1–7; Rogan, *The Arabs*, 184–86.

external attack, confirmed that Transjordan would not be absorbed into Palestine, and ceded to Abdallah's government the power to legislate. Importantly, it also allocated funds to strengthen the emirate's military—the famous Arab Legion. Under the command of a British officer, John Glubb (1897-1986), it soon became the most powerful military force in the Arab world.⁹¹

Abdallah's continued willingness to back Britain and his success in sustaining internal stability paid further benefits following World War II. In response to Transjordan's steadfast support during the conflict, London agreed to grant it full independence in 1946—though it remained in an informal colonial relationship with Britain. Thus, as a result of Abdallah's adroit management of the British and his establishment of close ties with the powerful tribes, Transjordan overcame its difficult beginnings to become an independent and comparatively stable state.⁹²

The Historical Debate: The Last Ottoman Generation

Scholars have long viewed the interwar era as a critical period in shaping the development of the Middle East. Not only did the imperial powers establish the region's modern borders between 1920 and 1938, but nearly all of its current-day inter- and intra-national conflicts have their roots in the events of that period. In keeping with this understanding, historians and nationalists alike have long interpreted instances of armed resistance to European rule in the interwar Arab states such as the Battle of Maysalun as the seminal events—indeed, as the defining myths—in the establishment of particularist nationalisms by peoples expressing their newly roused sense of national identity for the first time.⁹³

Recently however, the historian Michael Provence has challenged this understanding of the revolts and anticolonial movements of the 1920s and 1930s. He rejects the view that the early challenges to European hegemony in the interwar Arab states constituted the first stirrings of nationalism by people exhibiting “recently awakened national consciousness.” On the contrary, he contends that those resistance efforts were components of a unified movement orchestrated by the final generation of people who self-identified as Ottoman and whose intent was not to establish new nations but instead to replace the peace settlement and the division of the region with a unified, modern, and autocratic state laid out along lines similar to the late Ottoman Empire.⁹⁴

Provence maintains that the central figures in this effort were men born in the 1880s and 1890s who graduated from the Ottoman Empire's elite military academies

⁹¹ Nowar, *The Development of Trans-Jordan 1929-1939*, 9–13.

⁹² Rogan, *The Arabs*, 186–87.

⁹³ See for example Robert Fisk, *The Great War for Civilisation: The Conquest of the Middle East* (New York: Vintage, 2005), 813.

⁹⁴ Michael Provence, “Ottoman Modernity, Colonialism, and Insurgency in the Interwar Arab Middle East,” *International Journal of Middle East Studies* 43, no. 2 (2011): 207.

and who served in its armies during World War I. Imbued with the militarist, authoritarian-modernist ideas of the late empire and inspired by Kemal Mustafa's successful effort to overturn the postwar settlement in Anatolia, these officers organized a broad resistance movement aimed at ejecting the European imperial powers from Iraq, Syria, Lebanon, and, especially, Palestine. Theirs was an avowedly transnational effort, one that rejected the borders that the Europeans had applied to the Middle East following World War I and that sought to free the entire region of Western control. Accordingly, movement leaders did not limit their resistance efforts to the states in which they resided, but instead crossed frontiers to fight against the colonial powers. The career of Fawzi al-Qawuqji illustrates this point. A graduate of the Ottoman imperial military academy, he participated in the Battle of Maysalun in 1920, helped start the Great Syrian Revolt in 1925, led a militia force during the Arab Revolt in Palestine in the 1930s, fought against the British in Iraq in 1941, and battled Zionists in 1948. As Provence notes, that someone with al-Qawuqji's background took an active role in these campaigns was not unusual. Indeed, as he points out, former Ottoman officers led *all* of the resistance efforts in the Middle East between 1918 and 1948 and did not typically limit their efforts to a single Arab state.⁹⁵

Ultimately, the officers who constituted what Provence calls "the last Ottoman generation" failed in their effort to overturn the peace settlement. In fact, despite multiple revolts and insurgencies, they were entirely unsuccessful in their attempts to break Western control of the Middle East. Only as a result of World War II, imperial exhaustion, and Great Power rivalry would the mandates gain full independence, by which point those with an Ottomanist perspective had largely passed from the scene. Still, the Ottoman-trained officers had an enormous influence over the development of the postwar Arab states. While they had failed to push out the imperialists, they nonetheless succeeded in passing the militaristic, undemocratic "authoritarian modernity" of the late-Ottoman Empire on to the next generation of nationalist military leaders—the very men who would rule many of the Arab states after the Second World War.⁹⁶

Conclusion

The undemocratic legacy of the last Ottoman generation was far-from obvious in 1945, however. Instead, to the casual observer, the Arab states that had experienced European rule in the 1920s and 1930s seemed poised to emerge from World War II as relatively stable, civilian-led countries. Lebanon appeared to have arrived at a formula for sharing power among its confessional groups, for example, while Transjordan had found stability by building close ties between the monarch and the nomadic tribes. More

⁹⁵ Provence, 207; Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, passim.

⁹⁶ Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, 3–7.

importantly, the larger, more populous states of Egypt, Iraq, and Syria all had democratic governments dominated by moderate, secular, political parties.

As we have seen, however, the appearance of stability in Egypt and the former mandates was a mirage. With the possible exception of Transjordan, those states had failed to develop either close links with their people or strong civic-political cultures. By the late 1940s, as a result, none of their governments and dominant political parties enjoyed meaningful legitimacy or broad support.

In a calmer era, the moderate elites who dominated those states might have managed to hold on to power and may have even been able to build the civic-political culture on which a state's legitimacy ultimately rests. Unfortunately for them, however, they did not enjoy the luxury of stable times. Instead, those states emerged from colonial dominance just as the contest between Arabs and Zionists for control of Palestine was reaching its culmination in 1948. Already on shaky ground, their inability to stop the establishment of the Jewish state completely discredited the moderate elites and the regimes they led and opened the door for the pan-Arabist officers who succeeded Provence's last Ottoman generation to seize power. It is to the events in Palestine that would set the stage for their rise to dominance that we shall now turn.

Chapter Thirteen, The Establishment of Israel, 1918-1948

To most people enduring the horrors of World War I, Britain's conquest of Palestine and proclamation of the Balfour Declaration in late 1917 were matters of only marginal consequence. The capture of Jerusalem in December 1917 was a boost to Allied morale, of course, and the Balfour Declaration did commit Britain to support the establishment of a Jewish homeland in Palestine. Neither seemed important, however, in comparison to the dramatic revolution then convulsing Russia or to the titanic military struggle rapidly approaching its culmination in France.

In fact, the wartime events related to Palestine would prove to be of enormous regional and even global significance owing to the fact that they permitted the Zionists to make tangible progress for the first time in their heretofore halting effort to create a Jewish state. In doing so, the Balfour Declaration and the conquest of Palestine would set in motion a bitter, zero-sum struggle between two nationalist movements, one Jewish and one Palestinian, over which would control the land that lay between the Jordan River and the Mediterranean Sea. That struggle would end in a Zionist victory in a conflict that Israelis call the War of Independence and that Palestinians refer to as *al-Nakba*, or the catastrophe. The result would be the establishment of the state of Israel, the permanent exile of the overwhelming majority of the Palestinian population, and the birth of the region's most enduring and damaging conflict.

Establishing the Mandate, 1918-1922

Zionism Triumphant?

The events of 1917 electrified the Zionists. Not only had Britain committed itself to supporting the establishment of a Jewish state—as leaders like Chaim Weizmann (1874-1952) chose to interpret the Balfour Declaration's call for a "national home"—but, critically, it had also taken possession of the territory in which the Zionists hoped to found that state. Subsequent events further boosted their optimism. At the San Remo Conference in April 1920, London secured Paris and Rome's formal endorsement of a British mandate over Palestine and thus ensured that it had the legal standing to put the Balfour Declaration into effect. To top it off, London announced that the prominent British Zionist Herbert Samuel (1870-1963) would serve as the territory's first high commissioner.¹

The Zionists' postwar euphoria peaked in 1919 and 1920. With Britain seemingly firm in its commitment to their movement, Zionist leaders felt confident in pressing aggressively for the rapid realization of their maximalist goals. Lobbying to secure

¹ Benny Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920* (New York: Vintage Books, 2001), 75, 78, 80, 97; quote from "The Balfour Declaration, November 2, 1917," accessed May 29, 2020, https://avalon.law.yale.edu/20th_century/balfour.asp.

international support for an expansive understanding of the new state's borders, for example, future Israeli Prime Minister David Ben Gurion (r. 1848-1954, 1955-1963) argued that it should encompass the territory controlled by the ancient Hebrew kingdom at its height; he thus called for it to include land on both sides of the Jordan River as well as parts of modern-day Lebanon, Syria, and Egypt. Even the cautious Weizmann, the driving force behind the Balfour Declaration, could not hide his excitement at the prospect of the rapid establishment of a Jewish state throughout the entirety of the mandate. Though he continued to use the amorphous term 'homeland' rather than the more inflammatory word 'state' while attending the Paris Peace Conference in 1919, he was so confident of the impending success of Zionism that he felt comfortable publicly expressing his desire to "make Palestine as Jewish as England is English."²

The Arab Riots of 1920 and 1921

Ultimately, Zionist confidence following World War I was badly misplaced. Not only would the establishment of a Jewish state prove to be far from imminent, but Britain would begin to back away from its commitment to the Balfour Declaration just a few years after the conflict had ended. Indeed, more than a quarter century would pass before the movement finally succeeded in establishing a Jewish state in Palestine. What accounted for Zionism's postwar failure? Why was it unable at that time to achieve a goal that had seemed so tantalizingly close?

Ironically, the movement's inability to secure a Jewish state at that time was a function not of its weakness but instead of its strength. With every success, Zionism engendered a growing fear among members of the indigenous, Arab-Palestinian population that a powerful and alien minority was in the process of displacing them from the land on which they and countless generations of their ancestors had lived. The issuance of the Balfour Declaration, intemperate statements by Zionist leaders, rising Jewish land purchases and the concomitant displacement of Arab peasants, or fellahin, and, especially, immigration—18,500 Jewish people arrived in the territory between 1919 and 1921—stoked those concerns. In response, the Palestinians began to espouse an incipient, proto nationalism. The formation of Muslim-Christian associations throughout the territory in 1918 and the establishment of the First Palestinian Arab Congress in 1919 reflected this new outlook. Expressing widespread opposition to Zionism, these organizations called on Britain to cede the entirety of Palestine to Faysal I's (1883-1933) Arab Kingdom of Syria. They also made clear that they viewed the contest with the Zionists as an existential, winner-take-all fight. As one Arab bluntly

² Weizmann quoted in Avi Shlaim, *The Iron Wall: Israel and the Arab World, Revised and Expanded* (New York: W. W. Norton, 2014), 10; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 75, 81.

told the fact-finding King-Crane Commission, “[w]e will push the Zionists into the sea—or they will send us back into the desert.”³

Palestinian frustration with the British and fear of displacement finally boiled over in early 1920. Egged on by mayor Musa Kazi al-Husayni (1853-1934), a crowd of Arabs began attacking Jewish people and looting Jewish-owned shops in Jerusalem’s Old City following a Muslim religious procession on Easter Sunday. The rioting quickly got out of hand. Indeed, it was so intense and involved so many people that the British lost control of Jerusalem for three days and only managed to restore order after they declared martial law and deployed troops. By the time the soldiers had suppressed the riot, five Jewish people had been killed and 216 had been injured while four Arabs had died and twenty-three had been wounded.⁴

Palestinian fear of Zionism and mounting anger toward the British produced a second round of intercommunal clashes the following year. Stories of rising immigration and Jewish land purchases had left the Arabs primed for a fight; ironically, however, it was a May Day brawl between rival Jewish communist and socialist groups in Jaffa that finally sparked the return of intercommunal violence. Initially, the fighting remained confined to the leftist Zionist groups. Once it spilled over into a nearby Arab neighborhood, however, it quickly transformed into a vicious pogrom in which Arab mobs attacked Jewish people and shops in both Jaffa and the surrounding countryside. The British once more responded aggressively to the rioting. Declaring a state of emergency, they sent sizeable contingents of troops and armored cars into battle against armed bands of Arabs and even dropped bombs on villages. When the smoke finally cleared, British officials were shocked by the scale of the Jaffa riots: forty-seven Jewish people had died and one-hundred-and-forty-six had been wounded, while forty-eight Arabs had been killed and seventy-three injured.⁵

Dominated by the Labor Zionists—so called because they came from the socialist parties and labor unions—the Zionist leadership found the riots of 1920 and, especially, 1921 deeply concerning. They questioned the capacity and willingness of the British to protect Jewish people in the mandate, wondered whether their movement would be able to overcome Arab resistance, and worried that the Yishuv, the Jewish community in Palestine, was vulnerable to further attacks. Above all, they were concerned that the scale of the violence might lead either prospective Jewish immigrants or the British to arrive at the truth: that the Zionists faced a coherent, proto-nationalist Palestinian opposition. If Jewish Europeans drew such a conclusion,

³ Eugene Rogan, *The Arabs: A History – Revised and Updated Edition* (New York: Basic Books, 2012), 198; quote from Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 89–91.

⁴ Tom Segev, *One Palestine, Complete: Jews and Arabs Under the British Mandate* (New York: Metropolitan Books, 1999), 127–44.

⁵ Michael Provence, *The Last Ottoman Generation and the Making of the Modern Middle East* (Cambridge: Cambridge University Press, 2017), 127–28; Segev, *One Palestine, Complete: Jews and Arabs Under the British Mandate*, 173–201.

immigration would slow to a trickle and the Zionist share of the population in Palestine would remain tiny; if the British did, London would almost certainly respond by weakening or even abandoning its commitment to the Balfour Declaration in order to appease the numerically superior Arabs. In either case, the goal of Zionism—the establishment of a Jewish state in Palestine—would be doomed to failure.⁶

To counter such conclusions, the Zionists deployed two arguments. First, contending that the Arab and Jewish peoples shared common interests, they rejected the idea that ethnic tensions caused the riots and instead located the roots of the violence in class conflict between the exploited fellahin, and the Palestinian notables who owned the land they worked. The latter, the Zionists continued, had stirred up the otherwise apolitical Arab masses with lurid tales of Jewish aggression and land purchases in a successful effort to divert the peasants' anger—creating, in the process, a misplaced fury that soon manifested itself in the form of anti-Jewish rioting. Second, the Labor Zionists not only disputed the idea that Zionism was a threat to the well-being of the Arab peasants but argued that the continued inflow of Jewish immigrants and capital would transform Palestine's feudal economy into a dynamic, modern one that would dramatically lift Arab living standards. Once the economic benefits of Zionism began to manifest themselves, they added, any lingering resentment among the Palestinian masses toward the movement would rapidly dissipate as they refocused their anger on the landlords who had for so long exploited them.⁷

The Churchill White Paper of 1922

These arguments may have reassured Jewish Europeans thinking about immigrating to Palestine, but they failed to move British officials. Administrators in the mandate possessed enough experience managing colonies to tell the difference between transitory unrest and serious problems, and what they saw in Palestine in 1920 and 1921 deeply concerned them. Indeed, the sheer scale of the violence suggested not only that Palestinian opposition to the Balfour Declaration was far more serious than the Zionists claimed but that the disorder in the mandate was likely to continue unless the British addressed at least some of the Arabs' objections.⁸

The riots also laid bare for the first time the existence of significant tension between policymakers in London and those serving in Palestine over the question of which community the British should favor. Officials in the capital lamented the recent violence in Palestine but, insulated by distance from the disorder and in close contact with key Zionist leaders, remained committed to fulfilling Britain's promise to establish a Jewish homeland in the territory. To administrators in Jerusalem, in contrast, the Jaffa riots had exposed the contradictions inherent in London's simultaneous promise to

⁶ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 104.

⁷ David Fromkin, *A Peace to End All Peace: The Fall of the Ottoman Empire and the Creation of the Modern Middle East* (New York: Avon Books, 1989), 516–27; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 104–5.

⁸ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 102–4.

establish a national homeland for the Jews and its pledge to uphold the rights of the “existing non-Jewish communities.” They argued that since there was simply no feasible way that Britain could meet both of those obligations and since the Palestinians constituted the vast majority of the territory’s population, the mandatory government should pursue policies that favored the Arabs. In keeping with this line of thinking, officials in Jerusalem moved to appease the Palestinians immediately after the Jaffa riots by placing temporary restrictions on immigration and, more importantly by pressing London to back away from the commitment it had made in the Balfour Declaration to help the Zionists establish a Jewish national home in Palestine.⁹

Reconciling these policy decisions fell to Secretary of State for the Colonies Winston Churchill (1874-1965), who did so in 1922 in a statement colloquially referred to as the Churchill White Paper. It included three main provisions designed to ease Arab concerns. First, while the document restated Britain’s commitment to help the Zionists establish a national home in Palestine and confirmed that Jewish people had a right to immigrate to the territory, it also reassured the Arabs that the Balfour Declaration did not mean that all of Palestine “should be converted into a Jewish national home, but [only] that such a home should be founded in Palestine.” Second, it put restrictions on immigration for the first time by limiting the number of Jewish people who could move to the mandate to the “economic capacity of the country.” Finally, the paper formally detached Transjordan from Palestine.¹⁰

The British moved quickly to put the Churchill White Paper into effect. To ensure Zionist interests, London had the text of the Balfour Declaration formally incorporated into the mandate charter that the League of Nations issued in July 1922, and it arranged for the world body to accredit an important Yishuv organization, the Zionist Executive, as the official representative of the Jewish community in Palestine. Meanwhile, to ease Arab concerns, British officials in the mandate began enforcing limits on immigration in accordance with the white paper.¹¹

Churchill was pleased with the results of his work. In his eyes, London had addressed each party’s concerns and had done so in a way that resolved the contradiction inherent in the Balfour Declaration. With limits on immigration now in effect, the Palestinians could be certain that they would not lose their livelihood to Jewish settlers or see the entire mandate swallowed up by a Zionist state. Meanwhile, with the declaration having been formally incorporated into the mandate charter, the

⁹ Charles D. Smith, *Palestine and the Arab-Israeli Conflict: A History with Documents*, Sixth Edition (Boston: Bedford/St. Martin’s, 2007), 115–16; quote from “The Balfour Declaration, November 2, 1917.”

¹⁰ “‘Churchill White Paper’ - UK Secretary of State for the Colonies (3 June 1922),” accessed October 2, 2020, <https://unispal.un.org/DPA/DPR/unispal.nsf/0/F2CA0EE62B5680ED852570C000591BE> B.

¹¹ Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, 87–88.

Zionists could rest assured that they would get the national home that they sought. For a time, events in Palestine appeared to sustain Churchill's favorable assessment of London's new approach. Indeed, the absence of intercommunal violence in the mandate over the next few years suggested strongly that the colonial secretary really had resolved the conflict between Palestinian Arabs and Jewish immigrants.¹²

In reality, neither the Palestinians nor the Zionists were in the least bit pleased with the terms of the Churchill White Paper. The Arabs disliked it because they did not want *any* further Jewish immigration and because they believed—correctly—that the Zionists remained intent on establishing a Jewish state throughout the entirety of the mandate. For their part, meanwhile, the labor Zionists who dominated the Zionist Executive were unhappy with it because it appeared to indicate that Britain was backing away from the commitment it had made in the Balfour Declaration.¹³

The Simmering Conflict, 1922-1936

Zionism's New Approach

Nonetheless, the Zionist Executive opted to accept the terms of the Churchill White Paper. What explains this decision? Why did it agree to abide by a deal that most Zionists believed was contrary to their movement's interests? It did so for two reasons. First, the Zionist leadership understood that Jewish people accounted for only a small minority of Palestine's population—just 11 percent according to the 1922 census—and were thus not in a position to achieve the dream of a Jewish state without the continued backing of their British patron. Second, they agreed to abide by the new limits that Churchill had imposed because they viewed those restrictions as merely temporary ones that they would be able to discard after they had founded the Zionist state. That is, once they had established control of a part of Palestine—however small—they could then bring in as many immigrants as they liked and could expand their fledgling state to include additional territory.¹⁴

Accordingly, the Labor Zionists effected a significant—if tactical—policy change following the riots of 1920 and 1921 and the issuance of the Churchill White Paper. For the time being, they abandoned the strategy they had adopted after World War I of trying to secure the rapid establishment of a large Jewish state. Instead, in the words of the historian Avi Shlaim, the Zionists adopted a more pragmatic, gradualist policy of building up their position in Palestine “step by step, immigrant by immigrant, settlement

¹² Provençe, 87–88.

¹³ Shlaim, *The Iron Wall*, 10.

¹⁴ “Palestine: Report and General Abstracts of the Census of 1922,” n.d., <https://users.cecs.anu.edu.au/~bdm/yabber/census/PalestineCensus1922.pdf>; Smith, *Palestine and the Arab-Israeli Conflict*, 116.

by settlement” until such time that they had the international support and domestic strength needed to declare the establishment of the Zionist state.¹⁵

Revisionist Zionism and the Iron Wall

The Zionists were not unanimous in supporting the shift to a gradualist policy, however. Instead, under the leadership of a dissident member of the Zionist Executive, the Russian-born Ze’ev Jabotinsky (1880-1940), a minority rejected the pragmatic strategy of the mainstream Labor Zionists in favor of a far-more aggressive and militaristic approach. Called revisionists because they wanted to revise the mandate’s terms so that Transjordan was again subject to the Balfour Declaration, Jabotinsky and his supporters argued that the incremental strategy would not succeed and that the Zionist Executive could achieve that goal only by abandoning gradualism in favor of the immediate establishment of a Jewish state that included territory on both banks of the Jordan River. Jabotinsky repeatedly laid out his position before the Zionist Executive in the early 1920s; with the leadership firmly committed to gradualism, however, his objections fell on deaf ears. Frustrated, he resigned his seat in 1923 and threw himself into the task of organizing a political opposition movement, later known as the Revisionist Party, and an associated paramilitary youth organization called Betar. The new party enjoyed only modest support at first and remained very much in the minority during the interwar period; nonetheless, gaining a small-but-committed following, it provided an influential ideological challenge to the Labor Zionists.¹⁶

Jabotinsky’s embrace of militarism stemmed from his rejection of the Labor Zionists’ belief that a Jewish state could be established peacefully in Palestine. Their position was an absurd one, he maintained; given the zero-sum nature of the dispute over the mandate, the question of which people would control Palestine could only be settled through force. No promise of economic gain or assurance that a Jewish homeland would not harm the territory’s non-Jewish residents would placate the Arabs or obscure the fact that Zionism was, at heart, a “colonizing venture”—one that the Palestinians, whom he openly acknowledged were a national group, were bound to oppose. As he wrote,

“[e]very indigenous people . . . will resist alien settlers as long as they see any hope of ridding themselves of the danger of foreign settlement. This is how the Arabs will behave and go on behaving so long as they possess a gleam of hope that they can prevent ‘Palestine’ from becoming the Land of Israel.”¹⁷

In light of this fact, he continued, the Zionists could ensure their security only through the erection of a metaphorical “Iron Wall:” a powerful military that would first

¹⁵ Shlaim, *The Iron Wall*, 11.

¹⁶ Shlaim, 11–13.

¹⁷ Rashid Khalidi, *The Iron Cage: The Story of the Palestinian Struggle for Statehood*, Reprint Edition (Boston: Beacon Press, 2006), 72; quote from Shlaim, *The Iron Wall*, 14.

establish the Jewish state through brute strength and then protect it by responding to every Arab provocation—however minor—with overwhelming force. Eventually, after repeated defeats had exhausted the Palestinians of any hope that they could destroy the Jewish state, the Zionists would finally be able to engage in productive negotiations with the Arabs aimed at granting them limited rights in a Jewish-controlled country. Accordingly, in Jabotinsky's eyes, efforts to placate the Palestinians were a waste of time. Only overwhelming force could establish the Jewish state, and only disproportionate strength could ensure its survival.¹⁸

Jabotinsky's differences with the labor Zionists thus appeared to be enormous. Where he acknowledged that Palestinian resistance was inevitable and reflected nationalist sentiment, his labor opponents instead publicly clung to the view that the real conflict in the mandate was a class-based one that would disappear as the economy improved. Likewise, while he asserted that the Zionists could only establish a Jewish state through force, they continued to declare that the movement could do so through negotiation and compromise. At heart, however, the disagreements between the Labor Zionists and the Revisionists were more about tactics than substance. In truth, the moderates shared many of Jabotinsky's views and pursued their seemingly more conciliatory approach not because it reflected deeply held views but because they believed it would avoid alienating Britain: the patron on whose assistance the achievement of their goal ultimately depended.¹⁹

Their confidential statements made clear just how similar their views were to the Revisionists' position. For example, the prominent Zionist H. M. Kalvaryski admitted not long after the Jaffa riots that the violence was not the product of class conflict; as he wrote, "[i]t is pointless to consider this a question only of effendis [elites] . . . we should realize that we have to reckon with an Arab national movement." More importantly, despite their public commitment to peaceful coexistence, the Labor Zionist leaders who dominated the Zionist Executive—renamed the Jewish Agency in 1929—privately agreed with Jabotinsky that their movement would need to use force to establish and maintain a Jewish state. As Ben Gurion wrote confidentially in the 1930s, only after repeated military defeats had produced "total despair on the part of the Arabs . . . may the Arabs acquiesce in a Jewish Eretz [greater] Israel." Thus, though the Labor Zionists continued to declare their opposition to Jabotinsky in public, they privately accepted the broad thrust of his thinking and disagreed with him far more about timing and rhetoric than method or outcome.²⁰

Zionist Institution Building

In the meantime, the mainstream Labor Zionists moved to strengthen the Yishuv's position. With the Jewish community already possessing a mature national consciousness, they focused on creating a series of interlocking political, cultural, and

¹⁸ Shlaim, *The Iron Wall*, 13–15.

¹⁹ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 104.

²⁰ Kalvaryski quoted in Morris, 104; Ben Gurion quoted in Shlaim, *The Iron Wall*, 18–20.

economic structures that would serve as the foundation of the coming Jewish state and society. They established a parliament, the Haganah militia, an executive body known as the Jewish National Council, and the Histadrut, a federation of labor unions that represented workers but that also provided health care and pursued industrial development. These organizations worked closely with the existing Jewish National Fund, which purchased land for Jewish immigrants, and with the Zionist Executive, which, in addition to its role as the Yishuv's League of Nations-sanctioned representative to the mandatory power, also promoted immigration to Palestine. Tying these institutions together was the Ahdut HaAvoda, or Labor Unity Party, which, under the leadership of Ben Gurion and his allies, had been instrumental in founding the Haganah and Histadrut. In 1930, it merged with another socialist party to become Mapai—the political party that would dominate Israeli politics during the country's first three decades. Able to draw on the experience that highly educated European Jewish immigrants brought with them, these institutions grew rapidly in scale and sophistication. By 1930, as a result, they functioned as a de facto government for the Yishuv—one, as we shall see, that left it well prepared for eventual independence.²¹

Jewish economic dynamism matched this institutional growth. Operating separately from the Palestinian economy—and thus implicitly undermining the claim that the growth of the Jewish community would raise Arab living standards—the Yishuv economy underwent rapid growth during the mandate period. Benefitting from the continued arrival of highly educated Jewish European immigrants and the ongoing influx of huge quantities of capital, it expanded at a blistering average annual rate of 13.2 percent during the 1920s—roughly twice the rate at which the mandate's Arab economy grew. Jewish economic development was impressively diversified, moreover. It included not only industrial growth in urban areas but also the establishment of dozens of agricultural settlements in the countryside.²²

Palestinian Institution Building

The Palestinians, too, engaged in institution and nation building in the 1920s. The Palestinian Arab Congress continued to meet regularly and, in 1920, elected a governing board called the Arab Executive to coordinate resistance to Zionism. The Palestinians also gained control of two new, British-created institutions: the Supreme Muslim Council (SMC), which administered *awqaf* revenue, and the post of Grand Mufti of Jerusalem and Palestine, which afforded its first holder, the young nationalist Hajj Amin al-Husayni (c. 1897-1974), with substantial authority thanks to the power it vested in him to name Islamic judges, or qadis, to their posts. Meanwhile, following France's destruction of Faysal's Arab Kingdom of Syria in July 1920, the Palestinians abandoned their earlier pan-Arabist attempts to attach Palestine to greater Syria in favor of a new

²¹ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 109–10.

²² Khalidi, *The Iron Cage*, 13–15.

effort to transform the mandate into an independent, Arab state—a change that abetted the development of a strong, Palestinian nationalist identity.²³

While these institution- and nation-building efforts may appear significant at first glance, they were anemic in comparison to the Zionists' accomplishments in the interwar era. In stark contrast to the Yishuv's efficiently run bodies, Palestinian political structures were weak, riven by debilitating factional disputes, and incapable of providing the direction that modern states and societies require. What accounted for the ineffectiveness of these institutions? Why were the Palestinians unable to keep pace with the Yishuv in establishing the sort of quasi-state political, military, and economic structures that could, when the time was right, be transformed into state institutions?

Traditionally, historians have argued that three shortcomings intrinsic to Palestinian society explain the failure to develop governing structures comparable to those of the Yishuv. First, they have maintained that sharp class, religious, and political differences prevented the Palestinians from presenting a unified front against either the Zionists or the British. Most notably, the intense political rivalry between the grand mufti al-Husayni's dominant faction and the loose opposition party controlled by his family's traditional rivals, the Nashashibi clan, diverted Arab energies away from confronting Zionism and into internecine battles. Second, adherents of the traditional view have noted that the Palestinians were far-less developed politically than the Zionists and lacked their rivals' organizational sophistication and diplomatic tact; as a result, the Zionists were able to outmaneuver them at every turn. Finally, emphasizing historical contingency, proponents of this interpretation have asserted that Palestinian leaders were peculiarly ineffective and frequently inclined, as the historian Benny Morris puts it, "to shoot themselves in the foot."²⁴

Considerable evidence seems to support this interpretation. For example, the Arabs initially rejected High Commissioner Samuel's proposal in 1922 to establish a joint legislative council for the mandate based on a proportional system of representation that would have given them more seats than the Zionists. The Palestinians eventually warmed to the offer, but, as would seemingly happen again and again, the proposal was no longer on the table when they belatedly changed their minds. Such episodes have long shaped interpretations of the Palestinian nationalist effort in the interwar era and have led many scholars to conclude that there is more than

²³ Rashid Khalidi, "The Palestinians and 1948: The Underlying Cause of Failure," in *The War for Palestine: Rewriting the History of 1948*, ed. Eugene L. Rogan and Avi Shlaim (New York: Cambridge University Press, 2001), 21–23.

²⁴ Quote from Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 109; Khalidi, *The Iron Cage*, 65–66.

a grain of truth to Israeli Foreign Minister Abba Eban's (1915-2002) famous quip that "[t]he Palestinians never missed an opportunity to miss an opportunity."²⁵

The traditional view has not gone unchallenged, however. While acknowledging the manifest failings of the Palestinian leadership, a number of historians have recently taken issue with the conventional interpretation by arguing that it fails to capture adequately the significant constitutional constraints under which the Palestinians operated or to note the very different political circumstances in which the Arabs and Zionists found themselves. Indeed, they maintain that trying to explain Palestinian failures by focusing on their shortcomings vis-à-vis their Zionist rivals is of little analytical value. As the historian Rashid Khalidi notes, contrasting the Palestinians with sophisticated immigrants from modern Europe amounts to "comparing the incomparable." Far more useful, he contends, is to contrast them with the Arabs who lived in other parts of the region. As he points out, the Arabs of Egypt, Syria, and Iraq were no more sophisticated than the Palestinians yet were able to make significant progress toward the achievement of independence in the face of vigorous resistance from the powerful British and French empires. Given Arab achievements against those mighty states, in other words, the Zionists' relative sophistication cannot alone explain the Palestinians' remarkable lack of success in the interwar period.²⁶

Today, a growing number of historians argue that the Palestinians' lack of institutional success was largely a product of the mandate's peculiar political structure. Scholars like Khalidi, the historian Michael Provence, and others now conclude that the Palestinians were unable to secure their goals not because they were less sophisticated than the Zionists or because they suffered from peculiar amounts of factional infighting but instead because they found themselves caught in a constitutional "iron cage" from which they were never able to escape. This trap was a brutal catch-22: to get formal political representation from the British, the Palestinians had to consent to the terms of the mandate; accepting the mandate, however, meant acknowledging the right of the Zionists to establish a national home in Palestine—the precise turn of events that the Palestinians sought political standing to avert. To Provence, this dilemma explains why the Arabs pursued seemingly counterproductive courses such as their refusal to go along with Samuel's offer to establish a legislative council based on proportional representation. As he notes, the Palestinians did not reject that proposal because they were shortsighted or caught up in pointless factional infighting; they did so instead because it would have required them to formally accept the creation of the very Jewish homeland that they were fighting to prevent.²⁷

²⁵ Khalidi, *The Iron Cage*, 65–66; Eban quoted in Douglas Little, *American Orientalism: The United States and the Middle East since 1945*, second edition (Chapel Hill: University of North Carolina Press, 2004), 267.

²⁶ Khalidi, *The Iron Cage*, XLV, 22–30.

²⁷ Khalidi, 31–48; Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, 196–97.

These scholars also argue that British-imposed constitutional constraints accounted for many—though not all—of the fault lines that split the Palestinians. Here, they center their critique on the Supreme Muslim Council and the Grand Mufti of Jerusalem. The most powerful Arab institutions in the mandate, these structures were, in Khalidi's words, "invented traditions" that Britain had established to define the Palestinians in religious rather than national terms. Coming with attractive salaries and patronage opportunities, seats on the SMC and the post of grand mufti also amounted to de facto bribes that ensured that Palestinian elites were beholden to London and would thus toe the imperial line and encourage acceptance of British rule. Finally, these scholars contend that Britain used these institutions to divide the Palestinians. By carefully appointing members from both the dominant al-Husayni group and the al-Nashashibi opposition to seats on the SMC, the British ensured that the body served as an arena for factional skirmishes among the Palestinian leadership rather than as a center of resistance to Zionism or imperial rule. In other words, the infighting that put the Palestinians at a disadvantage was not a function of flaws intrinsic to their society, but instead the product of a deliberate British policy of divide and conquer.²⁸

Immigration and Land Purchases

The Zionists' institutional advantage over the Palestinians proved central to the rapid expansion of the Yishuv in the 1920s. Thanks to the efficient work of the Zionist Agency, more than 70,000 Ashkenazi Jews would immigrate to Palestine over the course of the decade. Many of those immigrants settled on the more than 240,000 acres of arable land that the Jewish National Fund purchased for them in the 1920s—land that not only dramatically increased the Jewish community's holdings but that also made possible the establishment of more than fifty new agricultural settlements.²⁹

Stepped up land purchases and immigration deeply troubled the mandate's Palestinian population in the decade after World War I. While immigration had been a concern since the 1880s, it gained new salience in the 1920s owing to the rapid growth of the Jewish community, which, thanks to the aforementioned arrival of Ashkenazi Jewish people more-than doubled between 1922 and 1931—an ominous trend that fed growing concerns among the Arabs that the Zionists would soon have the critical mass needed to declare the establishment of a Jewish state. Land transactions were equally worrisome. The Zionists were not only buying large tracts, but, in many cases, were also evicting Palestinians whose families had worked that land as tenant farmers for generations. Reduced thereafter to seasonal agricultural work or day labor in the mandate's towns and cities, the dispossessed saw the security and stability that they had once enjoyed replaced by a painful precarity.³⁰

²⁸ Khalidi, "The Palestinians and 1948: The Underlying Cause of Failure," 21–25.

²⁹ Rogan, *The Arabs*, 198.

³⁰ "Palestine: Report and General Abstracts of the Census of 1922"; "Census of Palestine 1931," n.d.,

Like all issues related to the Zionist-Palestinian conflict, the impact of Jewish land purchases on Arab tenant farmers remains hotly debated. Based on studies of mandate records, Israel's backers contend that property transfers to the Zionists resulted in the displacement only a small number of Arab farmers. Supporters of the Palestinians counter by arguing that the documentary evidence masks the true extent of evictions and that the actual figure is far higher. Regardless, the political impact was clear: concerns about being dispossessed radicalized many tenant farmers and—to the delight of those who feared being force off their land—compelled the Arab leadership in the mandate to loudly call for an outright ban on further land transfers. The fellahin would likely have been much-less enthusiastic about those demands had they known that many of the loudest voices publicly denouncing land transfers were at the same time quietly profiting from the practice. Indeed, a significant share of the Palestinian leadership—including fully one quarter of the men who served on the Arab Executive—were secretly selling land to the Zionists at the very time that they were vocally attacking the practice.³¹

Still, despite the Palestinians' apprehensions about continued immigration and land sales, the absence of significant conflict following the Jaffa riot in 1921 led British officials to conclude that the Churchill White Paper had been successful in restoring peace to the mandate. Benefitting from good economic times for much of the 1920s, Britain's efforts to conciliate both parties did, indeed, appear to have eased intercommunal tensions and to have led the Palestinians and Zionists to come to an accommodation, even if they did so grudgingly. On top of that, a decline in immigration in the late 1920s—the Jewish population remained essentially unchanged from 1926 to 1932 and even fell slightly in 1927 when net emigration exceeded immigration—suggested to many that the struggle between Zionism and Palestinian nationalism had peaked and was winding down.³²

The 1929 Palestine Riots

A new round of intercommunal conflict at the end of the decade made clear, however, that the struggle for control of Palestine was as white hot as ever. Rising Arab worries about Jewish immigration and land purchases set the stage for renewed clashes, but it was a religious issue related to the Wailing, or Western Wall that ultimately sparked the violence. A place where devout Jewish people had prayed for centuries, the Wailing Wall is all that remains of the Second Temple that sat on the Temple Mount from 516 BCE until its destruction in 70 CE. While holy to Jewish people, the wall is also part of the enclosure to Islam's third-holiest site, the Haram al-Sharif, or

<https://users.cecs.anu.edu.au/~bdm/yabber/census/PalestineCensus1931.pdf>; Smith, *Palestine and the Arab-Israeli Conflict*, 130–31; Rogan, *The Arabs*, 198.

³¹ Segev, *One Palestine, Complete: Jews and Arabs Under the British Mandate*, 274–75.

³² Khalidi, "The Palestinians and 1948: The Underlying Cause of Failure," 24; Kirsten E. Schulze, *The Arab-Israeli Conflict* (New York: Pearson, 2008), 8.

holy sanctuary, on which sits the Dome of the Rock shrine and the al-Aqsa Mosque, and is thus significant for Muslims as well.³³

The Haram al-Sharif, the Wailing Wall, the alley that runs along the wall, and the adjoining neighborhood were all part of a *waqf*. However, the people who directed those endowments did not control entry to the sites. Instead, it was the Status Quo, a formal arrangement that the Ottoman government had adopted in the nineteenth century and that the British continued to observe, that governed access to the religious sites. By its terms, Jewish people were entitled to pray at the Wailing Wall; to ensure free passage along the alley, however, they were banned from placing furnishing of any kind in front of it.³⁴

A dispute over the observance of the Status Quo that began in 1928 would turn the Wailing Wall into an intercommunal flashpoint and spark renewed violence the following year. The conflict began on the Day of Atonement in September 1928 when conservative Jewish worshippers erected screens to separate men from women praying at the wall. Fearing that the dividers were part of a Zionist plot to take control of the Wailing Wall and the Temple Mount, Muslim Palestinians demanded that the British restore the Status Quo and remove the screens; eager to uphold the peace, the mandatory government complied. Restoration of the Status Quo did not put the issue to rest, however. Instead, extremists in both camps issued evermore provocative statements and took increasingly inflammatory actions culminating in members of Betar holding a demonstration in front of the wall in mid-August 1929 at which they chanted “the wall is ours.” This affront was too much for Muslim Palestinians. On August 23, 1929, enraged Arab peasants poured out of the Haram al-Sharif following Friday prayers and attacked Jewish people and businesses. As news of the attacks travelled, the violence intensified and spread throughout the mandate. By the time British reinforcements finally managed to regain control on August 29, 133 Jewish people—including nearly the entire Jewish community of Hebron—and 116 Arabs had been killed.³⁵

The Passfield White Paper

Britain’s response revealed the intractable nature of the conflict. Shocked by the extent of the violence, London sent a delegation to Palestine called the Shaw Commission to ascertain the causes of the riots. It determined that the religious dispute was merely a symptom of a larger issue: “the twofold fear of the Arabs that by Jewish immigration and land purchases they may be deprived of their livelihood and in time pass under the political domination of the Jews.” In response, the Labor Party government of Prime Minister Ramsey MacDonald (r. 1929-1935) moved to formally alter Britain’s approach to Palestine through the issuance of the Passfield White Paper

³³ Smith, *Palestine and the Arab-Israeli Conflict*, 131.

³⁴ Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, 197.

³⁵ Provence, 197–200.

in 1931. It marked a significant change in policy. By limiting further immigration to the mandate's "economic 'absorptive capacity'" and by placing stringent restrictions on land sales to the Zionists, it effectively abandoned the Balfour Declaration's open-ended commitment to a Jewish homeland in Palestine in favor of a more even-handed approach.³⁶

Unsurprisingly, the Passfield White Paper created intense feelings among both Palestinians and Zionists. The Arabs were thrilled when they read its terms and concluded that their appeals to justice and majority rule had finally made an impression on the British. In contrast, Jewish leaders viewed it as an existential danger to the Zionist project—one that threatened to prevent the Yishuv from ever having the numbers needed to establish a Jewish state. Accordingly, in conjunction with British conservatives, they responded by mounting an intense and highly effective political pressure campaign against the Labor Party's vulnerable, minority government. Fearing for the survival of his cabinet, MacDonald soon capitulated. In February 1931, he issued a public letter addressed to Chaim Weizmann in which he repudiated the Passfield White Paper and recommitted Britain to the terms of the Balfour Declaration. Drafted in its essentials by Weizmann, the letter explicitly declared that London would place no limits on immigration or land purchases. It was now the Palestinians turn to be furious. With their hopes first raised and then dashed, they bitterly criticized the MacDonald government and denounced his note to Weizmann by calling it the "Black Letter."³⁷

The riots of 1929 and their aftermath laid to rest any illusion either that the situation in Palestine was peacefully resolving itself or that it could be settled through negotiations. In response, Zionists, British officials, and Arabs alike reexamined their assumptions and approaches to the evolving situation in the mandate. For the Palestinians, London's quick retreat from the Passfield White Paper indicated that reasoned arguments based on appeals for justice and self-determination would neither arrest the continued growth of the Jewish community nor stop the establishment of a Zionist state; as a consequence, many younger, more extreme Arabs began to organize armed militias. For the Zionists, meanwhile, the ferocity of the Arab attacks in 1929 and Britain's shocking inability to protect the Yishuv ended any lingering hopes that they might secure a Jewish state through peaceful means. After the riots, accordingly, the Labor Zionists transformed the Haganah from a weak, decentralized militia into a powerful proto army. Even that move did not go far enough for a small group of Haganah officers, however. Looking to retaliate aggressively to any further Arab attacks, they formed a separate militia group called Irgun. It soon became the military arm of the Revisionist Party. For the British, finally, the events of 1929 laid bare for the first time that Palestine was a poisoned chalice, a perpetual trouble spot that defied easy—or, for that matter, difficult—solutions. They would spend the next decade

³⁶ Smith, *Palestine and the Arab-Israeli Conflict*, 131–33.

³⁷ Rogan, *The Arabs*, 199.

seeking desperately to find some formula that could reconcile competing Arab and Zionist demands while permitting continued imperial control. They never succeeded.³⁸

The Great Depression and Hitler's Rise to Power

Meanwhile, surging antisemitism in Europe in the early 1930s and the German dictator Adolf Hitler's (r. 1933-1945) rise to power dramatically altered the situation in Palestine. Firmly in control of Germany by early 1933, Hitler and his supporters wasted no time in stepping up violence against Jewish people and in imposing ever-more stringent restrictions on them. The Nazis forbade them from marrying Gentiles, banned them from the professions and from teaching, and organized boycotts of Jewish businesses. The Nazis' campaign against German Jews—and similar efforts in nearby states—profoundly frightened Central and Eastern European Jewish people and led tens-of-thousands to consider emigration. Their options were far-more limited than they had been in the recent past, however. Rising anti-Semitism and widespread opposition to immigration during a time of unprecedented global mass unemployment led traditional havens such as Canada and the United States—the fabled *Golden Medene*—to maintain strict limits on Jewish immigration. As a result, European Jews had nowhere to go but Palestine.³⁹

The ensuing flood of refugees remade the mandate's demographic makeup and fundamentally altered the balance between the Jewish and Arab communities. In contrast to the average of 5,000-6,000 Jewish people who had arrived annually between 1929 and 1931, the mid-1930s saw tens of thousands of refugees pour into the territory each year. 30,000 immigrated in 1933, 42,000 more in 1934, 62,000 in 1935, and 30,000 in 1936. All told, this new wave of immigration—by far the largest yet—more than doubled the Jewish population of the mandate, increasing it from 164,000 in 1930 to 370,000 in 1936 and raising the Jewish share of Palestine's population from 18 percent in 1930 to 27 percent in 1935. Predictably, it created enormous alarm among the Arabs and fed a sharp rise in fears that the Zionists would, with British help, displace them from Palestine.⁴⁰

Land purchases in the early 1930s further contributed to Palestinian discontent. The Zionists acquired fewer total acres over the course of that decade than they had in the 1920s, but the land that they purchased tended to be cultivated rather than fallow and thus had a greater impact on the fellahin. Some Arabs tried to organize boycotts and extract pledges from landowners to prevent further property transfers. However, vigorous efforts to enforce such agreements failed in the face of surging land values. Indeed, with prices quadrupling between 1929 and 1935 even those deeply opposed to Zionism could not resist the temptation to make a windfall profit. As the German consul noted, "Arab nationalists 'in daylight were crying out against Jewish immigration and in

³⁸ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 118–20.

³⁹ Ian Kershaw, *To Hell and Back: Europe 1914-1949* (New York: Penguin Books, 2015), 207–16.

⁴⁰ Schulze, *The Arab-Israeli Conflict*, 8–9; Rogan, *The Arabs*, 201.

the darkness of night were selling land to the Jews.” The continued transfer of property proved enormously corrosive to Palestinian solidarity. It created an atmosphere of suspicion and helped sow divisions along lines of class that would hamstring Arab efforts to oppose the British and the Zionists for the duration of the mandate.⁴¹

The Great Arab Revolt, 1936

By the mid 1930s, land sales and the surge in immigration had left the mandate on a knife's edge. With Jewish refugees pouring into Palestine in unprecedented numbers and rumors of impending displacement rife, Arab anger and resentment finally boiled over. The result was the outbreak of the Great Arab Revolt in the spring of 1936: a mandate-wide rebellion against Zionism, British rule, and, to a lesser extent, the politically dominant Arab notables. Begun with high hopes and great determination, the rebellion was, for a time, a serious challenge to the Zionists and a huge financial drain on London. Ultimately, however, it failed to achieve its goal of compelling London to end its support for Zionism; as important, it left the Palestinians too weak, divided, and demoralized to thereafter mount an effective resistance to Zionism.

Revolt and General Strike, April-September 1936

The revolt began in early 1936. In April, a band of Palestinian militants murdered two Jewish men and wounded a third at a nighttime roadblock. The incident set off a spiral of retaliatory attacks and violence throughout the mandate. Loosely organized groups of urban Arabs threw rocks at police officers while their rural counterparts harassed British soldiers and engaged in acts of sabotage. The situation quickly escalated. In short order, increasingly sophisticated Arab militia bands were sniping at imperial troops, ambushing patrols, laying mines, sabotaging railways, and attacking isolated Jewish settlements. Meanwhile, a loosely connected network of nationalist groups moved to support the insurgency by organizing an economic boycott patterned after the recent Syrian General Strike. Designed to stop the flow of tax revenue that underwrote British control of the mandate, it rapidly attracted widespread support within the Palestinian community. By the summer of 1936, as a result, Britain found itself in a difficult spot—desperately trying to suppress a full-scale insurrection and to overcome a mandate-wide strike at the very moment that its limited military resources were already stretched thin by the increasingly aggressive actions of fascist Germany and Italy.⁴²

Immigration and land purchases established the conditions for the revolt and the strike, but what accounted for its timing? Two factors stand out. The first was geostrategic. Britain's failure to contest Benito Mussolini's (r. 1922-1945) invasion of Ethiopia in 1935 and its unwillingness to act when Hitler remilitarized the Rhineland in March 1936 suggested to Arab leaders that London lacked the fortitude to deal with a costly insurrection and would quickly capitulate in the face of continued pressure so

⁴¹ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 122–23.

⁴² Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, 231–37.

near to the Suez Canal. The second reason was economic in nature. The sudden collapse of the boom that the mandate had enjoyed between 1932 and 1935 left thousands of embittered Arabs confronting the prospects of unemployment—some because their Jewish employers had responded to the downturn with a “Hebrew labor” policy that involved replacing Palestinian workers with recent immigrants from Europe. By early 1936, as a result, the mandate was brimming with young, resentful Arab men who, having little to lose, were only too willing to take up arms against the British or to enforce the economic boycott.⁴³

The rebellion caught the Palestinian notables flat footed. Lulled into complacency by their dependence on the British for their salaries and patronage powers, the leading Palestinians in the al-Husayni and Nashashibi factions had quietly acquiesced to continued immigration and British rule during the 1920s and early 1930s and were thus on the sidelines when the younger, more-radical nationalists launched the insurgency in April 1936. Once the revolt began, however, the notables moved quickly to assume control. On April 25, the factions agreed to put down their differences and formed the Arab Higher Committee (AHC) to coordinate the economic resistance. The grand mufti served as its chairperson. Its first act was to announce that the boycott, thereafter known as the Arab General Strike, would continue until the British agreed to meet three demands: the immediate cessation of immigration, a blanket ban on Jewish land purchases, and the establishment of a representative government for the mandate. Publicly, the AHC focused exclusively on coordinating the economic resistance and denied that it had any connection to the militants; privately, however, it moved to support the insurgents and sought with mixed success to assume control of the rebellion.⁴⁴

Likewise caught off guard, the British initially responded as the Palestinians had anticipated. Desperate to avoid a costly anti-insurgency campaign, mandate authorities pursued a passive appeasement policy that included new limits on immigration in hopes that such an approach would undercut support for the rebellion. By the end of the summer, however, London reluctantly came to conclude that conciliation was not working and shifted to a more aggressive response. It deployed a division of troops to Palestine, imposed martial law, and instituted a system of collective punishment wherein British soldiers demolished houses and made mass arrests in Palestinian villages suspected of supporting the revolt. The new tactics were ugly, but they were also effective. By mid-autumn, it had become abundantly clear that the rebellion had failed.⁴⁵

The strike proved similarly unsuccessful. Despite imposing depression conditions on the mandate’s Arab population, it had been unable to compel the British to negotiate.

⁴³ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 127–29.

⁴⁴ Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, 232–33; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 129–33.

⁴⁵ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 130–34.

Why had the economic protest failed? More pertinently, why had it been unable to match the success of the concurrent Syrian General Strike? The explanation lies in the peculiar, bifurcated nature of Palestine's economy. Over the course of the prior two decades, rapid Zionist commercial, industrial, and agricultural development had produced a largely self-contained Jewish economic sector in the mandate. Far-more advanced than the Palestinian economy, it accounted for the bulk of the taxes that the British collected. As a result, though the strike had all-but shut down the Palestinian sector of the economy, it had failed to depress tax receipts enough to force the British to the table.⁴⁶

Faced with the failure of the insurgency and the prospect of impoverishment and even famine in the Palestinian community, the AHC began to consider bringing the uprising to a close. Ending the rebellion and the strike would prove tricky, however. Complicating the AHC's efforts to secure a negotiated settlement, the guerrilla commanders in the field made clear that they would accept nothing less than a total victory. As a result, the Arab Higher Committee would need to find—or somehow generate—political cover if it wished to push forward with a negotiated end to the strike and the revolt. It managed to do so in the fall of 1936 with the help of the emir of Transjordan and the kings of Saudi Arabia, Iraq, and Yemen. At the AHC's indirect behest, those rulers issued statements appealing for an immediate end both to the strike and the revolt. The gambit worked. In conjunction with London's pledge to send a committee to assess the situation in Palestine, their entreaties were sufficient to give the AHC the pretext it needed to bring the strike and the rebellion to an end in October. There was a cost, however. In inviting the monarchs to assume a direct role in their affairs, the Palestinians had set in motion a process that would soon lead to the Arab heads of state taking charge of their cause—a change in leadership that would, as we shall see, have baleful consequences for them.⁴⁷

The Peel Commission Report, 1937

Named for its chair, Lord William Peel (1867-1937), the commission that London had promised to send to the mandate arrived in November 1936. Its visit would prove to be a crucial turning point for Palestine—one that would fundamentally alter the terms of the debate over the disposition of the mandate. The Zionists and Palestinians approached the commission very differently. Yishuv leaders viewed it as a golden opportunity to make their case and seized the opportunity to do so. Presenting Zionism as a moderate movement that was more-than willing to compromise, Weizmann, Ben Gurion, and others promoted a plan before the commission that called for a partition of the territory into separate Jewish and Arab states as a way to end intercommunal violence. The Palestinians, in contrast, once again fumbled the opportunity to present their position. While the Zionist leadership was outwardly solicitous and claimed to be

⁴⁶ Provence, *The Last Ottoman Generation and the Making of the Modern Middle East*, 231, 237.

⁴⁷ Khalidi, *The Iron Cage*, 111; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 134–35.

eager to compromise, the AHC instead assumed a defiant attitude and initially refused to meet with the commission unless London first imposed an outright ban on further immigration. Indeed, it was only just before the commission returned to Britain that outside pressure finally succeeded in compelling the Palestinians to present their case.⁴⁸

Released in July 1937, the Peel Commission Report outlined a dramatic new approach to the mandate that would set the stage for its division in 1948. Admitting openly for the first time that “an irrepressible conflict has arisen between two national communities” and that a unitary state was doomed to failure, it proposed partitioning the mandate between the Jewish and Palestinian peoples. According to its plan, the Zionists would establish a small state in territory that stretched northward from Tel Aviv to the Lebanese border and then east to the Sea of Galilee, and Britain would retain control of Jerusalem and a corridor of land connecting it to the coast. The remaining territory, about eighty percent of the mandate, would be appended to Transjordan. It seemed like a Solomonic solution: a wise division of the mandate that would satisfy both parties’ interests. Resolving the clash between rival nationalist movements through partition raised a new issue, however: what was to happen to the Arabs who lived in the area assigned to the Jewish state and to the Zionists who resided in the territory allotted to the Palestinians? The Peel commission proposed to solve this issue through “population transfers” similar to the exchange of people that occurred between Greece and Turkey in the Treaty of Lausanne. In other words, a process of ethnic cleansing would take place in which Arabs would relocate from the Zionist state, and Jewish people would move out of the territory assigned to the Palestinians.⁴⁹

At first blush, the plan to forcibly move people to create ethnically homogenous states appeared to be even handed. As a result of the population transfers, both the Jewish state and the Arab one would have uniform populations and would thus be free of ethnic tension. Upon closer inspection however, the commission’s proposal revealed itself to be entirely one sided. While the plan called for a mere 1,250 Jewish people to relocate from the territory assigned to the Arabs, it required no fewer than 225,000 Palestinians to depart from the new Zionist state—a forced transfer that the Arabs condemned as unacceptable. The Peel Commission’s proposed population relocation was a critical component of the partition plan, however. Without it, the Jewish state would retain an Arab minority that was nearly as large as its Jewish population—thus rendering moot the entire point of Zionism.⁵⁰

The Zionist movement reacted with mixed emotions to the specifics of the partition plan but ultimately came to endorse the idea. Unsurprisingly, Jabotinsky and the revisionists outright rejected the commission’s recommendations on the grounds that they did not immediately grant the Zionists the entirety of both Palestine and Transjordan. In contrast, though they were disappointed that the plan allotted the

⁴⁸ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 138.

⁴⁹ Quotes from Rogan, *The Arabs*, 204–5.

⁵⁰ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 138–39.

Jewish state a mere 5,000 square kilometers of territory—a bit larger than six times the size of New York City—prominent Labor Zionists like Weizmann, Ben Gurion, and Moshe Sharett (1894-1965) were enthusiastic about the proposal and lobbied aggressively for its acceptance. They articulated two arguments in its favor. First, they contended that sovereignty would mean the lifting of immigration limits and a concomitant rapid increase in the Jewish state's population. Second, and more importantly, they maintained that a state along the lines that the Peel Commission had outlined would not be Zionism's end product but would instead constitute a critical opening bid that its people could enlarge through negotiation or conquest. As Ben Gurion privately wrote, "I am certain [that] we will be able to settle in all the other parts of the country, whether through agreement with our Arab neighbours or in another way." These arguments proved persuasive. Accordingly, while the Twentieth Zionist Congress that met in Zurich in August 1937 rejected the specific borders that the Peel Commission had drawn up, it voted in favor of partition.⁵¹

Resumption, 1937-1939

The Palestinians, in contrast, were furious about the report. Viewing the partition plan as unjust, they railed against the commission's proposal to expel hundreds-of-thousands of Arabs from Galilee and expressed their fear that the creation of a Jewish state—however small—would result in a dramatic increase in immigration that would lead, inevitably, to the Zionist state's expansion. More importantly, they also moved to resume the revolt. Beginning in the fall of 1937, Palestinian irregulars were once again harassing patrols, blowing up railroad tracks and viaducts, sniping at British soldiers, cutting telephone and telegraph wires, attacking Zionist settlements, and assassinating officials and suspected collaborators. For a time, they enjoyed substantial success. Indeed, by the summer of 1938, rebel forces controlled large swaths of rural Palestine and even several Arab towns.⁵²

Britain responded by returning to the aggressive tactics that had ended the first phase of the rebellion in 1936. It imposed martial law and followed by flooding the mandate with reinforcements; by 1939, as a result, the British army in Palestine numbered more than 100,000 troops—equivalent to one quarter of the population of Palestinian men. Those soldiers employed brutal counterinsurgency tactics to suppress the rebellion. They put Arab hostages on the fronts of trains and vehicles to discourage mine laying, blew up houses that belonged to rebels or their family members, arrested thousands without bringing them to trial, and summarily executed Palestinians suspected of engaging in insurrection.⁵³

⁵¹ Quote from Shlaim, *The Iron Wall*, 20–23; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 139.

⁵² Smith, *Palestine and the Arab-Israeli Conflict*, 144–45.

⁵³ Rashid Khalidi, *The Hundred Years' War on Palestine: A History of Settler Colonialism and Resistance, 1917–2017* (New York: Metropolitan Books, 2020), 44–45.

This time, moreover, the British did not limit themselves to a military response. Instead, using the assassination of a British official in September 1937 as a pretext, the mandate government went after the Palestinians' political leadership by declaring the AHC illegal and by arresting and deporting many of its members to the Seychelles Islands. The British executed their dragnet with characteristic skill; nonetheless, they failed to apprehend their main target: the increasingly militant grand mufti. Slipping out of Palestine at night by boat, he joined a handful of other prominent Palestinians in Lebanon, where, with little success, he tried to direct the rebellion.⁵⁴

The Zionists also reacted forcefully to the renewed violence. This new policy marked a dramatic shift away from the approach that they had adopted in response to the first phase of the revolt. Eager to cultivate a moderate image on the world stage, the Yishuv had been cautious at that time in its response to Arab attacks on Jewish settlements. A series of high-profile Palestinian terrorist incidents in October 1937 precipitated a dramatic change in that approach, however. From that point forward, the Labor Zionists abandoned their earlier restraint and fully embraced Jabotinsky's Iron Wall strategy. Using arms and training that the British had provided, the Haganah shifted from a strictly defensive strategy to "an 'aggressive defense' doctrine" that involved both frequent patrolling outside isolated settlements and retaliatory killings of Arab villagers. Irgun went much further. Beginning in November 1937, it began responding to Palestinian attacks by bombing bus stations, marketplaces, and commercial districts that Arabs frequented. These attacks were deadly. One in Haifa in July 1938 killed thirty-nine Palestinians and wounded seventy more, while another in Jaffa the following month killed twenty-four Arabs and injured thirty-nine. Thereafter, such violence—executed by both Arabs and Zionists—became commonplace throughout the mandate.⁵⁵

Thanks to the aggressive efforts of the British and the Zionists, the revolt finally petered out in 1939. It had been a bitter failure for the Arabs. By the time imperial forces had brought the rebellion to heel, 5,000 Palestinians had died, 10,000 had been wounded, and thousands more were languishing in prison. Meanwhile, valuable orchards had been levelled, crops had been torched or uprooted, and thousands of homes had been demolished. Perhaps worst of all, the Palestinians had become hopelessly divided along lines of class and faction as the revolt dragged on. Indeed, by late 1938, the rebellion had almost completely transformed from an anticolonial revolt into an Arab civil war replete with firefights, assassinations, and terrorist attacks.⁵⁶

It is nearly impossible to overstate the significance of the failure of the Great Arab Revolt to the outcome of the contest for Palestine. Put simply, it set the Zionists and the Arabs on very different trajectories. For its part, the Yishuv emerged from the rebellion

⁵⁴ Khalidi, 44–45.

⁵⁵ Segev, *One Palestine, Complete: Jews and Arabs Under the British Mandate*, 426–32; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 147–48.

⁵⁶ Khalidi, *The Hundred Years' War on Palestine*, 46–47.

much stronger than it had been at the insurrection's start. Largely unaffected by the conflict, its dynamic economy had maintained its rapid expansion while immigration had continued to fuel a Jewish population boom. Boosted by British weapons and training, meanwhile, the Yishuv's armed forces had taken a huge step in their transition from militia to conventional military. Finally, continuing their rapid development, Zionist political bodies had matured to the point that the Yishuv now constituted, in Rashid Khalidi's words, nothing less than a "para-state."⁵⁷

For the Arabs, on the other hand, the Great Arab Revolt had been an unmitigated disaster. Having played for broke, they had come up empty. The failed rebellion had grievously weakened the Arab community's economy, badly strained its social cohesion, and gutted its military capabilities. Politically, it had left the Palestinians hopelessly divided and had resulted in the death or exile of so many prominent figures that it created a debilitating leadership vacuum that the incompetent and rash grand mufti filled. The failure of the rebellion had been so total, in fact, that it rendered the Palestinians incapable of organizing either an effective resistance during World War II or a response to the Zionists when the question of who would control the territory was finally resolved in 1948. Thus, as Rashid Khalidi correctly notes, it was the revolt of 1936-1939 and not the war of 1948—a conflict that he considers to be no more than a "postlude"—that constituted the critical turning point in the contest for Palestine.⁵⁸

The White Paper of 1939

Astonishingly, in light of the defeat they had suffered, the Palestinians received one last opportunity to recover their position. They got this chance thanks to the rapidly worsening international situation. By the late spring of 1939, global war seemed inevitable. Germany had occupied Czechoslovakia in violation of an earlier promise to respect its borders, Italy had seized Albania, and Japan had begun casting covetous looks at Britain's valuable colonies in Southeast Asia. With hopes for global peace rapidly fading, London suddenly came to perceive its support for Zionism as a huge liability. Not only had it sparked a costly rebellion in Palestine that continued to divert scarce troops desperately needed elsewhere, but it had also alienated Arabs throughout the Middle East and thus threatened both access to critical oil supplies and the empire's control of the vital Suez Canal. Accordingly, London undertook a sweeping reappraisal of its approach to Palestine. The result was the release of a new policy statement: the White Paper of 1939. The latest in a long string of important documents regarding Palestine, it signaled a fundamental change in London's approach to the mandate. It imposed severe restrictions on land sales, limited immigration to 15,000 Jewish people per year until 1944—after which further immigration could only occur with Arab consent—and promised the establishment in ten years of a multiethnic, democratic,

⁵⁷ Khalidi, 47–54.

⁵⁸ Quote from Khalidi, *The Iron Cage*, 123; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 122.

government in the mandate that would be structured “in such a way as to ensure that the essential interests of each community are safeguarded.”⁵⁹

The White Paper came as a tremendous shock to the Zionists. It constituted a rejection not only of the partition proposal articulated in the Peel Commission Report but, far more importantly, of the Balfour Declaration itself: the legal bedrock on which Zionism rested. More immediately, it also threatened to severely restrict immigration at a time when growing Nazi violence made leaving Europe an imperative for Jewish people. The Labor Zionists consequently rejected its terms and responded by once more mounting an aggressive public relations campaign aimed at pressuring London into reverting to its earlier, pro-Zionist policy. This effort bore some fruit. Thanks to Britain’s increasingly single-minded focus on war preparation, however, it ultimately proved unsuccessful in reversing London’s new policy toward Palestine.⁶⁰

Characteristically, the Irgun went much further than the Labor Zionists in expressing dissatisfaction with the White Paper. Furious at what it considered to be London’s abandonment of its earlier commitment to a Jewish home in Palestine, the group responded by initiating a campaign of sabotage and terrorism against the British aimed at forcing them from the mandate. That effort only lasted until the start of the war, however, at which point the Irgun suspended its attacks out of respect for London’s need to focus on combatting Hitler. Not everyone in the Irgun supported that decision. Instead, a small faction split to form an even-more extreme splinter group called Lehi, or the Stern Gang. Single-mindedly focused on pushing London out of Palestine, it continued to launch terrorist attacks until British police killed its founder, Avraham Stern (1907-1942), in 1942.⁶¹

The terms of the White Paper were far more favorable to the Palestinians than any previous British policy had been. Indeed, its provisions might well have given them a genuine shot at preventing both the establishment of a Jewish state in the mandate and their eventual displacement. Nonetheless, the Palestinians dismissed the new British policy outright. What explains this choice? Why did they reject the opportunity that the White Paper seemed to present? Publicly, they did so because it failed to implement an immediate ban on immigration and because it pushed independence too far into the future. These were no doubt genuine concerns, but they did not represent the real reason that the Arabs rejected the White Paper. Instead, it was the Palestinians’ old problem, factionalism, that led them to renounce its terms. While many members of the AHC wanted to respond favorably to the new British policy, the unremitting hostility to the White Paper of the few remaining rebel bands still in the field and, more importantly, the strident opposition of the grand mufti tied the higher committee’s hands. As a result, the Palestinians turned down what British Colonial Secretary Malcolm

⁵⁹ Rogan, *The Arabs*, 206–7.

⁶⁰ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 158–59.

⁶¹ Morris, 158–59.

MacDonald (1901-1981) called a “golden opportunity” and refused to embrace the last chance—admittedly a slender one—that they would have to rescue themselves.⁶²

World War II

Picking Sides

The war that the British had feared finally came in September 1939. For the Labor Zionists who dominated the Yishuv’s leadership, the start of the conflict posed a difficult dilemma. On the one hand, there was no question that they would fight against Hitler’s brutal, anti-Semitic regime; on the other, they also bitterly opposed the White Paper of 1939 and did not want to take any action that might suggest acquiescence to its terms. How, then, could they help the British defeat the Nazi regime without simultaneously accepting—implicitly or explicitly—the terms of the hated White Paper? David Ben Gurion provided a way of resolving this predicament when he proposed that the Zionists adopt a bifurcated approach to the war and to Britain’s new policy position. “We shall fight the war against Hitler as if there were no White Paper,’ he declared in 1939, ‘and we shall fight the White Paper as if there were no war.’”⁶³

Ben Gurion’s two-track approach of contesting both Hitler and the British thereafter defined the mainstream Zionists’ approach to the war. Most focused on the first track. Seeking to help defeat Hitler, they either enlisted in Allied armies or joined the British-trained Jewish Brigade that served in Italy in 1945. Meanwhile, leaders like Ben Gurion pursued the second track through a skillfully executed diplomatic campaign that centered on winning Washington’s support for Zionism. Central to that effort was a meeting they held with a group of prominent Jewish Americans in New York City’s Biltmore Hotel in May 1942 at which they drafted a new Zionist policy statement. Known as the Biltmore Program, it formally replaced the movements’ earlier endorsement of partition with a call for the establishment of a Jewish commonwealth that would occupy the entire mandate.⁶⁴

Both tracks paid enormous dividends for the Yishuv. The creation of the Jewish Brigade and the enlistment of Zionists in Allied armies provided the Haganah with a cadre of experienced combat veterans who would later turn raw recruits into effective soldiers. Endorsed by both the Democratic and Republican parties in their election

⁶² Robert Fisk, *The Great War for Civilisation: The Conquest of the Middle East* (New York: Vintage, 2005), 359; Khalidi, *The Iron Cage*, 114–1245.

⁶³ James L. Gelvin, *The Israel-Palestine Conflict: One Hundred Years of War* (New York: Cambridge University Press, 2005), 119–22.

⁶⁴ Gelvin, 119–22; Walker Robins, *Between Dixie and Zion: Southern Baptists and Palestine before Israel* (Tuscaloosa: University Alabama Press, 2020), 131; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 166–67.

platforms in 1944, meanwhile, the Biltmore Program helped to secure critical American diplomatic support for the Zionist effort to overturn the White Paper.⁶⁵

The Palestinian leadership proved unable to match the Zionists' deft approach to the war. Once more, the main problem was the exiled grand mufti. Living in Baghdad during the first two years of the conflict, he made no effort to hide his fervent hope that Germany would win the war and, in the process, destroy the British Empire in the Middle East. He did not limit himself to anti-British diatribes, moreover. Instead, putting words into action, he helped to orchestrate Iraqi Prime Minister Rashid Ali's (1892-1965) ill-fated, pro-Axis coup in 1941, after which he spent the remainder of the war in Germany making propaganda broadcasts and recruiting Muslim POWs for service in Axis armies. These actions may have won al-Husayni accolades among anti-imperialists in the Arab world, but they also damaged the Palestinians' reputation and thereby weakened their case in the court of international public opinion—a venue that would prove to be of vital importance in the critical years immediately following the war.⁶⁶

The Holocaust

While al-Husayni was striving to recruit Muslims, his hosts and their many collaborators were profoundly if indirectly shaping Palestine's future through the Holocaust: the German-led effort to completely eradicate Europe's Jewish community. Resulting in the systematic murder of 5,600,000 to 6,900,000 Jewish people, it came perilously close to achieving its goal before the Allied conquest of Poland and Germany finally closed the death camps in 1945. Also known as the Shoah, the Holocaust powerfully influenced the contest for control of Palestine in three ways. First, it won enormous public sympathy for Zionism throughout Western Europe and especially in the pivotally important United States. Believing that the Holocaust amounted to a collective, civilizational moral failing, many Europeans and Americans came to conclude that helping the Jewish people establish a state in Palestine was the only way that the West could atone for its culpability in the genocide. Second, reinforcing the conviction that the Jewish people could never be truly secure until they had their own state, the Shoah hardened the Zionists' resolve and left them increasingly unwilling to compromise. Finally, the Holocaust created a huge population of Jewish refugees following the war who were, unsurprisingly, disinclined to live among people who had so recently collaborated in the Nazi effort to kill them. Many of them wished to follow the example of earlier generations and immigrate to the United States; however, a large proportion preferred the security that life in a Jewish state seemed to offer and demanded that they be permitted to move immediately to Palestine.⁶⁷

⁶⁵ Smith, *Palestine and the Arab-Israeli Conflict*, 175–76; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 188–89.

⁶⁶ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 165–67.

⁶⁷ Schulze, *The Arab-Israeli Conflict*, 10–11; Shlaim, *The Iron Wall*, 29; Little, *American Orientalism*, 77, 87.

Partition

The refugees and their aspiration to relocate to the mandate would play a central role in the Zionist leadership's postwar strategy. With the conclusion of the fighting in Europe in May 1945, the leadership of the Yishuv believed that it was on the cusp of finally achieving the Zionist dream. To secure it, they planned to first bring hundreds of thousands of refugees to Palestine after which they would declare the establishment of a Jewish state encompassing the entire mandate. Britain stood in the way of that dream, however. Eager to court the newly independent Arab states so as to retain its dominant position in the region, London continued the more even-handed policy toward Palestine that it had adopted with the White Paper of 1939. It made clear that it opposed the establishment of a Jewish state, and it limited immigration to the rate of 1,500 Jewish people per month—more than the White Paper had called for but only a fraction of the number of refugees that the Zionists were determined to bring to Palestine. London's approach thus put it on a collision course with the Labor Zionists and their plan to flood the mandate with immigrants and quickly declare a Jewish state.⁶⁸

Further alienating the Zionists from Britain, the conditions that had shaped David Ben Gurion's two-track response to the White Paper of 1939 no longer existed. Hitler had been defeated, and the Holocaust had afforded the Zionists new support in the United States that freed their movement from its earlier dependence on Britain. Accordingly, in the summer of 1945, Ben Gurion and the Labor Zionists abandoned the cautious, bifurcated strategy that had guided their movement since the outbreak of World War II in favor of a militant approach aimed at forcing London to quit Palestine as a prelude to the establishment of a Jewish state.⁶⁹

The Zionist Insurgency

The Haganah was late to the party, however. Furious about London's refusal to ease immigration limits in response to the Holocaust, Irgun and Lehi had already resumed their terror campaign against the British. Leadership changes in both organizations precipitated this new approach. Yitzhak Shamir (1915-2012), who would later serve as prime minister, revived the Stern Gang's dormant terror campaign against the British soon after he took charge of the group in 1943. In similar fashion, the Polish-Jewish immigrant Menachem Begin (1913-1992)—also a future prime minister—restarted the Irgun insurgency against the British just days after he assumed control of the group in 1944. Focusing their efforts on weakening Britain's civil authority, both organizations soon made their presence felt. Beginning in 1944, the Irgun carried out a spectacular string of bombings against government offices and police stations. Lehi, in contrast, focused on assassinations—including the murder of the British High

⁶⁸ Smith, *Palestine and the Arab-Israeli Conflict*, 186–87.

⁶⁹ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 173–74.

Commissioner to the Middle East, Lord Moyne (1880-1944), whom they targeted for his role in enforcing the blockade on immigration to Palestine.⁷⁰

The Haganah formally if secretly joined the insurgency in October 1945 when it entered into an alliance with Lehi and the Irgun called the Hebrew Rebellion Movement. It called for the Haganah to complement the revisionist groups' campaign of bombings and assassinations by engaging in non-lethal acts of sabotage. As such, while Lehi and Irgun forces pressed on with their terrorist attacks, Haganah units in late 1945 freed immigrants from detention centers, bombed coast-guard installations in retaliation for the seizure of ships carrying refugees, and sabotaged rail lines.⁷¹

Britain's aggressive reaction to the stepped-up insurgency heightened the standoff between London and the Zionists. The British initially responded to the rise in violence by deploying 100,000 soldiers to the mandate in early 1946. Seeking to pressure the Yishuv leadership into standing down, they followed by detaining hundreds of prominent Zionists including four members of the Jewish Agency at the end of June. This last move rankled Irgun's leadership. Committed to responding forcefully to any slight, it retaliated a few weeks later by detonating a huge bomb in the seat of the mandate government: the King David Hotel in Jerusalem. The resulting explosion was massive. Collapsing an entire wing of the hotel, it killed ninety-one people and wounded another one hundred.⁷²

Immigration and Diplomacy

The bombing was too much for the Labor Zionists. While the Irgun and Lehi pressed on with their terror attacks, the mainstream Zionists instead abandoned their sabotage campaign in favor of a return to a diplomatic approach. Their new strategy had two components. The first involved using the plight of Jewish Holocaust survivors languishing in camps in Europe to turn American public opinion against British policy in Palestine. To do so, the Zionists arranged for thousands of Holocaust survivors to travel to the mandate aboard chartered steamers—thereby compelling Britain either to permit them to disembark or to divert thousands of them to internment camps on the island of Cyprus. The second centered on solidifying Zionism's standing in the United States by presenting the movement as a moderate one. In support of that effort, the Labor Zionists abandoned the Biltmore Program's demand for a Jewish state in the entirety of the mandate and instead reembraced the more politically palatable idea of partition.⁷³

The decision to pursue a diplomatic strategy soon bore fruit. By the middle of 1946, the Labor Zionists' illegal-immigration campaign had made the predicament of Holocaust survivors suffering in refugee camps a major cause célèbre in the US and

⁷⁰ Rogan, *The Arabs*, 247–49.

⁷¹ Gelvin, *The Israel-Palestine Conflict*, 123.

⁷² Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 176–79.

⁷³ Shlaim, *The Iron Wall*, 25–28; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 176.

had turned public opinion in the US sharply against Great Britain. Among those who empathized with the Jewish survivors of the Holocaust was the president of the United States, Harry Truman (r. 1945-1953). Inclined for his whole life to root for the underdog, Truman was genuinely moved by the plight of the refugees and increasingly sought a favorable resolution of their situation. Empathy for the Jewish people was important in shaping Truman's position with regard to Palestine, but practical motives also played a critical role in spurring him to challenge Britain's approach to the mandate. A shrewd politician who had cut his teeth working for the Prendergast political machine in Kansas City, he keenly grasped the electoral benefits that would accrue to the Democratic Party in the 1946 elections in states like Illinois and New York that had substantial Jewish populations if he expressed his support for Zionist goals. "I have to answer to hundreds of thousands who are anxious for the success of Zionism," he told an aide. "I do not have hundreds of thousands of Arabs in my constituents." Truman was thus more-than-receptive to the Zionists' blandishments—particularly after they dropped their demand for the entire mandate. Accordingly, he announced on Yom Kippur in October 1946 that he favored both the immediate admission of 100,000 Holocaust survivors into Palestine and the partition of the mandate into Jewish and Palestinian states.⁷⁴

Truman's announcement that he supported the establishment of a Jewish state was the final straw for the British in Palestine. Demoralized by the Irgun and Stern Gang's ongoing terror campaign and now at odds with Washington, London began to give serious consideration to abandoning what had become a thankless and expensive commitment. Accordingly, after making one last futile effort to get the Zionists and Arabs to agree to a compromise solution, the British government announced in February 1947 that it was leaving Palestine the following year and dumped the seemingly insoluble problem onto the two-year-old United Nations (UN).⁷⁵

UNSCOP

Now tasked with resolving the Palestine crisis, the UN General Assembly responded by sending yet another fact-finding mission to the mandate. Called the United Nations Special Committee on Palestine (UNSCOP), it spent five weeks touring the mandate and conducting interviews in order to determine a path forward that was acceptable to all parties. Enjoying the considerable advantage of having bugged the committee's meeting rooms, the Zionists went to great lengths to make their case for partition. They presented their movement as moderate, compromising, and modern, and they worked diligently to ingratiate themselves with the UNSCOP representatives. They even went so far as to arrange for committee members to meet with Jewish settlers who spoke their native language. In contrast, in what one British official described as "exceedingly inept' diplomacy," the recently reconstituted AHC refused to present its

⁷⁴ Michael Oren, *Power, Faith, and Fantasy: America in the Middle East: 1776 to the Present*, Reprint edition (New York: W. W. Norton & Company, 2008), 483–92.

⁷⁵ Smith, *Palestine and the Arab-Israeli Conflict*, 192–93.

case and instead opted to boycott the committee entirely. It is true that officials from the independent Arab countries did present an Arab plan to the UNSCOP representatives; however, their proposal for a single, Palestinian-dominated state in which Jewish people would enjoy no political rights contrasted unfavorably with the Zionists' seeming reasonableness.⁷⁶

The committee released its report in September 1947. By an eight-to-three vote, it recommended that the UN partition the mandate into Palestinian and Zionist states bound together—as a function of practical necessity—in an economic union. To minimize dividing the Arab and Jewish communities, each state would have three, non-contiguous sections connected to each other at tangential crossing points. Jerusalem would join neither state. Instead, owing to its importance to the three Abrahamic faiths, it would become a UN-administered international trusteeship.⁷⁷

Britain was delighted with the report. Palestine had become a booby prize for London—an expensive obligation that sapped the army's morale and diverted resources desperately needed to rebuild the British economy following World War II. In addition, the detention of Jewish refugees seeking to flee persecution in Europe for the relative safety of the mandate had done serious damage to Britain's international reputation. Accordingly, London wasted no time in using the report's publication to announce that it was withdrawing from the mandate effective May 14, 1948.⁷⁸

Critics were quick to point out the many serious flaws and inequities inherent in the UNSCOP plan. The Arab League, which the Arab states had established in 1945 to coordinate their foreign policies, complained that while Zionists accounted for only one-third of the population in Palestine, the report had assigned them 56 percent of the mandate including much of the best farmland. They also noted that while the share of the territory allocated to the Palestinians possessed negligible numbers of Jewish people, the portion granted to the Zionist state contained 407,000 Arabs. Other observers focused on the Rube Goldberg-esque division of each state into three segments that enjoyed only tangential connections with each other. As they noted, the two crossing points linking the sections of the Jewish state would be coincident with those connecting the Palestinian territories—an arrangement that all-but guaranteed friction. More generally, they pointed out that the division of the proposed states into multiple segments gave each of them long, nearly undefendable borders, and left them, in one British academic's words, “entwined in an inimical embrace like two fighting

⁷⁶ Quote from William Roger Louis, *The British Empire in the Middle East, 1945-1951: Arab Nationalism, the United States, and Postwar Imperialism* (Oxford: Oxford University Press, 1984), 470; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 182–83.

⁷⁷ Gelvin, *The Israel-Palestine Conflict*, 124.

⁷⁸ Gelvin, 124.

serpents.” In part because of these criticisms, the formal partition proposal, Resolution 181, faced an uphill fight as it headed toward a vote in the UN General Assembly⁷⁹

Resolution 181, November 29, 1947

Ultimately, however, it was a pair of related challenges that posed the greatest threat to the passage of the resolution. First, the resolution required the support of two-thirds of the member states to pass. Second, it confronted substantial opposition among officials in the US Departments of State and Defense and in the Central Intelligence Agency who warned that support for Zionism could imperil America’s increasingly important economic interests in the Middle East and might even create an opening in the Arab world that Moscow could exploit. Highly influential in shaping American foreign policy, those officials could potentially turn the US against Resolution 181—a position that would sway many of Washington’s allies to vote against the partition plan and thus doom it to failure.⁸⁰

Fortunately for the Zionists, President Truman was unreceptive to the American foreign-policy establishment’s objections. With his eyes firmly locked on the upcoming 1948 presidential election, he rejected its recommendations. Instead, he not only instructed the US ambassador to the UN to vote in favor of partition but also aggressively strong-armed wavering states into supporting the resolution. His effort worked. Thanks to his pressure tactics, Resolution 181 narrowly passed on November 29, 1947, with thirty-three states in favor—including both the US and the USSR—thirteen opposed, and ten abstaining.⁸¹

Unsurprisingly, the Jewish and Arab communities in the mandate responded very differently to the resolution’s passage. To the Palestinians, the vote marked yet another demoralizing injustice imposed by imperialist Westerners that called on them to cede the land of their ancestors to a foreign invader. To the Zionists, in contrast, the passage of Resolution 181 was a moment of triumph. At long last, their dream of a Jewish state appeared to be at hand.⁸²

The Yishuv, particularly its Labor Zionist leadership, was well-prepared for independence. Politically, it had strong parties, a sense of unity, and a sophisticated parastate replete with a strong connection to its people and a suite of sophisticated and well-run ministries. Indeed, as the events of 1948 would show, its well-developed administrative structure needed only a name change and international recognition to

⁷⁹ Ussama Makdisi, *Faith Misplaced: The Broken Promise of U.S.-Arab Relations: 1820-2001* (New York: PublicAffairs, 2010), 192–93; Quote from Louis, *The British Empire in the Middle East, 1945-1951*, 369.

⁸⁰ Gelvin, *The Israel-Palestine Conflict*, 124–25.

⁸¹ Smith, *Palestine and the Arab-Israeli Conflict*, 198.

⁸² Hugh Wilford, *America’s Great Game: The CIA’s Secret Arabists and the Shaping of the Modern Middle East* (New York: Basic Books, 2013), 56–63; Smith, *Palestine and the Arab-Israeli Conflict*, 198.

become a modern state. The effort to establish a powerful army had been equally successful. While the Haganah may have suffered from a shortage of heavy weapons, it was nonetheless well-positioned to defend the Yishuv. It fielded 35,000 soldiers including many battle-hardened veterans of World War II, and it maintained factories capable of producing mortars, small arms, and ammunition. Perhaps most importantly, it possessed a highly capable general staff that was able to coordinate complex operations and to provide logistical support for units in the field.⁸³

The contrast with the Palestinians could hardly have been starker. Militarily, the Arabs' modest militia forces possessed only a handful of modern weapons and had no cadre of World War II veterans on which they could call. Worse, the Palestinians did not merely lack a coordinating body comparable to the Haganah's general staff but in fact enjoyed no central direction of their military forces whatsoever. Politically, meanwhile, the AHC did not function like a parastate along the lines of the Yishuv's government. Instead, it was an organization largely focused on infighting and factional maneuvering—one that lacked functionally defined ministries and that was incapable of offering the sort of effective leadership and oversight that are the hallmarks of modern states.⁸⁴

Even when the AHC made wise decisions, its endemic internal dissension ensured poor outcomes. In 1946, for example, it established the Arab Office to serve as a quasi-foreign ministry. Staffed with highly educated and professional people, it showed great promise and appeared poised to give the Palestinians an effective international presence for the first time. Unfortunately for the mandate's Arab population, however, it soon fell victim to factionalism—in this case, the grand mufti's insistence that he personally control all of the Palestinians' quasi-governmental institutions. As a result, by mid 1947—a peculiarly fateful moment in the contest for Palestine—the Arab Office had been weakened to the point that it was no longer able to provide the effective international representation that the Palestinians so desperately needed.⁸⁵

The Palestinians' abject lack of preparedness in 1947 had multiple sources but was primarily a legacy of the events of the interwar period and the manifestly unequal nature of the mandate. From the start, the "iron cage" that London had imposed on them in the early 1920s had prevented the Palestinians from establishing the foundations of an effective government or from gaining the international legitimacy that the mandate charter had accorded the Zionists. Britain's defeat of the Great Arab Revolt in the 1930s had been even more damaging to the Palestinians. It left them bitterly divided and desperately short of the experienced troops, arms, and capable military commanders that they would need to counter the Zionists in 1948. As a result, as the moment of reckoning approached, the Palestinians found themselves without either

⁸³ Smith, *Palestine and the Arab-Israeli Conflict*, 200; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 188.

⁸⁴ Khalidi, *The Hundred Years' War on Palestine*, 62–68.

⁸⁵ Khalidi, 62–68.

effective governing institutions or a capable military—deficits that would leave them thoroughly outmatched by the Zionists.⁸⁶

The First Arab-Israeli War, 1947-1949

The final contest to determine whether the Zionists or the Palestinians would dominate the territory began the day after the UN vote and continued until January 1949. It had two phases. The first was an unofficial civil war between the Zionists and Palestinians that occurred between November 1947 and May 1948. The second, in contrast, was a conventional conflict between the newly declared Zionist state of Israel and the militaries of five Arab states. Both would end in decisive Zionist victories that took place between May 1948 and January 1949.

Civil War, November 30, 1947-May 14, 1948

Physical geography and Zionist strategic choices gave the Palestinians the upper hand during the early part of the civil war. Focused at that time on reorganizing the Haganah into a conventional military force, the Zionists initially adopted a defensive strategy that centered on securing population centers and on keeping isolated settlements and the 100,000 Jewish people who lived in West Jerusalem supplied. This strategy would pay significant benefits in the long run, but, in the immediate term, it played to the Arabs' strength. The Palestinians were joined in January 1948 by Fawzi al-Qawuqji's (1890-1977) Arab League-sponsored Arab Liberation Army (ALA), and were able to take advantage of their many well-located towns and villages to interdict the roads on which Zionist supply trucks travelled. With most Haganah units tied down either defending scattered settlements or undergoing reorganization, these operations produced a number of early Arab victories and succeeded in preventing Zionist convoys from bringing supplies to Jerusalem—a situation that soon left the Jewish section of the city worryingly short of food. Even more concerning for Zionist leaders like Ben Gurion, the Haganah's lack of military success also created a dangerous diplomatic environment for the Zionists. Most notably, in an effort to end the violence, the US State Department began working on a new proposal that called for shelving the partition plan in favor of a UN trusteeship over the entire mandate.⁸⁷

By March, as a result, the Yishuv's leaders were desperate. As they well understood, their hopes of establishing a Jewish state required not only that they promptly resupply Jewish West Jerusalem but also that they secure a quick military victory that could stave off the trusteeship proposal. Accordingly, they moved to abandon the defensive orientation they had adopted at the start of the civil war in favor of a mandate-wide conventional offensive called Plan Dalet slated to begin in early April. Plan Dalet was a bold and ambitious operation designed to create an enlarged and contiguous Zionist state. It called for the now-fully reformed Haganah to secure

⁸⁶ Khalidi, 62–70.

⁸⁷ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 196–204.

lines of communication to outlying settlements and Jerusalem before seizing control of those Palestinian towns and villages that lay within the territory that the UN had assigned to the Zionists. Once those areas had been secured, the Haganah would mount an offensive aimed at knitting the three Jewish territories into a unified whole by taking control of land that Resolution 181 had allocated to the Arabs. Importantly, it authorized field commanders to destroy hostile Palestinian villages.⁸⁸

Plan Dalet was an enormous success right from the start. While the Palestinians' decentralized command structure and hit-and-run attacks had been well-gearred to the early part of the conflict, they proved poorly suited to the conventional warfare that characterized the Haganah offensive. As a result, Zionist units quickly achieved their objectives in all parts of the mandate—particularly on the Jerusalem front. There, they focused on capturing the strategic hilltop-village of al-Qastal, which commanded the road to the Holy City. In heavy fighting that saw the settlement change hands several times, Zionist troops not only secured control of the village in early April, but, in the process, also killed the Palestinians' only competent military commander, Abd al-Qadir al-Husayni (1907-1948).⁸⁹

An event of even greater significance occurred in the nearby village of Deir Yasin on the same day that Zionist forces took al-Qastal. Though the settlement had not taken part in any of the fighting, Lehi and Irgun decided to launch a joint operation to seize control of it so as to ensure the security of the road to Jerusalem. Despite heavy resistance and substantial Zionist casualties, they managed to secure the village by noon. It was what happened next that made the attack on Deir Yasin so critical to the outcome of the war. Following the battle, the attackers brutally massacred 115 of the survivors—the majority of whom were women, children, or the elderly. Some were paraded through West Jerusalem before they were killed; others had their mutilated corpses shoved down the village's wells. While the incident would soon become infamous, it was not entirely unique in the fighting for control of the mandate. Indeed, just four days after the massacre, Palestinian irregulars retaliated by ambushing a Zionist medical convoy outside Jerusalem and killing seventy nurses and doctors.⁹⁰

The capture of al-Qastal, the loss of Abd al-Qadir al-Husayni, and, especially, the massacre at Deir Yasin profoundly altered the course of the war. Permanently tipping the military balance on the Jerusalem front in favor of the Zionists, Al-Husayni's death and the Haganah victory at al-Qastal were of enormous military significance. From that point forward, supplies could flow unimpeded to Jewish West Jerusalem—a situation that greatly eased the Yishuv's heretofore-challenging strategic situation.⁹¹

In contrast, the massacre at Deir Yasin was of little military consequence, but proved to be of decisive psychological significance. Immediately following the incident,

⁸⁸ Morris, 204–6.

⁸⁹ Khalidi, *The Iron Cage*, 132–34.

⁹⁰ Rogan, *The Arabs*, 258–59; Smith, *Palestine and the Arab-Israeli Conflict*, 202–3.

⁹¹ Khalidi, *The Hundred Years' War on Palestine*, 72–74.

both the Arab states and the Jewish terrorist organizations that had perpetrated the massacre aggressively publicized the incident—the former to damage the Zionists' reputation, and the latter to weaken the Palestinians' will to fight and to encourage them to flee. Lehi and Irgun received a vastly greater return on their effort than did the Arab states. While news of the attack on Deir Yassin did tarnish Israel's reputation, its primary impact was to so thoroughly demoralize the Palestinian people that they abandoned all organized resistance to the Zionists and fled for safety. Most notably, when the Haganah and Irgun attacked Arab cities such as Haifa, Jaffa, and Tiberias during the latter stages of Plan Dalet, tens of thousands of Palestinians, fearing for their lives, took to the road in desperate flight. As a result, by mid-May at least 300,000 had become refugees.⁹²

With the success of Plan Dalet as a backdrop, David Ben Gurion formally declared the birth of the Jewish state—which the Zionists named Israel—on May 14, 1948, the date of Britain's official departure from the mandate. The Yishuv was well prepared for this momentous day. In sharp contrast to the Palestinians, who had failed to establish effective governing bodies and whose society had crumbled under the weight of Plan Dalet, the Zionists had succeeded in building strong administrative structures that they could easily convert into the institutions of a modern state. The Jewish National Council seamlessly transformed into the provisional government, and the Haganah, with little more than a name change, became the new state's military: the Israeli Defense Force (IDF). Meanwhile, Yishuv leaders had skillfully laid the groundwork for Israel's formal diplomatic recognition. As a result, almost immediately after Ben Gurion declared the establishment of Israel, the world's two most powerful countries—the United States and the Soviet Union—extended diplomatic recognition.⁹³

Still, the day was not an entirely sunny one. In the wake of Plan Dalet, the Arab League states had decided to reverse an earlier decision to send only volunteers and arms in support of the Palestinians and declared that they were preparing to mount a coordinated assault against the new state. This was grim news for the Israeli government, and it cast the future of the new state in doubt. Indeed, when a concerned Ben Gurion asked IDF leaders for an assessment of the new strategic situation, they informed him that Israel had no better than fifty-fifty odds of surviving a concerted Arab League attack.⁹⁴

Conventional War, May 15, 1948-January 9, 1949

Thus, just as the civil war between Zionists and Palestinians was coming to an end, the first formal Arab-Israeli War began. The conflict was a start-and-stop affair that involved three rounds of often-fierce fighting interrupted by two lengthy, UN-sponsored ceasefires during which diplomats desperately sought a peaceful resolution of the conflict. Like the civil war that preceded it, the conventional contest between Israel and

⁹² Smith, *Palestine and the Arab-Israeli Conflict*, 203.

⁹³ Shlaim, *The Iron Wall*, 34–35; Oren, *Power, Faith, and Fantasy*, 498–500.

⁹⁴ Oren, *Power, Faith, and Fantasy*, 498–500; Shlaim, *The Iron Wall*, 34–35.

the Arab states ended in a decisive victory for the Zionists—one that resulted in the permanent expulsion from Israel of hundreds of thousands of Palestinian refugees.⁹⁵

Despite its eventual victory, Israel did not seem to be in a commanding position during the first round of fighting. The day after British troops evacuated the mandate, forces from Egypt, Transjordan, Syria, Lebanon, and Iraq—later joined by token contingents from Yemen and Saudi Arabia—invaded Palestine just as their governments had promised. The IDF deployed units to contain these advances in the south and in the Galilee while it launched a concurrent offensive aimed at seizing East Jerusalem and its holy sites. In response, Transjordan's powerful Arab Legion occupied positions in and around the Old City. Disciplined and well equipped, it stopped the IDF assault on East Jerusalem cold and then defeated a series of fierce attacks on the nearby fortified hilltop town of Latrun, which commanded the road between the Holy City and Tel Aviv.⁹⁶

Exhausted by the fighting, both sides agreed to a United Nations-brokered ceasefire that went into effect on June 11, 1948. During the truce, the UN imposed an arms embargo on the region and tasked the Swedish diplomat Count Folke Bernadotte (1895-1948) with finding a way to end the fighting. A skilled mediator, he put forward a proposal in late June that centered on jettisoning the UN partition plan in favor of a federated union of Palestine and Transjordan that had autonomous Jewish and Arab sections. According to his proposal, Transjordan would annex the territory near Jerusalem that Resolution 181 had assigned to the Palestinians as well the Negev Desert in the south that the UN had allotted to the Jewish state; the Arab territory would thus be a contiguous block. Meanwhile, Israel would gain control of the fertile Western Galilee region that the original partition plan had granted to the Palestinians and that its forces had already overrun. As a result, the Jewish state would also possess a unified territory. It was a sound idea given the circumstances—and a much-more realistic proposal than the UNSCOP plan. Nonetheless, setting a pattern that would torpedo other peace proposals for decades to come, both sides rejected it out of hand in hopes that they could instead achieve a total victory.⁹⁷

In the meantime, Israel took advantage of the ceasefire to strengthen its military position. It smuggled in previously purchased heavy weapons, integrated new recruits into its formations, and completed a secret arms purchase with Czechoslovakia. By early July, as a result, the balance of forces had tipped dramatically in favor of the IDF.⁹⁸

⁹⁵ Shlaim, *The Iron Wall*, 35, 56.

⁹⁶ Avi Shlaim, "Israel and the Arab Coalition in 1948," in *The War for Palestine: Rewriting the History of 1948*, ed. Eugene L. Rogan and Avi Shlaim (New York: Cambridge University Press, 2001), 89–94.

⁹⁷ Schulze, *The Arab-Israeli Conflict*, 16.

⁹⁸ Schulze, 16; Smith, *Palestine and the Arab-Israeli Conflict*, 205.

The second round of fighting reflected that shift. Taking advantage of an Egyptian violation of the truce, the IDF used its new weapons to launch offensives on all fronts. Aimed at securing Jerusalem, the main effort faltered in the face of continued stiff resistance on the part of the Arab Legion. The enormous success that the Israelis enjoyed on all other fronts more than offset that failure, however. IDF troops simultaneously took the undefended towns of Lydda and Ramle southeast of Tel Aviv, drove the Lebanese and Syrians back in the north, and badly bloodied the Egyptians in the south. As a result, by the time the second round of fighting had stopped, Israel had dramatically improved its military position and had—by seizing control of land that the UN had assigned to the Palestinians—substantially increased the amount of territory that it possessed.⁹⁹

With both sides once again looking to pause the fighting, the UN succeeded in imposing a second ceasefire on July 18. Taking advantage of the lull, Bernadotte reworked his earlier proposal in hopes of finding a formula that could serve as the basis for a lasting peace. Announced on September 16, his new plan followed the broad outlines of his previous proposal in terms of territorial adjustments but discarded the federal structure in favor of two, distinct states: Israel and an expanded Transjordan. Importantly, it also included a provision guaranteeing that the hundreds of thousands of Palestinians who had recently fled their homes would enjoy the right of return. Bernadotte's proposal was once again a reasonable one given the circumstances. It offered both a means of ending the immediate conflict and a way of establishing a durable peace. However, it was also highly unpopular among both Arabs and Zionists—especially the Israeli revisionists, who viewed it as a threat to their maximalist territorial goals. Indeed, they perceived it as such a serious risk that Lehi's leaders—including future Prime Minister Yitzhak Shamir—had Bernadotte assassinated just a day after he announced the new plan, thereby rendering his revised peace proposal stillborn.¹⁰⁰

The final round of fighting that began on October 15 confirmed Israel's dominant position. It began in the south. Seeking to take undisputed possession of the Negev Desert, IDF forces drove the now-thoroughly demoralized Egyptian army back toward the border. The Israelis then followed with Operation Hiram, which pushed Lebanese troops and Qawuqji's ALA out of central Galilee and thereby secured full control of the northernmost territory that Resolution 181 had allotted to the Palestinians. In the process, IDF troops expelled the Palestinian civilians who lived in the area, with the result that the land Israel had acquired was, in the words of one officer, "'clean' and 'empty' of Arabs." Finally, after one last Israeli offensive that pushed the Egyptians across the frontier, the UN secured a ceasefire on January 7, 1949, that brought the 1948 War to an end. By that point, the scale of Israel's victory was undeniable. The Zionist state had not merely survived the conflict that accompanied its birth, but had, in

⁹⁹ Shlaim, "Israel and the Arab Coalition in 1948," 94–96.

¹⁰⁰ Rogan, *The Arabs*, 266.

the process, secured control of the vast majority of the territory that had composed the mandate of Palestine.¹⁰¹

Armistice Agreements

Peace talks commenced on January 13, 1949, on the island of Rhodes. With the help of Bernadotte's successor, the African American statesman and civil-rights leader Ralph Bunche (1904-1971), Israel concluded a series of bilateral understandings with the Arab League states that brought the fighting to an end. These agreements did not formally terminate the hostilities between Israel and the Arab states, however. Instead, since the documents were ceasefires rather than peace treaties, Israel remained in a technical state of war with its neighbors. As such, the Jewish state could not trade with those countries, exchange ambassadors with them, or establish civil-aviation links between its airports and theirs; Egypt even barred ships bound for Israel from transiting the Suez Canal.¹⁰²

The armistice agreements also formalized the division of the former mandate between Israel, Egypt, and Jordan. Israel retained the territory it held at the end of the fighting, which totaled 78 percent of the mandate rather than the 56 percent that the UN had allotted it. For its part, Egypt assumed control over the twenty-five-mile-long, five-mile-wide ribbon of territory that extended from the Egyptian border northward along the Mediterranean that has been known since 1948 as the Gaza Strip. Finally, Abdallah's kingdom—renamed Jordan in 1949—formally annexed what came to be called the West Bank: a slightly reduced version of the block of territory around Jerusalem that the UN had originally assigned to the Palestinians. Finally, Israel and Jordan split the city of Jerusalem proper between them, with West Jerusalem joining the Jewish state, and the Old City, which included the Haram al-Sharif, falling under Jordanian rule.¹⁰³

The Birth of the Palestinian Refugee Crisis

The talks may have ended the fighting, but they did not resolve the situation of the refugees that the war had created. During the conflict, the overwhelming majority of the Palestinians who lived in the territory that became Israel had fled their homes. As a result, while a small minority of them, roughly 160,000, remained in Israel at the time the war ended, the vast majority, some 700,000, found themselves living as displaced persons either in the West Bank, the Gaza Strip, or in the adjoining Arab states. They would never see their homes again. Instead, thanks to Israel's adamant refusal to permit them to return after the fighting had ceased, the displaced Palestinians remained refugees following the conflict.¹⁰⁴

¹⁰¹ Shlaim, "Israel and the Arab Coalition in 1948," 98–100; Shlaim, *The Iron Wall*, 56.

¹⁰² Shlaim, *The Iron Wall*, 43–49.

¹⁰³ Shlaim, 43–49.

¹⁰⁴ Khalidi, *The Hundred Years' War on Palestine*, 82–83.

What accounted for this situation? How could Israel justify its refusal to permit the Palestinians the right to return to their homes—a right enshrined in international law? Ben Gurion's government advanced two claims in defense of its position. First, it asserted that the refugees' flight was, in itself, a hostile act in support of Arab military operations against Israel; as such, the Palestinians had forfeited any right to return to their homes. Second, after 1950, it argued that the Arab states' retaliatory expulsion of 300,000 Middle Eastern and North African Jews obviated Israel of any obligation to contribute to a solution to the Palestinian refugee crisis. After all, it had its own refugee situation to deal with.¹⁰⁵

The treatment that the Palestinians received from the Arab states to which they had fled was not much better. With the notable exception of Jordan, which granted citizenship to refugees living in its territory, states such as Lebanon, Egypt, and Syria imposed severe restrictions on the Palestinians. They confined them to the refugee camps that the UN had set up, tightly controlled their movements, and denied them the right to employment.¹⁰⁶

Why did the Arab states refuse to absorb the displaced Palestinians after Israel had made clear that it was not going to permit them to return to their homes? To a substantial degree, they did so out of fear that the Palestinians might cause unrest either by engaging in political activity or by competing for jobs. More cynically, countries such as Egypt refused to integrate the Palestinians into their societies because they perceived a propaganda advantage in keeping them in the camps: images of destitute refugees dwelling in tents provided them with stark visual evidence that could be deployed to damage Israel's reputation in the increasingly important court of public opinion. Thus, despite a UN resolution explicitly calling for the refugees to have the right to return to their homes and despite Arab claims of solidarity with the Palestinians, hundreds of thousands of refugees and millions of their descendants would languish for decades in dismal, overcrowded refugee camps in Egypt, Syria, Jordan, Lebanon, the West Bank, and the Gaza Strip. Many remain in them to the present day.¹⁰⁷

The Historical Debate: The New Historians

The cliché that history is written by the victors accurately characterizes the enduring historiographical influence of the unabashedly pro-Zionist traditional interpretation of Israel's founding. First articulated by Zionist leaders immediately after the 1948 War, that perspective has thoroughly dominated early Western and Israeli views of the establishment of the Jewish state and continues to shape popular understandings to the present day. Its broad outlines are straightforward. It maintains that the Zionists hoped to peacefully establish a small state in the Holy Land that could

¹⁰⁵ Smith, *Palestine and the Arab-Israeli Conflict*, 230–32.

¹⁰⁶ Shlaim, *The Iron Wall*, 51–52.

¹⁰⁷ Edward W. Said, "Afterword: The Consequence of 1948," in *The War for Palestine: Rewriting the History of 1948*, ed. Eugene L. Rogan and Avi Shlaim (New York: Cambridge University Press, 2001), 209; Shlaim, *The Iron Wall*, 51–52.

serve as a refuge for the war-weary Jewish people. Though the Zionists made numerous efforts to arrive at a mutually beneficial understanding with the Palestinians, the Arabs remained obdurately opposed and rejected their entreaties out of hand. Instead, military forces from seven Arab states invaded Israel just after it declared its independence in an effort to kill the Jewish state in the cradle. In support of that effort, Arab leaders ordered hundreds of thousands of Palestinians to leave their homes so as to ensure that the invading armies could maneuver freely against the IDF. Israel appeared to be doomed in the face of this onslaught. Despite being badly outnumbered, however, the plucky Zionists somehow managed to pull off a miraculous, David-versus-Goliath victory against the attackers that secured their new state. Meanwhile, once the fighting had ended, the Israelis made clear that they were willing to make substantial concessions in order to secure a just peace for all parties; unfortunately, the Arabs remained as implacably hostile as ever and refused to negotiate. Thus, despite its yearning for peace, Israel remained in a technical state of war with the Arab countries for the next few decades.¹⁰⁸

Based on memoirs, interviews with participants, and, in many cases, lived experience, this account would quickly come to dominate Western and Israeli understandings of the events of 1947-1948. In Israel, people from across the political spectrum accepted it as an accurate account of their nation's founding, and it served as the basis for the country's secondary-school history curriculum. In America, it informed the storyline of Leon Uris's 1958 novel, *Exodus*, as well as the plot of the film adaptation that director Otto Preminger released in 1960. Both the movie and the novel were huge hits. Uris's book was the most successful work of fiction since Margaret Mitchell's *Gone with the Wind* while Preminger's adaptation was the third highest grossing movie of 1960. As such, they firmly established the traditional narrative as the dominant understanding of Israel's founding in the United States.¹⁰⁹

The traditional interpretation stood largely without question in Israel and the West for several decades after the 1948 War. Beginning in the 1980s, however, a group of younger Israeli scholars known as the New Historians began to raise questions about its accuracy. Drawing on recently declassified Israeli documents,¹¹⁰ New Historians like Avi Shlaim, Simha Flapan, Benny Morris and Ilan Pappé argue not only that the conventional narrative—or, as Shlaim calls it, the “heroic-moralistic version”—grossly distorts the events of 1948, but that the people who popularized it did so as part of a deliberate effort to create a sense of national cohesion in Israel and to secure international support for the Jewish state. These scholars challenge nearly every aspect of the conventional narrative, but their most pointed differences with the traditionalists

¹⁰⁸ Avi Shlaim, “The War of the Israeli Historians,” *Annales* 59, no. 1 (2004): 161–69.

¹⁰⁹ Melani McAlister, *Epic Encounters: Culture, Media, and U.S. Interests in the Middle East since 1945, Updated Edition* (Berkeley: University of California Press, 2005), 159–65; Shlaim, “The War of the Israeli Historians,” 166.

¹¹⁰ Israel, to its credit, follows the British thirty-year rule and has thus long since declassified diplomatic documents pertaining to the 1948 War. None of the Arab states follows this procedure. See Shlaim, “The War of the Israeli Historians,” 161–62.

center on three key topics: the relative disparity in forces, the cause of the Palestinian refugee crisis, and the failure of the combatants to conclude peace and normalized relations following the war.¹¹¹

The New Historians argue, first, that far from being at a military disadvantage, Israel held the upper hand throughout the fighting. They note, for example, that the IDF enjoyed a substantial advantage in firepower—particularly after it received a large shipment of Czech arms during the first ceasefire. They also demonstrate that Israel not only possessed a numerical advantage from the start of the fighting, but that the balance of forces tipped further and further in the IDF's favor as the war progressed. As Avi Shlaim points out, its strength rose from 35,000 in May 1948 to 65,000 in July to 96,000 in December. In contrast, the Arab states deployed only 25,000 troops in Palestine at the start of the war and lacked the logistical capabilities to match the IDF's growth.¹¹²

The Arabs were not just outnumbered and outgunned, moreover, but were sharply divided among themselves. As Shlaim notes, Egypt, Syria, and Saudi Arabia suspected that Abdallah was scheming with the Israelis to annex the West Bank as a first step toward the creation of a greater Syria under his rule; accordingly, they sent forces into Palestine less to attack Israel than to check his ambitions. Their suspicions were correct. Meeting secretly with future Israeli Prime Minister Golda Meier (1898-1978) before the war, Abdallah had promised to respect Israeli territory in exchange for Meier's pledge that Israel would accept his annexation of the West Bank. As Shlaim points out, it is true that the deal broke down over the question of which state would possess Jerusalem and that Israel and Transjordan subsequently engaged in some of the heaviest fighting of the war on the outskirts of the Holy City. Nonetheless, in the broadest sense, the outline of the agreement held: Israel accepted Transjordan's annexation of the West Bank, and Abdallah kept his troops out of Israeli territory. Happy to have one of his Arab rivals taken down a notch, the emir even permitted the IDF to shift the bulk of its forces to the Egyptian front in October 1948 unmolested. Thus, Israel benefited not merely from having more and better armed troops, but also from the fact that it faced divided enemies that were as focused on weakening their rivals as on attacking the Jewish state.¹¹³

Second, the New Historians take issue with the conventional interpretation's assessment of the refugee crisis. While acknowledging some instances in which Arab leaders encouraged Palestinians to flee the area of battle, they maintain that those efforts had only a very minimal impact. Instead, they argue that it was a combination of Israeli attacks on Arab cities during Plan Dalet, orders instructing IDF commanders to expel Palestinians from villages—often couched in euphemistic phrases such as one that encouraged them “to assist the Arabs to depart”—and, above all, the rapid spread

¹¹¹ Avi Shlaim, “The Debate about 1948,” *International Journal of Middle East Studies* 27, no. 3 (1995): 287–88.

¹¹² Shlaim, “Israel and the Arab Coalition in 1948,” 79–101.

¹¹³ Shlaim, 79–101.

of news about the Deir Yasin massacre that impelled hundreds of thousands of Palestinians to flee beginning in April 1948. In other words, they assert that it was Zionist actions rather than Arab directives ultimately produced the Palestinian exodus.¹¹⁴

Finally, the New Historians debunk the idea that Arab intransigence was solely responsible for the inability of the two sides to establish a lasting peace and normalized relations following the war. Relying on Israeli diplomatic documents, they instead demonstrate that the Arab states bucked popular anti-Israeli sentiment and made a series of conciliatory peace proposals following the conclusion of the fighting. For example, they show that Egypt's king Faruq (r. 1936-1952) had sent out peace feelers as early as September 1948 and that Abdallah had long signaled his openness to a deal. They also point out that Syrian Colonel Husni Zaim (r. 1949), who had seized power in March 1949, offered not only to recognize Israel but to permanently resettle 300,000 Palestinian refugees in his country in exchange for a small strip of territory and control of the eastern half of the Sea of Galilee. Still, as substantial as this offer was, Ben Gurion rejected it out of hand. Believing that Israel had time on its side and that it had gotten everything it needed through the armistice arrangements, he refused to countenance yielding even a fraction of an acre of Israeli territory in exchange for peace. "I am not in a hurry and I can wait ten years," he told one reporter, "We are under no pressure whatsoever." As a result, the two sides failed to secure a durable peace at the conclusion of the 1948 War.¹¹⁵

Peace?

More problematically, thanks to the increasingly inflexible positions that each side adopted following the conflict, stability and normalized relations became less rather than more likely as the years passed. For the Israelis, hardening attitudes were a function of recent Jewish history. The contrast between the Jewish experience in the Holocaust and the 1948 War had driven home the lesson that they could never again let hostile acts go unpunished; as a result, both revisionists and Labor Zionists alike redoubled their commitment to the Iron Wall strategy of securing peace through overwhelming strength. For Arab leaders, meanwhile, the hardening of attitudes was a product of public opinion. Repeatedly, the Arab people had made clear that, at a minimum, there could be no peace with Israel that did not involve the return of the Palestinians to their homes—a possibility that the Israelis had declared, with equal insistence, to be off the table. Unsurprisingly, given the persistence of these manifestly incompatible positions, the two sides would remain locked in a state of war for decades. Indeed, Israel only ended its formal state of war with Egypt in 1979 and with Jordan in 1994; remarkably, it remains in a technical state of war with Syria and Lebanon at the time of this writing.¹¹⁶

¹¹⁴ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 257.

¹¹⁵ Shlaim, *The Iron Wall*, 40–42, 47–55.

¹¹⁶ Shlaim, 43–49.

The fact that the *Nakba* and the establishment of Israel had thoroughly destabilized the Arab world did nothing to improve the prospects for peace. Indeed, the totality of the setback—and its startling contrast with the breezy confidence of prewar government propaganda—spurred a vast political sea change in the Arab states. It swept new regimes into power, fueled the dramatic growth of radical pan-Arab nationalism, contributed to the precipitate decline of Britain's position in the Middle East, and threw open the door to significant Soviet involvement in the region.¹¹⁷

Conclusion

Few grasped in the years leading up to 1948 that Israel's creation would have such a disruptive impact on the Middle East. Perhaps the most far sighted among them was a young, British-born member of the AHC's Arab Office named Albert Hourani (1915-1993), who would go on to be one of the greatest historians of the region. While giving testimony to a fact-finding group in the late 1940s, he predicted that the establishment of a Jewish state in Palestine would completely destabilize the region. It would, he argued, “involve a terrible injustice and could only be carried out at the expense of dreadful repressions and disorders, with the risk of bringing down in ruins the whole political structure of the Middle East.” Hourani's statement proved painfully prophetic—accurately predicting the sweeping changes and embrace of radicalism that the region would experience over the course of the 1950s and 1960s. It is to this tumultuous period and to the man, Gamal Abdel Nasser (r. 1954-1970), who would dominate it that we shall now turn.¹¹⁸

¹¹⁷ Shlaim, 43–49.

¹¹⁸ Khalidi, *The Hundred Years' War on Palestine*, 61–62.

Chapter Fourteen: The Age of Nasser, 1949-1967

Introduction

By the late 1940s, the core of the Arab Middle East—Egypt, the Levant, and Iraq—had secured formal independence from the European imperial powers that had controlled the region during the interwar period. Having finally achieved the long-sought goal of sovereignty, the people of those states expected rapid economic growth, political populism, unity, the defeat of Israel, and an assertive, independent foreign policy. Despite a startling degree of instability in the region, they appeared to be well on their way to attaining those expectations under the dominant political figure of the age: Egypt's charismatic president, Gamal Abdel Nasser (r. 1954-1970). From the mid 1950s through the late 1960s, the Egyptian strongman stood astride the Middle East like a colossus. To the delight of the Arab people, he defied the imperial powers, imposed sweeping economic and political reforms, threatened Israel, and even, for a brief period, seemed poised to erase the frontiers that Europe had imposed on the region following the First World War. It was an exhilarating period for nationalists—a time of hope and anticipation for a brighter future for the Arab people.

Framing the Postwar Middle East

The Middle East that Nasser dominated was, even by the standards of the times, a complex and volatile place. Between 1949 and 1967, the region experienced a dizzying array of coups, assassinations, purges, covert operations, conspiracies, terrorist incidents, interstate conflicts, Great Power interventions, radical economic development schemes, and diplomatic crises. What explains the Middle East's volatility during this period? Why did it experience such a high degree of instability? Many local, regional, and international forces played a role in shaping the region's development in the two decades following World War II. Ultimately, however, it was a combination of four overlapping currents—decolonization, the ideology of Arab nationalism, the Cold War, and the Arab-Israeli conflict—that, collectively, constituted the wellsprings of regional instability during this period.

in

Decolonization

The first of these began in the interwar period. Decolonization is an often-contested process by which subject peoples used diplomacy, force, and collective popular actions to challenge and then dissolve their subordinate relationships with the imperial powers. It played a central role in framing the development of the Middle East in the 1950s and 1960s. Indeed, decolonization shaped nearly all of the conflicts and political disturbances that the region experienced during those years. That it did so is perhaps surprising in light of the fact that the Arab countries at the core of the region—Iraq, Syria, Lebanon, Jordan, and Egypt—had all achieved political independence by 1946. Thus, they exercised the prerogatives of sovereign members of the international

community such as maintaining diplomatic relations with other states and holding seats in the United Nations (UN).

Appearances, however, very much obscured the reality of continued Western dominance of the Arab Middle East. While France was no longer a major player in the region, Britain retained enormous influence well into the postwar era. Indeed, through treaties, ties with political clients, economic concessions, and the maintenance of military bases, London continued to informally dominate a number of Arab states. In Jordan, a \$35 million annual subsidy was coupled with continued control of the country's powerful Arab Legion commanded by a British officer, John Bagot Glubb (1897-1986). This afforded London enormous leverage. Meanwhile, in Iraq, Britain's possession of two airbases, its dominant position in the consortium that held the concession to exploit oil, and its close relations with King Faysal II (r. 1939-1958) and the country's leading politician, Nuri al-Said (1888-1958), assured it of substantial influence. In Egypt, finally, Britain's ownership of the strategic Suez Canal, its control of Sudan—which Cairo hoped to retain as an Egyptian colony—and its maintenance by way of treaty of a huge military complex in the canal zone ensured that it continued to enjoy a commanding position.¹

London and the Arab states' divergent goals in the region ensured that the decolonization endgame in the Middle East would be fractious and bitterly contested. Britain, for its part, was disinclined to withdraw from the region or even to moderate its involvement. Weakened by World War II and by the loss of its valuable colony in India in 1947, London believed that its position in the Middle East was essential to its continued status as a Great Power; as a result, it vigorously resisted demands that it loosen its grip on the region. The Arabs were no more inclined to compromise than was London. Groups of nationalist army officers, Islamic organizations, and popular Arab political movements alike found Britain's ongoing indirect control of the Middle East to be utterly intolerable and loudly insisted on an immediate end to its influence in the region. As a result, the British government and anti-imperialist forces in the Arab world would find themselves locked in a bitter struggle over decolonization in the immediate postwar era—a contest that would come to color nearly all aspects of the Middle East's development.²

Arab Nationalism

The ideology of Arab nationalism also played a critical role in shaping the postwar Middle East. Maturing during the interwar period and especially in the 1940s, Arab nationalism held that the Arabs constituted a single people linked by a common history, culture, and language, and that their present subordinate situation stemmed

¹ Salim Yaqub, *Containing Arab Nationalism: The Eisenhower Doctrine and the Middle East*, First Edition (Chapel Hill: The University of North Carolina Press, 2004), 121; Arthur Goldschmidt, *A Brief History of Egypt* (New York: Checkmark Books, 2008), 140.

² Eugene Rogan, *The Arabs: A History – Revised and Updated Edition* (New York: Basic Books, 2012), 277–318.

largely from the fact that the Europeans had deliberately partitioned the region following the First World War in order to weaken them. Their prescription for the ills of the Arab world followed from this predicate. By overthrowing collaborationist governments such as Jordan's and Iraq's and by replacing the European-imposed fragmentation of the Arab world with political unity and shared purpose, the Arabs could once again stand up and restore themselves to the glory of their ancestors.³

Two schools of Arab-nationalist thought—moderate and Ba'athist—dominated in the two decades after World War II. Moderates acknowledged the reality of local nationalist sentiments and thus accepted the political division of the Arab world into separate states. As such, rather than pushing for the unification of existing governments into a single state, they called for the Arab countries to work collectively to coordinate their economic programs, defense policies, and, especially, diplomatic strategies. It was this line of thinking that led to the development of the Arab League in 1945.⁴

Sharply critical of the moderates, the Ba'athists instead championed a more radical and far-reaching construction of Arab nationalism. Two Syrians, the Christian Michel Aflaq (1910-1989) and the Muslim Salah al-Din al-Bitar (1912-1980), developed the basic tenets of the ideology in the 1940s and founded the transnational Ba'ath or "Renaissance" Party in 1947. With branches in a number of Arab countries, the movement focused on the achievement of two overarching goals: the establishment of an equitable society and, more importantly, the replacement of the existing, Western-created Arab countries with a single, unified state. Achieving those goals would not be easy. The leaders of the existing Arab monarchies and republics not only had a vested interest in the continued division of the Middle East, but, deeply embedded as they were in the imperialist powers' exploitive economic system, could be relied upon to oppose the Ba'athists' call for equity and social justice. Thus, party leaders concluded, the Arab people could only achieve a true rebirth rooted in equity and social justice by violently overthrowing the collaborationist governments.⁵

The Cold War

The third current that shaped the region in the postwar era was the diplomatic, economic, and military competition between the US and the USSR known as the Cold War. For three reasons, the global contest between the superpowers would have a peculiar focus on the Middle East. First, though both produced enormous quantities of oil, the Soviet Union and the US alike correctly perceived the strategic and economic significance that petroleum would play as the world continued to industrialize. Accordingly, each sought to control the Middle East's oil reserves to meet its future needs and, in America's case, to ensure that its allies in Europe and Japan could

³ Ira M. Lapidus, *A History of Islamic Societies*, 2nd ed. (Cambridge: Cambridge University Press, 2002), 546–48.

⁴ Albert Hourani, *A History of the Arab Peoples* (Cambridge, Mass: Belknap Press of Harvard University Press, 2002), 405–7.

⁵ Hourani, 404–5.

access the cheap energy on which the postwar reconstruction of their shattered economies depended. Second, both understood the geostrategic significance of the Middle East in a future superpower conflict. Situated just to the south of the Soviet border, it was the only territory from which long-range American B-29 bombers could strike Soviet industrial targets in the Ural region; later, it would provide the Soviets with bases from which they could attack American strategic, ballistic-missile submarines operating in the Mediterranean Sea. Finally, both Moscow and Washington correctly understood the critical role that the Suez Canal and the region's airbases would play in facilitating east-west communications in any future superpower conflict.⁶

In light of these views, it is no surprise that the region was one of the earliest theaters in the Cold War. As we saw in chapter eleven, the Soviet Union had sought to secure an oil concession in northern Iran and had attempted to strongarm Turkey into giving it control over the Bosphorus and Dardanelles Straits immediately following the end of World War II. Two of the most important opening salvos in the Cold War, these events powerfully influenced Washington's evolving views of Soviet behavior. In particular, they played an important part in shaping the containment policy that the American diplomat George Kennan (1904-2005) first articulated in 1947. Predicated on the assumption that the USSR would collapse if it could no longer expand, containment called for the US to counter Moscow's efforts to increase its influence outside the Eastern Bloc. The crises in Iran and Turkey not only shaped Kennan's thinking but also proved critical in leading President Harry Truman (r. 1945-1953) to put containment into practice through the issuance of the Truman Doctrine in 1947. One of the most critical initiatives of the Cold War, it justified a request for \$400 million in economic and military aid in support of Greek and Turkish efforts to resist communist aggression in part on the need to prevent "confusion and disorder [from] spreading throughout the entire Middle East."⁷

The Middle East's importance as a theater in the Cold War grew substantially over the course of the 1950s and 1960s. In the process, the superpower contest powerfully—and adversely—affected the region. Both the USSR and the US frequently interpreted events in the Middle East through a Cold War lens, and both sought to use their power and patronage to shape the region to their advantage in the global struggle. Proxies and allies within the Middle East did not always share their patrons' Cold War perspective; however, they were more than happy to exploit the superpower conflict to achieve their own ends—an approach that often resulted in the global struggle intensifying existing regional conflicts. Thus, in sharp contrast to Europe where the superpower contest would come to inhibit intraregional conflicts, the Cold War would

⁶ Rashid Khalidi, *Sowing Crisis: The Cold War and American Dominance in the Middle East* (Boston: Beacon Press, 2010), 15–16, 44–46, 111.

⁷ "The Truman Doctrine," March 12, 1947,

<https://www.ourdocuments.gov/doc.php?flash=false&doc=81&page=transcript>.

interact with other forces in the Middle East, such as decolonization and Arab nationalism, to deepen fault lines and intensify regional conflicts.⁸

The Arab-Israeli Conflict

The Cold War played a particularly prominent role in distorting and worsening the fourth current that shaped the postwar Middle East: the Arab-Israeli conflict. The battle lines here were hard and fast. With many of its citizens having recently experienced the horrors of the Holocaust or been expelled from their homes in the Arab world, there existed in Israel an overwhelming preference for an aggressive, Iron Wall-informed approach to national security over one that sought peace through compromise or negotiation. For their part, the Arabs, saw what they termed the “Zionist entity” as an affront to their nationalist aspirations and as a European settler-colonial venture—“the fruit,” in Nasser’s words, “of imperialism.” As a result, Arab governments refused to recognize Israel and loudly and publicly called for its destruction.⁹

Exacerbated by the ongoing plight of the Palestinian refugees, the standoff between Israel and its Arab neighbors dramatically shaped both the internal political evolution of the Middle Eastern states and their foreign policies. Domestically, it inflamed popular sentiments, generated friction between leaders committed to a hardline approach and those willing to consider a negotiated settlement, and, in the Arab states, justified violence and extraconstitutional actions that would destabilize countries for decades. In the realm of international relations, meanwhile, the Arab-Israeli conflict led both Israel and the Arab states to become enmeshed in the Cold War—even when they explicitly rejected the Manichaean logic on which it rested—because affiliation with one camp or the other provided the only reliable way to get the diplomatic support and sophisticated weapons needed to keep pace with the other side. In sum, in the 1950s and 1960s, the Arab-Israeli conflict would—in conjunction with decolonization, Arab nationalism, and the Cold War—militarize the region, distort its political, social, and economic development, and ensure that it failed to establish a stable international order.¹⁰

The Legacy of the 1948 War, 1949-1955

The interplay between the Arab-Israeli conflict and the forces of decolonization, Arab nationalism, and the Cold War was especially strong in the years immediately after Israel’s founding. As we saw in chapter thirteen, the 1948 War that had accompanied the Jewish state’s birth was a watershed moment that dramatically reshaped the political and ideological contours of the Middle East. It affected Israel and its Arab enemies very differently, however. For the Arab states, defeat at the hands of Israel—particularly after their governments had promised an easy victory—was a bitter pill that

⁸ Khalidi, *Sowing Crisis*, 101–58.

⁹ Benny Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920* (New York: Vintage Books, 2001), 267.

¹⁰ Khalidi, *Sowing Crisis*, 114–40.

delegitimated the ruling elites and initiated a period of instability and sweeping political change that lent strength to both the decolonization process and Arab nationalism. For Israel, in contrast, success in 1948 buoyed public spirits, legitimated its government, and solidified support for the hardline, Iron Wall approach that Ze'ev Jabotinsky (1880-1940) had first articulated and that David Ben Gurion (r. 1948-1954, 1955-1963) and other Labor Zionists had come to adopt as their own.¹¹

Israel: Infiltration and the Iron Wall

Israelis emerged from the 1948 War with a newfound sense of confidence and a reaffirmation of the belief that they could best ensure the survival of their country through military force. Celebrating the “New Zionist Man” who refused to let any slight pass unpunished, they looked favorably on their army’s aggressive and successful defense of their infant country and contrasted it with the seeming helplessness with which Europe’s Jewish people had responded to the Holocaust. The country’s leadership shared this view. Despite overtures from some Arab leaders, they refused to negotiate after the victory in 1948. Instead, aware that a Western arms embargo called the Tripartite Agreement had frozen Israel’s dominant military position in place, David Ben Gurion’s Labor Zionist coalition government opted to rely on an Iron Wall strategy of military intimidation to maintain Israel’s security.¹²

Palestinian infiltration into Israel following the 1948 War soon presented a test of that commitment. Even before the fighting had ceased in early 1949, refugees had begun crossing the frontier into Israel. The vast majority of them did so for benign reasons. Most merely sought to harvest crops on their former land, to recover property, to visit relatives, or to smuggle goods. However, a small proportion entered Israel to engage in sabotage or to launch terrorist attacks. Though disorganized and infrequent, those incursions nonetheless imposed a tangible and painful cost on the Jewish state—resulting, as they did, in the deaths of nineteen Israelis in 1950, forty-eight in 1952, and forty-two in 1953. They were especially demoralizing to the North African, or Maghribian, and Middle Eastern, or Mizrahim immigrants whom the government had settled in *moshavim*: communal agricultural settlements established along the frontier for the purpose of asserting control over the border. In fact, infiltration attacks had led so many Mizrahim and Maghribian Jewish people to flee the *moshavim* by the early 1950s, that officials in the Jewish Agency were coming to fear that the effort to use the settlements to control the border was teetering on the edge of failure.¹³

With its experience in the 1948 War having affirmed its belief that only overwhelming military force could alter Arab behavior, David Ben Gurion’s Labor Zionist

¹¹ Charles D. Smith, *Palestine and the Arab-Israeli Conflict: A History with Documents*, Sixth Edition (Boston: Bedford/St. Martin’s, 2007), 232–34.

¹² James L. Gelvin, *The Israel-Palestine Conflict: One Hundred Years of War* (New York: Cambridge University Press, 2005), 167.

¹³ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 271–72; Kirsten E. Schulze, *The Arab-Israeli Conflict* (New York: Pearson, 2008), 24.

government opted to respond to rising infiltration with an aggressive, Iron Wall approach. Initially, it moved to deter incursions through actions on the Israeli side of the border. Israeli Defense Force (IDF) soldiers set booby traps, laid mines, organized ambushes, and adopted a deadly “free fire” policy in which they shot at suspected infiltrators without warning. All told, this approach resulted in the deaths of between 2,700 and 5,600 Palestinians between 1949 and 1956.¹⁴

Somewhat later, the IDF shifted to an even-more aggressive policy centered on launching disproportionate retaliatory attacks against the villages from which the infiltrators had crossed the Israeli border. This new approach had three interrelated goals: it sought to deter future infiltration, to spur the Arab states to prevent further incursions, and, in keeping with the Iron Wall approach’s emphasis on military intimidation, to demonstrate the futility of contesting Israel’s overwhelming military superiority. Collective punishment of Arab villages suspected of aiding infiltration began in 1951 and crested in October 1953 when Unit 101, an elite special-operations force commanded by future prime minister Ariel Sharon (1928-2014), attacked the Jordanian-ruled West Bank Village of Qibya. Resulting in the demolition of forty-five homes and the deaths of sixty-nine Palestinians—two thirds of whom were women and children—the operation proved popular in the Jewish state, but inflamed public opinion in the Arab world and sparked international condemnation of Israeli militarism.¹⁵

The Arab States: Assassination, Instability and Leadership Changes

Just as Israel’s victory in 1948 had affirmed public support for its government, so did defeat delegitimize the ruling regimes in many Arab states. As we saw in chapter twelve, many of those governments were unpopular even before the fighting had begun in 1948. Corrupt, incapable of solving domestic problems, and unable to break free of Western dominance, the Arab republics and constitutional monarchies enjoyed little popular support and already faced growing challenges from pan-Arab Ba’athists, Islamic modernists, communists, and nationalist military officers. Ultimately, however, it was their defeat in 1948 that doomed them. Their abject failure to prevent the establishment of Israel—particularly, their inability to provision their armies with adequate food and ammunition—cost them what little support they had managed to retain and left them vulnerable, isolated, and weak.¹⁶

The result was a series of rapid and often bloody changes in government that ended with a new group, the officer corps, assuming a dominant position in the Arab states. This process began in Syria. Frustrated by the civilian government’s poor support of the war effort, Colonel Husni al-Za’im (r. 1949) seized control of the government in March 1949 in a coup so well executed that CIA officers subsequently

¹⁴ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 272–74.

¹⁵ Avi Shlaim, *The Iron Wall: Israel and the Arab World, Revised and Expanded* (New York: W. W. Norton, 2014), 85–93.

¹⁶ Rashid Khalidi, *Resurrecting Empire: Western Footprints and America’s Perilous Path in the Middle East* (Boston: Beacon Press, 2005), 58.

studied it as a model of effective regime change. Za'im proved considerably less successful in maintaining his grip on power than on assuming it, however. Opposed to his willingness to pursue peace with Israel, a hardline faction of officers overthrew and executed him in August 1949; they in turn fell from power in a third coup that December. Za'im was not the only Arab leader to perish at that time as a result of a willingness to negotiate with Israel, moreover. In what amounted to a declaration of popular opposition to peace with the Jewish state, a Palestinian tailor's apprentice assassinated Jordan's Abdallah (r. 1921-1951) in 1951 at the entrance to the al-Aqsa Mosque shortly after news leaked that the king had engaged in negotiations with Israeli leaders. Following the brief and unhappy rule of Abdallah's mentally unstable son, Talal (r. 1951-1952), his sixteen-year-old grandson, Hussein (r. 1952-1999), succeeded to the throne.¹⁷

The Egyptian Revolution, 1952

The next change in government in the region was far-more consequential. Launched in 1952, the Egyptian Revolution not only put the officer corps in firm charge of the largest Arab state but also set the stage for the military's dominance of the entire Arab world for the following two decades. Dissatisfaction with civilian rule among the Egyptian officers had grown rapidly in the years before 1948. Many had long been frustrated with King Faruq's (r. 1936-1952) venality and with the Wafd Party's corruption and inability to end Western dominance; some had even toyed with the idea of a coup in the past. It was Egypt's defeat at the hands of Israel, however, that turned them irrevocably against the old order. Blaming their failure in 1948 on the government's inability to provide adequate supplies to the units deployed to Palestine, a group of junior officers hailing largely from the lower-middle-class resolved that Egypt could only achieve real independence and national greatness under a new, reform-minded, military-dominated, Arab-nationalist government. Led by a thirty-four-year-old son of a postal worker, Lieutenant Colonel Gamal Abdel Nasser, they organized a secret, revolutionary society immediately after the war called the Free Officers and began plotting not merely the overthrow of Faruq but the end of the entire liberal order. It was a bold objective for a group of young men of comparatively low rank; accordingly, the Free Officers recruited a sympathetic senior figure, the fifty-one-year-old General Muhammad Naguib (1901-1984) to serve as a figurehead who could, once they assumed power, lend them the gravitas that they lacked.¹⁸

Government missteps soon gave them their chance to act. Seeking to restore the Wafd Party's popular standing after the war by negotiating a revision of the Anglo-Egyptian Treaty of 1936, Prime Minister Mustafa al-Nahhas (1879-1965) initiated talks with London in early 1950. Begun with great fanfare, the negotiations soon deadlocked—just as they had in the 1920s and 1930s. This time, the discussions

¹⁷ Hugh Wilford, *America's Great Game: The CIA's Secret Arabists and the Shaping of the Modern Middle East* (New York: Basic Books, 2013), 100–108; Philip Robins, *A History of Jordan*, Second Edition (Cambridge: Cambridge University Press, 2019), 77–85.

¹⁸ Goldschmidt, *A Brief History of Egypt*, 144–45.

foundered over London's insistence on retaining control of the massive base complex it possessed in the Canal Zone. Frustrated, al-Nahas adopted a more forceful approach in October 1951. Seeking to pressure the British, he unilaterally abrogated the treaty and encouraged popular guerrilla attacks and acts of sabotage against imperial troops and installations. It was an ill-considered move. Paramilitary assaults on imperial positions along the canal spurred British soldiers to respond with a series of escalating military actions that culminated in January 1952 in a massive assault on a police barracks in Isma'iliyya that killed forty-six Egyptian police officers. Infuriated, mainstream nationalists scheduled a general strike in Cairo for the following day to protest the attack. Unfortunately for them, the demonstration rapidly degenerated into an orgy of violence and looting. By the time the smoke cleared on what came to be known as Black Saturday, fifty Egyptians and seventeen foreigners lay dead, and dozens of businesses associated with the West and the Egyptian upper class—department stores, automobile dealerships, night clubs, and high-end hotels—had been reduced to ashes in a series of well-planned arson attacks.¹⁹

Those buildings were not all that lay in ruins that day. The fighting in Isma'iliyya and the destruction and looting on Black Saturday also marked the end of the line for Egypt's old regime. Unable either to negotiate an end to Britain's presence in the canal zone or to maintain order in the capital, Faruq, the Wafd Party, and the entire parliamentary system had lost all legitimacy in the eyes of the Egyptian people and were, from that point forward, living on borrowed time.²⁰

The Free Officers' revolution finally came on July 23, 1952. Thanks to the old order's unpopularity, Nasser and his allies were able to seize control of the government with little difficulty in a nearly bloodless coup. They followed by establishing a new executive organ, the Revolutionary Command Council (RCC), nominally headed by Naguib, to oversee the political changes that they intended to implement.²¹

Despite the Free Officers' later reputation as militant revolutionaries, they did not immediately implement a radical agenda. Instead, having no firm ideological positions at that time beyond adherence to moderate Arab nationalism and an ambiguous commitment to social progress, they focused initially on securing three objectives that would strengthen their position: the elimination of rival sources of political power, the passage of a series of populist reforms, and the signing of an agreement with London that would—at last—secure the removal of British troops from Egypt.²²

¹⁹ Rogan, *The Arabs*, 280–81.

²⁰ Goldschmidt, *A Brief History of Egypt*, 142.

²¹ Arthur Goldschmidt Jr., *Modern Egypt: The Formation of a Nation State* (New York: Routledge, 2004), 144–46.

²² Alain Roussillon, "Republican Egypt Interpreted: Revolution and Beyond," in *The Cambridge History of Egypt. Volume 2, Modern Egypt, from 1517 to the End of the Twentieth Century*, ed. M. W. Daly, 1st ed. (Cambridge University Press, 1998), 336–38.

Consolidation of Power

Nasser and the officers first moved to eliminate all threats to their control of Egypt. They deposed Faruq and replaced him with his infant son, Fuad II (r. 1952-1953), dissolved the legislature, suspended the constitution, and arrested key political leaders on corruption charges. They also outlawed all political parties including the Egyptian Communist Party. To replace them, they established a new mass movement called the Liberation Rally. Renamed the National Union in 1956 when the Free Officers reorganized it along more explicitly pan-Arabist lines, the Liberation Rally was not a political party in the conventional sense. Instead, it was a device that the officers created to connect with, control, and mobilize the masses in support of the RCC and its initiatives. Seeking to blunt leftist criticisms, finally, they formally abolished the monarchy in June 1953 and declared Egypt a republic. For lack of a better candidate, Nasser and the RCC decided to elevate Naguib to the newly created office of president.²³

Nasser and his supporters' efforts to assume a monopoly over political power were not yet entirely complete, however. Two challenges remained to be addressed. First, while initially cordial, relations between the officers and the Muslim Brotherhood had grown strained over the course of 1953 owing to the fact that neither wished to share power. Increasingly viewing the brothers as a threat, Nasser and his allies consequently began to take steps to curb the organization in early 1954. This task proved to be beyond their capabilities at that time, however. The brotherhood possessed a large following, a robust apparatus of schools and social welfare organizations, and a potent paramilitary wing. Accordingly, the officers decided to back down for the time being. Second, to the surprise of many, Naguib had emerged as a genuine threat to Nasser and his allies. A supporter of the pluralistic, democratic system that the RCC had been systematically dismantling, the general had become popular with the masses and, more dangerously, had secured a substantial following among the officers. Unlike the brotherhood, Naguib was a danger that could not be put off. Accordingly, Nasser and his allies engineered a bureaucratic showdown in the spring of 1954 in which he and his supporters overcame early setbacks and purged Naguib from the government. Thereafter, Nasser enjoyed unquestioned control of the RCC and, by extension, the country Egypt.²⁴

Populism

In the meantime, Nasser and his RCC allies were also moving to secure the revolution by implementing a series of populist economic and social reforms. On the surface, these changes appeared to be far reaching. They banned the use of traditional titles such as pasha and bey, capped rents, established wage minimums, and legalized

²³ Steven A. Cook, *The Struggle for Egypt: From Nasser to Tahrir Square*, Reprint edition (Oxford University Press, 2011), 91–92; Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 104–13.

²⁴ Goldschmidt, *A Brief History of Egypt*, 147–52.

unions. Most importantly, they included a package of land-reform laws designed to improve the lot of the poorest peasants. The program placed limits on the amount of property any single individual could own, called for the confiscation of excess amounts—including the royal family's substantial agricultural holdings—and then redistributed the appropriated land to poor peasants.²⁵

A closer look reveals, however, that the RCC's early reform package was in fact quite tepid. Reflecting the fact that the officers remained as yet uncommitted to a particular ideological position, the changes they implemented tended to be modest in terms of both method and impact. The officers' conservatism was most evident in their signature initiative: land reform. That program did not simply take land from absentee owners and directly redistribute it to the fellahin. Instead, it reimbursed landowners for the property they had lost and required the destitute peasants who had obtained land through the program to shoulder part of the cost. As a result, it did little to alter the distribution of wealth and power in Egypt.²⁶

Still, the broader reform program did pay two important dividends for the officers. First, by improving the lot of the impoverished masses—however modestly—it turned the fellahin and urban poor into vociferous supporters of Nasser and the officers. Second, the land redistribution component of the officers' reform program weakened the economic and political power of a critical component of the Wafd Party: the landed elite. In so doing, it left the Wafd too weak to mount a serious challenge during the RCC's vulnerable early years.²⁷

The Anglo-Egyptian Treaty of 1954

Meanwhile, Nasser entered negotiations with London in 1953 that finally secured the evacuation of British troops from the canal zone. Two changes made an agreement possible. First, the development of the hydrogen bomb rendered large base complexes such as the one Britain maintained along the Suez Canal obsolete; As such, London was more inclined to make concessions than it had been in the past. Second, unlike earlier Egyptian leaders, Nasser was willing to drop Egypt's claims to Sudan. With these stumbling blocks out of the way, the two sides were able to quickly reach an agreement. Signed in July 1954, the resulting Anglo-Egyptian Treaty called for Britain to remove its remaining troops from Egypt no later than June 1956; to ease London's security concerns, however, the deal permitted it to reoccupy the base complex in the event of a war in the Middle East.²⁸

The agreement set the stage for Nasser's final consolidation of power. It was the weakness of the treaty, ironically, that made his success possible. Sensing an

²⁵ Cook, *The Struggle for Egypt*, 44–47.

²⁶ Goldschmidt, *A Brief History of Egypt*, 146.

²⁷ Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 107.

²⁸ Douglas Little, *American Orientalism: The United States and the Middle East since 1945*, second edition (Chapel Hill: University of North Carolina Press, 2004), 58–59.

opportunity to challenge the Egyptian leader on nationalist grounds, the Muslim Brotherhood launched a coordinated campaign against Nasser's government. Focusing on the provisions permitting Britain to reoccupy the canal-zone bases in the event of war, the brothers argued that the deal was a bad one for Egypt that merely perpetuated its neocolonial relationship with London. For a time, the criticism was effective and eroded support for Nasser's government. In October 1954, however, the organization pushed its luck too far. At a government-sponsored rally in support of the agreement in Alexandria's Manshiya Square, a member of the brotherhood attempted to assassinate Nasser. The Egyptian leader's cool response to the attack—he remained unflustered and finished his speech as if nothing had happened—created a surge of support for Nasser that gave him the political capital to mount an aggressive move against the Muslim Brotherhood. The ensuing campaign was a success. The regime arrested thousands of brotherhood members, imprisoned its leadership, and dismantled its institutional structure. Severely weakened, it would not recover until the 1970s. Thus, by the end of 1954, Nasser had eliminated the last serious threat to his power. Thereafter, he ruled Egypt without challenge.²⁹

Nasserism

Nasser had risen to power without fixed views or a formal ideology. By the time he consolidated his position in 1954, however, he had come to develop a guiding philosophy, Nasserism, that would shape his subsequent approach to governance, regional relations, and global affairs. Nasserism was very much a product of its times. Calling for social justice, an end to imperialism, neutrality in the Cold War, and Arab unity, it reflected the main ideological currents that shaped the Arab world in the 1950s and 1960s. For a time, it was enormously popular both inside and outside Egypt—constituting, from the mid 1950s through 1967, the dominant ideology in the Arab world.³⁰

Domestically, Nasser made rapid economic modernization and the improvement of living standards for the neglected lower classes his central objectives. Like many other rulers in the Developing World, he did not attempt to achieve these goals through a conventional, market-oriented approach. Leery of the capitalist system of development that the imperialist powers championed, he instead sought to attain his objectives through a state directed, Import Substitution Industrialization (ISI) model of development based loosely on the program that Ataturk (r. 1923-1938) had implemented in Turkey. Nasser's new modernization drive focused on agricultural expansion and, especially, the creation of state-owned industrial concerns. Central to both efforts was a planned hydroelectric project, the Aswan High Dam, in upper Egypt. Though fabulously expensive—its estimated cost was a then-staggering \$1 billion—it promised to revolutionize life in Egypt. Once completed, it was expected to expand the amount of cultivable land in the country by more than one-million *feddan*—equivalent to

²⁹ Cook, *The Struggle for Egypt*, 60–61.

³⁰ Yaqub, *Containing Arab Nationalism*, 31–32; Cook, *The Struggle for Egypt*, 111.

about one-million acres—and to supply sufficient electricity not merely for the country's planned industrial expansion but for *all* of the country's needs.³¹

Internationally, meanwhile, Nasserism focused on securing complete Arab independence. A moderate Arab nationalist, he did not insist as the Ba'athists did that the Arab countries needed to come together into a single political union in order to end colonial dominance. Instead, he believed that the Arab states could remain independent of each other yet still achieve that goal so long as they advanced a shared foreign policy and worked collectively.³²

Despite his ambivalence toward formal Arab political unity, Nasser nonetheless believed that the Arab people should pursue an aggressively independent and anti-imperialist foreign policy. This outlook manifested itself in four ways. First, arguing that they were tools of continued imperial control, he vociferously criticized the continuation of Britain's defensive treaties with Jordan and Iraq. Second, he steadfastly resisted American-sponsored efforts to draw Egypt and the Arabs into the Cold War on the Western side through collective-security alliances such as the Baghdad Pact that Iraq, Britain, Turkey, Iran, and Pakistan established in early 1955. As he noted acidly, Israel and Britain—the latter a signatory to the Baghdad Pact—were far-more immediate threats to the Arab people than were the distant Soviets. Third, he called for the Arabs to use the Cold War to their advantage through a policy he called Positive Neutrality, in which they would exploit their independence in the global struggle to cut favorable deals with both superpowers. Finally, to the consternation of the moralistic US Secretary of State John Foster Dulles (1888-1959), who viewed the Cold War in starkly black-and-white terms, he became a major leader in the emerging Non-Aligned Movement.³³

Nasserism rapidly gained influence in the Arab world in the mid 1950s. Its success stemmed from a combination of humble technological developments and the Egyptian leader's high-profile achievements in foreign relations. Diplomatically, Nasser's ability to finally secure the evacuation of British troops from Egypt dramatically raised his stature among anti-imperialist Arab nationalists throughout the region who were impressed by his ability to finally compel London to resolve the longstanding dispute on terms favorable to Egypt. Technologically, the widespread adoption of new, low-cost transistor radios permitted the Egyptian leader to bypass other Arab governments and to speak directly to their peoples via Radio Cairo's powerful transmitters. The new medium made for a potent propaganda tool—one that Nasser used with consummate skill to influence politics in other Arab states. In February 1955, for example, he used Radio Cairo to spark a coup in Syria that resulted in a neutralist, pro-Nasserist government seizing power from one that had close ties with Iraq's conservative, Hashimite monarchy. A concerted radio campaign later that year against the Baghdad Pact not only compelled King Hussein to abandon plans to join the alliance, but also spurred Palestinians—who constituted a majority of the population—to hold violent

³¹ Goldschmidt, *A Brief History of Egypt*, 157.

³² Yaqub, *Containing Arab Nationalism*, 31–32.

³³ Rogan, *The Arabs*, 289–91; Yaqub, *Containing Arab Nationalism*, 31–32.

demonstrations against his government. Desperate to mollify the protestors so he could hang onto his throne, Hussein responded by dismissing the British officers who led the Jordanian army and by agreeing to hold free elections in 1956—a vote that a strongly pro-Nasserist party handily won.³⁴

The Suez Crisis, 1956

Nasserism reached dizzying heights of popularity as a result of one of the defining international crises of the 1950s: the brief war between Egypt on the one hand and an alliance of Israel, Britain, and France on the other that occurred in late 1956. Known as the Suez Crisis in the West and as the Tripartite Aggression in Egypt, the conflict was the culmination of the interplay between the Arab-Israeli conflict, the Cold War, Arab nationalism, and decolonization. The last, desperate gasp of classical, nineteenth-century style imperialism in the region, the contest would remake the geostrategic outlines of the Middle East and would vault Nasser to a new level of influence that would solidify his status as the region's dominant figure for the next decade.

The Gaza Raid, February 1955

While the war had a complex genesis, its proximate cause lay in the ongoing issue of infiltration into Israel. As we have seen, Tel Aviv's commitment to the Iron Wall had led it to launch retaliatory raids aimed both at deterring Palestinians from crossing the border and at encouraging the Arab states to curb infiltration. Unsurprisingly, those punitive actions did not sit well with the nationalistic Free Officers. Furious about Tel Aviv's repeated violations of Egyptian territory, they began sending small groups of soldiers into Israel in early 1954 where they mounted guerrilla attacks and collected intelligence.³⁵

Now it was Israel's turn to react. Wedded to Jabotinsky's Iron Wall, Tel Aviv was deeply committed to responding swiftly and overwhelmingly to such attacks. Accordingly, after one such operation ended in the death of a cyclist in February 1955, the Israeli government decided to retaliate by attacking an Egyptian military installation in the Gaza Strip. Designed to embarrass Nasser by revealing Egypt's military weakness, the assault was a stunning success. Commanded by Ariel Sharon, the force of 150 elite paratroopers that executed the Gaza Raid killed thirty-seven Egyptian soldiers and left their base a smoldering ruin.³⁶

³⁴ Yaqub, *Containing Arab Nationalism*, 38–41.

³⁵ Shlaim, *The Iron Wall*, 131–37.

³⁶ Shlaim, 131–37; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 282–83.

The Czech Arm Deals, September 1955

The Gaza Raid infuriated Nasser. As Tel Aviv had intended, the attack demonstrated that he lacked the power to defend Egypt and thus threatened to undermine his still-tenuous grip on power. As such, Nasser felt that he had no choice but to respond. Accordingly, he followed the Gaza Raid with two actions—one designed to harm Israel and the other to redress the military balance. First, to strike back at the Jewish state, he tightened Egypt's four-year-old blockade of the Strait of Tiran that controls access to the Israeli port of Eilat that sits at the head of the Red Sea. In conjunction with the earlier closure of the Suez Canal to vessels bound for Israel, this action ensured that ships travelling from the Far East could only access the Jewish state by making a costly circumnavigation of Africa. Second, he made an aggressive push to modernize the Egyptian army by purchasing advanced weapons abroad.³⁷

This latter task appeared to be a daunting one. The Tripartite Agreement had made arms purchases all-but impossible, and Washington, Nasser's first choice for weapons, had indicated that it would only consider making an exception for Egypt if it joined an anti-Soviet security treaty like the Baghdad Pact. Determined to pursue an independent foreign policy, Nasser found that condition unacceptable and pressed on with his search. He finally met with success behind the Iron Curtain. In September 1955, he worked out an agreement with the Kremlin to purchase arms from the Soviet client state of Czechoslovakia. It was a good deal for Nasser. According to its terms, Egypt would receive a massive package of weapons over the following year including submarines, 230 T-34 tanks, 200 armored personnel carriers, 50 Il-28 medium bombers, more than 100 advanced MiG-15 and MiG-17 fighter jets, and 600 artillery pieces. In exchange, Nasser agreed to provide Czechoslovakia with 100,000 tons of cotton valued at \$86 million.³⁸

The Czech arms agreement completely disarranged the strategic balance in the Middle East. It did so in three ways. First, it effectively nullified the Tripartite Agreement that had banned arms sales to the region. Indeed, France, which sought to curb Nasser's power in response to his support for Arab rebels in its Algerian colony, promptly followed the deal by selling AMX-13 light tanks, artillery, and advanced Mystère IV fighter jets to Israel. Second, despite the French arms deal, the agreement threatened the preponderant military position that Israel had enjoyed since 1950. Once Egypt had integrated the arms into its military, it would—to Tel Aviv's chagrin—possess a military that was at least as powerful as the IDF. Finally, the deal dramatically altered the Cold War balance in the Middle East. With the conclusion of the agreement, the

³⁷ Yaqub, *Containing Arab Nationalism*, 39–40.

³⁸ Wilford, *America's Great Game*, 190–91; Guy Ziv, "Shimon Peres and the French-Israeli Alliance, 1954-9," *Journal of Contemporary History* 45, no. 2 (2010): 418; Little, *American Orientalism*, 168.

USSR had not merely leapfrogged over America's efforts to contain its influence in the region but had established close ties with the largest and most important Arab state.³⁹

Financing the Aswan High Dam

Unsurprisingly, the Eisenhower administration found the Czech arms deal deeply troubling. It feared that the agreement would, like the proverbial camel's nose under the tent, be the pivotal first step in a longer Soviet campaign of subversion that would end with Egypt and perhaps other Arab states absorbed into the Eastern Bloc—an outcome that would, in zero-sum fashion, shift the balance in the Cold War in the USSR's favor. Determined to prevent Moscow from consolidating its influence in Egypt before it was too late, the administration let Nasser know in December 1955 that it was prepared to help finance the centerpiece of Egypt's economic development scheme: the Aswan High Dam. The American proposal included a \$200 million World Bank loan, as well as grants of \$56 million from the US and \$14 million from Britain to help fund the first phase of the project. Washington and London also pledged to provide loans for later phases of the dam's construction. In exchange, they insisted that Egypt forgo future arms purchases from the Eastern Bloc and that it open its economy to the West.⁴⁰

Despite what seemed like a promising offer, the deal gradually fell apart over the spring and early summer of 1956. Nasser had been eager to secure foreign aid to kickstart the dam's construction and was pleased when he first learned of the Eisenhower administration's proposal. His enthusiasm cooled, however, when he became aware of the conditions that the Americans had attached to their offer; to Nasser, they smacked of the colonialist limits that Britain had imposed on Egypt in the late-nineteenth and early-twentieth centuries. Accordingly, Nasser refrained from immediately accepting the US offer; instead, in keeping with his policy of Positive Neutrality, he moved in the spring of 1956 to see if he could shop it to get a better deal from Moscow. That initiative appeared promising for a time; in the end, however, the USSR was unable to provide a financing package as attractive as Washington's. As a result, in July, Nasser instructed the Egyptian ambassador to the US to let Secretary of State Dulles know that Egypt had agreed to Washington's terms.⁴¹

By that point, however, the administration had decided to pull its funding offer. It made this about face for three reasons. First, Dulles and President Dwight Eisenhower (r. 1953-1961) believed that the Egyptian government would need to impose crippling austerity measures to service the loans and worried that its people would come to hate the countries from which Egypt had borrowed the funds. Better that they vent their spleens at the Soviet Union—which Dulles and Eisenhower presumed would step in as the project's primary foreign backer—than at the US. Second, Washington had grown

³⁹ Khalidi, *Sowing Crisis*, 182; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 284.

⁴⁰ Little, *American Orientalism*, 168–69; Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 121.

⁴¹ Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 121.

weary of Nasser's efforts to play both sides of the street. Dulles's Manichean belief that humanity was locked in an epochal struggle between the Free World and the totalitarian Soviet Bloc accorded no legitimacy to fence sitters. Given this outlook, Nasser's continued efforts to play Washington off Moscow through Positive Neutrality and his propaganda broadcasts against pro-Western states rankled the secretary of state and soured him on the Egyptian leader. Finally, in May 1956, Nasser extended formal diplomatic recognition to the People's Republic of China (PRC)—a state that the US had been working aggressively to keep isolated. That move was the final straw for Dulles and Eisenhower. Furious with Nasser for breaking with the US on such an important issue, they decided to pull the plug on the aid package. Thus, on July 19, the secretary of state informed the disappointed Egyptian ambassador that the US was withdrawing its offer to finance the dam.⁴²

Nationalization of the Suez Canal, July 1956

If Cold War considerations shaped Washington's decision to withdraw its funding offer, it was the ideology of decolonization that guided Nasser's response. The Egyptian leader was livid about the Eisenhower administration's decision to pull its funding offer. In his eyes, Washington's action was not merely a threat to the foundation of Egypt's economic development program but also an effort to embarrass and weaken his government. He could not let it go without a response. In short order, he came up with one that was as bold as it was ingenious: he would nationalize the Western-owned and controlled Suez Canal and use the toll revenues it generated to finance the dam.⁴³

With a dramatic flair, he announced his new plan in a live radio address he gave in Alexandria just eight days after Dulles had withdrawn the American financing offer. Upon hearing Nasser utter a codeword in his speech—de Lesseps, the name of the French engineer who had overseen the construction of the waterway—Egyptian forces took possession of the canal and assumed control of its operations. The announcement thrilled people in Egypt and throughout the Developing World. To their delight, Nasser had not merely found a means of offsetting the loss of American funding for the Aswan High Dam but had done so in a way that challenged lingering Western imperialist influence in Egypt.⁴⁴

His action had the added benefit of being perfectly legal. In taking control of the waterway, Egypt promised to reimburse the Suez Canal Company's shareholders based on the closing value of its stock the day Nasser had ordered the waterway nationalized. As such, the seizure of the canal was consistent with international law. Indeed, as Egyptians delightedly pointed out in response to London's howls of protest, it

⁴² Steven Z. Freiburger, *Dawn Over Suez: The Rise of American Power in the Middle East, 1953-1957* (Chicago: Ivan R. Dee, 1992), 153–57; Yaqub, *Containing Arab Nationalism*, 45–46.

⁴³ Rogan, *The Arabs*, 299–300.

⁴⁴ Rogan, 299–300.

was no different than the Labor government's nationalization of Britain's railroads and steel industry immediately following World War II.⁴⁵

The Tripartite Aggression, October-November 1956

Legal or not, Nasser's move elicited anger in London and Paris. On the surface, their concerns were narrowly economic in nature. With the British government owning 44 percent of the canal and French investors holding the balance, both London and Paris were insistent that Egypt restore their property—promises of reimbursement be damned. British pique had a strategic component to it as well. Ownership of the canal had meant possession of an asset of global significance—of which London increasingly wished to regain control. On a more fundamental level, however, French and British rage was a function of the symbolism of Egypt's actions. Nasser, an upstart from the Developing World, was challenging what they saw as the natural order of things in which Western Great Powers dictated to the weak and not the reverse. As such, from the moment Egypt seized the canal, London and Paris committed themselves to overturning his action—peacefully or otherwise.⁴⁶

Washington took a more temperate approach to the crisis. Well aware of the serious damage that an attack on Egypt would do to the West's reputation among the world's many recently decolonized people, they were determined to see it resolved through diplomacy rather than force. Accordingly, they floated a succession of plans to settle it that revolved around the idea of placing the canal under international control. In every case, their efforts proved to be stillborn. Committed to keeping the canal in Egyptian hands, Nasser repeatedly rebuffed Washington's proposals and made emphatically clear that his government would reject any solution that did not involve his country retaining untrammelled sovereignty over the canal. Egypt's success in operating the waterway following nationalization strengthened his hand—and even raised questions about whether Nasser's action constituted a problem. Indeed, as several American congresspeople pointed out after a briefing on the situation, if the Egyptians could run the canal and if they were willing to compensate the Suez Canal Company's shareholders, then there was no crisis to resolve.⁴⁷

Egypt's ability to operate the waterway may have satisfied American congressional leaders, but it failed to meet French Prime Minister Guy Mollet's (r. 1956-1957) and British Prime Minister Anthony Eden's (r. 1955-1957) insistence that France and Britain regain control of the canal. Accordingly, ignoring Washington's clear desire to see the situation resolved peacefully, they began hatching a plot in conjunction with Israel to regain control of the canal through force. Designed to disguise what was, at heart, an act of naked imperialist aggression, the scheme was as preposterous as it was out of step with the times. It called for Israel—which joined the conspiracy on

⁴⁵ Little, *American Orientalism*, 172.

⁴⁶ Cook, *The Struggle for Egypt*, 68; Yaqub, *Containing Arab Nationalism*, 47.

⁴⁷ Tony Judt, *Postwar: A History of Europe Since 1945* (New York: Penguin Books, 2005), 295; Little, *American Orientalism*, 173–74.

October 21—to set the scheme in motion by first invading Egypt through the Sinai Peninsula. Next, pretending that they were impartial parties acting only to ensure that trade could continue to flow through the canal, France and Britain would use the fighting as a pretext to issue an ultimatum demanding that both Israeli and Egyptian forces withdraw immediately from the canal zone. When Nasser refused to comply—as the conspirators knew he would—French and British troops would then seize control of the waterway.⁴⁸

London, Paris, and Tel Aviv were confident in this plan. They not only believed that they it would result in Egypt's military defeat, but that it would also allow them to achieve all of their goals vis-à-vis Egypt. Britain would regain control of the canal, France would end Nasser's support for the Algerian rebels and recover its shareholders' property, and Israel would destroy the military equipment that Nasser had acquired from Czechoslovakia before the Egyptian army could integrate it into its forces—thus ensuring the perpetuation of the IDF's regional military dominance. Finally, though they were less certain that they could achieve this outcome, they were also hopeful that a quick military victory over Egypt might prove so embarrassing for Nasser that the Egyptian people would overthrow him.⁴⁹

The conspirators wasted no time in putting their scheme into action. As per their plan, Israeli forces started the war with a surprise invasion of the Sinai Peninsula. Well led and equipped, three columns of IDF troops quickly steamrolled their way through Egypt's ill-prepared defenses on October 29 and began to advance rapidly westward toward the canal. A fourth swung south to take control of Sharm al-Shaykh, the town that commanded the Strait of Tiran. London and Paris issued their ultimatum the next day. When Egypt refused to comply by the deadline they had set, they joined the war—first by bombing Egyptian airbases and then, on November 5, by dropping paratroopers in the canal zone. The next day, Royal Marines stormed ashore near the northern entrance of the waterway. Brushing aside Egyptian forces, they quickly seized control of the city of Port Said.⁵⁰

Despite this military success, however, it had already become quite clear that the political side of the operation was not unfolding in accordance with the conspirator's plan. From the very start, the entire world had seen right through London and Paris's ludicrous pretense that they were acting for the good of the international community and quickly intuited the broad outlines of the conspiracy. As a result, the parties to the collusion found themselves squeezed from all quarters. Wildly inconsistent with decolonizing spirit of the times, their imperialistic actions faced near-universal condemnation in the UN General Assembly and in the world press. Worse, they confronted economically devastating fuel shortages thanks to concerted Arab actions. Not only did Egypt block the flow of oil through the canal by scuttling cement-filled ships

⁴⁸ Judt, *Postwar*, 295–96.

⁴⁹ Donald Neff, *Warriors at Suez: Eisenhower Takes America into the Middle East* (New York: The Linden Press, 1981), 335–36; Judt, *Postwar*, 295–96.

⁵⁰ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 291–97.

in it, but the Syrian government responded to the war by cutting the pipeline that connected the oil fields of Iraq to the terminals on the Mediterranean. As a result, the flow of oil to Europe slowed to a trickle. Making matters all-the grimmer, the Kremlin piled on by issuing a statement on November 5 threatening to launch rockets on Paris and London if they did not cease their aggression.⁵¹

Far more important, however, was the economic and diplomatic pressure that they faced from across the Atlantic. Eisenhower was apoplectic with Eden, Mollet, and Ben Gurion for organizing the attack on Egypt. He resented that they had acted without Washington's approval, expressed concern that their invasion might fatally damage the West's reputation in the Arab world, and fumed that they had undertaken such a shabby and transparent act of imperialism so close to the American presidential election. Above all, Eisenhower was furious that the attack had permitted Moscow to crush the Hungarian Revolution—a concurrent uprising aimed at ending Soviet control of Hungary—without suffering significant damage to its international reputation.⁵²

Eisenhower moved quickly to ratchet up the pressure on Tel Aviv, Paris, and London. On the diplomatic front, he secured passage of a UN resolution regarding the Suez Crisis on November 2. It called for an immediate ceasefire, the withdrawal of French, British, and Israeli troops from Egypt, and the Installation of United Nations Emergency Force (UNEF) peacekeepers in the Sinai. Concurrently, he took two economic actions that grievously imperiled France and, especially, Britain. First, he accelerated a devastating war-induced run on Britain's currency by instructing the Federal Reserve Board to dump America's holding of pounds and by blocking London's attempt to secure an emergency loan from the International Monetary Fund (IMF). Second, the president barred efforts to redirect oil from the Western Hemisphere to France and Britain so as to make up for the loss of Middle Eastern petroleum. As he told an aide, he was leaving Paris and London to “boil in their own oil” until they came to their senses.⁵³

Faced with this pressure, France and Britain quickly buckled. On November 6, they agreed to the terms of the UN ceasefire. Facing continued American pressure, they began withdrawing their troops from Egypt almost immediately. Israel was more reluctant to pull out but finally agreed to remove the last of its soldiers from the Sinai in March 1957. UNEF peacekeepers followed their withdrawal by taking up positions at Sharm al-Shaykh and along the Egyptian side of the frontier with Israel to ensure against the outbreak of a new conflict.⁵⁴

⁵¹ Yaqub, *Containing Arab Nationalism*, 52–53.

⁵² Wilford, *America's Great Game*, 259.

⁵³ Quote from George C. Herring, *From Colony to Superpower: U.S. Foreign Relations since 1776* (New York: Oxford University Press, 2008), 676; Freiburger, *Dawn Over Suez*, 192–95.

⁵⁴ Peter L. Hahn, *Crisis and Crossfire: The United States and the Middle East Since 1945* (Washington, D.C.: Potomac Books Inc., 2005), 32.

Consequences

Momentous in significance, the Suez Crisis fundamentally restructured the power dynamics of the Middle East. The biggest winner by far was Nasser. Though Egypt had suffered a clear military defeat, his diplomatic triumph over the despised imperial powers and the hated Israelis massively raised his stature in the Arab world and instantly turned him into the leader of the anti-Zionist cause. Israel also strengthened its position as a result of the war. It is true that Washington had compelled Tel Aviv to yield the territory its army had conquered in the Sinai. However, the stationing of UNEF troops at Sharm al-Shaykh opened the Gulf of Aqaba to Israel-bound shipping while the success of its soldiers further enhanced the IDF's military reputation. Washington, too, emerged in a stronger position. Thanks to its support of Egypt, the US gained prestige in the Arab world—at least for a time. The Kremlin saw its reputation rise much further, however. Its promise in October 1958 to provide critical funding for the Aswan High Dam and, more importantly, the widespread belief in the Middle East that Moscow's ultimatum to France and Britain had been decisive in bringing the war to a swift conclusion had substantially raised its standing with people throughout the Arab world.⁵⁵

In contrast, Paris and, especially, London emerged from the crisis the clear losers. The conflict demolished their reputations in the region, hastened Britain's departure from the Middle East, and grievously undermined the position of Arab states like Iraq that continued to maintain close ties with London. More broadly, Suez marked the swan song of London's pretensions as a global imperial power. As the *Sunday Times* later wrote, Anthony Eden had been “the last prime minister to believe Britain was a great power and the first to confront a crisis which proved beyond doubt that she was not.”⁵⁶

The Eisenhower Doctrine

The US responded aggressively to the new power dynamic that existed in the Middle East following the Suez Crisis. Fearing that the sudden decline of British influence would create a vacuum that the USSR could fill, the president announced a new policy toward the region, the Eisenhower Doctrine, in an address to congress in January 1957. Echoing the language of the Truman Doctrine, it declared that the United States would provide military and economic aid to Middle Eastern countries that faced “overt armed aggression from any nation controlled by International Communism” and would, if necessary, even commit its forces to their defense.⁵⁷

⁵⁵ Schulze, *The Arab-Israeli Conflict*, 30–31; Yaqub, *Containing Arab Nationalism*, 63–64.

⁵⁶ Avi Shlaim, *War and Peace in the Middle East: A Concise History, Revised and Updated*, Revised ed. edition (New York: Penguin Books, 1994), 29–31.

⁵⁷ Dwight Eisenhower, “The Eisenhower Doctrine, January 5, 1957,” accessed March 28, 2022, <https://sourcebooks.fordham.edu/mod/1957eisenhowerdoctrine.asp>.

Despite the rhetorical emphasis on Marxism, Moscow was not the main focus of the Eisenhower Doctrine. Instead, as the historian Salim Yaqub has demonstrated, Eisenhower and Dulles designed the new policy primarily to contain the radical Arab nationalism that Nasser championed. In their estimation, the Egyptian president's promotion of Positive Neutrality made him an unwitting pawn whose actions threatened to give the USSR a backdoor way into the Middle East. The conservative Arab states could check him, Eisenhower and Dulles believed, but would not act directly against such a powerful and popular figure unless they enjoyed clear American backing. Thus, the Eisenhower Doctrine was designed less to directly target the USSR than to indirectly limit its influence by giving Western-leaning Arab states the support they needed to stand up to Nasser.⁵⁸

Approved by congress in March 1957, the Eisenhower Doctrine appeared to pay immediate dividends. Strengthened by the prospect of American assistance, Jordan's King Hussein purged his government of Nasserists in April and successfully staved down a challenge by insubordinate pan-Arabist army officers. Buttressed by American aid money and the deployment of the Sixth Fleet to the Eastern Mediterranean, he then solidified his grip on power by cracking down on the Nasserist opposition and by imposing martial law. Events in Lebanon reinforced the perception that the doctrine was a success. With clandestinely provided American funds, pro-Western Maronite President Camille Chamoun (r. 1952-1958) orchestrated a decisive electoral victory over the Nasserist United National Front (UNF) opposition—which had itself received covert aid from Egypt—in parliamentary elections held in June. Thus, by the early summer of 1957, the Eisenhower Doctrine seemed to be well on its way to achieving its purpose of containing Nasserism.⁵⁹

The Lebanese elections marked the policy's high point, however. For the next twelve months, Washington's effort to limit or rollback Nasserism failed completely—most notably, in Syria. Two events in that country in 1957 had brought it to the attention of American officials. First, deftly exploiting the post-Suez rise in anti-Western sentiment in Syria, the Ba'ath Party and Syrian Communist Party (SCP) had succeeded in compelling the government to increase the number of leftists in the cabinet in December 1957. Second, Moscow's influence had increased significantly when it agreed earlier that year to provide Soviet military hardware and development loans in exchange for Syrian wheat.⁶⁰

Determined to prevent Syrian pan-Arabists from delivering their country to Moscow, American officials responded by sponsoring a pro-Western coup and by prodding the conservative Arab regimes to put diplomatic pressure on Damascus. Begun with high hopes, these efforts went nowhere. The coup plot quickly fizzled thanks to the swift actions of the very capable head of the Syrian security service, Abd al-Hamid Sarraj (1925-2013). Meanwhile, the conservative Arab states' willingness to

⁵⁸ Yaqub, *Containing Arab Nationalism*, 1–2.

⁵⁹ Yaqub, 119–45.

⁶⁰ Wilford, *America's Great Game*, 268–76.

support the administration's hardline approach collapsed after Nasser launched a blistering, radio-based propaganda campaign against them that resonated with their people and thus threatened their grip on power. Seeking to end the Egyptian leader's attacks, they soon began to urge Washington to moderate its position toward Damascus. Lacking either supportive regional allies or proxies inside Syria, the administration had no choice but to comply. Thus, by the end of 1957, the once promising Eisenhower Doctrine lay in tatters.⁶¹

Nasserism at High Tide, 1958-1961

Emblematic of the doctrine's failure was the fact that Nasser's popularity reached new heights the following year. As a result of his success in the Suez Crisis, he had established himself as the unquestioned leader of Arab nationalism and as the top champion of Arab unity. Thanks to his savvy use of Radio Cairo in the Syria crisis, meanwhile, he had attained an unparalleled degree of influence over the people of the Arab world. The result was a dramatic enhancement of his soft power and, for a time, the assumption that he would bring all the Arabs together under his rule in a unified state.

The United Arab Republic

That expectation stemmed from a stunning turn of events in early 1958: the political merger of Egypt and Syria. While Cairo would ultimately dominate the union, unification had its genesis in the aspirations of the Syrian Ba'ath Party. During the mid-1950s, it had entered into a mutually beneficial marriage of convenience with the rival SCP that had allowed the two comparatively small parties to assume dominant positions in the government. To the Ba'athists' alarm, however, the country's increasingly close relationship with Moscow had boosted the SCP's standing to such a degree that it appeared likely that the communists would soon eclipse the Ba'ath and assume sole control of the state. Desperate, party leaders concluded that the only way that they could forestall that outcome was by securing the political merger of Syria and Egypt. Talk turned into action in early 1958 thanks to a precipitate move by a cabal of Ba'athist army officers. Acting without prior approval, a group of them—later joined by Ba'ath Party cofounder Salah al-Din al-Bitar—travelled to Egypt and pitched Nasser on the idea of unifying their states under his rule.⁶²

The Egyptian leader initially expressed skepticism. A moderate Arab nationalist, he had long interpreted pan-Arabism to mean that the Arab states would closely coordinate their foreign policies rather than formally join together in a political union. He was also leery of taking on the challenges intrinsic in running a country that was not territorially contiguous, and he blanched at the idea of becoming involved in Syria's byzantine political arena. At the same time, however, he grasped that union would substantially strengthen his position in the Arab world and afford him greater clout on

⁶¹ Wilford, 268–76.

⁶² Rogan, *The Arabs*, 304–6.

the global stage. He consequently set aside his doubts and agreed to the proposal subject to three conditions designed to ensure his total control of the unified state: Syria's political parties had to immediately dissolve, its military officers needed to abandon any involvement in politics, and its state institutions had to be absorbed into—and thus subordinated to—Egypt's existing government. The expansiveness of Nasser's terms surprised the Ba'athists but did not faze them. Confident that their organizational and ideological sophistication would allow them to dominate the unified government, they accepted Nasser's conditions with little debate. As a result, on February 1, 1958, Nasser proclaimed that Egypt and Syria had joined together to form a new state: the United Arab Republic (UAR).⁶³

The union of Egypt and Syria seemed to mark a watershed moment in the region every bit as significant as the Suez Crisis or the establishment of Israel. To Arab nationalists, the erasure of the borders that the European powers had imposed at the conclusion of World War I heralded nothing less than the end of imperial dominance and the beginning of an Arab revival. It was a heady moment—one that emboldened pan-Arabists throughout the region and raised expectations even in the conservative Gulf monarchies that the entire Arab world would imminently enter the union.⁶⁴

The pressure to join fell most heavily on Lebanon and, especially, Jordan. Excited by the merger of Syria and Egypt, many Jordanians joined opposition groups to demand that King Hussein abdicate as the first step in a process that would end with their country joining the UAR. This pressure was substantial and, for a time, seemed likely to sweep the Jordanian king from power. Hussein was not prepared to go quietly, however. Instead, taking advantage of the fact that he had solidified his control of the military the prior year, he responded with a proposal that used the language of unity and nationalism to fortify the Hashimite position. It called for Jordan and Iraq to come together in a federal union wherein each would continue to enjoy domestic autonomy but would fully integrate their militaries and would pursue a joint foreign policy. Equally concerned about the danger of Nasserism, Faysal II's government was receptive. Accordingly, just thirteen days after Syria and Egypt had merged, the two Hashimite kingdoms announced the birth of a rival superstate, the Arab Union.⁶⁵

Lebanon, too, felt the heat. The establishment of the UAR appeared not only to have reinvigorated the opposition to Camille Chamoun, but to have created the expectation that the country would soon join Nasser's superstate. For a time, in fact, Lebanon seemed to be teetering on the edge of civil war between the president's mostly Christian, pro-Western supporters and the largely Sunni Nasserists over the issue of whether it should merge with the UAR. Appearances were deceiving, however. In reality, the question of joining Nasser's state was of only secondary importance in spurring the political crisis that occurred in Lebanon in 1958. Far more critical in dividing

⁶³ Rogan, 306–8.

⁶⁴ Yaqub, *Containing Arab Nationalism*, 189.

⁶⁵ Charles Tripp, *A History of Iraq* (Cambridge: Cambridge University Press, 2002), 145–46.

the country was President Camille Chamoun's effort to remain in power beyond his constitutionally limited six-year term. Indeed, the country remained tense but peaceful in the months following the creation of the UAR and only descended into civil war when Chamoun made clear in May 1958 that he intended to have the constitution amended to permit him to serve a second term.⁶⁶

The Iraqi Coup, July 1958

Ultimately, the state most dramatically affected by the merger of Syria and Egypt was Iraq. The announcement of the establishment of the UAR had electrified Iraqis opposed to the old order and led many to conclude that their country's future lay in joining Nasser's superstate. Such sentiments were particularly prominent among a group of pro-Nasserist army officers who had earlier formed a secret organization called the Free Officers. Excited about the pan-Arabist implications of the UAR, they began plotting a coup in 1958 under the leadership of General Abd al-Karim Qasim (r. 1958-1963) and Colonel Abd al-Salam Arif (r. 1963-1966) aimed at ending the monarchy and finally freeing Iraq of Western dominance.⁶⁷

Their chance came in July 1958. Taking advantage of the movement of an infantry brigade through the capital—ordered, ironically, for the purpose of deterring a rumored coup against Hussein—the Free Officers made their move. In a well-planned operation, they seized government buildings, the palace, and, critically, the radio station, from which Arif urged people to take to the streets to support the new republic they had established. The subsequent, rapid destruction of the old order proved to be anarchic and bloody. The soldiers sent to detain the king instead massacred the royal family and hastily buried them in shallow graves on the palace grounds. The savvy Nuri Said managed to avoid capture for a time but he, too, was executed after he was discovered trying to slip out of Baghdad disguised as a woman. Meanwhile, the street protests that Arif had called for to protect the new government from counterrevolution rapidly got out of hand. Unleashing three decades of pent-up fury, the mob sought vengeance in an orgy of violence against anything and anyone associated with the previous regime. They tore down a statue of Faysal I (r. 1921-1933), torched the British Chancery, dug up and mutilated the body of the former regent, Abd il-Alah (1913-1958), and killed a number of officials by dragging them to death behind cars.⁶⁸

It was a grim day for Washington. Qasim and the other coup plotters followed the revolution by immediately withdrawing Iraq from the Baghdad Pact. Worse, not long after, they established close relations with the USSR. In other words, the coup had not only thoroughly gutted the alliance system that Eisenhower and Dulles had so carefully constructed for the purpose of preventing Moscow from extending its power into the

⁶⁶ Yaqub, *Containing Arab Nationalism*, 192–96, 206–11.

⁶⁷ Phebe Marr, *The Modern History of Iraq*, Second Edition (Boulder: Routledge, 2004), 83–84.

⁶⁸ Tripp, *A History of Iraq*, 146–50.

Middle East but had also given the Soviets immediate influence over an important Arab state.⁶⁹

1958 Lebanese and Jordanian Crises

Panic gripped the leaders of the conservative Arab regimes following the Iraq coup. Blaming the tumultuous events on Nasser, they frantically called on the Western states to protect their governments lest they be the next to fall. Jordan was the most vulnerable of the conservative states. Keenly aware of his country's exposure, King Hussein immediately requested the deployment of Western troops to his capital to protect his government. London was receptive. Eager to prevent Nasserism from sweeping through the entire Middle East, it quickly acceded to his request. As a result, the first of 3,700 British paratroopers began arriving on July 17.⁷⁰

Chamoun also demanded Western intervention. Claiming that Nasser would follow the Iraq coup by seeking to subvert his government, he invoked the Eisenhower Doctrine just hours after he learned of the events in Baghdad and called for the US to deploy troops to his country. The request put the administration in a difficult position. By that point, the president and his advisors had long since determined that the 1957 Lebanese political crisis was the product of confessional divisions and Chamoun's overweening ambitions rather than Nasser's interference. More broadly, having concluded from the abysmal failure of their effort to halt Syria's leftward drift that the Eisenhower Doctrine had failed to achieve its goals, they had begun to pursue a quiet rapprochement with the Egyptian leader. Nonetheless, the day after the Iraq coup, the administration acceded to Chamoun's request and began landing a force that eventually numbered more than 15,000 soldiers.⁷¹

What explains Eisenhower's apparent change of heart? Why did he agree to Chamoun's request? According to the historian Douglas Little, the driving force behind the president's decision to deploy troops was his desire to uphold American credibility. In Little's view, Eisenhower believed that the US had no choice but to act if it wished to maintain its standing in the region. As the president reasoned, if Washington failed to answer Chamoun's request for help, friendly governments in the Middle East would first begin to doubt Washington's commitment and then come to terms with Nasser to the detriment of Western interests. Thus, regardless of Chamoun's intent, the president felt he had to act.⁷²

In the end, the intervention had at best equivocal results for the US. On the positive side, Washington managed to secure a promise from Nasser to respect Jordanian and Lebanese independence in exchange for the withdrawal of British and American troops from those states. It also brokered a deal in Lebanon that resulted in

⁶⁹ Rogan, *The Arabs*, 312–14.

⁷⁰ Yaqub, *Containing Arab Nationalism*, 222, 231.

⁷¹ Wilford, *America's Great Game*, 286–87.

⁷² Little, *American Orientalism*, 134–35.

the election of the popular head of the army, the Maronite General Fuad Shihab (r. 1958-1964), to the presidency at the end of Chamoun's term. Respected by all the factions, he pursued a policy of reconciliation and moderation that soon restored order—thus permitting Eisenhower to honor his deal with Nasser and withdraw American troops. At the same time, however, the intervention also imposed significant costs on Washington. Most notably, having more-than-a whiff of neo-colonialism to it, the deployment had stripped away any residual goodwill that the United States had earned in the Arab world as a result of its support for Nasser during the Suez Crisis.⁷³

Karim Qasim

Nasserism seemed unstoppable in the summer of 1958. Lebanon and Jordan had managed to avoid being drawn into the UAR, it is true, but those setbacks were surely temporary. Far more important, the coup in Iraq seemed to suggest not only that Nasserism was the wave of the future, but that Iraq's territory, its substantial population and, most critically, its enormous oil wealth would soon be under the Egyptian leader's control. Few doubted this outcome. The US, the Soviet Union, the European states, and Arab nationalists alike agreed that it was only a matter of time before the Iraqi Free Officers petitioned Cairo to add Iraq to the UAR. The future appeared bright indeed for Nasser.⁷⁴

Those assessments proved to be off the mark. Despite the widespread belief that union was inevitable, the new government in Baghdad ultimately chose not to join the UAR. That decision reflected Qasim's victory in an internal dispute with Arif in the weeks immediately following the coup. On one level, the conflict between them was a simple power struggle over which man would rule. On another, however, it reflected a larger debate over Iraq's national identity—one that revolved around the question of whether it should remain an independent, multi-ethnic, multi-confessional country, or whether it should instead be a component of a larger, progressive, secular-but-predominantly Sunni Arab state. Arif favored the latter position, and, with the support of Iraq's Ba'athists, Arab nationalists, and many Sunni Arabs, pressed for the country to immediately join the UAR. Qasim disagreed. Uninterested in ceding national sovereignty or Iraq's ample oil revenue to Egypt, he countered by promoting an "Iraq first" policy that enjoyed the support of heretofore marginalized groups such as the Shi'i and Kurdish communities as well as the Iraq Communist Party (ICP), which was leery that Nasser would compel them to disband. With their backing, Qasim was able to force Arif out of the government in late 1958 after which he consolidated his position by purging Ba'athists and pan-Arabists. Furious, Nasser responded by launching a vigorous propaganda campaign against Qasim and by seeking to isolate Iraq in the

⁷³ Yaqub, *Containing Arab Nationalism*, 244–45, 251–53; Wilford, *America's Great Game*, 288.

⁷⁴ Little, *American Orientalism*, 202.

Arab world. Aware that the ICP was a key pillar of Qasim's rule, he also vented his frustration by clamping down on the communist party in Egypt.⁷⁵

Nasser did manage to find a silver lining in his defeat. A savvy political strategist, he believed that he could use the crackdown on the communists to improve relations with the US and to secure much-needed American food aid. He was correct on both counts. Having already decided to seek a rapprochement with Nasser following the failure of the Eisenhower Doctrine, the administration found the Egyptian leader's move against the communists encouraging. It consequently agreed to permit Cairo to purchase \$110 million worth of wheat through the PL-480 "Food for Peace" program. It was a good deal for Nasser. Payable in UAR pounds rather than in scarce hard currency, the agreement provided Egypt with much-needed aid at a heavily discounted price.⁷⁶

Nasser was nevertheless right to be upset about Qasim's decision to stay out of the UAR. Iraq's refusal to join the union was a major diplomatic defeat for him and the larger Nasserist movement he led. It demoralized nationalists throughout the region, undermined the Egyptian leader's claim to speak for all Arabs, and strongly suggested that pan-Arabism was not, in fact, the wave of the future. Despite this defeat, Nasserism remained a powerful force in the Arab world after 1958. The wind had come out of its sails, however, and few continued to view Arab unification as inevitable.⁷⁷

Demise of the UAR

Despite these diminished expectations, most observers were still shocked when Syria, in a sudden move, seceded from the UAR in September 1961. Its departure was the result of growing disenchantment in Damascus with the union. As we saw, despite Nasser's insistence that Egyptians would control the upper echelon of the government, the Ba'athists had wagered in 1958 that they would be able to outmaneuver their less-ideologically sophisticated partners and would thus assume a dominant voice in the UAR. That bet proved to be a bad one. Instead of manipulating officials in Cairo, they found themselves unhappily shunted to the side and thoroughly dominated by the imperious Egyptians. Syrian landlords and businesspeople similarly chafed under Nasserist rule. They railed against the imposition of a land-reform program that Cairo had initiated in 1958 and—capitalist to the core—fumed at the Egyptian government's heavy-handed intervention in the economy.⁷⁸

The UAR collapsed under the weight of these differences in the summer of 1961. Two events led to the final break. First, the promulgation of the July Laws—a sweeping new package of economic regulations that called for the nationalization of most businesses—destroyed what little enthusiasm for union remained among Syria's

⁷⁵ Tripp, *A History of Iraq*, 151–54.

⁷⁶ Hahn, *Crisis and Crossfire*, 44; Yaqub, *Containing Arab Nationalism*, 257–58.

⁷⁷ Rogan, *The Arabs*, 316–17.

⁷⁸ Rogan, 320.

influential landowner and business classes. Second, Cairo's decision to force Sarraj from his post as chief of intelligence in Syria infuriated many military leaders. In fact, this move was the final straw for Syria's Ba'athist army officers. Just a week after Sarraj stepped down, they responded with a coup that resulted in Syria immediately seceding from the UAR. Caught off guard, an initially furious Nasser vowed to compel Syria to return to the union. Cooler heads soon prevailed, however. Realizing that he could not regain control of Syria if its military opposed him, he abandoned the effort and agreed to a peaceful separation. Egypt formally retained the UAR name until 1971 and the two states periodically entered into perfunctory reunification talks, but, from that point forward, the united Arab state—the cherished dream of Arab nationalists—was dead.⁷⁹

Political, Social, and Economic Development, 1958-1967

While the officers who dominated Egypt, Syria, and Iraq gave a great deal of attention to foreign relations, they by no means neglected domestic concerns in the late 1950s and 1960s. On the contrary, they promoted a populist vision in all three states designed both to improve the lot of the impoverished masses and to consolidate their grip on political power. The broad outline of their efforts was similar though the specifics differed from country to country. In each state, the officers sought to foster rapid economic growth, to distribute income more equitably, and to establish political institutions designed to mobilize popular support behind the government. For a time, their efforts seemed to be quite successful.

Egypt: Arab Socialism

Egypt had already taken tentative steps down this path in the years following the Free Officers' coup. As we have seen, its government had adopted an economic development model based loosely on Mustafa Kemal's ISI system and had implemented a series of socio-economic changes such as land reform aimed at improving life for the poorest Egyptians. In the early 1960s, however, Egypt effected a rapid shift to a far more aggressive, centrally-directed economic development model dubbed Arab Socialism. Unsurprisingly, Nasser was the driving force behind this change. Increasingly influenced by Soviet central-planning ideas, he believed that it was the only way that Egypt could industrialize rapidly and equitably.⁸⁰

The transition to Arab Socialism took place with great speed in the early 1960s. Nasser initiated the process at the start of the decade when he ordered the nationalization of the largest banks and decreed the start of Egypt's first Five Year Plan (1960-1964). He went substantially further the following year when he promulgated the July Laws. Sweeping in extent, they nationalized many industries including insurance, textiles, steel production, and shipping, and resulted in the government taking a 10-50 percent stake in nearly all other firms. The July Laws also instituted a minimum wage, capped the salaries of highly compensated people, adopted a highly progressive tax

⁷⁹ Cook, *The Struggle for Egypt*, 76.

⁸⁰ Goldschmidt, *A Brief History of Egypt*, 168–69.

regime that set the top marginal rate at a confiscatory 90 percent, and effected a far-reaching land-reform program that limited individuals to no more than 100 *feddans*—a little more than 100 acres—and redistributed the excess property to landless peasants.⁸¹

Nasser completed the move to Arab Socialism in 1964. That year, the state dramatically expanded the public educational system in order to ensure that the country had the trained workers needed to manage its ongoing industrial development; as a result, the share of the population that held a college degree began to increase substantially. More importantly, the Egyptian president also issued a new decree in 1964 that resulted in the nationalization of nearly all of the firms over which the government had assumed partial ownership three years earlier.⁸²

Concurrently, Nasser orchestrated a series of changes designed both to radically restructure Egyptian society and to solidify his control of the country's people. He first reorganized the National Union into the Arab Socialist Union (ASU) and used it to mobilize the people in support of the regime and its economic program. To ensure against the emergence of organized criticism of his regime and its economic policies, he followed by having the state assume control over previously autonomous groups such as labor unions, student organizations, and universities. Finally, he oversaw the issuance of a new National Charter in 1962 that formalized Egypt's commitment to Arab Socialism and radical change. It was a truly revolutionary document. Most notably, it called for women to enjoy equal rights and acknowledged for the first time the need for contraception to help control the country's rapidly growing population.⁸³

Nasser's economic program appeared to be quite successful for a time. Indeed, the benefits of Arab Socialism appeared to be substantial. Workers enjoyed enhanced benefits, many peasants acquired land, and the distribution of income became increasingly equitable. Even the middle class benefited from the shift to Arab Socialism. While some blanched at the adoption of a state-centered development scheme, most were delighted when Nasser had guaranteed work for the middle class by decreeing in 1962 that all university graduates who could not find jobs on their own would be given government employment.⁸⁴

Economic and demographic data sustained the belief that Arab Socialism was a success. Gross domestic product rose at an impressive clip during the early and mid 1960s, for example, while Egypt's average life expectancy increased by more than three years between 1954 and 1969. It is true that per capita income only rose by 2 percent a year during the decade thanks largely to Egypt's robust population growth. It was also the case that many landlords and members of the bourgeoisie had left Egypt in

⁸¹ Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 135–36.

⁸² Cook, *The Struggle for Egypt*, 77–78; Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 135–36.

⁸³ Goldschmidt, *A Brief History of Egypt*, 171–72; Cook, *The Struggle for Egypt*, 78–82.

⁸⁴ Cook, *The Struggle for Egypt*, 71–72.

response to the imposition of Arab Socialism—an emigration that deprived it of capital and entrepreneurs. Still, on balance, the country appeared to be much better off. Nasser, it seemed, could justifiably take pride in having quickly improved Egypt's economy.⁸⁵

Iraq—Political Instability and Autocracy

In the broadest sense, Iraq's social, political, and economic development between 1958 and 1967 paralleled the course that Nasser had charted. Following the pattern set in Egypt, Iraq had a populist, military-dominated regime. Like Nasser's state, the Iraqi government also moved to distribute income more equitably and sought to break the power of the established elite. In several critical ways, however, Iraq's experience differed substantially from Egypt's development—most notably in terms of its political instability and in its failure to develop meaningful connections between the people and the state.

Iraq's difficult history after 1958 reflected the country's awkward early years. Stitched together to serve British interests, it emerged from the mandatory period with a weak civic-political culture, a powerful, coercive state, and a political elite that ruled on the basis of patronage and personal ties rather than popular support. The subsequent brief dominance in the 1930s of Sunni officers who had absorbed the late Ottoman Empire's authoritarian, militarized conception of the state only served to reinforce and institutionalize that political system. The result would be the hardening of ethnoreligious fault lines and the institutionalization of a state that conferred near-unchecked powers of coercion and patronage upon those who controlled it—a legacy that ensured that the country's post-coup governments would focus far-more on buttressing the state that granted them such overwhelming authority than on reforming society.⁸⁶

Qasim's government certainly reflected and perpetuated the pattern that had been established during the mandate and the monarchy. That it did so was far from what the coup plotters' public declarations had suggested when they took power. Indeed, immediately following the coup, they had indicated that they intended not to establish an autocratic state but instead to liberalize Iraq. The new regime promised to end the former government's corruption and exploitation, to democratize the country politically and economically, and to grant rights to minority groups—particularly the Kurds. Unsurprisingly, as a result, many of the Free Officers' civilian allies and supporters anticipated that the coup would mark a fundamental turning point for Iraq—one that would result in both its state and society undergoing a fundamental restructuring along genuinely liberal lines.⁸⁷

They were to be bitterly disappointed. Intoxicated by the enormous power that he had inherited from the Hashimite government, Qasim instead focused on quickly

⁸⁵ Goldschmidt, *Modern Egypt: The Formation of a Nation State*, 136–37.

⁸⁶ Tripp, *A History of Iraq*, 148–49, 152.

⁸⁷ Marr, *The Modern History of Iraq*, 83–84; Tripp, *A History of Iraq*, 148.

consolidating power in his hands. He did so with great skill. He first used his personal ties to other figures in the military and the enormous patronage powers that the state afforded him to cement his control of the officer corps. Next, he took advantage of fighting between the communist party and troops commanded by pan-Arab officers in Mosul in March 1959 to institute a sweeping purge of those in the army and the administration who might oppose him. Shortly thereafter, he used ICP demonstrations in Kirkuk that had devolved into violence between the communists and the city's more conservative, middle- and upper-class Turkoman people as a pretext to crack down on the now-emboldened party. Finally, in October 1959, he exploited a failed assassination attempt against him to crush Iraq's small Ba'athist party.⁸⁸

Qasim's focus may have been on exploiting the patronage powers of the state to solidify his position, but that is not to suggest that he completely abandoned the idea of advancing social progress. To the contrary, one of the new regime's first acts was the Agrarian Reform Law, a land-reform program that divided large estates into small plots and redistributed them to landless peasants. While limited—the fellahin had to pay for the land that they acquired—it nonetheless improved the position of many peasants. Qasim also instituted several reforms aimed at placating the Kurds. Shortly after taking power, he issued a new constitution that accorded them equal status in Iraq, and he welcomed the Kurdish leader Mustafa Barzani (1903-1979) back from exile in the Soviet Union. For a time, he also acknowledged the far-left Kurdish Democratic Party (KDP) as one of the country's few legitimate parties.⁸⁹

Other reforms were more-far reaching. For example, the revised Personal Status Law extended to women far-greater rights than they had previously enjoyed. Among other provisions, it ended child marriage, limited polygamy, and, most importantly, gave women equal inheritance rights. Focused on the oil industry, meanwhile, Public Law 80 substantially redressed the imbalance in power between the Iraq Petroleum Company (IPC) and Baghdad. Issued in October 1961, it returned over 99.5 percent of the IPC concession to the state including the valuable Rumaila field and thus gave the country much greater control over its oil reserves.⁹⁰

While people were universally enthusiastic about Public Law 80, Qasim's reform package as a whole engendered substantial criticism. It was especially unpopular among the Shi'i clerics. They condemned the revisions he had made to the Personal Status Law, attacked his regime's embrace of leftist ideas and tolerance of communists, and condemned his land-reform program—the latter because it reduced the tithes that Shi'i estate owners had paid. The Shi'i clergy were so put off by Qasim's reforms, in fact, that some—most notably, the young cleric Muhammad Baqir al-Sadr (1935-

⁸⁸ Tripp, *A History of Iraq*, 156–58; Marr, *The Modern History of Iraq*, 94.

⁸⁹ Marr, *The Modern History of Iraq*, 98–100, 104–5.

⁹⁰ Marr, 100–102.

1980)—abandoned their traditional quietism and became active in political groups such as the Islamic *Da'wa* Party.⁹¹

For Barzani and the Kurds, in contrast, it was the *absence* of meaningful reforms that fueled dissatisfaction with Qasim's government. The primary issue for them was autonomy. Barzani demanded that the central government permit the use of Kurdish in education, that it devolve substantial power to local authorities, and, above all, that it grant the Kurds immediate autonomy. Qasim resisted those conditions. Concerned that Barzani would try to transform autonomy into full independence, he sought to walk a fine line in which he appeared to move toward granting autonomy even as he dragged his feet on ceding any real power to the Kurds. No stranger to Iraq's hardball political tactics, Barzani was quick to deduce that Qasim did not intend to meet Kurdish demands. Accordingly, after quietly building up his military power in the rural north, he issued an ultimatum in August 1961 demanding immediate autonomy. When Qasim refused to meet it, he ordered his forces to revolt.⁹²

By that point, Qasim's grip on power was beginning to slip. Three events contributed to his increasingly vulnerable position. First, his mishandling of the Kurdish situation resulted not only in the revival of the previously moribund Kurdish insurgency, but, to the frustration of his fellow officers, soon mired the Iraqi army in an expensive and seemingly unwinnable war. Second, Baghdad's growing isolation in the Arab world increasingly troubled key members of the Free Officers. Already bitter toward Qasim for sidelining them after the 1958 coup, these pan-Arabist, Sunni officers resented his repudiation of Nasser and bitterly opposed his "Iraq first" foreign policy. Finally, Qasim's ill-thought-out attempt to annex Kuwait following that country's independence from Britain in June 1961 further weakened his position. This effort revived a 1938 claim to the emirate and furthered what the historian Charles Tripp calls an Iraqi "national myth": the idea that Kuwait fell within Iraq's "natural" boundaries and that the British had maliciously hived the emirate off to weaken it. Bold but amateurish, Qasim's effort to annex Kuwait went nowhere. The quick deployment of British troops to the emirate and Nasser's equally rapid extension of a security guarantee to it compelled the Iraqi leader to hurriedly back down. The diplomatic equivalent of shooting oneself in the foot, the failed attempt cost him a great deal of prestige and, critically, undermined his standing among his fellow officers.⁹³

As dissatisfaction with Qasim surged, a coalition consisting of the reconstituted Ba'ath Party and a group of pan-Arabist officers led by Qasim's former partner, Abd al-Salam Arif, organized a conspiracy to overthrow the Iraqi dictator. They moved against Qasim in February 1963. Taking advantage of his failure to establish the connections between state and society that could be used to mobilize the masses to support him, the plotters overwhelmed regime loyalists after two days of bitter fighting in Baghdad and then arrested and summarily executed Qasim. They followed by establishing a

⁹¹ Marr, 103–7.

⁹² Marr, 103–7.

⁹³ Marr, 107, 111–15.

Ba'athist state ruled by a National Council of the Revolutionary Command (NCRC) that included a mix of party members and pan-Arab officers; Abd al-Salam Arif, who was not himself a Ba'athist, served as president.⁹⁴

Sharply divided along ideological lines between the radical Ba'athists and the more conservative pan-Arabist military officers, the regime was not destined to last. The two factions clashed over a number of points including the broader question of the country's future direction and the more immediate issue of which group would dominate their increasingly shaky coalition. A particular sticking point for the officers was the Ba'ath Party's development of a militia—the Ba'athist National Guard—that it used to bloodily suppress the rival ICP. The more-moderate pan-Arabist officers were delighted that their governing partners had destroyed the communists as a political force; however, they also viewed the Ba'ath Party's growing radicalism and, especially, the rapid expansion of its militia with mounting trepidation. Fearing that the Ba'athists would become unassailable if their paramilitary force were permitted to grow unchecked, Abd al-Salam Arif decided to mount a pre-emptive move against them in 1963. Launched in November, the operation succeeded both in destroying the National Guard and in purging the Ba'athists from the NCRC.⁹⁵

Abd al-Salam Arif had indicated to his supporters that he planned to liberalize Iraq once he had taken control of the government. As he came to understand fully the unchecked authority he commanded, however, he became intoxicated with power and instead opted to establish an authoritarian regime that bore a striking resemblance to Qasim's government. Like his predecessor, he made no effort to legitimize his rule by fostering a civic-political culture in Iraq or by securing the consent of the governed; instead, he chose to rule through a combination of the state's enormous powers of coercion, patronage, and the personal relationships that he had established with key members of the officer corps. The central role that ties of mutual obligation and patron-client relationships played in the functioning of his government was made clear in the immediate aftermath of his death in a helicopter crash in 1966. As it was well understood that continued stability in Iraq rested on the perpetuation of the network of personal connections, mutual obligations, and patronage that Arif had created, the officers decided to replace him with the only person who could be trusted to maintain that system: his brother, Abd al-Rahman Arif (r. 1966-1968).⁹⁶

Syria—the Ba'ath in Power

Though the Iraqi Ba'ath party had lost power in 1963—temporarily, as we shall see in the next chapter—its sister branch managed to assume control of Syria the same year. Its return to power was largely a function of the weak nature of the government in Damascus following the demise of the UAR. The conservative military officers who had orchestrated Syria's secession from Egypt had opted to reestablish the parliamentary

⁹⁴ Tripp, *A History of Iraq*, 169–71.

⁹⁵ Marr, *The Modern History of Iraq*, 117–23.

⁹⁶ Tripp, *A History of Iraq*, 171–85.

system that had existed prior to the 1958 union and had entrusted the government to the same class of urban notables that had earlier dominated the country. Out of step with the times, those men were unable to provide Syria with effective leadership or to control the cliques that had emerged within the officer corps. As future Syrian President Hafiz al-Asad (r. 1970-2000) aptly put it, “[i]t was a government without a people and without an army, the rule of a class which had had its day.” Taking advantage of the regime’s lack of popular support, Ba’athism’s founders, Aflaq and al-Bitar, set about reconstituting the Ba’ath Party and began scheming with dissatisfied army officers to overthrow the government. They did not have to wait long. Taking advantage of the regime’s lack of support, a coalition of Nasserist officers and the Ba’ath Party launched a successful coup that ended parliamentary government in Syria once and for all.⁹⁷

The new government proved every bit as wobbly as the one it had supplanted. Its instability stemmed from two significant problems. First, with ostensibly like-minded regimes in power in Syria, Iraq, and Egypt, the new regime felt compelled to join in unification talks aimed at bringing the three states into a re-formed UAR. On the surface, these discussions were amicable ones between like-minded leaders seeking a common end. In reality, however, Nasser, the Ba’athists in Syria, and those in Iraq were bitter rivals, and each would only accept union if their regime enjoyed complete dominance. As a result, the negotiations were little more than a Potemkin exercise in which diplomats politely discussed the finer points of political unity while their governments and parties fired off propaganda broadsides at each other. Unsurprisingly, the unity talks soon collapsed.⁹⁸

Second, the new government found itself paralyzed by a split in the Ba’ath Party between a civilian faction and a group of officers called the Military Committee. Dominated by Aflaq, the civilians stressed pan-Arabism and downplayed Ba’athism’s calls for dramatic economic reforms. In contrast, their more radical opponents in the Military Committee deemphasized Arab unity in favor of a focus on imposing a quasi-Marxist-Leninist system on Syria. A theoretician rather than a political infighter, Aflaq was unable to slow either the radicals’ full-throated assault on Syria’s traditional market-based economy or their increasing dominance of the party and government. Finally, in 1966, the Military Committee launched a coup that formalized its growing power by forcing Aflaq out and by putting the radical General Salah Jadid (r. 1966-1970) in control of the government.⁹⁹

The 1966 coup had three important ramifications for Syria. First, the purge of Aflaq created an open rift with the Iraqi branch of the Ba’ath Party, which had remained supportive of the movement’s founder. As a result, following the Ba’athist seizure of power in Baghdad in 1968, Syria and Iraq found themselves locked in a bitter cold war

⁹⁷ Patrick Seale, *Asad: The Struggle for the Middle East* (Berkeley: University of California Press, 1989), 67–80.

⁹⁸ John McHugo, *Syria: A History of the Last Hundred Years* (New York: The New Press, 2014), 144; Tripp, *A History of Iraq*, 174.

⁹⁹ Seale, *Asad*, 81–83, 86–104.

as each vied for dominance of the broader Ba'athist movement. Second, the change in government resulted in Syria pursuing a radical economic restructuring plan aimed at replacing the country's traditional capitalist economic system with a socialist one. Forced through on an aggressive timetable, it created significant economic problems and social dislocation for the country even as it succeeded in imposing state dominance of the economy. Finally, the coup resulted in Syria pursuing a radically aggressive foreign policy. Under Jadid, it developed close ties with the Soviet Union, expressed hostility toward the conservative Arab states, and, most importantly, adopted recklessly bellicose policies toward Israel that would—as we shall see presently—help to spark a third Arab-Israeli war.¹⁰⁰

The Arab Cold War

Saudi-Egyptian Proxy War

Meanwhile, even as Egypt, Syria, and Iraq competed with each other for dominance of pan-Arabism, a broader, ideological conflict was emerging between them and the conservative, monarchical Arab states. Dubbed the Arab Cold War by the political scientist Malcolm Kerr, it took shape as a result of a coup in Yemen in the early 1960s. Inspired by Egyptian radio propaganda, Nasserist army officers led by General Abdallah al-Sallal (r. 1962-1967) deposed the Zaidi Shi'i Imam Muhammad Badr (r. 1962) in September 1962 and declared the establishment of a Nasserist state: the Yemen Arab Republic (YAR). For a brief time, it seemed like another victory for Nasserist pan-Arabism. To the chagrin of the officers, however, the imam had managed to slip out of the palace before al-Sallal's troops could arrest him and had escaped to the mountainous north. There, he set to work organizing a pro-royalist tribal insurgency against the new government.¹⁰¹

The civil War in Yemen soon expanded into a broader ideological contest between Nasserist Egypt and a Saudi-led coalition of conservative Arab monarchies. The regionalization of the contest between the imam and al-Sallal started almost immediately after the coup. Fearing that radical republicanism would spread throughout the Arabian peninsula if the Yemeni Nasserists were able to consolidate their position, the Saudi regent, Faisal bin Abdulaziz al Saud (r. 1964-1975), began to supply the rebels with arms in the fall of 1962. Nasser felt compelled to respond. The Kingdom's actions not only challenged his dominance of the Arab world but, as important, also presented him with a golden opportunity to restore his tarnished reputation following Syria's embarrassing secession from the UAR. Accordingly, he quickly agreed to provide aid to the republicans—first by supplying them with weapons and then by deploying 20,000 Egyptian troops to Yemen. As a result, what had begun as a local,

¹⁰⁰ Seale, 104–24.

¹⁰¹ Cook, *The Struggle for Egypt*, 77; See also Malcolm Kerr, *The Arab Cold War: Gamal 'Abd al-Nasir and His Rivals, 1958-1970*, 3rd ed. (New York: Oxford University Press, 1971).

internal dispute in Yemen had, by the end of 1962, become a full-fledged proxy fight between Egypt and Saudi Arabia.¹⁰²

Begun with high hopes, the Egyptian intervention in Yemen quickly turned into a debacle. Despite eventually sending 70,000 of its best-equipped troops to the YAR, Egypt found itself incapable of defeating the imam's increasingly sophisticated guerrilla fighters. Indeed, "Nasser's Vietnam" quickly turned into a major drain on Egypt. The war sapped the country's economy, weakened morale in its military, and damaged its leader's reputation—particularly after his troops employed chemical weapons against the rebels. It also produced enormous numbers of casualties. All told, an estimated 26,000 Egyptian soldiers died in Yemen between 1962 and the withdrawal of Nasser's expeditionary force in 1968.¹⁰³

Nasser's intervention in Yemen also brought his country into conflict with the US. Initially, President John F. Kennedy (r. 1961-1963) hoped that the war would not interfere with recently improved ties between Washington and Cairo. Indeed, in part to forestall such an eventuality, he pressed both the royalists and the republicans to accept a UN-brokered end to the conflict. Nasser's increasingly bellicose rhetoric and his decision to deploy Egyptian troops to Yemen were too much for Kennedy, however. As a result, the US began to abandon its effort to cultivate Nasser in 1963 in favor of a return to its earlier containment policy. The Kennedy administration made this shift emphatically clear in 1963 when it responded to an Egyptian attack on royalist guerrilla bases inside Saudi Arabia by deploying a squadron of F-100 fighter-bombers to the kingdom. Far less eager to conciliate Nasser, Kennedy's successor, Lyndon Johnson (r. 1963-1969), accelerated the pivot back to a policy of containment—most notably by ending PL-480 wheat sales to Egypt in 1965. Nasser was furious about what he viewed as an attempt to blackmail his country. In response, he launched a fresh wave of propaganda attacks against the West, held a prominent meeting with the Argentine Marxist revolutionary, Che Guevara (1928-1967), and—in an action especially nettlesome to Johnson—allowed the communist Vietnamese National Liberation Front (NLF) to establish an office in Egypt.¹⁰⁴

The Middle Eastern Arms Race

The United States and Israel

Washington answered Nasser's growing hostility in part by seeking to build up Israel as a regional counterweight. Kennedy initiated this new policy in 1962. Believing that Egypt's Soviet-supplied bombers gave Nasser a first-strike advantage that could destabilize the region, the president broke with America's longstanding refusal to provide weapons to Israel and agreed to sell it powerful Hawk anti-aircraft missiles.

¹⁰² Warren Bass, *Support Any Friend: Kennedy's Middle East and the Making of the U.S.-Israel Alliance* (New York: Oxford University Press, 2003), 102–3.

¹⁰³ Quote from Bass, 143; Cook, *The Struggle for Egypt*, 77.

¹⁰⁴ Little, *American Orientalism*, 184–86.

Johnson went considerably further. After reassessing the strategic balance in the Middle East in 1965, he moved to counter the Arab states' growing arsenals of Eastern Bloc weapons by supplying offensive American arms to Israel for the first time. That year, he agreed to sell to the Jewish state 210 powerful M-48 tanks. He followed in 1966 by providing the IDF with forty-eight advanced A-4 Skyhawk jets.¹⁰⁵

Rising Soviet Influence

Moscow's arms sales to the republican Arab states had indeed been substantial. Between 1961 and 1965, Egypt alone had acquired from the USSR 530 tanks, 170 MiG-19 and MiG-21 fighters, and 25 Tu-16 bombers. Syria and Iraq received similarly substantial deliveries of advanced Soviet weapons. Why did the USSR lavish these arms on its Arab allies? What did it hope to gain? Moscow had two primary reasons for supplying the weapons. First, it did so as part of a quid pro quo designed to achieve one of its key Cold War-related security goals. In exchange for advanced weapons, the USSR received the right to develop bases at Alexandria in Egypt and at Tartus in Syria from which its aircraft and ships could hunt American Polaris-missile-armed strategic submarines operating in the Eastern Mediterranean. Second, providing arms to the Arab republics left those states dependent on Moscow for spare parts, ammunition, and technical support and thus dramatically enhanced Soviet influence in the region.¹⁰⁶

The Six Day War

The flood of high-tech weapons into the region helped set the stage for another enormously consequential conflict: The Six Day War. Ending in an overwhelming Israeli victory, the confrontation reshaped the Middle East's political and diplomatic structure and fundamentally altered the nature of the Arab-Israeli conflict. It is difficult to overstate the importance of this brief war. Indeed, in a century replete with turning points, the Six Day War stands alongside the post-World War I division of the region, the establishment of Israel in 1948, and the events of 1979 as one of the defining moments in the creation of the modern Middle East.

Rising Tensions

Rising tensions in the mid 1960s set the stage for the conflict. After a period of relative quiet in the late 1950s and early 1960s, Arab-Israeli relations took a turn for the worse as a result of a conflict over a resource that is peculiarly scarce in the region: water. In late 1963, Tel Aviv stood on the verge of completing the National Water Carrier, a major infrastructure scheme designed to divert water from the Sea of Galilee in the north to the arid Negev Desert in the south. Israelis were excited about the project

¹⁰⁵ Bass, *Support Any Friend*, 148–50, 155–57; Zach Levey, “The United States’ Skyhawk Sale to Israel, 1966: Strategic Exigencies of an Arms Deal,” *Diplomatic History* 28, no. 2 (April 2004): 255–76.

¹⁰⁶ Levey, “The United States’ Skyhawk Sale to Israel, 1966: Strategic Exigencies of an Arms Deal,” 259; Khalidi, *Sowing Crisis*, 29–30, 37.

and the growth that it promised to bring. The Arabs, in contrast, found it deeply troubling. They feared that it would give Israel the arable land needed to absorb another three-million Jewish immigrants and would thus permit it to become so populous that its defeat and the restoration of the Palestinian refugees to their homes would become nearly impossible.¹⁰⁷

Many Arab leaders saw not just crisis in the Israeli project, however, but opportunity as well—the opportunity, that is, to take Nasser down a peg or two. Tired of the Egyptian leader’s interference in their internal affairs, radical republicans and conservative monarchs alike had long had their eyes out for any means of chipping away at his dominant position in the Arab world. The gap between Nasser’s vocal anti-Zionism and his caution with regard to the water diversion plan seemed to provide just the opportunity they been looking for. Accordingly, taking advantage of a well-known fact in Arab political circles—that Arab leaders could weaken their rivals by denouncing them for being soft on Israel even when they themselves were privately opposed to any action that risked bringing down the wrath of the IDF—states like Jordan and Syria began firing off a series of withering propaganda attacks in which they chided Nasser for his failure to match his anti-Israeli rhetoric with commensurate action against the water carrier. As a result, the Egyptian leader’s standing in the Arab world began to fall.¹⁰⁸

Smarting from the criticism, Nasser moved to restore his anti-Zionist bona fides at an Arab League summit in Cairo called to address the water diversion scheme. Meeting in January 1964, the attendees agreed to put down their differences and to unify in opposition to Israel and its plans to use Galilee water to reclaim land in the Negev. The meeting proposed two concrete actions. First, the attendees established the United Arab Command (UAC) to coordinate their militaries and placed it under Egyptian control. Second, with financial support from the other Arab states, they called on Syria to divert the headwaters of the Jordan River in order to deny Israel the water needed to irrigate the Negev. The summit thus appeared to have been a great success for Nasser. It had put a stop to the propaganda attacks against him and it had done so in a way that neither required Egypt to fight a war with Israel that he believed—correctly—it would lose nor cost him his standing as the leading anti-Zionist figure.¹⁰⁹

Begun with high hopes, the effort to redirect the Jordan proved to be entirely unsuccessful in the face of unusually fierce Israeli opposition. The intensity of the resistance to the scheme was a function of what Brigadier General Israel Lior called “Syrian Syndrome”: the tendency of IDF officers stationed on the northern front to hold a particularly strong commitment to the ideology of the Iron Wall. Enthusiastically employing disproportionate force against the bulldozers, tractors, and dredges involved in the scheme, those officers were able to prevent Syria from making any progress in

¹⁰⁷ Michael B. Oren, *Six Days of War: June 1967 and the Making of the Modern Middle East* (New York: Oxford University Press, 2002), 18–19.

¹⁰⁸ Oren, 19.

¹⁰⁹ Oren, 18–21.

1964 and 1965. The failure of the diversion effort frustrated and embarrassed Syria's Ba'athist government and led it to seek some way of hitting back at the Israelis. Aware that its military could not hope to stand against the might of the IDF, it opted for an alternative, asymmetric response in which Syrian-supported Palestinian guerrillas mounted raids into Israel either directly across the border or through Jordanian territory.¹¹⁰

These incursions and the fighting in the Galilee put Nasser in a difficult position. That Syria rather than Egypt was leading the charge against Israel and that Damascus rather than Cairo was risking retaliation by permitting Palestinian raids from its territory was embarrassing to the man who had been the self-declared leader of the anti-Zionist cause since his diplomatic victory in the Suez Crisis. Even worse, his decision to accept command of the UAC—which seemed at the time to insulate him from censure—had in fact left him vulnerable to renewed criticism from his Arab rivals. Most notably, after Israel dramatically raised tensions by retaliating for the Palestinian raids with a devastating attack on the West Bank village of Samu in 1966, they tore into him for his failure as UAC leader to retaliate. Long the subject of the Egyptian leader's propaganda broadsides, Jordan's King Hussein was particularly unrelenting in his criticism. Where, he asked bitterly, was Nasser's vaunted air force? Why, he added, did Egypt continue to hide behind the UNEF peacekeepers that separated its forces from Israel? By the end of 1966, such criticism had put Nasser squarely on the defensive. Already weakened by his country's seemingly endless entanglement in Yemen, he was now in danger of losing his influence in the Arab world.¹¹¹

The Road to War

In the meantime, the situation on the Israeli-Syrian frontier intensified dangerously in early 1967. The proximate cause of the rise in tensions was the change that had occurred in Syria's government in 1966. More extreme and aggressive than the government it replaced, Salah Jadid's radical Ba'athist regime had encouraged the Palestinian guerrillas to step up the frequency of their attacks into Israel. IDF officers afflicted with "Syrian syndrome" had little patience for the resulting guerrilla incursions or for Syria's tolerance of them and responded by seeking to incite a clash that could justify punitive action. Their efforts bore fruit on April 7 when Syrian forces fired on an armored tractor that the IDF had, as a provocation, sent into the Demilitarized Zone (DMZ) that had divided the two countries since 1948. Escalating rapidly, the incident ended with Israeli aircraft shooting down six Syrian MiGs and—to the intense embarrassment of the Syrian government—overflying Damascus.¹¹²

The clash not only dramatically ramped up tensions, but also led directly to an episode that seemed to give Nasser the opportunity to restore his battered prestige as

¹¹⁰ Shlaim, *The Iron Wall*, 242–47.

¹¹¹ Oren, *Six Days of War*, 24–27, 31–38.

¹¹² Tom Segev, *1967: Israel, the War, and the Year That Transformed the Middle East* (New York: Henry Holt and Company, 2005), 195–204.

the leader of the Arab cause against Israel. On May 13, 1967, Moscow—likely trying to bolster its Syrian ally—falsely informed Egypt that Israel was massing troops in the north for an invasion of Syria. Nasser quickly determined that the Soviet warning was untrue; nonetheless, he followed it by taking a very public action that appeared to be designed to deter Israel. Shortly after he received the Soviet warning, he ordered the Egyptian army to call up its reserves and to deploy offensive units into the Sinai. If Israel invaded Syria, his action appeared to say, it would face an immediate Egyptian attack.¹¹³

What impelled Nasser to undertake such a seemingly aggressive action if he knew that the Soviet report of an impending Israeli attack on Syria was false? He did so not because he wanted a war—he understood all too well that the Arabs lacked the power to defeat Israel, particularly with the cream of the Egyptian army hundreds of miles away in Yemen—but because he saw in it an opportunity to restore his standing as the leading anti-Zionist. That is, he acted because he believed that he could, in doing so, make a credible claim that his troop deployment had intimidated the Israelis and had thus saved Syria from an IDF attack; best of all, since Tel Aviv was not in fact preparing to attack Syria, he could do so without running any real risk of war. In other words, as the historian Avi Shlaim notes, the entire operation was “an exercise in brinkmanship”: an elaborate charade designed to impress Arab public opinion and to silence his critics without drawing Egypt into a conflict it could not win. As such, it had been a huge success. Nasser was once again the toast of the Arab world and the undisputed leader of the anti-Zionist cause.¹¹⁴

It was at that point, however, that events began to get ahead of Nasser. Carried along by the accolades he had received in the Arab world, he followed the deployment in the Sinai by demanding that UN Secretary General U Thant (1909-1974) remove the UNEF peacekeepers that had been stationed along the Egyptian side of the border with Israel since the Suez Crisis. Nasser’s request posed a serious dilemma for the secretary general. On the one hand, protocol demanded that he withdraw the peacekeepers immediately; after all, Egypt was a sovereign state and thus had the right to order the UNEF forces to leave. On the other, the pursuit of a peaceful resolution of the crisis suggested that he should slow-walk Nasser’s request and thus give the Egyptian leader ample opportunity to reverse course.¹¹⁵

Acting correctly but unwisely, U Thant went with protocol. Stressing that the UN had to respect the wishes of a sovereign state, he agreed to remove the UNEF troops. Worse, he not only insisted that the UN had to withdraw *all* of the peacekeepers—including those stationed in the Gaza Strip and, crucially, those positioned at Sharm al-Shaykh—but ordered them to leave on an accelerated timetable. By the end of May, as

¹¹³ Laura James, “Egypt: Dangerous Illusions,” in *The 1967 Arab Israeli War: Origins and Consequences*, ed. William Roger Louis and Avi Shlaim (Cambridge: Cambridge University Press, 2012), 58–60.

¹¹⁴ Shlaim, *The Iron Wall*, 252–53.

¹¹⁵ Oren, *Six Days of War*, 69–75.

a result, no barrier separated Egyptian troops from Israeli soldiers, and nothing prevented Nasser from reoccupying the entire Sinai Peninsula. A victim of his own success, the Egyptian president now faced enormous pressure in the Arab world to deploy troops along the Israeli border and to reimpose the blockade of the Strait of Tiran. Unable to resist, he bowed to the might of popular opinion and did so—ordering his soldiers forward and declaring the strait once again closed to ships bound for Israel.¹¹⁶

In the immediate term, Nasser's actions were a huge success. Having apparently intimidated Israel, which had made no response to the Egyptian army's redeployment, he had reclaimed his position as the leading anti-Zionist and had regained the adulation of the Arab world. Indeed, Baghdad, Beirut, Jerusalem, Hebron, and many other Arab cities saw massive, spontaneous rallies in support of him and his aggressive actions. That public boost paid substantial and immediate diplomatic dividends for Nasser. Most notably, support for the Egyptian leader in Jordan impelled King Hussein—typically at odds with Nasser—to travel to Cairo and sign a mutual defense treaty with Egypt at the end of the month.¹¹⁷

While Nasser and Egypt were flying high at the end of May, Israel was gripped with indecision and anxiety. The deployment of Egyptian forces to the Sinai had not initially alarmed most Israelis; after all, the UN continued to maintain peacekeepers along the border. Complacency had given way to shock and fear, however, when U Thant had ordered the removal of UNEF forces and, especially, when Nasser had reinstated the blockade of the Strait of Tiran. From that point, the country experienced “the waiting,” a time of searing anxiety during which its people—many of them survivors of the Holocaust—dug trenches, filled sandbags, and prepared for the worst. The mood was little better in the government. Indeed, gripped by fear that the Jewish state was in mortal peril, a quiet but vigorous debate was taking place within the cabinet. On one side were those like Chief of Staff Yitzhak Rabin (1922-1995) who, contending that the IDF's deterrent value would wither away if it failed to respond to the closure of the strait, called for an immediate, preemptive attack. On the other, were those like Prime Minister Levi Eshkol (r. 1963-1969) who urged caution and who argued that Israel could only act if it could first convince a reluctant Washington to support such a move.¹¹⁸

It was a peculiarly difficult few weeks for the cabinet. Caught between American reluctance and Egyptian bellicosity, Israeli leaders found themselves operating under crushing stress. Finally, in early June, a path forward opened up: American Secretary of Defense Robert McNamara (1916-2009) informed the head of Mossad, the Israeli intelligence service, that Johnson had experienced a change of heart and was now

¹¹⁶ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 306; Oren, *Six Days of War*, 81–84.

¹¹⁷ Segev, *1967: Israel, the War, and the Year That Transformed the Middle East*, 299–300.

¹¹⁸ Oren, *Six Days of War*, 61–63, 75–82, 114–15.

willing to support an IDF preemptive attack. The news came as an enormous relief to Eshkol's government and ended the indecision that had gripped the cabinet. Now enjoying a clear green light from Washington, it promptly voted to launch a surprise attack scheduled to begin at daybreak on June 5.¹¹⁹

The Conflict, June 5-10, 1967

Ironically in light of the handwringing in Tel Aviv that preceded it, the war was an Israeli rout from start to finish. The IDF initiated the conflict on June 5 by mounting a surprise, dawn air attack that annihilated nearly the entire Egyptian air force as it sat defenseless on the ground. Desperate for assistance, Nasser lured Jordan, Iraq, and Syria into launching a series of air strikes against Israel later that morning by claiming that Egypt was in the process of inflicting a decisive defeat on the IDF. After easily parrying those attacks, Israel responded with a sequence of devastating, late-afternoon retaliatory raids on Jordanian and Syrian air bases. Israel's first-day air offensive had been a total success. All told, the IDF had destroyed more than 400 Egyptian, Syrian, and Jordanian planes and had, in doing so, secured total air superiority.¹²⁰

The ground war was similarly one-sided. Israeli troops launched a series of devastating and well-executed offensive drives beginning on June 5 that sent Arab armies reeling. As a result, by the time the war ended with a UN-brokered ceasefire on June 10, the IDF had seized the Golan Heights from Syria, the West Bank and Jerusalem from Jordan, and the Gaza Strip and Sinai Peninsula from Egypt. In doing so, it had tripled the territory that Israel controlled.¹²¹

Aftermath

The sheer totality of Israel's victory over the Arabs had consequences both profound and replete with irony, but four stand out. First, the defeat thoroughly discredited many of the Arab governments. Indeed, much as the 1948 War had delegitimated earlier regimes and generated a wave of political change, the Six Day War produced a series of coups that altered the political landscape in the Arab states of the Middle East and North Africa. Ironically, however, none of the changes in government that followed the Six Day War involved the overthrow of the officer-dominated states that had failed so spectacularly in the Six Day War. Instead, the wave of coups that began in 1967 either reshuffled the leadership of the existing military-controlled governments—as happened in Syria and Iraq—or resulted in radical Arab nationalists toppling conservative regimes such as occurred in Libya.¹²²

Second, the war altered and complicated the relationship between Tel Aviv and Washington. Following the conflict, the US pursued a two-pronged approach to Israel.

¹¹⁹ Shlaim, *The Iron Wall*, 252–57.

¹²⁰ Shlaim, 257–67.

¹²¹ Shlaim, 257–67.

¹²² Rogan, *The Arabs*, 340–41.

On the one hand, it pressed Tel Aviv to accept a comprehensive settlement with the Arab states on the basis of the exchange of the land it had conquered in 1967 for peace and recognition. On the other, it abandoned its earlier pretense of neutrality in the Arab-Israeli conflict and thereafter assumed the role of Israel's military, diplomatic, and economic patron—a relationship to which Washington lent substance by providing Tel Aviv with fifty advanced, F-4 Phantom fighter-bombers in 1968. Though it was not entirely obvious at first, these policies worked at cross purposes. Ironically in fact, the more that the US provided Israel with military and diplomatic support, the more that Tel Aviv could safely ignore Washington's efforts to press it to trade land for peace.¹²³

Third, the war set in motion a process that would fundamentally change the nature of the Arab-Israeli conflict. At first, the totality of the IDF victory seemed to suggest that the contest was effectively over. Indeed, Israel's triumph had been so total that it appeared to have accorded the Jewish state complete dominance in the struggle. However, even as Israel's success had given it a commanding position vis-à-vis the Arab states, its assumption of control over the West Bank, Jerusalem, and the Gaza Strip had revived the older conflict between Israelis and Palestinians. Thereafter, Israeli rule created a new, shared experience of oppression among the Palestinians that would quickly reinvigorate their sense of national identity and their demands for self-rule. In other words, as total as the Israeli victory over the Arab states may have been, it merely created a more visceral and immediate clash between Zionists and Palestinians and, thus, as we shall see in chapter sixteen, a return of the struggle to its origins as a contest between nationalisms.¹²⁴

The final irony was that neither side was prepared—despite the enormity of Israel's military victory—to engage in peace negotiations. In Tel Aviv, a sense of hubris almost instantly replaced the ennui that had gripped the nation during “the waiting”; as a result, while the government theoretically endorsed the idea of trading land for peace, it felt no pressure to negotiate with the Arab states or to yield any of the strategic territory it had acquired. For Arab leaders, in contrast, both the sheer extent of their defeat and the vast degree of Israel's continued military superiority rendered productive talks impossible. How could they negotiate when Israel held all the cards? As a result, even as it was clear that their armies lacked the strength to regain the territory lost in June, the Arab governments issued the famous “three nos” at the Khartoum Conference in September 1967: no negotiations, no recognition of Israel, and no peace.¹²⁵

Resolution 242

In November, meanwhile, the UN Security Council established the parameters for a broad settlement of the Arab-Israeli conflict through the passage of Resolution

¹²³ Yaacov Bar-Siman-Tov, “The United States and Israel since 1948: A ‘Special Relationship’?,” *Diplomatic History* 22, no. 2 (1998): 241–42; Rogan, *The Arabs*, 340–41; Hourani, *A History of the Arab Peoples*, 413–15.

¹²⁴ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 338–43.

¹²⁵ Shlaim, *The Iron Wall*, 274–77.

242. On the surface, its call for Israel to withdraw “from territories occupied in the recent conflict” and its declaration that “every state in the area had the right to “live in peace within secure and recognized boundaries” appeared to establish the basis for a straightforward settlement in which Israel would exchange the land it had taken in the Six Day War for peace and recognition. In reality, Resolution 242 was a complex and slippery document—“a masterpiece of deliberate British ambiguity,” in Avi Shlaim’s words—that employed intentionally vague language that left it subject to competing interpretation. Pointing to the French version, which used the definite article, The Arabs argued that it required Israel to withdraw from *all* of the territories it had taken in 1967; Tel Aviv, in contrast, maintained that the reference to secure borders left open the possibility that it could retain Jerusalem and those portions of the West Bank that it felt were necessary for its national security. Thus, in the short term, the resolution designed to establish the basis of a durable peace merely served to create a new arena of rhetorical dispute. Still, Resolution 242 did mark an important milestone in the Arab-Israeli conflict and remains to the present day the starting point for all negotiations between Israel, its Arab neighbors, and the Palestinians.¹²⁶

The Historical Debate: Israel’s Motives

Unsurprisingly, given its significance, the Six Day War has proven to be an enormously contentious subject. Indeed, nearly every aspect of the war has produced intense debate. One of the most hotly contested issues related to the conflict was the question of Tel Aviv’s motives in launching the surprise attack on June 5. Calling attention to the speed and extent of the IDF’s achievement on the first day of the war, many Arab governments contended even before the fighting had ceased that Israel had initiated the war not out of fear or for defensive purposes, but instead out of a desire to seize territory. In support of this argument, they have noted that the Israeli government was well aware that the most experienced Egyptian formations remained mired in Yemen and have pointed out that the US had shared intelligence with Tel Aviv indicating that the formations Nasser had sent into the Sinai had assumed a defensive posture and were thus in no position to attack Israel.¹²⁷

Other officials and scholars—even many who are often critical of Israel—dispute this view, however. Noting the deep concern that gripped the Jewish state in the weeks leading up to the conflict, for example, the historian Avi Shlaim contends that Tel Aviv acted out of a fundamental concern for its security and that it gained land through opportunity and military exigencies rather than through a preconceived strategy. As he writes, the Eshkol government had no “master plan for territorial aggrandizement. Its territorial aims were defined not in advance but in response to developments on the battlefield.” Others emphasize the role of credibility in Israeli thinking. As the historian and former Israeli Ambassador to the United States Michael Oren argues, Eshkol’s

¹²⁶ “United Nation’s Resolution 242,” November 22, 1967; Shlaim, *The Iron Wall*, 277–78; Oren, *Six Days of War*, 136.

¹²⁷ Shlaim, *The Iron Wall*, 257–58.

government went to war not to gain land but instead because it feared that failing to do so would suggest weakness and would thus invite a devastating Arab attack.¹²⁸

Conclusion

Though the pan-Arabist officer-dominated states survived and even flourished in the years following the Arabs' stunning defeat, the Age of Nasser itself came to an abrupt end the moment Israeli bombs began to strike their targets on the morning of June 5. Indeed, the enormity of Nasserism's defeat in the Six Day War had, with shocking suddenness, laid bare the movement's fundamental failings. It had promised political populism but had delivered dictatorship and corruption; had championed pan-Arabism but had left the Arab states locked in petty ideological and political feuds; and had pledged independence from the imperial powers but had rendered Egypt dependent on the USSR for military and economic aid. Above all, its adherents had declared loudly and repeatedly that Nasserism would bring Israel to heel and permit the Palestinians to return to their homes; as the smoking remains of the Egyptian air force made emphatically clear, however, Nasserism had failed utterly to do so and had, in the process, left the Palestinians even further from statehood than they had been at the end of the 1948 War. Over the following decade, as a result, proponents of new, more extreme variants of Arab nationalism and, later, Islamism would gain popular support at the expense of Nasserism by claiming that they could achieve the ends that the Egyptian leader had promised but had failed to deliver: Israel's defeat, the return of the Palestinian refugees to their homes, and sustained economic growth.

Those new ideologies did not emerge in isolation, however. Instead, events occurring at the same time outside the core of the Middle East would influence them in critical ways. It is to the post-World War II experience of those geographically peripheral areas—Turkey, the Gulf States, and, especially, Iran that we will next tur

¹²⁸ Oren, *Six Days of War*, 62; Shlaim, *The Iron Wall*, 258–62.

Chapter Fifteen: Arabia, Turkey, and Iran, 1949-1979

Introduction

While pan-Arabism and the Arab-Israeli conflict gave coherence to events in the core of the Middle East during the 1950s and 1960s, they had little impact on the region's peripheral states. Instead, countries like Iran, Turkey, Saudi Arabia, and the Gulf States largely followed their own paths. Turkey struggled to transition to democracy, Iran pursued a top-down modernization effort under the direction of its ruler, Shah Muhammad Reza (r. 1941-1979), and Saudi Arabia and the Gulf emirates fought to get a greater share of the revenue that their vast oil holdings generated. As such, the region's peripheral states exercised only limited influence over the rest of the Middle East during the 1950s and 1960s.

That situation changed dramatically in the 1970s. While Turkey remained largely outside of the region's mainstream during that decade, Iran, Saudi Arabia, and the Gulf emirates vaulted to prominence and began to exert substantial influence over the other states of the Middle East. To a considerable degree, their growing significance reflected important economic changes. Most notably, oil revenue stemming from surging demand in the industrialized world transformed Iran, Saudi Arabia, and the Gulf States into important regional and even global economic players in the 1970s. However, political changes also played a critical part in giving the peripheral states greater influence—particularly Iran. There, the cataclysmic Iranian Revolution of 1979 would put the first modern, Islamist government in power in the Middle East and would, in so doing, encourage further Islamist challenges throughout the region.

Saudi Arabia and the Gulf States

At the end of World War II, the Arabian peninsula was an impoverished backwater that fell within Britain's sphere of influence. By 1979, in contrast, it had become a fabulously rich region that occupied a prominent place within the American economic and strategic order—all thanks to its vast petroleum reserves and to the success of the producing states in gaining control of that resource. Indeed, the oil that they possessed not only raised living standards to previously unimagined heights but gave the Arab producer states substantial regional and even global influence. Petroleum was not an absolute good for the Arab producer states, however. Even as it provided vast economic and political benefits, it would also spark significant domestic and international security issues for the Gulf States and, importantly, would help to midwife the birth of the violent, destabilizing Jihadist ideology that would begin to gain prominence in the 1980s.

The Hydrocarbon Age

Oil loomed large in the years immediately following World War II for the developed states of the West. Petroleum had emerged as the most vital resource of the

age—one that fueled both the ongoing process of industrialization and the steady rise in living standards that millions had come to expect. A key raw material in the manufacture of many goods, it was essential to the production of fertilizer, chemicals, and plastics. More importantly, it had become the essential commodity on which nearly all forms of transportation had come to depend. This was especially true in the wealthy United States. There, inexpensive gasoline not only facilitated the movement of goods and people, but, more broadly, rendered the social and spatial revolution of mass suburbanization possible. Already in motion before World War II, suburbanization accelerated rapidly after the conflict's conclusion as Americans took advantage of inexpensive, government-subsidized loans and the construction of limited-access highways to buy new, detached homes in auto-dependent greenfield developments outside the central city. Automobile ownership rates made the sheer scale of this trend clear. Between 1945 and 1950 alone, the number of registered cars in the US nearly doubled, rising from twenty-six million to fifty million.¹

The affluence of the postwar United States was not a foregone conclusion, however. The suburbanization of the US that created so many jobs and that produced so much wealth depended on the continued flow of cheap, plentiful fuel; inexpensive gas, in other words, was the literal oil that kept the machine of American prosperity going. As we saw in chapter eleven, however, policymakers had come to conclude that domestic sources could not produce enough petroleum to meet future demand and that the US would need to secure new supplies from other parts of the world if it hoped to keep the postwar boom going. With seemingly limitless reserves, Saudi Arabia and the Gulf States appeared to be the obvious place from which to obtain that oil.²

American policymakers were hardly alone in linking cheap petroleum to postwar abundance or in connecting continued economic growth to the Middle East's vast reserves, moreover. Instead, nearly all of the industrialized states made clear that they saw inexpensive Middle Eastern oil as the foundation of continued prosperity. For example, British Foreign Minister Ernest Bevin (1881-1951) declared after World War II that continued access to the region's reserves was necessary if Britain hoped "to achieve the standard of living at which we are now aiming."³

The Middle East's petroleum reserves were also essential to the achievement of vital geostrategic goals. In the late 1940s, American policymakers feared that the tremendous postwar economic dislocation in Western Europe could lead the region's impoverished people to turn in desperation to communism and, in so doing, to bring their states into the Soviet Bloc—thereby irrevocably tipping the global balance of power in Moscow's favor. To forestall such a train of events, Washington moved aggressively to rebuild Western Europe's shattered economy through the passage of the Marshall Plan in 1948. Substantial in scale, it provided the critical injection of funds needed to

¹ Daniel Yergin, *The Prize: The Epic Quest for Oil, Money & Power* (New York: Touchstone Books, 1991), 409, 542, 551–53.

² Yergin, 410, 427.

³ Yergin, 410, 427.

jumpstart the region's reconstruction and to spark a revival in living standards. American grants could not achieve a recovery by themselves, however. Implicit in Washington's economic plans for Europe was the steady availability of inexpensive and plentiful Middle Eastern oil; without it, Marshall Plan aid—however generous—would be unable either to turn Western Europe's moribund economy around or, by extension, to contain communism in the region.⁴

The Fifty-Fifty Deal

Deeply committed to opposing Marxism, King Abdulaziz Ibn Saud (r. 1902-1953) was only too happy to contribute to the American effort to contain it. At the same time, however, he was increasingly dissatisfied with the royalty arrangement that Saudi Arabia had negotiated with Aramco. This change marked a significant shift from the king's earlier outlook. As we saw in chapter eleven, he had initially been more-than content with the terms of the agreement and with the share of oil revenue he was receiving. Not only had the Aramco concession put all of the risk on the petroleum companies, but, since the late 1930s, it had also provided the kingdom with heretofore unimaginable amounts of wealth through the sale of its oil.⁵

By the end of the 1940s, however, Saudi views of the concession had begun to change. The sheer scale of Aramco's profits—they were three times the amount that the Saudi government received in royalties—and the fact that the US Treasury had collected \$43 million in tax revenue from the sale of Saudi oil in 1949 while the kingdom had received only \$39 million in royalty payments led Saud and his advisors to conclude that the terms of the concession needed to be revised. News of Venezuela's success in securing a fifty-fifty arrangement wherein it received half of the revenue from the sale of its oil confirmed such thinking and spurred the king to demand a new agreement based on similar terms. In doing so, however, Riyadh did not want to put Aramco at a competitive disadvantage by raising its costs. Accordingly, Saudi officials focused less on getting a larger share of Aramco's revenue than on capturing some or all of the substantial tax receipts that the American government collected on Aramco's Saudi operations.⁶

With the assistance of high-priced American tax attorneys, Saud's advisors found a way to divert that revenue. The solution lay in exploiting a component of the US tax code, the Foreign Tax Credit, that permitted American companies to deduct from their US returns any taxes—but not royalties—that foreign governments assessed on their operations. In other words, by agreeing to pay taxes to Riyadh that would then be offset

⁴ Melvyn P. Leffler, *For the Soul of Mankind: The United States, the Soviet Union, and the Cold War*, First Edition (New York: Hill and Wang, 2007), 57–70; Madawi al-Rasheed, *A History of Saudi Arabia*, Second Edition (Cambridge: Cambridge University Press, 2010), 118.

⁵ Eugene Rogan, *The Arabs: A History – Revised and Updated Edition* (New York: Basic Books, 2012), 356.

⁶ Yergin, *The Prize*, 445–46.

dollar for dollar on its American return, Aramco could permit Saudi Arabia to enjoy a fifty-fifty split of its profits at no cost to the company's bottom line. Implemented in 1951, the new arrangement worked as advertised. Riyadh's share of oil revenue grew from \$60 million in royalties in 1950 to \$110 million in royalties and tax receipts in 1951, while, over the same period, the US Treasury's saw the revenue it received on the sale of Saudi oil in the United States fall from \$50 million to just \$6 million. Company profits remained unaffected.⁷

Surprisingly, officials in President Harry Truman's (r. 1945-1953) administration were not only tolerant of this arrangement, but they were also outright enthusiastic about it. They recognized that the arrangement provided a way for them to advance an important but stalled regional agenda. During the late 1940s, Truman had sought to provide the kingdom with economic and military assistance designed to aid in the containment of the USSR; that effort had faltered, however, in the face of congressional hostility to the allocation of funds to states hostile to Israel. The Foreign Tax Credit provided a way around that opposition. Not requiring action on the part of congress, it permitted the administration to indirectly provide Riyadh with aid and thus gave a delighted Truman administration a backdoor means of securing its regional objectives.⁸

The new deal that Saud and Aramco officials negotiated had implications that went well beyond the kingdom. Other states in the Middle East quickly followed suit and renegotiated the terms of their concessions along similar lines with the result that the fifty-fifty deal became the new standard. The Arab producer states were delighted with the terms of the new arrangement—after all, it meant that they were now receiving a much larger share of the profits on the extraction and sale of their oil.⁹

At the same time, however, they quickly realized that the new profit-sharing agreement they had won was a more qualified achievement than it appeared at first blush. The issue was one of control. While the petroleum states received a larger share of the profits, the oil companies continued to make all decisions related to the price and amount of crude extracted. In other words, while the producing countries earned more money as a result of the fifty-fifty deal, they had moved no closer to securing what was gradually becoming a critical goal for them: winning control of the petroleum that constituted the primary source of their wealth.¹⁰

Saudi Arabia, 1953-1970

During the two decades following Saudi Arabia's adoption of the fifty-fifty agreement, the oil-producing countries of the Middle East focused on nation building

⁷ Rachel Bronson, *Thicker Than Oil: America's Uneasy Partnership with Saudi Arabia* (Oxford: Oxford University Press, 2006), 56.

⁸ Douglas Little, *American Orientalism: The United States and the Middle East since 1945*, second edition (Chapel Hill: University of North Carolina Press, 2004), 54–55.

⁹ Yergin, *The Prize*, 447–49.

¹⁰ Bronson, *Thicker Than Oil*, 55–56.

while simultaneously dealing with the challenge of a global oil market that was frequently unfriendly to the producer states. Unsurprisingly, given the enormity of its petroleum holdings, Saudi Arabia took the lead in these efforts. Initially, it was not an easy path for the kingdom. Poor leadership, power struggles within the ruling family, and Nasserist-inspired instability combined to frustrate efforts to modernize Saudi Arabia in the 1950s and early 1960s. Under Ibn Saud's son Faisal (r. 1964-1975), however, the Saudi state would find its footing and would establish a stable regime that would provide a model for the other oil-producing states located along the Persian Gulf.

Upon Ibn Saud's death, his son, Saud (r. 1953-1964), became the kingdom's second monarch. The new king faced three daunting challenges upon acceding to the throne. First, as Saudi Arabia lacked a functioning bureaucracy and did not yet even adhere to a budget, he needed to quickly develop the state capacity needed to administer an increasingly complex economy. Second, thanks to Ibn Saud's free spending, he had to address a substantial budget deficit and to find a way to impose fiscal discipline on the Saudi state and the country's very large royal family. Finally, he was compelled to parry the increasingly dangerous ideological challenges that Nasserism and Ba'athism posed to his state and to his family's rule.¹¹

Saud enjoyed at best mixed success in addressing these problems. To his credit, he did manage to establish the foundations of a modern government by expanding the size and reach of the bureaucracy and by creating a number of new departments such as the Ministries of Communications and Education. At the same time, however, his personal vices and diplomatic inexperience left him ill equipped to resolve the other issues that the kingdom confronted. Rather than addressing Saudi Arabia's budget deficit, for example, he made it worse. Wasteful to an almost comedic degree, it took him just five years to turn the \$200 million debt that he inherited in 1953 into a whopping \$480 million deficit.¹²

His efforts to meet the third challenge, fending off pan-Arabism, likewise fell flat. Saud initially sought to ensure the kingdom's security by developing a close relationship with Gamal Abdel Nasser (r. 1954-1970). He felt compelled to abruptly change course in 1956, however, when he learned that the Egyptian leader had gained a worryingly large following in Saudi Arabia and, worse, had begun plotting a coup attempt against him. Determined to push back, the king responded in 1957 by offering the apparently receptive head of the Syrian security service, Abd al-Hamid Sarraj (1925-1913), a £1.9 million bribe conditioned on the intelligence chief preventing the union of Syria and Egypt. Unfortunately for the king, Sarraj had no intention of blocking the creation of the United Arab Republic (UAR) and promised to do so only so that he could lure the inexperienced and overconfident Saudi king into a trap. Sarraj's plan worked to perfection. Releasing unimpeachable evidence of Saud's plot to the press, he

¹¹ James Wynbrandt, *A Brief History of Saudi Arabia*, Second Edition (New York: Checkmark Books, 2004), 208–13.

¹² Wynbrandt, 211, 216, 218.

thoroughly embarrassed the Saudi king—destroying his credibility both internationally and, more critically, within the ruling family.¹³

Accordingly, though Saud retained the formal title of king, his half-brother, Crown Prince Faisal, assumed control of the Saudi government in 1958 at the behest of the other princes. The change in leadership was very much to the kingdom's benefit. Pious, fiscally prudent, shrewd, and diplomatic, Faisal was everything that his brother was not. He immediately launched a successful anticorruption campaign and, with the assistance of the International Monetary Fund (IMF), quickly restored the kingdom's fiscal position. Despite Faisal's success, however, Saud was not yet fully out of the picture. As a result, there followed a six-step forward, five-step back dance in which the two brothers traded power over the next few years. That struggle only ended in 1964 when Saud's threat to start a debilitating civil war led sixty-five 'ulama' and one-hundred princes to depose him and to formally elevate Faisal as the kingdom's third ruler.¹⁴

Now king in his own right, Faisal moved aggressively to legitimate the rule of the Saudi family and to protect the country from subversive ideologies and the threat of Arab nationalism. He approached this task with a three-pronged strategy. First, in what amounted to an updated version of the traditional Arab practice of building loyalty through generosity, the new king devoted an increasing share of the state's oil revenue to the provision of patronage in the form of healthcare and welfare benefits, educational opportunities, and a steadily rising number of well-paying jobs in the expanding army and central bureaucracy. Second, he also devoted substantial sums to development programs aimed at providing Saudi Arabia with modern infrastructure—particularly, as we shall see, after the price of oil exploded in the 1970s. Finally, he secured the support of the Wahhabi religious scholars by giving them the perquisites and salaries of state employees, ceding control of religious education in the universities to them, and extending to them greater authority over religious observation and public morality.¹⁵

Religion was also central to Faisal's efforts to combat Arab nationalism outside the kingdom. Personally pious and politically conservative, he disapproved of the secular and socialist ideas of the Ba'athists and Nasserists and viewed those ideologies as dangerous to the well-being of the Arab people. Determined to blunt their appeal, he promoted pan-Islamism as an alternative form of identity and sought to bolster the kingdom's standing with more-conservative, religious Arabs by emphasizing Saudi Arabia's role as guardian of the Holy Places. More concretely, he gave sanctuary and support to Egypt's Muslim Brotherhood and, as we saw in chapter fourteen, challenged Nasser by providing support and bases to the royalist guerrillas who were waging an insurgency against the Egyptian-backed republicans in Yemen.¹⁶

¹³ Salim Yaqub, *Containing Arab Nationalism: The Eisenhower Doctrine and the Middle East*, First Edition (Chapel Hill: The University of North Carolina Press, 2004), 196–97.

¹⁴ Bronson, *Thicker Than Oil*, 75–77, 82–83, 91–92.

¹⁵ al-Rasheed, *A History of Saudi Arabia*, 120–28.

¹⁶ Bronson, *Thicker Than Oil*, 70; Yaqub, *Containing Arab Nationalism*, 123, 133–34.

The Organization of the Petroleum Exporting Countries (OPEC)

Meanwhile, continued dissatisfaction with the amount of revenue the producing countries were receiving from the sale of their oil spurred those states to demand a say in pricing and production decisions in the early 1960s. Their unhappiness stemmed directly from a growing oil surplus that had emerged over the course of the prior decade. A product of the imbalance between a sharp increase in the supply of petroleum and a far-more modest rise in demand, the glut put strong downward pressure on prices. The entry of the USSR into the global oil marketplace in 1955 dramatically exacerbated this trend. Flooding the world with cheap petroleum, Moscow made an already-bad situation even worse for oil-producing states.¹⁷

At first, the glut affected the oil companies rather than the producer states. That it did so was a function of corporate policy. While the companies continued to pay royalties based on the official posted rate, they increasingly had to discount the price at which they sold their oil on the open market—meaning, in other words, that they were forced to eat the difference between the market rate and the official one. The companies were willing to continue doing so for a time but, by the end of the 1950s, found the practice no longer tenable. British Petroleum (BP), formerly AIOC, was the first to act. To the shock of the producing states, it cut the posted price of oil by a substantial eighteen cents per barrel in 1959. For a time, it stood alone. In August 1960, however, Standard Oil of New Jersey announced that it, too, was reducing the amount paid to the producing states for oil—in its case chopping the per barrel price by fourteen cents, a drop of 7 percent. This move broke the dam. In short order, the rest of the petroleum companies followed suit with cuts of their own.¹⁸

The leaders of the oil producing states were enraged. The August 1960 reduction of the posted price meant a corresponding drop in royalty payments to states that derived nearly all of their revenue from the sale of petroleum. Worse, the oil companies had announced the change unilaterally; indeed, they had not even offered so much as a by-your-leave to the countries that actually owned the oil.¹⁹

The producers responded with collective action. Already unhappy about the failure of revenue to rise sufficiently in the years since the fifty-fifty agreements went into effect, the governments of Iran, Iraq, Kuwait, Saudi Arabia, and Venezuela met in Baghdad in September 1960 to explore the creation of a collective organization that could better represent them in negotiations with the oil companies. Under the direction of the head of the Saudi Directorate of Oil and Mining Affairs, Abdullah Turayqi (1919-1997), the attendees established the Organization of Petroleum Exporting Countries (OPEC) and tasked it with returning oil prices to their pre-cut level by limiting the

¹⁷ Yergin, *The Prize*, 514–15.

¹⁸ Yergin, 521–23.

¹⁹ Rogan, *The Arabs*, 82.

amount produced. In other words, Turayqi and his fellow ministers were trying to establish a cartel that could raise prices by limiting supply.²⁰

The organization would later assume a powerful position in global economic and political affairs. Initially, however, it enjoyed at best mixed success. On the positive side, OPEC rapidly gained many new member states over the course of the 1960s and, by dint of its mere existence, compelled the oil companies to think twice about cutting prices or taking further unilateral action. At the same time, however, the ongoing oil glut—made worse by the entrance into the market of major new producers like Libya—ensured that prices remained low and that the organization could not raise them by limiting production.²¹

The attempt by the Arab oil states to embargo the sale of petroleum to pro-Israeli countries during the Six Day War demonstrated the producers' continued weakness. Seeking to help Egypt, Syria, and Jordan, the oil ministers of Libya, Kuwait, Iraq, Saudi Arabia, and Algeria announced on June 6, 1967, that they were deploying what they called the "oil weapon": an immediate embargo of the sale of petroleum to countries friendly to Israel including the US, Britain, and West Germany. The embargo's impact on Middle Eastern production was immediate and substantial. Coupled with the closure of the Suez Canal, it depressed oil exports from the region to a mere 40 percent of their prewar level.²²

Despite this significant drop in supply, however, the oil weapon ultimately proved unsuccessful in preventing countries friendly to Israel from receiving adequate supplies of petroleum. How was it that a substantial drop in Middle Eastern production did not produce a concomitant shortfall in the consuming states? Put simply, its failure was a function of America's excess domestic capacity. Taking advantage of the country's ability to quickly ramp up production, the administration of President Lyndon Johnson (r. 1963-1969) ordered American oil companies to expand domestic production and directed them to restructure supply chains so that the US could make good West Germany and Britain's petroleum deficits; as a result, the embargo proved to be little more than a costly failure for the Arab oil producers. Indeed, the lesson was clear: so long as America retained the excess capacity needed to serve as the world's swing producer—meaning that it constituted a supplier of last resort that could raise production to offset shortfalls elsewhere—the oil weapon would remain ineffective.²³

The Gulf States: "East of Suez"

Meanwhile, a significant changing of the guard took place in the Persian Gulf in the late 1960s and early 1970s. During the nineteenth and early-twentieth centuries, as we have seen, Britain had assumed a colonial relationship with Kuwait, Bahrain, Qatar,

²⁰ Rogan, 357–458.

²¹ Yergin, *The Prize*, 523–30.

²² Bronson, *Thicker Than Oil*, 99–102.

²³ Bronson, 101–2; Quote from Yergin, *The Prize*, 554–58.

Oman, and the seven smaller emirates known collectively as the Trucial States. In contrast to Britain's diminished role in other parts of the Middle East, that arrangement remained essentially unchanged following the Suez Crisis. By the mid 1960s, however, the breakup of the empire had rendered the arrangement between London and the Gulf principalities anachronistic and had made the stationing of the 6,000 British troops that garrisoned the emirates a needless burden on the exchequer—one that seemed increasingly incompatible with Britain's more-and-more challenging fiscal situation. As such, it was hardly surprising that Prime Minister Harold Wilson (r. 1964-1970, 1974-1976) responded to the onset of a severe balance-of-payments crisis in January 1968 in part by announcing the liquidation of Britain's remaining military obligations "east of Suez."²⁴

Wilson's announcement set off a panic in the Gulf States. The leaders of those oil-rich but vulnerable countries had at times chafed under British dominance, but they also valued the protection that it had afforded them against both broader ideological dangers like pan-Arabism and against more acute military threats like the radical Popular Front for the Liberation of Oman and the Arabian Gulf (PFLOAG) guerrilla movement that had emerged in the 1960s. Indeed, the sheikhs of Dubai and Abu Dhabi were so concerned about the loss of British security that they offered to defray the cost of the garrisons if Britain agreed to retain them. Ultimately, London opted not to take them up on this offer. Instead, it helped the seven Trucial States join together to form the United Arab Emirates (UAE) so that they could better protect each other, and it provided Oman with military assistance in its campaign to defeat PFLOAG. These initiatives were successful. As a result, despite the anxiety that some of the leaders of the Gulf States had expressed at the time of Wilson's announcement, the Persian Gulf region remained stable following the withdrawal of British troops in 1971.²⁵

The Turning Point

Like all of the producing countries, the Gulf States benefited enormously from the rapid increase in oil revenue that occurred during the 1970s. That growth stemmed from a sudden and stark turnaround in the economic relationship between the consuming and producing states. Thanks largely to suburbanization in the US and increased automobile ownership in Western Europe, North America, and Japan in the 1960s, consumption of oil in the non-communist world exploded—rising from nineteen million barrels per day in 1960 to a then-staggering forty-four million barrels per day in 1972. Supply also increased during this period, but, in a turnaround from the situation in the 1950s, it failed to match pace with the now-galloping surge in demand. Irresistibly, as a result, the price of crude oil began to tick upward. The rate of increase was slow at first,

²⁴ Kristian Coates Ulrichsen, *Insecure Gulf: The End of Certainty and the Transition to the Post-Oil Era* (Oxford: Oxford University Press, 2015), 21; Yergin, *The Prize*, 565–66.

²⁵ Ulrichsen, *Insecure Gulf*, 21.

but it began to accelerate dramatically after 1970—doubling between that year and 1973 alone.²⁶

Changes in the American oil market contributed significantly to the rise in prices and made possible an important shift in power from consuming to producing states. While the United States remained the world's largest producer into the 1970s, it reached its peak, pre-fracking production total of 11.2 million barrels per day in 1970. The US subsequently made up the rapidly-growing difference between the amount of oil it produced and the amount it consumed by increasing the quantity it imported, which surged from 2.2 million barrels per day in 1967 to a massive 6 million barrels per day in 1972. The mounting shortfall between US production and consumption alone spurred a rise in world oil prices. More importantly, though few noted it at the time, the American petroleum deficit had a secondary impact of equal consequence. With its spare capacity having disappeared thanks to the plateauing of domestic production and the seemingly insatiable rise in global demand, the United States was thereafter no longer able to offset supply shocks or to prevent price surges by pumping more oil as it had in the past. With enormous consequences for the price of oil, in other words, the US had ceded its role as the global swing producer—a position that Saudi Arabia would henceforth occupy.²⁷

Meanwhile, as a result of the changing balance between producers and consumers, the OPEC states were now finally in a position to begin shifting control over the supply and price of oil from the companies to the producer countries. That process began in 1971 with the Tripoli and Tehran Agreements. Made possible by the growing imbalance between supply and demand, those deals included both substantial price increases and a move from the fifty-fifty arrangement that had obtained for the prior two decades to a new system that granted the oil states 55 percent of the profits from the sale of their oil.²⁸

This success merely whetted the OPEC states' appetite for further concessions. The following year, they pushed for what Saudi Minister of Petroleum and Mineral Resources Ahmed Zaki Yamani (1930-2021) called "participation": partial ownership of the companies that held the oil concessions in their states. This process followed two paths. Intent on avoiding substantial disruptions to the market, some states like Saudi Arabia agreed on a gradual process of nationalization. In its case, it began with a 25 percent ownership position that gradually increase to 51 percent in 1983. Others such as Algeria and Libya pursued a far-more aggressive path and assumed majority stakes immediately. Regardless of the approach, participation succeeded in substantially increasing the producer states' share of what were, thanks to rising demand, rapidly growing profits. It is important to keep in mind, however, that revenue was not the main reason that the OPEC states pursued ownership. On the contrary, while they certainly enjoyed the added income they received, they sought partial ownership primarily as a

²⁶ Bronson, *Thicker Than Oil*, 111; Yergin, *The Prize*, 567.

²⁷ Yergin, *The Prize*, 567–68.

²⁸ Yergin, 582–83.

way for them to challenge the concession system and, critically, to secure control—and, thus, sovereignty—over their countries' most valuable resource.²⁹

The transformation of the relationship between the producing and consuming states reached its culmination during the Yom Kippur War between Israel and the Arab republics of Egypt and Syria in October 1973 (we will explore the conflict itself in chapter sixteen). Rapidly increasing prices were the first sign of the changed environment. OPEC had already succeeded in raising the price of a barrel of oil by 12 percent to \$2.90 in June and was preparing at the time the war started to demand another big increase at a meeting with oil company representatives scheduled to begin in Vienna on October 8. Already fraught owing to the crisis atmosphere created by the conflict, the talks collapsed when oil company representatives balked at the producers' shocking demand for a doubling of prices. The failure of the negotiations did not stop the Arab oil ministers from raising prices, however. On the contrary, emboldened by the fighting, they adjourned to Kuwait City where they unilaterally raised the cost of a barrel of oil to \$5.12.³⁰

More importantly, they once again deployed the oil weapon. The day after they raised prices, the Arab ministers announced significant production cuts scheduled to take effect immediately. A few days later, they followed by imposing an embargo on the sale of petroleum to countries like the US that supported Israel.³¹

This time, the oil weapon worked. With Saudi Arabia having replaced the US as the world's swing producer, the embargo and, especially, the production cuts created oil shortages and severe economic pain in the industrialized world. The result was steep inflation, biting recessionary conditions, and wrenching social discontent among Western consumers long conditioned to expect cheap energy.³²

More broadly, the embargo and the production cuts signaled that the producing states, rather than the consuming ones, now called the shots when it came to the supply and price of petroleum. The OPEC states were quick to take advantage. While motorists in the US endured gas lines and fumed about rationing, the producers cheerfully exploited their newfound power to complete the process of nationalizing their oil industries that had begun with the "participation" agreements in 1972. Henceforth, as a result, those states would hold title to their most valuable natural resource. Western consumers were largely unaware of that shift, but they were intimately familiar with another change that OPEC put in place at the shah's urging. Meeting in Tehran in December 1973, the organization increased the price of oil to \$11.65 a barrel—a four-fold increase in less than six months.³³

²⁹ Quote from Bronson, *Thicker Than Oil*, 112; Rogan, *The Arabs*, 360–61.

³⁰ Rogan, *The Arabs*, 369–70.

³¹ al-Rasheed, *A History of Saudi Arabia*, 136–37.

³² Yergin, *The Prize*, 615–17, 626–30.

³³ Bronson, *Thicker Than Oil*, 112, 122.

The Good Times

The next four years were, in the words of one of the organization's executives, "OPEC's Golden Age." They were indeed good times for producer countries like Saudi Arabia and the Gulf States. The massive influx of petrodollars—American money earned through the sale of oil—provided the means for them to thoroughly transform and modernize their economies. The amounts were mind boggling. Between 1965 and 1975, the surge in oil revenue increased Saudi Arabia's GDP rose from 10.4 billion Saudi Riyals to over 164 billion Riyals thanks almost entirely to the massive surge in oil prices that occurred in the early 1970s. No one benefited more from this flood of wealth than the thousands of princes who made up the royal family. Drawing little distinction between state revenue and dynastic wealth, they waged an arms race of conspicuous consumption in which they spent extravagant sums on fleets of cars, new palaces, huge yachts, and European luxury properties. All told, the royal family alone skimmed off an estimated 30 to 40 percent of the kingdom's oil revenue.³⁴

While the princes certainly enjoyed their newfound income, the scale and openness of their spending also threatened to undermine the legitimacy of the royal family. After all, their gambling, drinking, and naked materialism were hardly consistent with the strictures of Wahhabi Islam. Aware of the potential danger that the gap between the kingdom's professed values and the behavior of its rulers posed to the dynasty's legitimacy, Faisal took action to prevent the emergence of discontent. To limit any objections that the 'ulama' might have raised to the princes' lifestyle, he increased the subsidies that he provided to mosques and religious colleges and lavished funds on the religious scholars' effort to bring Wahhabi ideas to Muslims outside the kingdom. Meanwhile, though he refused to consider any genuine move toward democratic government, he took two steps designed to retain the support of the emerging middle class. First, he steadily increased the number of people directly employed by the state. Second, he took advantage of the surge in oil revenue to dramatically accelerate a series of ambitious, five-year modernization programs that had begun at the start of the 1970s. Thanks to the quadrupling of oil prices in 1973, for example, he was able to raise the amount allocated to the first five-year plan (1970-1975) from \$9.2 billion to \$21 billion. Later programs were even larger. Most notably, as a result of the windfall that followed the Yom Kippur War, the Saudi government was able to assign an order of magnitude more money—a staggering \$200 billion—to the second five-year plan (1976-1980).³⁵

As a result of these programs and a third completed in the early 1980s, Saudi Arabia experienced a stunning transformation in a remarkably brief period of time.

³⁴ Kim Ghattas, *Black Wave: Saudi Arabia, Iran, and the Forty-Year Rivalry That Unraveled Culture, Religion, and Collective Memory in the Middle East* (New York: Henry Holt and Co., 2020), 52–53; quote from Yergin, *The Prize*, 634; Lawrence Wright, *The Looming Tower: Al Qaeda and the Road to 9/11* (New York: Vintage Books, 2006), 100.

³⁵ al-Rasheed, *A History of Saudi Arabia*, 124–27, 138.

Between 1970 and 1985, it completed a huge array of modern infrastructure projects including highways, ports, electric power networks, universities, powerplants, housing developments, and hospitals, and it massively increased the social welfare benefits provided to the Saudi people. Faisal did not survive to see the success of the economic-development plans that he had launched, however. Shot by one his nephews as revenge for the death of a sibling during demonstrations against the introduction of television in the 1960s, he died in 1975. Thus, it was his brothers and successors, Khalid (r. 1975-1982), and Fahd (r. 1982-2005), who, having continued his policies, reaped the fruit of his modernization program.³⁶

Differences over Israel notwithstanding, Saudi Arabia and the Gulf principalities retained strong relations with the US in the 1970s. To a substantial degree, the perpetuation of close ties with Washington was a function of the oil producers' vulnerability. Ironically, the flood of wealth that those states had begun receiving at that time had not rendered them secure but had instead transformed them into tempting targets for foreign enemies. Acutely aware of this situation, Faisal was determined to retain good relations with Washington in order to ensure his kingdom's security. Even during the Yom Kippur War when ties between the two countries were at a low ebb, for example, he acted to minimize the breach between Washington and Riyadh and even went so far as to secretly provide oil to the US Navy. Likewise, it was the Saudi king who persuaded the other Arab producer states to bring the oil embargo to an end in March 1974.³⁷

More broadly, aware that recessionary conditions in the developed world would ultimately hurt the producing states, the Saudi king also agreed to help the US resolve the balance-of-payments problem that the sudden rise in petroleum prices had caused. Accordingly, over the next few years, his government undertook a series of actions designed to repatriate a huge share of the petrodollars that it had acquired from the sale of oil in the US. It purchased American goods for its modernization effort, acquired \$34 billion worth of arms from Washington, invested in businesses and real estate in the United States, and bought US government securities. These latter purchases were substantial. Indeed, by the end of the 1970s, the kingdom held more US debt than any other entity.³⁸

While the close relationship with America may have helped secure Saudi Arabia from foreign dangers, it contributed to significant tensions within the kingdom. The problem was the newly empowered Wahhabi 'ulama'. As previously noted, Faisal had sought to legitimate the regime and to mute opposition to modernization in part by embracing religion. Central to this campaign was his effort to secure the support of the religious scholars. To do so, he had put them on the state payroll, ceded domestic power to them, given them control of higher education, and lavished revenue on their efforts to spread Wahhabism abroad. For a time, this strategy seemed to work.

³⁶ Wynbrandt, *A Brief History of Saudi Arabia*, 233–35.

³⁷ Yergin, *The Prize*, 606–32.

³⁸ al-Rasheed, *A History of Saudi Arabia*, 135–36; Bronson, *Thicker Than Oil*, 118–22.

Supported by ample state funds, the 'ulama' largely refrained from criticizing the royal family and instead focused on spreading their severe construction of Islam both to people within the kingdom—particularly the hundreds of thousands of Arab guest workers who lived in Saudi Arabia—and, more broadly, to Muslims across the globe.³⁹

By the mid 1970s, however, Faisal's effort to secure the support of the religious scholars had begun to boomerang. Though they appreciated the state's largesse, the Wahhabi 'ulama' increasingly recoiled at the kingdom's close relationship with the US and at the Saudi princes' over-the-top lifestyles and began to doubt the regime's commitment to the Islamic values it claimed to espouse. As a result, even as the Wahhabi 'ulama' were using the government's substantial financial assistance to spread extreme Islamist and jihadist ideas throughout the Arab world—particularly, as we shall see, in Egypt—they were also quietly fostering domestic opposition to a state that many religious scholars had come to conclude was not adhering to the tenets of Wahhabism. Indeed, unbeknownst to the Saudis, Wahhabi ideas had given birth to a secretive, insurgent Islamist movement within the kingdom. Led by a man who claimed to be the Mahdi, the group had emerged out of the Religious University of Medina—where, not coincidentally, members of Egypt's uncompromising Muslim Brotherhood had found sanctuary—and had, in the late 1970s, quietly begun to plot the overthrow of the Saudi government.⁴⁰

Turkey: Democracy and Its Discontents

For the Republic of Turkey, the period between 1945 and 1980 was one of both tumult and continuity. During that time, the country experienced dramatic change including the transition from a single-party state to a multiparty democracy, the rise of political extremism on both the right and left, substantial, if uneven, economic growth, and tremendous social and demographic change including rapid urbanization. At the same time, however, the military's assumption of an extraconstitutional role as the guarantor of Kemalism, political stability, and secularism ensured that postwar Turkey enjoyed substantial continuity with the values and ideas of the early republic.

The Transition to Multiparty Democracy, 1945-1980

At the end of the Second World War, Turkish President Ismet Inonu (r. 1938-1950) and the Republican People's Party (RPP) that had ruled the country since the early 1920s found themselves in a difficult position. They had successfully navigated the challenging diplomatic waters of the war years but had done so only by pursuing policies that alienated three key components of the Kemalist political coalition: the urban bourgeoisie, government administrators, and the large landowners of rural Turkey. The bureaucrats had seen inflation destroy the purchasing power of their salaries thanks to the expansion of the money supply required to pay for a bigger military, the large

³⁹ Fawaz A. Gerges, *Making the Arab World: Nasser, Qutb, and the Clash That Shaped the Middle East* (Princeton: Princeton University Press, 2018), 330–32.

⁴⁰ al-Rasheed, *A History of Saudi Arabia*, 135–36, 142–44.

landowners had endured costly price caps on agricultural products, and the bourgeoisie had suffered under confiscatory wartime taxes. By 1945, as a result, these influential parts of the ruling party's base had quietly begun to demand fundamental political changes.⁴¹

This discontent set the stage for the liberalization of Turkey's political system immediately following World War II. The process started when four dissident members of the assembly including future prime minister Adnan Menderes (1899-1961) and future president Celal Bayar (r. 1950-1960) openly opposed a land-reform bill, arguing that it would weaken faith in the government's commitment to private property and thus discourage investment. Though the law passed over their objections, Menderes and his colleagues remained undaunted. They followed their defeat by calling for political liberalization and full democracy. These demands were too much for the RPP. It responded by kicking Menderes and two fellow dissidents out of the party—a move that appeared to doom their push to liberalize Turkey. Salvation came from an unexpected source. In a dramatic speech at the opening of the assembly in November 1945, President Inonu rescued the liberalization drive by calling for a gradual transition to full democracy. Having been given a second political life, Menderes and Bayar followed by officially establishing the republic's first meaningful opposition party, the Democratic Party (DP), in January 1946.⁴²

The creation of the DP shook up the political system. Appealing directly to groups like the bourgeoisie, the landlords, and the peasants, the new party soon amassed an enthusiastic following among those disaffected with the RPP and thus made a good showing in elections held in May 1946. The DP's success deeply alarmed the leaders of the RPP. Worried that they might continue to lose support, they responded by coopting the Democratic Party's main policy proposals. They jettisoned the latest five-year plan in favor of a free-market approach designed to appease the bourgeoisie, devalued the Turkish currency—the lira—and moved to weaken support for the DP among the religiously minded peasantry by allowing schools to teach religion once again. These changes seemed to work. Coupled with the defection of a number of DP assembly members to the newly formed Nation Party in 1948, the new policies appeared to put the RPP in a commanding position going into the 1950 election.⁴³

The DP's upset victory thus came as a startling surprise. In what amounted to an electoral revolution, the party won 54 percent of the vote and an overwhelming majority in the assembly. It was a shocking upset—one that led the RPP and its allies to consider a variety of responses to its rival's success. Indeed, the military leadership even quietly proposed staging a coup. To his credit, however, Inonu decided to respect the wishes of the voters and declined the offer. As a result, for the first time since its

⁴¹ Erik J. Zürcher, *Turkey: A Modern History*, Third Edition (London: I.B. Tauris, 2004), 206–8.

⁴² Zürcher, 209–12.

⁴³ Douglas A. Howard, *The History of Turkey* (Westport, Conn: Greenwood Press, 2001), 118–19.

formation, the party that had secured the country's independence and territorial integrity after the Great War had lost its monopoly on power. Turkey had become a genuine multi-party democracy.⁴⁴

The Democratic Party in Power, 1950-1960

Under Menderes, who served as prime minister, the Democratic Party enjoyed widespread popularity during its first few years in power, thanks to its promotion of policies that appealed to its business and especially rural constituencies. Its key reforms fell in two areas. First, while the DP retained a broad commitment to secularism, it also eased some of Kemalism's more extreme restrictions on religion. It ended the ban on making the call to prayer in Arabic, endorsed teaching Islam in public schools, and permitted religious broadcasts on the radio. Second, it instituted a series of sweeping economic changes designed to liberalize the economy and to encourage the agricultural sector. It established the Turkish Industrial Development Bank to loan state funds to capitalists, gave farmers easy credit from the Agricultural Bank, maintained high wheat prices, and, with the help of the Marshall Plan, facilitated the "tractorization" of agriculture through the importation of American farm equipment. The results were impressive. Benefitting from good weather and a 50 percent increase in the amount of land under cultivation, the economy grew at a 12 percent clip during the early 1950s.⁴⁵

Economic growth translated into continued success at the polls for Menderes and the Democratic Party during the 1950s. Boosted by the booming agricultural sector, the DP cruised to an easy victory in the 1954 elections. It won 58 percent of the vote to the RPP's 35 percent—giving the Democrats 503 seats in the assembly as compared to a mere 31 for the Republicans. The DP did not fare as well in elections held in 1957, however. That year, it won just 47 percent of the vote while the RPP's share rose to 41 percent. Still, since the Turkish political system rewarded the leading party with a disproportionate share of assembly seats, the DP retained a tight grip on power.⁴⁶

The precipitous decline in support for Menderes and the DP in 1957 stemmed largely from Turkey's worsening economic circumstances. Resting as it did on the continued importation of tractors and other agricultural equipment, the growth that had boosted the party in the early 1950s had produced a steadily rising balance-of-payments deficit. By 1958, the bill finally came due. Desperate for an infusion of additional foreign cash, Prime Minister Menderes's government was forced to agree to implement a series of harsh austerity reforms in exchange for a \$359 million IMF stabilization loan. The resulting economic pain fueled growing opposition to the DP—particularly among urbanites, who saw the purchasing power of their salaries eroded by

⁴⁴ Zürcher, *Turkey*, 217–18.

⁴⁵ Carter V. Findley, *Turkey, Islam, Nationalism, and Modernity* (New Haven: Yale University Press, 2010), 307, 325; Howard, *The History of Turkey*, 120.

⁴⁶ Zürcher, *Turkey*, 223–24, 231–32.

inflation, and among farmers, who no longer enjoyed favorable treatment from the government.⁴⁷

Menderes's increasingly heavy-handed political repression also contributed to the DP's declining popular appeal. Taking a page from the RPP, his party had freely used the power of the state against its political rivals from the moment it had assumed power. Indeed, shortly after taking control of the government, the prime minister had ordered the treasury to confiscate property that the RPP had obtained when it had controlled the state in the 1930s. After its big victory in 1954, the DP went even further. It implemented changes to the electoral system that imposed new burdens on the opposition and passed a law that permitted the government to force long-tenured RPP bureaucrats, judges, and academics to take early retirement. Later, as its political position weakened, the DP embraced outright illiberalism—organizing a parliamentary commission to investigate the RPP and banning the party's representatives from sitting in parliament.⁴⁸

The Coup of 1960

This last action was too much for the military. Alarmed by the DP's increasingly autocratic actions and, more importantly, by its deviation from Kemalism, a group of military officers launched a coup on May 27, 1960. Citing the need “to prevent fratricide” and to “extricate the parties from the irreconcilable situation into which they had fallen,” the coup's leader, Colonel Alparslan Turkes (1917-1997), announced over the radio that the army had assumed control of the state. The junta promptly closed parliament, suspended the constitution, named a popular general as head of state, and established a new body, the National Unity Committee (NUC), to run the country until a new constitution could be drawn up and new elections held. They also outlawed the DP and arrested all of the party's assembly members including Menderes and Bayar.⁴⁹

Differing visions of Turkey's future soon split the NUC into radical and conservative factions. The key issues dividing the body were the duration of military rule and the extent of the reforms that the coup leaders intended to put in place. Led by Turkes, the radical faction wanted the military to rule for a long period so that it could impose a top-down political and cultural revolution designed to turn Turkey into an authoritarian state along the lines of Nasser's Egypt. The rival conservatives instead called for a quick end to military rule and for a limited series of reforms aimed at ensuring the perpetuation of Kemalism and at preventing future political parties from abusing power as the DP had. For a time, the two factions coexisted uneasily in the NUC. The balance tilted irrevocably toward the conservatives, however, after the radicals undertook a purge in the army in August. Alienating many moderate officers heretofore unaffiliated with either faction, this move left Turkes and his supporters so

⁴⁷ Zürcher, 228–30.

⁴⁸ Findley, *Turkey, Islam, Nationalism, and Modernity*, 306–9.

⁴⁹ Howard, *The History of Turkey*, 125–26, 134.

isolated that the conservatives were able to force them from the NUC in November 1960.⁵⁰

Having dealt with the radicals, the conservative faction then moved to revamp Turkey's political system and to implement a new constitution. They did so in hopes of achieving two overarching goals. First, they wanted to prevent future political parties from trampling basic rights as the DP had. Accordingly, they devoted fully one third of the new constitution to provisions assuring civil liberties such as the independence of the media and the right to collective bargaining. Second, they sought to ensure the perpetuation of the Kemalist revolution. To do so, they strengthened the judiciary, created a bicameral legislature, and established a new institution, the State Planning Organization (SPO), to coordinate economic development through a series of five-year plans modeled on the etatist practices of the 1930s. Most importantly, they created a new executive body called the National Security Council (NSC). Composed of the president, the chief of the general staff, and the heads of the service branches, it advised the government on national security issues. Its remit did not remain confined to matters of defense for long, however. Instead, through a process of bureaucratic mission creep wherein its members perceived a growing range of topics to be relevant to national security, the NSC soon became an enormously influential body—one that ensured a powerful voice for the Kemalist officers in the new government.⁵¹

Establishing the Second Republic, the coup of 1960 marked a watershed moment in Turkey's political development. Not only did it break the DP's grip on power, but it also set two lasting precedents. First, it established that the military had an extraconstitutional right to take control of the state if the politicians abused power or strayed too far from the Kemalist legacy. Second, the coup set the standard that future military involvement in politics would be temporary and that the officers would return power to the civilian government as soon as feasible.⁵²

The Second Republic, 1960-1980

The performance of the economy during the early years of the Second Republic seemed to validate the coup. Under the direction of the SPO, Turkey shifted away from the DP's emphasis on private enterprise and back to the etatist, Import Substitution Industrialization (ISI) approach that had shaped the Turkish economy during the heyday of Kemalism in the 1930s. There was one substantial difference, however. This time, the drivers of economic growth were private firms rather than state-controlled operations. Sheltered by protective tariffs and ensured access to imported raw materials through the maintenance of an artificially strong lira, the powerful, oligopolistic corporations that emerged in the 1960s quickly came to thoroughly dominate the Turkish economy. For a time, they led an impressive expansion. Indeed, to the satisfaction of the officers who had overseen the creation of the Second Republic,

⁵⁰ Howard, 134–35.

⁵¹ Zürcher, *Turkey*, 245; Findley, *Turkey, Islam, Nationalism, and Modernity*, 311.

⁵² Findley, *Turkey, Islam, Nationalism, and Modernity*, 310.

economic growth averaged an impressive 6.9 percent per year between 1963 and 1967.⁵³

The coup leaders were far-less pleased with the political system that emerged during the early years of the Second Republic, however. Their frustration began with the first post-coup election in the fall of 1961. While Inonu's RPP won the election as they had hoped, it secured the backing of just 37 percent of the electorate, while parties that drew support from the same groups that had backed the now-outlawed Democratic Party took 49 percent of the votes cast. This suggested both the continued appeal of the DP's platform, and the depth and breadth of popular opposition to the coup. Inonu did manage to oversee several RPP-dominated coalition governments in the early 1960s, but those governments were shaky and the RPP was never able to find its footing. Instead, to the chagrin of the military it was seven-time Prime Minister and future President Suleyman Demirel's (1924-2015), center-right Justice Party (JP) that would emerge as the dominant political force during the Second Republic.⁵⁴

More troublingly from the perspective of the military, the mainstream parties no longer enjoyed a monopoly on political power in Turkey. Instead, thanks to rapid demographic growth and, ironically, the liberal constitution that the officers had put in place, new, extremist political movements of both right and left proliferated in the decade following the coup. Inspired by Marxist writings and by the transnational, antiestablishment zeitgeist of the 1960s, leftist movements won a following among university students and Turkey's small-but-growing working class. These organizations grew rapidly during the late 1960s and early 1970s but failed to make substantial political inroads in part because they splintered along ideological fault lines. For example, the Marxist Federation of the Revolutionary Youth acrimoniously split at the beginning of the 1970s into competing Guevarist and Maoist factions. In contrast, the far-right ultranationalist groups that developed in the 1960s—including, most notably, Alparslan Turkes's National Action Party (NAP)—enjoyed substantial coherence and unity.⁵⁵

Predictably, the rise of extremist political organizations produced violence and instability. Inspired by the dramatic global revolutionary events of 1968 such as the Tet Offensive in Vietnam and the French student revolt in Paris, far-left organizations began to engage in aggressive political acts such as seizing control of Istanbul University and occupying a tire factory. They grew substantially bolder and more militant in the early 1970s. Over the course of the decade, far-left groups organized guerilla movements, robbed banks, and kidnapped American military personnel as part of their effort to overthrow the state. Furious, the ultranationalists did not permit these actions to go unchallenged. Instead, with the tacit support of Demirel's government, rightist militias

⁵³ Zürcher, *Turkey*, 264–66.

⁵⁴ Howard, *The History of Turkey*, 137–39.

⁵⁵ Findley, *Turkey, Islam, Nationalism, and Modernity*, 314–15.

such as the NAP-affiliated Gray Wolves joined the state's security apparatus in aggressively combatting leftist violence.⁵⁶

Fed up with the volatility and the Marxist threat to Turkey's political system, the military moved to restore order in 1971. In what became known as the Coup by Memorandum, the chief of staff gave Prime Minister Demirel a letter in March 1971 indicating that the army would once more assume power if the government did not implement sweeping reforms and "neutralize the current anarchical situation"—a threat that impelled Demirel to promptly resign. The officers followed by moving to curtail the leftist challenge. Declaring martial law, the NSC directed a brutal and far-reaching crackdown on leftist organizations and pressured the assembly to approve a series of constitutional amendments designed to prevent further disorder. They included curbs on the autonomy of the universities, restrictions on press freedoms, and limits on the independence and power of the courts. Ironically, in other words, the same officers who had insisted on enshrining civil liberties in the constitution of 1961 were now demanding that the government severely curtail those freedoms.⁵⁷

Turkish Foreign Policy

The story of Turkey's diplomatic relations between 1945 and 1980 was far-more straightforward than the course of its anarchic internal political system. Its foreign policy during that time centered on achieving two goals. First, in conjunction with the United States, it worked to prevent the USSR from expanding its influence into the Middle East. Second, it sought to ensure that Britain's majority-Greek colony of Cyprus would become a separate state rather than a province of Greece when it won independence and that its government would respect the rights of the island's Turkish minority. For the most part, these goals were in harmony. Coming into conflict in the 1970s, however, they briefly drove a wedge between Washington and Ankara.⁵⁸

Turkey was an early partner in America's effort to contain the Soviet Union. As we have seen, Washington had worked closely with Ankara to foil Joseph Stalin's (r. 1928-1953) effort to gain control of the Dardanelles and Bosphorus Straits in 1946 and had made Turkey a recipient of American military and economic aid through the Truman Doctrine the following year. The relationship became even closer when Turkey joined the North Atlantic Treaty Organization (NATO) in 1952. Assuming the role of an "unsinkable aircraft carrier" for the United States, it thereafter hosted combat jets, U-2 spy planes, and—until President John F. Kennedy (r. 1961-1963) removed them following the Cuban Missile Crisis—Jupiter nuclear missiles. Despite acrimony over the

⁵⁶ Findley, 315–16.

⁵⁷ Howard, *The History of Turkey*, 147–48.

⁵⁸ Zürcher, *Turkey*, 208–9, 234–37.

withdrawal of the rockets, Turkey remained a committed and faithful partner in Washington's containment policy.⁵⁹

A crisis stemming from the decolonization of the island of Cyprus would threaten that relationship, however. In the early 1950s, the Cypriots had entered into negotiations with London aimed at ending British rule of the island. Two factors complicated those talks. First, the population of Cyprus was divided uneasily between a Greek majority that lived largely in the south and a Turkish minority—about 20 percent of the population—that predominated on the island's north side. Second, under their nationalist leader, the Archbishop Makarios (r. 1960-1977), the Greek Cypriots were pushing for *enosis*, or union with Greece, rather than independence.⁶⁰

The possibility of Cyprus joining Greece was utterly unacceptable to the Turkish government. It feared that *enosis* would both imperil the rights of the Turkish Cypriot minority and present Ankara with a serious strategic dilemma should it find itself in a conflict with Athens. Accordingly, the Turkish government first successfully pressured Britain into making Cyprus an independent state rather than a part of Greece and then ensured that the new state's constitution guaranteed the civil and political rights of the island's Turkish minority. Despite a push among Greek Cypriots to curtail those assurances following the country's independence in 1960 and an outburst of intercommunal violence in 1964, relative calm prevailed during Cyprus's first decade. Thus, despite the dispute over *enosis*, the island seemed to have put ethnic conflict behind it and to be well on its way to becoming a stable, multi-ethnic nation-state.⁶¹

Events soon made clear that such appearances were far off the mark. Rather than coming together as a unified polity, Cyprus instead shattered along ethnic lines in 1974. The cause was the continued desire among Greek-Cypriot ultranationalists to achieve *enosis*. Intent on securing that goal, extremists in the Cypriot National Guard deposed Makarios in 1974 and—with the encouragement of the military junta then ruling Greece—declared their commitment to immediate union. The Turkish government reacted with fury to this precipitate action. Viewing the move as both a serious threat to its national security interests and as an unacceptable violation of the civil and political rights of the Turkish Cypriots, it responded by deploying troops to the predominantly Turkish northern half of the island in July. The landing dramatically enflamed the situation and sparked the start of a sustained campaign of ethnic cleansing across the island that resulted in the flight of Greek Cypriots to the south and Turkish Cypriots to the north. Ankara followed by formally partitioning the island and by establishing a Turkish Cypriot government in the territory its troops occupied. That state gained

⁵⁹ George C. Herring, *From Colony to Superpower: U.S. Foreign Relations since 1776* (New York: Oxford University Press, 2008), 609, 614–15; Howard, *The History of Turkey*, 128–29, 145; quote from Findley, *Turkey, Islam, Nationalism, and Modernity*, 309.

⁶⁰ Zürcher, *Turkey*, 237–38.

⁶¹ Zürcher, 237–38.

nominal independence as the Turkish Republic of Northern Cyprus in 1983; however, Ankara was the only government to recognize it.⁶²

The intervention in Cyprus may have been popular among Turkish nationalists, but it also sparked a growing rift between Washington and Ankara. Angry at Turkey's precipitate move, congress responded by promptly imposing a sweeping embargo on military aid. Furious at what it saw as unwanted American interference in a Turkish domestic issue, Ankara in turn retaliated by immediately shutting all American military facilities in Turkey save for the NATO airbase at Incirlik. The rift was not merely deep but also long lasting. Pressured by Greek-American and Armenian-American lobbying groups, congress did not lift its ban on the provision of military aid to Turkey until July 1978. Only then did bilateral relations begin a gradual thaw.⁶³

The 1980 Turkish Coup D'etat

Domestically, meanwhile, Turkey experienced mounting economic and political challenges over the course of the 1970s. The country's economic woes had two sources. First, while it had enjoyed high growth during the "easy" period of ISI development when domestic demand for consumer goods remained unmet, it had, by the early 1970s, entered the "hard" phase during which it was unable to offset plateauing internal demand by shifting to an export-oriented growth strategy because its protected industries were too inefficient to compete in global markets. Second, and worse, the revolution in oil prices that occurred in 1973 proved far-more challenging for petroleum-importing states in the Developing World than for economically advanced countries like the US. Indeed, thanks to the quadrupling of crude prices in 1973, Turkey was soon spending fully two-thirds of its foreign-currency earnings on imported petroleum. Unwilling to abandon its expansionary policies, the government accepted a dramatic increase in its external debt and dealt with its rapidly worsening balance-of-payment problems by increasing the money supply. This approach worked for a time, but, coupled with the rise in fuel prices that followed the Iranian Revolution, also pushed inflation to an unsustainable 90 percent in 1979.⁶⁴

Meanwhile, despite a declaration of martial law and an intensive crackdown on leftist organizations, the Coup by Memorandum had failed to restore stability to the political system. Two issues proved particularly troubling. First, thanks to the mainstream parties' unwillingness to work with each other and to their inability to secure outright majorities, small parties—particularly far-right ones—found themselves in a position to wield influence well beyond what their numbers justified. For example, Turkey's NAP was able to translate the paltry three seats it held in the assembly in 1975

⁶² Zürcher, 275–76.

⁶³ Howard, *The History of Turkey*, 151.

⁶⁴ Zürcher, *Turkey*, 267; Findley, *Turkey, Islam, Nationalism, and Modernity*, 329.

into two cabinet posts thanks to its ability to give Demirel the slim majority he needed to form a government.⁶⁵

Second, the late 1970s witnessed a return to the political instability and violence that had marked the early part of the decade. The resurgence of bloodshed began in May 1977 when clashes between police and demonstrators during a May Day parade resulted in thirty-nine deaths and sparked a retaliatory bombing campaign by leftists. The following year, the murder of the retired rector of Istanbul Technical University created fears of worsening political disorder, while five-days of confessional conflict between Sunnis and Alevis—a mystical variant of Twelver Shi'ism—raised the specter of sustained religious clashes. The situation deteriorated rapidly from that point. All told, political violence claimed the lives of 5,000 people between January 1978 and September 1980—a wave of violence that made clear that the civilian government had lost its ability to maintain order. In response, the military once again stepped in—orchestrating a bloodless coup in September 1980 that finally brought the tumultuous Second Republic to an end.⁶⁶

Iran: From Coup to Revolution

Iran had a far-more difficult time after World War II than did Turkey. That it did so is perhaps surprising in light of the country's enormous oil reserves and the significant growth it experienced under the rule of the shah. After all, a rising standard of living is typically associated with political stability. In Iran's case, however, the economic expansion not only masked but, ironically, contributed to a seething discontent with the regime that permeated nearly every group and class in the country. Tamped down for many years, that pent-up resentment would burst forth in a sudden political deluge in the late 1970s that would, in an astonishingly brief period of time, first sweep aside the shah's regime and then establish a purportedly traditional, but ultimately revolutionary new state under the firm dominance of the Shi'i clerical establishment.

The Nationalist Challenge, 1950-1953

The early years of Muhammad Reza's reign would prove pivotal in setting the stage for the revolution. Two issues dominated Iran at that time. The first was the question of who would rule. Would it be the young shah, whom the Allies had placed on the throne in 1941 and who enjoyed the backing of the armed forces? Or would it instead be the majlis, the Iranian parliament, which the rural notables dominated? The second issue revolved around the question of who would control the country's oil resources and earn the lion's share of the profits from its sale. Would it be the British-owned AIOC or would it be the Iranian government? By the late 1940s, this latter issue had become paramount in Iran. While the Iranian people could not see the exact figures because the AIOC refused to permit them to review the company's books, they were well aware that both the firm's profits and the British government's tax revenues from

⁶⁵ Findley, *Turkey, Islam, Nationalism, and Modernity*, 317–18.

⁶⁶ Howard, *The History of Turkey*, 153–54.

the sale of Iranian oil were vastly larger than the royalties paid to the government of Iran. It was a galling situation for Iranian nationalists—one that they were determined to change.⁶⁷

These two issues would come together at the end of the 1940s to profoundly shape the trajectory of Iran's political development. Initially, the shah appeared to hold the upper hand in the struggle for political dominance. Emboldened by a failed assassination attempt in February 1949, the heretofore weak monarch acted to centralize power in his hands by declaring martial law and by taking an active role in the direction of the country for the first time. He followed by pursuing negotiations aimed at replacing the 1932 agreement between Iran and the AIOC with a new one patterned after the far-more-generous fifty-fifty deal that Venezuela had obtained in 1943. It was a bold initiative—one that if successful would not only increase the amount of revenue that his government received each year but would also boost his nationalist credentials to such a degree that he would be able to fully restore the power that the monarchy had lost when the Allies had deposed his father.⁶⁸

The AIOC was unreceptive to the shah's proposal, however. On the contrary, the company indicated that it was willing to agree only to comparatively modest revisions to the existing system of royalty payments. Given the disparity in power between his state and the powerful, British-government-controlled AIOC, the shah felt that he had little choice but to accept the company's offer; as a result, he soon signed what came to be known as the Supplemental Agreement. The shah thus proved unable to secure the fifty-fifty split he had sought. He was far-from disappointed with the new arrangement, however. Instead, believing that the deal was enough for him to secure the nationalist credentials needed to consolidate his dominance of the government, he was more than satisfied with the new arrangement. There remained one substantial complication, however: the new understanding with the AIOC would not go into effect until the majlis had approved it.⁶⁹

Unfortunately for the shah, the emergence in 1949 of a powerful new opposition movement prevented him from securing the legislature's formal endorsement of the Supplemental Agreement. It was led by an elderly notable named Mohammad Mossadegh (1882-1967). A strong nationalist and proponent of democratic rule, Mossadegh had long been a prominent figure in Iran. He had participated in the Constitutional Revolution, held office as a majlis deputy, and served as a cabinet minister until Shah Reza Kahn (r. 1925-1941) forced him into a long internal exile.⁷⁰

⁶⁷ Nikki Keddie, *Modern Iran: Roots and Results of Revolution* (New Haven: Yale University Press, 2006), 123–24.

⁶⁸ Yergin, *The Prize*, 452–54.

⁶⁹ Ervand Abrahamian, *The Coup: 1953, the CIA, and the Roots of Modern U.S.-Iranian Relations* (The New Press, 2013), 46–48; Yergin, *The Prize*, 452–54.

⁷⁰ Ervand Abrahamian, *A History of Modern Iran* (Cambridge: Cambridge University Press, 2008), 114.

His time in the political wilderness ended in spectacular fashion just after the shah's government revealed the terms of the Supplemental Agreement. Opposed to the shah's power grab and to any deal with the AIOC that gave Iran less than 50 percent of the profits, he organized and launched a series of mass popular demonstrations in October 1949 that directly challenged both of the dominant power centers in the country at that time: the shah and the British. The protests immediately reshaped the balance of political power in Iran. Most notably, they nurtured the rise of a new, populist political party under Mossadegh's direction called the National Front. Drawing support from a broad range of groups that included the Westernized middle class, the bazaar merchants, many leftists, and the followers of the influential Ayatollah Abol Ghasem Kashani (1882-1962), the party gave Mossadegh the strength to block approval of the Supplemental Agreement in the majlis. It also provided him with the institutional basis needed to mobilize people in support of a series of goals that included free elections, the suspension of martial law, strict royal adherence to the 1906 constitution and, to—great popular acclaim—the nationalization of the AIOC.⁷¹

From that point, he had little trouble dealing with the inexperienced shah. Taking advantage of the chaos that followed the assassination of Prime Minister Ali Razmara (r. 1950-1951) in March 1951, he succeeded in persuading the majlis to approve his nationalization plan and to name him prime minister. He solidified his position the following summer by luring Muhammad Reza into a political blunder. It started when Mossadegh demanded that the shah cede to him the power to name the minister of defense—a change that would give him control over the military. The shah's refusal to do so prompted Mossadegh to resign; thus, for a brief period, it appeared that Muhammad Reza's steadfastness had allowed him to reassert royal control. In fact, the shah had fallen into a well-planned trap. At Mossadegh's behest, the National Front and the communist Tudeh Party responded by holding massive demonstrations in Tehran that quickly compelled the young monarch to back down. Thereafter, Mossadegh, who had resumed his post as prime minister and assumed the defense portfolio, ruled Iran almost without challenge.⁷²

Having subordinated the shah, the prime minister next set his sights on the British. Determined to take control of Iran's oil, he followed the passage of the nationalization bill in the majlis by seizing the oilfields and by announcing his desire to enter into talks with the British aimed at determining how much Iran would have to pay AIOC shareholders as compensation for the nationalization of the company's assets. The British government did not prove to be as easily intimidated as the young shah had been, however. Determined to retain control over Iran's oilfields, it steadfastly refused to accept nationalization or to discuss compensation and indicated that it was willing to negotiate only over the size of Iran's royalty payments and not over ownership or control. To show it meant business, it simultaneously undertook a series of aggressive actions designed to force Tehran to back down. Most notably, it brought suit against

⁷¹ Abrahamian, *The Coup*, 48–57; Abrahamian, *A History of Modern Iran*, 114.

⁷² Abbas Amanat, *Iran: A Modern History* (New Haven: Yale University Press, 2017), 525–41.

Iran in the World Court, beefed up its military forces in the Persian Gulf, froze Iranian assets, and imposed a blockade designed to prevent Mossadegh's government from selling oil. It also waged an aggressive and racist propaganda campaign designed to damage Mossadegh's reputation in the West by painting him as an emotional and irrational man.⁷³

The Iranian prime minister was taken aback by these nakedly imperialist actions and appealed to the United States for aid. He got precious little support for his troubles. Officials in the Truman administration were not unsympathetic to Iran and did take steps to try to bridge the divide between London and Tehran. At the same time, however, they were disinclined to undermine the position of a key Cold War ally and concerned that the nationalization of the oil concession in Iran might encourage other countries in the Developing World to follow suit. Accordingly, Washington focused not on getting London to accept Mossadegh's action, but instead on an ill-conceived—and ultimately unsuccessful—campaign to find the right mix of verbiage and enhanced profit sharing that would allow Mossadegh to accept continued British control over Iran's oil.⁷⁴

Operation Ajax, 1953

With America refusing to take Tehran's side, London's strategy of pressuring Iran appeared to be bearing fruit in early 1953. Increased unemployment, rising inflation resulting from the Iranian prime minister's decision to continue to pay the country's idle oil workers, and a general economic slowdown stemming from the British blockade eroded support for the National Front. Likewise, Mossadegh's growing reliance on extraconstitutional methods such as his move to suspend voting midway through elections in 1952 alienated many of his allies including Kashani—a loss of support that his increasingly close relationship with the Tudeh Party only partially counterbalanced.⁷⁵

These problems would grievously erode Mossadegh's support, but it was a shift in American policy that would ultimately do him in. The change was a function of electoral politics. As we have seen, the Truman administration had sought to maintain at least a public neutrality in the crisis. The administration of Truman's successor, Dwight Eisenhower (r. 1953-1961) had no such compunction. Instead, fearing that Mossadegh's reliance on the Tudeh had undermined his ability to resist Soviet subversion, it adopted an explicitly pro-British line. It urged Mossadegh to accept a compromise solution that came nowhere near to meeting Iran's minimum conditions and refused Tehran's request for a desperately needed loan. The president went much further in the late spring of 1953. Fearing the Tudeh Party's rising influence, he

⁷³ Abrahamian, *A History of Modern Iran*, 117.

⁷⁴ Herring, *From Colony to Superpower*, 672; Abrahamian, *The Coup*, 82, 113–17, 132–33.

⁷⁵ Michael Axworthy, *A History of Iran: Empire of the Mind*, Second Edition (New York: Basic Books, 2016), 235–36; Keddie, *Modern Iran*, 129.

instructed the CIA to foment a coup aimed at removing Mossadegh from power and at installing the pro-Western shah as the dominant political force in the country.⁷⁶

Direction of the coup—codenamed Operation AJAX—fell to former President Theodore Roosevelt's (r. 1901-1909) grandson, the CIA agent, Kermit "Kim" Roosevelt (1916-2000). A covert-operations veteran, he devised a plan in the early summer of 1953 to restore the shah and remove Mossadegh from power. It called for Imperial Guard troops to deliver a firman, or royal decree, to the prime minister dismissing him from office and naming General Fazlollah Zahedi (1892-1963) as his successor. Concurrently, other units loyal to the shah would arrest National Front leaders, seize key sites in Iran including the telegraph office and the radio station, and occupy strategic positions in the capital so as to prevent a repeat of the mass demonstrations that had thwarted the shah's attempt to dismiss Mossadegh in the summer of 1952. It was a well-developed operation—albeit a nakedly colonialist one—and the administration quickly approved it. Accordingly, in July Roosevelt entered Iran by car from Iraq with a can-do attitude and a briefcase full of cash.⁷⁷

Operation AJAX may have been a sound plan on paper, but it flopped badly when royalist forces tried to execute it on August 15. Tipped off by an informer, Mossadegh arranged for a large force of loyal troops to meet the unit sent to deliver the firman to him and had its commander arrested. Meanwhile, other soldiers loyal to the prime minister prevented royalist forces from leaving their barracks. As a result, the coup collapsed before it had even started.⁷⁸

Despair and euphoria coexisted in Iran in the immediate aftermath of the failed coup. On one side, Mossadegh's supporters were exuberant that the operation had flopped so badly. Huge crowds of National Front and, especially, Tudeh loyalists poured into the streets the following day to celebrate what they saw as a great victory over Iran's foreign and domestic enemies. On the other, the coup's demoralized supporters were in complete disarray. The shah and his wife had fled by plane for what they believed would be permanent exile as soon as they had learned of Operation AJAX's failure; many other coup plotters—including Zahedi—were forced to lie low to avoid arrest. The Eisenhower administration, for its part, could neither hide nor escape. Instead, it ordered Roosevelt to return home and glumly pondered how it could salvage its relationship with Mossadegh in hopes that he might continue to help the US contain Soviet influence in Iran.⁷⁹

Roosevelt, however, was not yet ready to abandon the operation. Instead, confident that his network of Iranian assets remained intact, he hurriedly put in motion a new plan designed to exploit Mossadegh's post-coup complacency. The operation had two phases. First, on August 18, the US ambassador persuaded Mossadegh to have

⁷⁶ Wilford, 160–64.

⁷⁷ Abrahamian, *A History of Modern Iran*, 180–85; Yergin, *The Prize*, 468.

⁷⁸ Wilford, *America's Great Game*, 165.

⁷⁹ Abrahamian, *The Coup*, 183–87.

security forces clear the Tudeh mobs from the streets of Tehran by hinting that the US might withdraw its recognition of his government if he could not maintain order. Confident that the royalists had been thoroughly routed, the prime minister agreed to do so immediately. Accordingly, later that day, army and police units forced thousands of Tudeh supporters to return to their homes. With the streets no longer controlled by Mossadegh's defenders, the stage was now set for the second phase of Roosevelt's improvised plan. Deftly exploiting London's contacts with key underworld figures, he used the cash he brought with him to Iran to arrange for a pro-shah mob composed of toughs from the athletic clubs of south Tehran to begin rampaging through the capital on the morning of August 19. Predictably, Mossadegh ordered the army to break up the riot. When he did, Roosevelt struck. He arranged to have pro-shah units occupy the telephone-telegraph office, the military staff headquarters, and the radio station. Most importantly, he directed a substantial force to lay siege to Mossadegh's home.⁸⁰

The American CIA agent's ad-libbed plan was a complete success. By the end of the day, pro-royalist forces held all key government ministries as well as the critical communications exchanges. More importantly, after a short and bloody firefight outside Mossadegh's residence, they had arrested the prime minister. Three days later, a delighted shah returned from Italy to reclaim his throne—determined, this time, not merely to reign but to rule Iran as an absolutist.⁸¹

Operation AJAX had momentous consequences for Iran, Britain, and the United States. For London, the results were mixed. On the surface, the AIOC—renamed British Petroleum (BP) in 1954—was in a significantly weaker position in Iran than it had been before the oil-nationalization dispute. It had not only been forced to agree to the fifty-fifty split in profits that had become standard throughout the world but had also been compelled to accept a consortium arrangement in Iran in which it held only 40 percent of the shares while Royal Dutch Shell, a group of American companies, and the Compagnie Française assumed the balance. Indeed, the outcome for Britain seemed to have proven former Secretary of State Dean Acheson's (1893-1971) famous quip about the oil standoff to be true: “[n]ever had so few lost so much so stupidly in so short a time.” In truth, though, London's losses were far-less severe than they appeared at first blush. Not only did it retain control over pricing and production decisions in Iran, but—thanks to the fact that Royal Dutch Shell was really a British company—it also continued to hold the dominant position in the consortium.⁸²

The results were also ambiguous for Washington, though they did not appear to be so at the time. Indeed, so far as the Eisenhower administration was concerned, the coup had been an enormous success. It replaced a figure whose Cold War bona fides were in doubt with one of impeccable anticommunist credibility and had done so at a

⁸⁰ Stephen Kinzer, *All the Shah's Men: An American Coup and the Roots of Middle East Terror* (New York: Wiley, 2003), 168–85.

⁸¹ Wilford, *America's Great Game*, 166.

⁸² Abrahamian, *The Coup*, 205–26; Acheson quote from Abrahamian, *A History of Modern Iran*, 46–48.

minimal cost. As such, it appeared to be a great victory—one that helped to institutionalize covert operations as a preferred tool of American foreign policy. Its success ultimately proved to be short lived, however. It permanently tarnished Washington's image in Iran and left many Iranians bitter at what they perceived to be a serious betrayal of both their country and American principles.⁸³

For Iran, in contrast, the results were almost entirely negative. Most obviously, by entrenching the autocratic shah in power, the coup severely set back efforts to institutionalize a system of representative government. It also grievously weakened the moderate, secular opposition movements and thus created an opening that more extreme forces would later exploit. It had another, subtler implication as well, however. Roosevelt's ability to pull victory from the sure jaws of defeat led Iranians of all social classes to conclude that the US—especially the CIA—had an almost-preternatural ability to manipulate political events in Iran. As we shall see, that belief would play a key role in shaping events during the Iranian Revolution.⁸⁴

Where Historians Disagree: The Iran Coup

Two complementary interpretations have long dominated the conventional Western narrative of the nationalization crisis and the roots of the 1953 coup. The first largely exonerates the West. Scholars and diplomats who support this view contend that the British were willing to compromise on the question of nationalization and that the US was making a good faith effort to resolve the dispute in a manner acceptable to both Tehran and London. Thanks to Mossadegh's highly emotional character and his country's peculiarly intense, Shi'i-based nationalism, however, Iran was unwilling to make concessions. Writing shortly after the revolution, for example, the Middle East expert Barry Rubin blamed Tehran for the failure of negotiations, arguing that "Mossadegh and his fellows had stirred passions" to such heights that "his supporters might, perhaps literally, tear him to pieces" were he to have compromised with London.⁸⁵

The second conventional interpretation of the nationalization crisis instead downplays the question of oil ownership. Locating the coup firmly within the structure of the Cold War, scholars and writers who argue from this perspective maintain that Washington went forward with Operation AJAX out of fear that the USSR would exploit the chaos in Iran if it failed to act. Reflecting this view, for example, the journalist Robert Fisk argues that the Eisenhower administration acted because it was "fearful that Mossadegh would hand his country to the Soviets" as a result of his efforts to take control of Iran's oil. In a similar vein, the political scientist Mark Gasiorowski maintains that Washington moved to topple the Iranian prime minister not for economic reasons but instead out of concern that his unwillingness to come to terms with London could create

⁸³ Herring, *From Colony to Superpower*, 672–73.

⁸⁴ Abrahamian, *The Coup*, 205–26.

⁸⁵ Barry Rubin, *Paved with Good Intentions: The American Experience and Iran* (New York: Oxford University Press, 1980), 60–61.

the proper conditions for the communist Tudeh Party to take control of Iran and bring it into the Soviet orbit.⁸⁶

In recent years, the historian Ervand Abrahamian has forcefully challenged both of these interpretations. He first rejects the view that London's willingness to make concessions and Washington's putative efforts to serve as an impartial "honest broker" had foundered in the face of extremist Iranian nationalism and Mossadegh's refusal to consider reasonable compromises. Instead, he contends that Britain and Iran were equally inflexible because the dispute over which of them would control Iran's oil was an intrinsically zero-sum contest that no amount of negotiation could possibly finesse. Likewise, he maintains that Washington was not seeking a genuine compromise but was instead focused on persuading Mossadegh's government to accept a "pseudonationalization" scheme that created the appearance of Iranian ownership while ultimately leaving decisions about production, price, and exploration firmly in British hands.⁸⁷

London's actions were self-explanatory in Abrahamian's interpretation—they were a matter of whether Britain would continue to get the lion's share of the revenue on the sale of Iranian oil—but how does he explain Washington's motivations? Here, Abrahamian repudiates the idea that concerns about potential communist inroads in Iran shaped either Washington's decision to support Britain in the dispute or its willingness to go forward with Operation AJAX. Instead, downplaying Cold War considerations, he contends that American inflexibility on the question of continued British control was a product of a desire to prevent the establishment of a precedent that could lead to the nationalization of Western oil concessions in other developing states and that could thus threaten American interests. As he asserts, Eisenhower was acting not out of concerns about Soviet inroads in Iran but instead out of fear of "the repercussions that oil nationalization could have on such faraway places as Indonesia and South America." In other words, in Abrahamian's conception, the coup that deposed Mossadegh was not, at heart, an event rooted in the Cold War struggle between the superpowers, but instead a fundamentally imperialist action rooted in a different contest: the struggle between the Developed World and the Global South.⁸⁸

Consolidation of the Shah's Regime

Having been granted a second political life, the shah moved aggressively upon his return to Iran to solidify his control of the country. He immediately outlawed the National Front and other opposition parties, arrested many of their leaders, and denied non-regime-approved candidates the right to run for office. Later, he established a pair

⁸⁶ Robert Fisk, *The Great War for Civilisation: The Conquest of the Middle East* (New York: Vintage, 2005), 95; Mark Gasiorowski, "US Foreign Relations Toward Iran During the Mussadiq Era," in *The Military and the United States: A Historical and Political Reassessment*, ed. David W. Lesch (Boulder: Westview Press, 2007), 59.

⁸⁷ Abrahamian, *The Coup*, 2–5.

⁸⁸ Abrahamian, 2–5.

of parties so subservient to the throne that cynical Iranians referred to them as the “‘yes’ and the ‘yes, sir’” parties. Making liberal use of the regime’s now-uncontested power, he also ordered the police to destroy the Tudeh Party. This effort was not entirely successful thanks to the party’s robust, organizational structure; still, the campaign did succeed in driving it underground. Later, with the assistance of the CIA and its Israeli equivalent, Mossad, the shah’s government established a ruthless secret-police organization known by its Persian acronym, SAVAK. Making liberal use of force, torture, and arbitrary arrest, it worked to prevent the emergence of any effective political opposition. Finally, the shah exploited Iran’s surging oil income—it rose by a factor of five between 1950 and 1960—to provide patronage to his key supporters and to fund substantial pay raises designed to ensure the loyalty of the bureaucracy and the army.⁸⁹

The shah’s effort to solidify his control of Iran in the years immediately following the coup was quite successful. He secured the loyalty of the military with high pay, pensions, generous benefits, and new equipment from the US, while his expansion of the bureaucracy substantially increased the number of people who were dependent on the regime for their livelihood. More broadly, the steady rise in oil revenue underwrote a growing prosperity that seemed to erode popular support for the opposition during the latter half of the 1950s. Indeed, coming on top of the support the regime enjoyed from conservative clerics such as Kashani and the Ayatollah Borujerdi (1875-1961), the economic good times seemed to have cemented the shah’s hold on power.⁹⁰

New Challenges, 1960-1963

Appearances were not what they seemed, however, and anti-regime sentiment continued to simmer just beneath the surface. Three events brought it into the open in the early 1960s. First, the onset of a severe economic contraction ended the good times and rekindled anger toward the government on the part of many poor and middle-class Iranians. Second, pressure from the Kennedy administration for democratic reforms compelled the regime to implement a series of liberalizing measures—including an end to the ban on the National Front—and thus created room for anti-shah political activity for the first time since 1953. Finally, the government’s naked vote rigging and ballot-box stuffing in successive majlis elections infuriated Iranians of all classes. The result was a wave of popular strikes and demonstrations in the early 1960s organized by the Westernized middle class to protest the shah’s autocratic rule.⁹¹

Muhammad Reza’s difficulties broadened considerably in 1963 when an influential cleric, the Ayatollah Ruhollah Khomeini (1900-1989), began giving sermons in the main madrasa in the religious city of Qom denouncing the government. A scholar of Aristotelian logic, Khomeini was an uncompromising figure who despised the shah for his embrace of secular, Western reforms and who advocated an extreme and austere

⁸⁹ Gene R. Garthwaite, *The Persians* (Malden, MA: Blackwell Publishing, 2005), 247; quote from Axworthy, *A History of Iran*, 240.

⁹⁰ Abrahamian, *A History of Modern Iran*, 125–27; Axworthy, *A History of Iran*, 240.

⁹¹ Keddie, *Modern Iran*, 140–47; Axworthy, *A History of Iran*, 242.

version of Shi'ism. Many of his key ideas—such as his opposition to constitutionalism, his hostility to the idea of female suffrage, and his insistence on gender separation even among children—were unpopular with the majority of Iranians. Khomeini was savvy enough to avoid making those positions public, however, and instead focused on three lines of attack that he grasped would win him broad popular support: criticizing the autocratic nature of the regime, condemning it for selling oil to Israel, and, especially, demanding an end to American dominance of Iran.⁹²

This last line of attack took on a new intensity in 1964. Eager to strengthen Iran's relationship with the US, the shah accepted a \$200 million loan for the purchase of military equipment and signed a status-of-forces agreement that granted diplomatic immunity to US military personnel stationed in Iran. Khomeini loudly condemned both deals. In a series of scathing sermons, he argued that the status-of-forces agreement “reduced the Iranian people to a level lower than that of an American dog” and likened the American loan to the one-sided concessions that the venal Qajar shahs had sold to Europeans in the nineteenth century. These criticisms were effective and quickly won him a substantial following.⁹³

The White Revolution

The shah responded aggressively to the challenges of the early 1960s. To signal to both the National Front and its more radical offshoot, the Freedom Movement, that there were limits to the amount of political dissent he was willing to tolerate, he instructed state security forces to violently suppress student riots at the University of Tehran in January 1962. To assert his power over the clerics, meanwhile, he ordered SAVAK agents and army paratroopers to storm the madrasa in Qom and arrest Khomeini in June 1963. The move against the ayatollah was bloody. Several seminary students died in the operation and hundreds more perished when the military brutally suppressed demonstrations in support of Khomeini in Tehran over the next few days. The regime soon released Khomeini, but twice rearrested him after he continued to preach against the shah's rule. Finally, in 1964, the government moved to blunt the ayatollah's criticism by exiling him from Iran.⁹⁴

The attempt to diminish Khomeini's influence was a complete failure, however. Settling in the Shi'i religious center of Najaf in Iraq, he began giving a series of influential sermons inveighing against the shah that soon drew an audience among the thousands of Iranians who visited the tomb of the First Imam, Ali Ibn Talib (601-661),

⁹² Axworthy, *A History of Iran*, 242–46.

⁹³ Quote from Peter L. Hahn, *Caught in the Middle East: U.S. Policy toward the Arab-Israeli Conflict, 1945-1961* (The University of North Carolina Press, 2004), 145; Keddie, *Modern Iran*, 147–48.

⁹⁴ Amanat, *Iran*, 594–601.

each year. Smuggled home by pilgrims, copies of the ayatollah's increasingly vitriolic addresses would win him a substantial following throughout Iran.⁹⁵

Meanwhile, though the regime relied heavily on repression to cow its critics, it also took steps to coopt them. Most notably, it paired its use of force with a reform program called the White Revolution. Intended to prevent either a "black" Shi'i revolution or a "red" communist one, it sought to rally the people to the state by modernizing Iran and by eliminating many of the country's worst inequities. It instituted profit sharing for industrial workers, legalized women's suffrage, and established an educational initiative based on a successful Cuban model designed to teach illiterate Iranians to read and write. Most importantly, it put in place an aggressive land reform program aimed at making the peasants independent and at weakening the political power of the rural notables. The regime went even further in 1967 when it passed the Family Protection Law. Designed to emancipate women, it raised the age of marriage to fifteen, made divorce more difficult for men, and barred husbands from taking more than one wife without the consent of his existing spouse or spouses. It also discouraged the practice of veiling, though it did not outright ban the practice. The White Revolution and the Family Protection Law were not without their detractors, and they did nothing to loosen the monarch's grip on power. Nonetheless, they proved to be broadly popular—sufficiently so that they helped to blunt the challenges that the shah faced in the early 1960s.⁹⁶

The Oil Boom

Equally important in taking the air out of the opposition was the robust, oil-fueled economic growth that Iran experienced in the late 1960s and, especially, in the early- and mid-1970s. As we have seen, surging global demand for oil beginning in the 1960s had given the producers the upper hand vis-à-vis both the consumer states and the oil companies and had dramatically raised the price of petroleum. The result was an absolute bonanza of foreign revenue for Iran. Its oil income shot up from a comparatively modest \$34 million in 1953-1954 to \$5 billion in 1973-1974 before hitting a mind-boggling \$20 billion in 1975-1976—a nominal increase of 59,000 percent in a little more than two decades!⁹⁷

Sustained by the massive inrush of petrodollars, the White Revolution and other modernization initiatives produced impressive results. In the countryside, the land reform program resulted in more than two-million peasant families acquiring legal title to their own farms while the construction of huge new dams provided the water needed to irrigate their fields. Meanwhile, in the cities, petrodollars and ISI policies produced a period of rapid industrialization. With government support in the form of tariffs and subsidized loans, entrepreneurs connected to the palace established thousands of industrial facilities including textile mills, machine tool factories, and soft-drink bottling

⁹⁵ Keddie, *Modern Iran*, 143–48, 192.

⁹⁶ Abrahamian, *A History of Modern Iran*, 134; Keddie, *Modern Iran*, 144–46, 167.

⁹⁷ Axworthy, *A History of Iran*, 247; Abrahamian, *A History of Modern Iran*, 123–24.

plants, while the regime itself underwrote the construction of larger industrial concerns such as steel mills, automobile assembly plants, and even a nuclear powerplant. Education and health programs enjoyed similarly rapid growth. Between 1963 and 1977, elementary-school enrollment rose by 250 percent while college registration quadrupled. Over the same period, the number of doctors in the country tripled.⁹⁸

It was in many regards a tremendous achievement. Fueled by the vast surge in oil revenue, the regime had succeeded both in rapidly developing Iran's economy and in modernizing much of the country's infrastructure—and it had done so in record time. It is certainly true that the shah's stated goal of making Iran one of the globe's top-five industrial powers by the end of the century was hyperbole. Still, sustained by the huge inflow of petrodollars, Iran experienced undeniable and enviable economic growth in the late 1960s and early 1970s.⁹⁹

At the same time, the shah misspent a huge share of Iran's oil windfall. In particular, his regime devoted an absurd amount of its new wealth to arms purchases designed to fulfill the shah's ambition to serve as the "gulf policeman." This aspiration meshed well with a new American foreign-policy initiative, the Nixon Doctrine, that President Richard Nixon (r. 1969-1974) had announced in 1969. Designed to prevent the US from ever again becoming bogged down in a conflict like the Vietnam War, it called for Washington to refrain from deploying American troops to the Developing World and to instead rely on powerful regional allies to contain Soviet adventurism. Staunchly anticommunist, the government of Iran appeared well suited to fill that role in the Gulf region.¹⁰⁰

The shah's regional ambitions also complemented America's now-pressing need to redress the balance-of-payments deficit that surging oil prices had created. As we saw in the context of Saudi Arabia and the Gulf States, Washington understood that the sale of complex weapon systems to the producer countries would allow the US to repatriate a substantial share of the petrodollars that had begun flowing to the Middle East in the early 1970s. Accordingly, after the Yom Kippur War, the Nixon administration removed nearly all restrictions on arms purchases to Iran and encouraged the shah and his generals to go on a shopping spree. Now free to do so, Iran bought billions of dollars' worth of advanced weapons including F-14 fighter jets, destroyers, and the latest anti-aircraft missiles. Washington was not alone in providing Tehran with high-tech arms. Also looking to redress its balance-of-payments deficit, Britain sold substantial quantities of weapons to the Iranian military. In fact, it provided so many Chieftain tanks to Iran in the 1970s that the shah's army ended up fielding more of them than did the British military.¹⁰¹

⁹⁸ Abrahamian, *A History of Modern Iran*, 133–34; Axworthy, *A History of Iran*, 246–47.

⁹⁹ Abrahamian, *A History of Modern Iran*, 133–34.

¹⁰⁰ Herring, *From Colony to Superpower*, 785, 798–99.

¹⁰¹ Keddie, *Modern Iran*, 163–64.

The shah also squandered a significant amount of oil revenue on spectacles designed to enhance the prestige of the monarchy. These included his formal coronation ceremony held in 1967 and, especially, his commemoration of the Iranian monarchy's two-and-a-half millennia anniversary. Attended by a vast assemblage of heads of state and other notables, the celebration was a bacchanalian occasion of epic proportions. Held at the ruins of the ancient Persian capital of Persepolis, this lavish and ahistorical event involved thousands of actors in period costumes, a series of bizarre speeches in which the shah linked himself to Cyrus the Great (r. 559-530 BCE), and a succession of almost comically indulgent feasts. Catered by Maxim's of Paris, the elaborate meals were served on the finest Limogès china and washed down with a staggering 25,000 bottles of wine including prized vintages such as Château Lafite Rothschild. To the shock of ordinary Iranians, the total bill for the event was \$21 million—an obscene figure for the time.¹⁰²

Renewed Discontent

The commemoration of the Iranian monarchy at Persepolis was not merely a conspicuously expensive affair, but also one that failed utterly to achieve its purpose. Designed to bind the shah's subjects to the regime through the construction of a new, national identity, the elaborate ceremony had little resonance with a people that perceived its history not in dynastic terms but instead in an Islamic framework. Indeed, rather than building loyalty to the regime, it contributed to a rising sense of alienation and helped to feed a growing opposition that, ironically, drew increasing strength from the same oil-fueled economic boom and reform program that had dissipated the protest movements of the early 1960s. What explains this conundrum? How had the reforms of the White Revolution and the rapid economic growth that had raised living standards for most Iranians come to produce widespread dissatisfaction with the government in the 1970s?¹⁰³

Three consequences of the economic expansion and the modernization effort explain this apparent contradiction. First, while the oil boom resulted in an absolute rise in per capita income, it also badly exacerbated Iran's already stark income inequality. That it did so was largely a function of the fact that the regime's economic-development strategy was a top-down program that centered on distributing oil revenue to the wealthiest Iranians so that they could invest it in businesses that would, it was hoped, raise living standards for the lower classes. As a result, the flood of oil money that poured into Iran in the 1970s went disproportionately to regime officials and their hangers on—a fact that the upper class's increasingly brazen displays of conspicuous consumption made obvious to all.¹⁰⁴

¹⁰² Amanat, *Iran*, 664–66; Michael Axworthy, *Revolutionary Iran: A History of the Islamic Republic* (New York: Oxford University Press, 2016), 76–77.

¹⁰³ Keddie, *Modern Iran*, 167–68.

¹⁰⁴ Abrahamian, *A History of Modern Iran*, 140–41; Garthwaite, *The Persians*, 251–52.

Second, while the oil boom and the shah's modernization program were popular, they also challenged the position of two highly influential groups in Iranian society: the bazaar merchants and the religious scholars. For the bazaar merchants, the period since the White Revolution had been a difficult one. The oil boom and the growing integration of Iran into the global economy had produced a rise in cheap imports that undercut the price of the artisanal handicrafts that were their livelihood, while the development of Western-style stores threatened their longstanding dominance of the retail sector. For the 'ulama', the problem was less the economic changes that Iran was experiencing than the reforms that the shah had pushed through. Most notably, they bitterly opposed the Family Protection Law's transfer of marriage, divorce, and inheritance proceedings from shari'a courts to secular ones. That change not only contradicted their deeply held understanding of how a properly functioning Muslim society should be organized but, as important, directly undercut their authority, influence, and income.¹⁰⁵

Finally, and most importantly, the reforms of the White Revolution and the rapid economic expansion that followed raised expectations that the shah's regime either could not—or would not—meet. The issue of unmet expectations existed up and down the social ladder. For rural Iranians, the White Revolution had led them to conclude that they would soon acquire land and see the benefits of modernity extended to their villages. Instead, to their mounting frustration, most continued to live in communities that lacked basic elements of modern life such as electricity, plumbing, and schools. Worse, while many peasants did gain title to their farms thanks to the land-reform program, the poorest—the rural laborers—received no land at all and frequently lost their jobs due to the mechanization of agriculture that the influx of oil wealth had made possible. Millions consequently found themselves with no choice but to move to the shantytowns that were springing up around the country's cities. There, they eked out a meager existence as low-wage, unskilled workers—a far cry from the life that the White Revolution and the oil bonanza had promised.¹⁰⁶

Members of the westernized middle class fared much better than rural Iranians during the economic expansion of the 1970s, but they too were bitter—not just about the regime's failure to meet their economic expectations but, as important, about its refusal to satisfy their political aspirations. Their economic frustration stemmed to a large degree from their inability to fill the best jobs or to access the housing to which they aspired. Instead, the top positions managing the shah's infrastructure and maintaining the military's new, high-tech military equipment went to highly skilled American contractors; thanks to the lavish salaries the regime had to offer to attract them, they were also well-positioned to outbid Iranians for increasingly scarce housing. Politically, meanwhile, the White Revolution had reinforced the expectation of middle-class Iranians that they should and would enjoy the right to participate in the

¹⁰⁵ Keddie, *Modern Iran*, 223; Axworthy, *A History of Iran*, 247; Garthwaite, *The Persians*, 250.

¹⁰⁶ Abrahamian, *A History of Modern Iran*, 141–43.

governance of the state in accordance with their economic importance. To their frustration, the shah instead continued to jealously maintain his tight grip on power.¹⁰⁷

As anger at the regime mounted in the 1970s, many workers, students, and members of the intelligentsia began to join underground political organizations or to adhere to radical new ideologies. Militant leftists gravitated toward a pair of underground guerrilla organizations: the *Fedayin-e Khalq*, an offshoot of the Tudeh Party that drew on the works of radical thinkers like Frantz Fanon (1925-1961), and the *Mojahedin-e Khalq*, a radical outgrowth of the Freedom Movement that promoted a mix of Shi'i and Marxist thought. Put off by the violence that those groups espoused, other, more-moderate leftists instead fell under the influence of the French-educated intellectual Ali Shariati (1933-1977). Drawing on both progressive Western thought and the spirit of Jamal al-Din al-Afghani (1838-1897), he articulated a new, radical conception of Shi'ism as a populist, revolutionary ideology and likened the Imam Husayn (626-680) to modern-day insurgents like Che Guevara (1928-1967). A prolific author, he criticized the 'ulama' for quietism and for collaborating with the upper classes, and he attacked the Pahlavi state for its autocratic rule and failure to adhere to the social-justice values that he maintained lay at the heart of what he called the "Red Shi'ism of the Imams." Shariati's death under suspicious circumstances while in exile in 1977 ended his criticism of the regime but did not diminish his popularity. On the contrary, his ideas continued to inform the views of millions of Iranians during the revolution.¹⁰⁸

While many leftist students and intellectuals gravitated toward Shariati, the large, westernized middle class instead coalesced around the National Front and Mehdi Bazargan's (1907-1995) Freedom Movement. Revived in 1977 by two veterans of Mossadegh's government, Shapur Bakhtiar (1914-1991) and Karim Sanjabi (1905-1995), the National Front was a liberal secular political party that demanded free elections and a return to the constitution of 1906. Itself an offshoot of the National Front, Bazargan's Freedom Movement similarly called for a return to democracy and promoted a moderate, Islamic-modernist form of government that integrated Western science and technology with Muslim values.¹⁰⁹

For their part, the urban poor, the bazaris, and the 'ulama' increasingly adhered to the views of the exiled Khomeini. Spread by cassette tapes that returning pilgrims smuggled into Iran, the ayatollah's weekly sermons offered a scathing critique of the Pahlavi state. He condemned the regime for its economic shortcomings, its corruption, its inability to extend the benefits of modernization to rural people, its refusal to protect the bazaris from foreign competition, and, especially, its subservience to the US.

¹⁰⁷ Garthwaite, *The Persians*, 243, 251–52; Keddie, *Modern Iran*, 164.

¹⁰⁸ Vali Nasr, *The Shia Revival: How Conflicts within Islam Will Shape the Future* (New York: W. W. Norton, 2006), 128–29; quote from Abrahamian, *A History of Modern Iran*, 143–46, 148–49.

¹⁰⁹ Keddie, *Modern Iran*, 198–200, 232.

Seemingly rooted in traditional Shi'i thinking, these critiques resonated deeply with millions of poorer and more conservative Iranians.¹¹⁰

In fact, Khomeini's ideas broke sharply with longstanding Shi'i teachings. He made this clear in a book he published in the early seventies called *Velayat-e Faqeh*, or *The Jurist's Guardianship*. Arguing that the idea of monarchical government was a relic of pre-Islamic times, it called for Iran to be restructured as a theocratic state in which the senior clerics ruled the country as the representatives of the Hidden Imam. Khomeini claimed in the book that his idea for religious rule was well grounded in longstanding Shi'i thought. Despite some precedent among late-Safavid-era clerics, however, it in fact marked a radical break with earlier Shi'i thinking—an innovation that had no basis in either Qur'anic teaching or the guidance of the Twelve Imams.¹¹¹

Khomeini was well aware that his guardianship idea would be unpopular with many Iranians and that it would engender fierce opposition from pro-Western reformers and advocates of democracy. Accordingly, taking advantage of the fact that the book was all-but-unknown beyond a small circle of seminary students, he shrewdly refrained from discussing it publicly and instead presented himself as a typical Shi'i conservative. Indeed, as we shall see, it was only after he had curbed or defeated his rivals in 1979 that he would begin to pursue openly the theocratic state that he had envisioned in *Velayat-e Faqeh*.¹¹²

Vacillation

Meanwhile, even as dissatisfaction with the shah's regime deepened, rising criticism from the international community put further pressure on the Iranian government in the mid 1970s and encouraged the opposition to act more boldly. Human-rights groups including the International Commission of Jurists and, especially, Amnesty International, brought global attention to the lack of due process and to the shocking prevalence of torture in the Iranian legal system. Far more important, however, was the criticism coming out of Washington. Implying that the US might limit or end arms shipments if Iran did not end its more egregious human rights' violations, the new American president, Jimmy Carter (r. 1977-1981), aggressively pressed the shah to ease his government's worst human-rights abuses. In response, the regime implemented an assortment of high-profile, but ultimately modest reforms including the relaxation of censorship laws and the release of some political prisoners.¹¹³

These reforms may not have been substantive, but because they signaled weakness they had the unintended consequence of spurring opposition groups to step up their criticism of the shah's government. Most notably, in 1977, newly formed professional associations issued petitions attacking the government while the National

¹¹⁰ Abrahamian, *A History of Modern Iran*, 147–48.

¹¹¹ Nasr, *The Shia Revival*, 125.

¹¹² Nasr, 125.

¹¹³ Keddie, *Modern Iran*, 214–17.

Front released a public letter that boldly demanded the restoration of civil liberties and free elections. Owing in part to the treatment that Muhammad Reza had begun for the cancer that would soon kill him, the government seemed unable to maintain a consistent line toward the opposition groups and instead veered erratically between conciliation and repression. Most notably, while state security forces violently attacked protestors at demonstrations held in Tehran in November 1977, the courts quickly acquitted and released those whom the police had arrested. This lack of consistency proved disastrous for the regime. Continued repression merely stoked anti-shah anger, while conciliatory gestures suggested not moderation or restraint but instead a degree of indecision and vulnerability that served only to embolden the regime's opponents to further action.¹¹⁴

The Iranian Revolution, 1978-1979

Indeed, the regime's vacillation in 1977 inspired the outbreak of a protest by seminary students in Qom in January 1978 that is conventionally understood to mark the beginning of the revolution. The demonstration had its origin in an inflammatory editorial in a government-controlled newspaper that accused Khomeini of engaging in immoral behavior. Concluding from the inconsistent response to the recent demonstrations in Tehran that they could protest without bringing the wrath of the state down on them, thousands of conservative seminary students in Qom reacted by taking part in a march organized to protest the attack on the ayatollah. Furious about the insults to Khomeini, they demanded that the shah's government issue a formal apology and that it permit him to return to Iran. The religious students had in fact badly misjudged the regime's willingness to tolerate dissent, however. This time—in sharp contrast to their more-measured handling of the protests in Tehran—the police used deadly force to break up the demonstrations.¹¹⁵

The violence in Qom was a watershed event in the growing opposition to the shah. Important in its own right, it set in motion a train of events that would result in the protests spreading rapidly throughout Iran. Key to that process was the deft exploitation of the Shi'i custom of holding a commemorative service forty days following a person's death. Shrewdly turning the memorial services into demonstrations against the regime, the religious scholars and bazaar leaders who took charge of the populist opposition to the shah ensured that the state would respond to the services with deadly force—thus guaranteeing that a series of even bigger protests would occur in an even-larger number of cities and towns forty days later. Still, even as the demonstrations grew in size and number, there was little sense at that time that the protests were going to produce a change of regime in Iran. Indeed, the situation began to stabilize in the early summer after a moderate ayatollah, Mohammad Shariatmadari (1906-1986), called for an end to the memorial demonstrations in order to prevent further bloodshed.¹¹⁶

¹¹⁴ Axworthy, *A History of Iran*, 252–53; Abrahamian, *A History of Modern Iran*, 156–58.

¹¹⁵ William R. Polk, *Understanding Iran* (Basingstoke: Griffin, 2011), 124.

¹¹⁶ Axworthy, *Revolutionary Iran*, 104–9, 112–13.

The lull proved to be very brief, however. Following a deadly arson attack on the Rex Theater in Abadan in August that resulted in the deaths of 377 people, the protests roared back to life with a vengeance. Widely—if incorrectly—blamed on SAVAK, the incident galvanized popular opinion against the shah. As a result, the demonstrations that followed the cinema fire were different in kind and size from those that had occurred in the spring. Considerably larger and involving a far-broader and more-moderate cross section of the population, they reflected surging dissatisfaction with the shah's regime among the technocratic middle class.¹¹⁷

The new round of protests culminated with a massive demonstration in Tehran on September 8. Known as Black Friday, the huge rally and the regime's violent response to it marked the beginning of the end for the shah. Though the government's use of force that day resulted in the deaths of far-fewer people than was commonly understood at the time, it nonetheless poisoned the waters and ended any chance that the shah could negotiate an agreement that permitted him to remain in power. Instead, within weeks, oil workers, bazaar merchants, bank employees, railroad engineers, and even the shah's own bureaucrats had crippled the economy by going on strike.¹¹⁸

Indeed, by December, the government was so weak that it had no choice but to permit a massive opposition march in Tehran involving an estimated two-million people. If the sheer size of the rally did not make adequately clear that the people of Iran wanted the shah to go, the protestors' approval of a series of resolutions demanding Khomeini's return and the transformation of the country into an Islamic republic surely did. With desertions from the military rapidly rising, even the shah began to grasp that his time as Iran's ruler was coming to an end. In desperation, he asked Shapur Bakhtiar to serve as prime minister. On December 30, the latter agreed to do so—an action that resulted in his prompt expulsion from the National Front—but only on the condition that Muhammad Reza depart Iran immediately. The shah dragged his feet for a time, but, on January 16, 1978, he and his family finally left the country. Just sixteen days later Khomeini returned to a hero's welcome.¹¹⁹

Khomeini's Victory, 1979-1981

The departure of the shah and Khomeini's triumphal homecoming marked the end of the first phase of the Iranian revolution and the beginning of its far-more divisive and complex second stage. The first phase had centered on the twin questions of whether the shah would continue to rule and whether Iran would remain a monarchy. On those points, nearly all Iranians agreed. The second addressed a more fundamental and divisive issue: whether the country would have a democratic state controlled by the people or a fundamentally religious one dominated by the Shi'i clergy.¹²⁰

¹¹⁷ Amanat, *Iran*, 716–18.

¹¹⁸ Abrahamian, *A History of Modern Iran*, 159–61.

¹¹⁹ Keddie, *Modern Iran*, 234–35.

¹²⁰ Garthwaite, *The Persians*, 255.

That there was a second phase to the revolution came as a surprise to the moderates in the Freedom Movement and the National Front. Since all political parties save for the far-left ones agreed on a shared vision of the country as an Islamic republic, the moderates had assumed that the departure of the shah marked the end of the revolution and the beginning of Iran's transformation into a democratic state. They soon discovered, however, that their idea of an Islamic republic was decidedly different from the conception that Khomeini held. Where they interpreted the term to mean that Iran would be a Western-style, multiparty democracy that rested on Shi'i values, Khomeini instead understood it to describe a state based on the concept of *velayat-e faqeh* that he had earlier developed—meaning that the clerics would hold the final say over political and social questions. In other words, the moderates were assuming that Iran would become a democratic state while Khomeini was calling for it to be, at root, a theocracy. Still, the leaders of the moderate parties were not particularly concerned about Khomeini's undemocratic conception of what an Islamic republic entailed. Despite the ayatollah's enormous popularity—three-million people greeted him upon his return from exile—he was politically inexperienced and appeared to be both naive and fundamentally backward looking. In other words, he seemed like a lightweight—someone the moderates were confident they would run rings around.¹²¹

For a brief time, events seemed to reinforce this view. It is true that Khomeini forced Bakhtiar from office by declaring that he would not recognize the legitimacy of any figure associated with the shah's regime. However, faced with the popular and savvy liberal politicians, a seemingly chastened Khomeini agreed to establish a liberal provisional government and named Freedom Movement leader Mehdi Bazargan as its prime minister. He likewise accepted an arrangement in which the liberal parties held two-thirds of the cabinet posts in the new government, and—critically—indicated that he no longer supported the idea of clerical rule. Accordingly, Bazargan and the other moderates turned their attention away from Khomeini following his return and focused on drafting a new constitution that would transform Iran into an Islamic republic dominated by secular political parties.¹²²

Events would soon make clear that the moderates had badly underestimated Khomeini. He was not, as they believed, either backward looking or naive. On the contrary, he was an unusually savvy and ruthless political figure—a skilled bureaucratic infighter and a charismatic leader who was capable of deftly manipulating public opinion. Shrewdly exploiting the rapidly shifting events of the revolution, he outmaneuvered the moderates over the next year and solidified his position as the dominant figure in Iran.¹²³

Khomeini began by quietly moving to strengthen his political position at the moderates' expense. First, to divide his opponents and to prevent the emergence of a broad coalition against him, he struck a temporary alliance with the Tudeh and the

¹²¹ Garthwaite, 225–29.

¹²² Keddie, *Modern Iran*, 240–43; Garthwaite, *The Persians*, 256.

¹²³ Abrahamian, *A History of Modern Iran*, 162–63.

Fedayin-e Khalq. He followed by establishing an increasingly powerful Islamist shadow state that paralleled Bazargan's provisional government. Dominated by the 'ulama', this extraconstitutional body was composed of the Revolutionary Council, which passed laws and checked the authority of the Provisional Government, and the Central Committee, which coordinated the many local Islamist groups and militias that had emerged in late 1978 and early 1979 as the shah's state was imploding. Other moves soon followed. Over the course of the spring and early summer of 1979, Khomeini assumed control over the judiciary, arrested and executed members of the shah's regime, and—most critically—created a new Islamist paramilitary force called the Revolutionary Guard that he could use to break up the moderates' political rallies. Undertaking all of these initiatives with consummate skill, Khomeini quickly solidified his political position. As a result, by the summer of 1979, he was ready to vie with the moderates for control of the revolution.¹²⁴

Khomeini focused his challenge on the constitution that the liberals had been writing. In June 1979, Bazargan and the moderates had completed a draft document based on the constitution of France's Fifth Republic. While rooted in Shi'i values, the new constitution was a strongly democratic document. It emphasized the rule of the people, eliminated the monarchy, and, importantly, assigned no meaningful power to the clergy. Satisfied with their work, the liberals were ready to present their draft to the people for formal ratification. Before they could do so, however, Khomeini's supporters objected. Calling for a formal review before the document went to a public vote, they demanded that an elected council of prominent Shi'i legal scholars, the Assembly of Experts, first examine and refine the draft. Stacked as it was with Khomeini's supporters, the assembly unsurprisingly released a revised draft that was very different from the one that the moderates had produced. Far more theocratic than democratic in both structure and character, it drew explicitly on the concept of *velayet-e faqeh* to assign the clerics and, especially, Khomeini, enormous power. Emblematic of these changes was its declaration that the ayatollah's rule was divinely ordained and that he answered to none but God.¹²⁵

Still, Khomeini's position and the passage of the revised constitution were far from assured. Many Iranians, perhaps even a majority, supported neither the ayatollah nor his idea of clerical rule and instead backed moderate parties like the Freedom Movement and the National Front. Women were particularly vocal in expressing their opposition to clerical rule. Fearing the loss of the rights that they had gained since the White Revolution, they organized mass demonstrations to make clear their refusal to accept the imposition of new restrictions. Even many of the 'ulama' opposed Khomeini. Most notably, a number of influential moderate clerics found his *velayet-e faqeh* idea to be troubling both because it lacked grounding in Shi'i thought and because it broke with the religious scholars' longstanding practice of acknowledging the government's authority over secular matters in exchange for the state respecting the clergy's dominance over religious and social affairs. In other words, Khomeini may have had

¹²⁴ Keddie, *Modern Iran*, 241–45.

¹²⁵ Abrahamian, *A History of Modern Iran*, 162–64; Keddie, *Modern Iran*, 246–48.

enormous influence, but he also continued to face substantial opposition that limited his capacity to force through the radical changes enumerated in the revised draft of the constitution.¹²⁶

Ultimately, it was Khomeini's deft exploitation of the Hostage Crisis with the US that assured both his victory in the struggle for power and the passage of the revised constitution. The conflict's genesis lay in President Carter's decision to permit the exiled shah to travel to New York City for cancer treatment in October 1979. While unsuccessful—Muhammad Reza died in 1980—the president's decision breathed life into an elaborate conspiracy theory in Iran that held that Washington was plotting to overthrow the revolution and to return the shah to power. Such an action was wildly beyond the capabilities of even the powerful CIA, but it appeared entirely plausible to a people well aware of the agency's role in overthrowing Mossadegh in 1953. As one hostage taker later declared, "In the back of everybody's mind hung the suspicion that, with the admission of the Shah to the United States, the countdown for another coup d'etat had begun." Attention quickly centered on the US embassy in Tehran. Fearing that the CIA was using it as a base from which to execute a change of government—just as it had in 1953—a group of Iranian students loyal to Khomeini seized it on November 4, 1979, and took a number of American State Department personnel hostage.¹²⁷

In two ways, the occupation of the embassy immediately and fundamentally altered the political balance in Iran. First, acute conflicts with a foreign power typically create a rally-round-the-flag effect in which patriotic sentiment and approval for a country's leader rise rapidly. Such was certainly the case in Iran in 1979 and 1980. Even among people who opposed his ideas, support for Khomeini soared following the onset of the Hostage Crisis. Second, just as the conflict with the US enhanced Khomeini's influence, so did it grievously damage the moderates' political standing. Championing an essentially American-style system of constitutional government, Freedom Movement and National Front leaders suffered a sharp loss of support due to their association in the public imagination with the US.¹²⁸

Now in a commanding position, Khomeini shrewdly exploited the crisis. He kept his moderate rivals on their heels by preventing Bazargan from securing the release of the hostages—a move that prompted the prime minister to resign—and by steadily releasing copies of embassy files detailing meetings between them and American officials that, while unremarkable, appeared ominous in the hothouse atmosphere of revolutionary Iran. He also muted dissent by equating criticism of the government during a time of crisis with treason. Most importantly, he took advantage of the upsurge in

¹²⁶ Garthwaite, *The Persians*, 259–61.

¹²⁷ Kinzer, *All the Shah's Men*, 203.

¹²⁸ Amanat, *Iran*, 821–25.

patriotic sentiment that the Hostage Crisis produced to secure popular approval of the revised constitution in a referendum held in December 1979.¹²⁹

The passage of the constitution was a major political victory for Khomeini that solidified his control of Iran. Based on the ideas that he had outlined in *The Jurist's Guardianship*, the document established an Iranian state that bore only a superficial resemblance to the kind of government that people like Bazargan had envisioned. While the constitution did assign the power to conduct the routine business of state to a secular president and majlis, it placed them firmly under the thumb of an interlocking apparatus of powerful, unelected institutions controlled by the clergy. At the top of that structure stood Khomeini: the supreme leader who officially ruled in the name of the Hidden Imam. The constitution accorded him vast powers. He held the authority to declare war, to grant amnesty, and to appoint many high-ranking officials including the commander of the military, the head of the Revolutionary Guard, and the chief justice. Just beneath him was an unelected, twelve-man body called the Guardian Council that had the power to determine who could—and who could not—stand for election to the majlis. More importantly, functioning like an Islamic supreme court, it could veto any legislation that it determined was contrary to Shi'i principles as it defined them; in other words, the Guardian Council held the unchecked power to determine the constitutionality of any law the majlis passed.¹³⁰

The constitution did not merely vest enormous power in the religious scholars, but also created a complex, self-reinforcing structure that ensured that any effort to chip away at their authority was doomed to failure. To start with, the supreme leader was not popularly elected; instead, the Assembly of Experts selected him from among their ranks. He, in turn, appointed half of the Guardian Council. The majlis had the power to name the other six members, but it had to choose them from a list of candidates created by a body called the Supreme Judicial Council, which was itself composed of clerics appointed by the supreme leader. To complete the series of closed loops that Khomeini and his supporters had put in place to ensure perpetual clerical dominance, the Guardian Council enjoyed the power to vet the legal scholars on the Assembly of Experts who chose the supreme leader. In other words, since the top-ranked clerics effectively named each other to the key posts and councils in the government and since one of those bodies, the—Guardian Council—determined the constitutionality of all laws, only the clerics could meaningfully alter the structure of government or, critically, shift power from the 'ulama' to the secular state. Unsurprisingly, they have vigorously resisted proposals to do so. Thus, despite periods of intense public dissatisfaction with religious rule, Iran's government has remained essentially unchanged since 1979.¹³¹

Khomeini claimed that the new constitution and the system of religious rule were firmly grounded in longstanding Shi'i religious traditions. In fact, it was a radical innovation with almost no past precedent. The very core of it—the concept of *velayat-e*

¹²⁹ Keddie, *Modern Iran*, 248–49.

¹³⁰ Amanat, *Iran*, 783–89; Abrahamian, *A History of Modern Iran*, 164–65.

¹³¹ Keddie, *Modern Iran*, 247–48.

faqeh—had traditionally formed the legal justification for the clergy’s oversight of those who could not care for themselves such as widows, children, and the developmentally disabled, and had never before constituted the basis for a government. Likewise, the Shi‘i clergy had at no point in the past exercised political power and had instead been content to adhere to a form of separation of church and state wherein the secular government dealt with the affairs of the world while the clergy held sway over religious and social matters. In other words, while carefully camouflaged in a conservative wrapper, the new state that Khomeini’s constitution established in fact constituted a fundamental break with past practice—one that vested nearly all power in the hands of the ayatollah and his closest allies.¹³²

The passage of the new constitution marked the end of the second phase of the revolution and the culmination of Khomeini’s nearly total political victory. By the time constitution went into effect, he had neutered the moderate parties that had been his most serious rivals and had assumed complete control of the government. It is true that he still had to deal with the leftist organizations with which he had entered into an alliance of convenience against the liberals in 1979. Likewise, he continued to face a spirited, if doomed, opposition from the first president of the republic, Abolhassan Bani Sadr (r. 1980-1981). By 1980, however, neither the president nor the leftists could overcome Khomeini’s commanding ideological and institutional position. Thus, despite a savage terror campaign on the part of the *Mojahedin-e Khalq*, he was able to brush aside his remaining political enemies and cement his total dominance of Iran—a reality made emphatically clear when Bani Sadr and the leader of the *Mojahedin-e Khalq* fled the country together aboard a commandeered jet in July 1981.¹³³

The Transformation of Iranian Society

Khomeini’s political victory had enormous ramifications for Iran. Most obviously, it permitted the ayatollah and his supporters to impose a sweeping Islamist, social revolution on the country. They wasted no time in doing so. The new government immediately banned alcohol, most Western movies, and nearly all music, quickly closed newspapers that did not hew to the official line, compelled schoolteachers to adopt a stilted curriculum developed by the state, launched a violent campaign to Islamicize the universities, and reimposed shari‘a law—including traditional criminal penalties such as flogging. No group felt the effects of the Khomeinist social revolution more acutely than did women. The regime required them to veil, revoked the rights they had received under the Family Protection Law, discouraged them from wearing Western clothing, forced them from the judiciary and teaching, reimposed gender segregation, and even revived practices such as child marriage. Implemented at breakneck speed, the changes that Khomeini had imposed were far reaching and constituted a true social

¹³² Abrahamian, *A History of Modern Iran*, 146, 176–77.

¹³³ Garthwaite, *The Persians*, 266–67.

revolution—one that transformed Iran into a visibly different country than it had been just a few years earlier.¹³⁴

Unsurprisingly, the Khomeinist social revolution was far from universally popular. While the ayatollah's reform program received a warm reception among rural Iranians and shantytown dwellers, it faced staunch opposition among members of the technocratic middle class who were bitterly disappointed that the overthrow of the shah had resulted not in a genuinely representative government but instead in a state that was even more autocratic than the one it had replaced. They were unable to effectively resist the new regime after Khomeini had solidified his grip on the country, however—especially, after his government began to make liberal use of brute force against their protests. Defeated, the Westernized middle class thereafter adopted a cynical view of the government that saw it as an alien and even hostile entity that they had to endure so long as they remained in Iran. Not all of them chose to do so. Weary of the repression they faced at home, more than one-million members of the middle-class voted with their feet in the years after the revolution and—at great economic cost to Iran—took their considerable education, skills, and entrepreneurial abilities with them to the West in search of a better life and greater freedom.¹³⁵

Conclusion

The Iranian Revolution was yet another of the watershed events that the Middle East experienced during the tumultuous twentieth century. It greatly weakened America's position in the region, stoked Sunni- Shi'i conflict, and, most importantly, announced the arrival of Islamism as a major political force in the Muslim world. It was not, however, the only major event to occur in the region in 1979. Indeed, while it was clearly the most important, it was but one of a series of cataclysmic changes that would, collectively, effect a fundamental restructuring of the Middle East that year. It is to these events and to their complex genesis that we will now turn.

¹³⁴ Amanat, *Iran*, 809–15; Keddie, *Modern Iran*, 257.

¹³⁵ Garthwaite, *The Persians*, 269.

Chapter Sixteen: The Middle East After Nasser, 1967-1979

Introduction

As we saw in chapter fourteen, the Six Day War was one of the seminal events in the making of the modern Middle East—a great before-and-after moment that reshaped the region. Its impact was pervasive but uneven. In the Arab states, it produced an ideological sea change, inspired a series of coups, and created pressure for another war that could restore Arab honor and make possible the resolution of the Arab-Israeli conflict. The changes were equally significant in Israel and the Occupied Territories. For a growing number of Israelis, the seizure of the Sinai, Golan Heights, West Bank, and Gaza Strip created a tempting opportunity to expand their country; for the Palestinians, in contrast, the war produced a new, shared experience of oppression that would quickly reinvigorate their sense of national identity. More broadly, the Six Day War fundamentally transformed the tenor of the Arab-Israeli conflict. Effecting changes that would result in the establishment of peace between Israel and Egypt and in the emergence of an organized Palestinian resistance movement, it set in motion a process that would transform the struggle between Israelis and Arabs from a state-to-state dispute back into what it had been before 1948: a contest between two peoples for control of the land of Palestine.

Not every major trend in the decade following the Six Day War was directly related to the conflict, however. Nurtured by Gamal Abdel Nasser's (r. 1954-1970) successor, Anwar Sadat (r. 1970-1981), and supported financially by the Arab oil states, a new movement, Islamism, emerged as a force in the Arab world. Focused initially on recruiting members and on building institutional structures, the Islamists and their Jihadi offshoots would burst forth at the end of the 1970s to mount a powerful and persistent challenge to Arab states across the political spectrum—one that would, in conjunction with a series of dramatic events in 1979, once more reshape the region.

The Legacy of 1967

The Decline of Nasserism

The Six Day War's biggest casualty by far was Nasserism: the pan-Arabist ideology that had so thoroughly dominated the Arab world during the prior decade. The deficiencies of Egypt's military leadership and the rash decision to engage in brinksmanship with Israel at a time when the cream of the Egyptian army was bogged down more than 1,000 kilometers away in Yemen thoroughly discredited Nasserism in both Egypt and the broader Arab world. Egypt's defeat was so total and crushing, in fact, that, for a time, Nasser himself appeared to be in danger of losing power. Shortly before the war had ended, a group of generals had blamed him for the defeat and demanded that he step down immediately. Chastened, Nasser announced his

resignation in a televised speech to the stunned Egyptian people on the evening of June 9, 1967.¹

His retirement ultimately proved to be temporary, however. The regime and the ideology of Nasserism may have suffered black eyes as a result of the war, but its leader continued to retain the affection of the masses. Accordingly, hundreds of thousands of Egyptians poured into the streets of Cairo following his address and demanded that he remain Egypt's leader. The impact of these demonstrations was immediate. Nasser stayed in power while the officers who had so recently tried to remove him instead found themselves purged or, in some cases, killed.²

Defeat may have discredited Nasserism, but it did not weaken support for the radical economic policies or Arab nationalist views that lay at its core. Despite the enormous scale of the Arab states' failure against Israel, in fact, popular support for the core ideas of Nasserism—pan-Arabism, secular nationalism, and socialism—remained strong. As a result, rather than shifting to the right in response to Israel's triumph, the political center of gravity in much of the Arab world lurched sharply leftward in the years immediately following the Six Day War. Indeed, both civic groups such as Egypt's heretofore-Nasserist Organization of Socialist Youth and states such as the People's Democratic Republic of South Yemen (PDRY) alike either outwardly embraced more radical Marxist-Leninist forms of Arab nationalism following the Six Day War or adopted policies and methods consistent with them. Thus, ironically, while the defeat in 1967 destroyed Nasserism's credibility in the Arab world, it produced not a repudiation of its underlying ideas but instead an embrace of more extreme versions of them.³

Political Changes

The changes in government that occurred in the wake of the war reflected this paradox. Just as had happened after 1948, the Arab world experienced a series of coups following their defeat at Israeli hands. This time, however, the new regimes did not turn against the prevailing ideology of the day as their predecessors had, but instead doubled down on it. For example, Libya's youthful Lieutenant Muammar el Qaddafi (r. 1969-2011) and Sudan's Colonel Ja'far al-Numayri (r. 1969-1985) launched successful coups following the Six Day War that replaced moderate governments with openly radical ones. Meanwhile, in Iraq, a loose alliance of army officers and Ba'athists led by Ahmed al-Bakr (r. 1968-1979) overthrew President Abd al-Rahman Arif (r. 1966-1968) and his Nasserist prime minister, Tahir Yahya (1916-1986) and established a radical Ba'athist government. Finally, in Damascus, Defense Minister Hafiz al-Asad (r. 1970-2000) pushed President Salah Jadid (r. 1966-1970) from power in a bloodless

¹ Steven A. Cook, *The Struggle for Egypt: From Nasser to Tahrir Square*, Reprint edition (Oxford University Press, 2011), 92, 98–101.

² Benny Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920* (New York: Vintage Books, 2001), 344–45.

³ Fawaz A. Gerges, *Making the Arab World: Nasser, Qutb, and the Clash That Shaped the Middle East* (Princeton: Princeton University Press, 2018), 284–313.

coup in 1970. In every case but Syria—which retained the same Ba’athist regime under new leadership—the change of government marked a substantial shift to the left.⁴

Israel, the Palestinians, and the Occupied Territories

While the Arabs grappled with the implications of defeat, the Israelis instead dealt with the challenges of victory. The Israeli Defense Force’s (IDF) staggering success in 1967 presented the young country with a dilemma. On the one hand, the war had given it valuable strategic depth and had ended with it holding territories, particularly the West Bank, that growing numbers of Israelis argued should—for historical and religious reasons—be incorporated into Israel. On the other, retention of the captured land raised several thorny problems beyond the obvious one that seizing territory through conquest is illegal. First, while the Labor Party coalition that ruled Israel intended from the start to retain the symbolically vital city of Jerusalem and the strategically important Jordan River valley, it had endorsed Resolution 242 and had thus committed Israel to swap most of the land it had captured in the war in exchange for peace and diplomatic recognition. Second, more than one-million Palestinians lived in the territories that advocates of annexation most coveted: the Gaza Strip and, especially, the West Bank. Annexing that land into Israel would thus confront the government with the unpalatable choice of either imposing a costly military occupation or granting citizenship to the Palestinians who lived in those areas—a move that would risk the Jewish demographic dominance of Israel that was intrinsic to the Zionist project. Prime Minister Levi Eshkol (r. 1963-1969) neatly captured the crux of this dilemma in a remark to his successor, Golda Meir (r. 1969-1974), after the war. “The dowry pleases you,” he quipped, “but the bride does not.”⁵

While the cabinet wrestled with these issues, popular pressure for annexation and the construction of settlements in the Occupied Territories increased rapidly. Leading the charge was a new movement, Greater Land of Israel, that formed shortly after the war ended. Drawing support from adherents of Revisionist Zionism but also from a growing number of Laborites, it loudly demanded that the cabinet move quickly to annex the conquered territories. Greater Land of Israel did not merely lobby, moreover. Instead, it also took direct action to ensure permanent control of the land that Israel had conquered. Most notably, in parallel with the government’s establishment of strategic military settlements, its members began constructing illegal settlements in the West Bank and Golan Heights in an effort to “create facts” that would, by their very existence, prevent future Israeli cabinets from ever yielding those territories. Those

⁴ Albert Hourani, *A History of the Arab Peoples* (Cambridge, Mass: Belknap Press of Harvard University Press, 2002), 417; Charles Tripp, *A History of Iraq* (Cambridge: Cambridge University Press, 2002), 188–91.

⁵ Charles D. Smith, *Palestine and the Arab-Israeli Conflict: A History with Documents*, Sixth Edition (Boston: Bedford/St. Martin’s, 2007), 307–8; quote from Paul T. Chamberlin, *The Global Offensive: The United States, the Palestine Liberation Organization, and the Making of the Post-Cold War Order*, Reprint edition (Oxford University Press, 2012), 35.

wildcat settlements posed a serious challenge for Eshkol's government. Caught between its commitment to Resolution 242 and the popularity of Greater Land of Israel, his cabinet equivocated before ultimately agreeing to retroactively approve the illegal settlements—a move that served only to encourage the annexationists to begin planning further wildcat settlements.⁶

Meanwhile, much as it had altered political views in Israel, so had the conquest of the territories profoundly reshaped the outlook and expectations of the Palestinians. Since 1949, they had, with some notable exceptions, patiently and passively waited for Arab leaders such as Nasser to put their bellicose rhetoric into action and crush Israel. The Arab defeat in 1967 had made a mockery of those expectations. Rather than gaining self-rule as a result of the war, the one-million Palestinians who resided in the West Bank and Gaza Strip instead found themselves under an oppressive occupation regime in which the despised Zionists—who exercised total control of key resources such as water—strove to prevent the emergence of any kind of organized opposition or nationalist movement. It was a dispiriting turn of events to say the least. The Israeli conquest of the territories did not end the Palestinians' aspirations for political independence, however. Instead, Tel Aviv's harsh rule and seeming desire to once again displace them from their land renewed the Palestinians' sense of shared identity and rekindled their nationalist hopes—just as the Zionist movement had sparked the emergence of Palestinian nationalism during the first half of the twentieth century.⁷

Egypt after the Six Day War

Meanwhile, Egypt, the most powerful Arab state, found itself in crisis following the Six Day War. The crushing defeat had imposed vast costs on the country and had come at a time when the regime's domestic promises were beginning to reveal themselves to be every bit as disastrous as its foreign-policy. Nasser was the person most responsible for Egypt's increasingly bleak situation, but, thanks to his untimely death in 1970, he ultimately did not shoulder the onus of resolving its many problems. That task instead fell to his successor, Anwar Sadat, who, realizing that he had inherited a mess, moved aggressively to chart a dramatically new course for Egypt.

The War of Attrition

Nasser may not have overseen the resolution of Egypt's many problems, but he did set that process in motion through his effort to force a resolution of the conflict with Israel. The imbalance in power between Israel and the Arab states made this task a peculiarly difficult one. As we have seen, Israel's military preponderance was so total that it felt no pressure to make concessions, while the Arab states were so weak that they had nothing with which to bargain. As a result, while Tel Aviv accepted in principle the idea that it would eventually have to trade land for peace with the Arab states, it felt

⁶ Quote from Smith, *Palestine and the Arab-Israeli Conflict*, 307–8; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 335.

⁷ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 336–43.

no pressure to do so quickly. In fact, it was happy to continue to hold some of the Occupied Territories including the Sinai Peninsula indefinitely and was increasingly interested in retaining other parts—particularly the West Bank—in perpetuity.⁸

Faced with this situation, Nasser devised a plan designed to force Israel to negotiate through a limited military operation. Realizing that Egypt desperately needed to improve its diplomatic and fiscal positions, he first pursued a rapprochement with Saudi Arabia and its Gulf allies. The two sides quickly reached a deal. In exchange for Nasser withdrawing his troops from Yemen, the Kuwaiti and Saudi governments agreed to provide Egypt with an annual subsidy of \$266 million. Nasser next moved to reequip his devastated army. With Moscow embarrassed by the poor performance of its proxies' Soviet-equipped militaries, he was able to persuade the Kremlin to quickly replace the arms that Egypt had lost in 1967 with more advanced weapons including powerful surface-to-air missiles (SAMs); in exchange he agreed to grant the USSR substantial additional basing rights.⁹

With his fiscal situation stabilized and his military reequipped, Nasser followed by announcing in March 1969 the start of the War of Attrition: an intense but limited campaign in which Egyptian artillery bombarded the Bar-Lev Line, a defensive complex consisting of a massive sand berm and a series of heavily fortified strongpoints that the Israelis had constructed along the east bank of the canal. He launched this new conflict to achieve two aims. First, he hoped to make clear to Tel Aviv that its continued occupation of the Sinai would come at an unacceptable cost. Second, he sought to create a crisis that would impel the international community to force Israel to withdraw from the land it had taken in 1967.¹⁰

Nasser's effort to compel negotiations through a limited-war strategy proved to be a failure. The international community did not intervene, and Israel responded with a retaliatory air offensive that not only mauled Egypt's anti-aircraft defenses but also—to Nasser's great embarrassment—struck targets deep inside the country. The Egyptian leader refused to back down, however. Desperate to restore the military balance, he asked the Kremlin in January 1970 to provide yet-more weapons that would give Egypt the strength to end Israel's dominance of the air. After initially balking, Moscow agreed to supply Egypt with squadrons of MiG-21 fighters and large numbers of advanced SAM missiles; more importantly, it also quietly sent 200 pilots to fly the jets and thousands of Soviet technicians to operate and support the anti-aircraft batteries. Reinforced in this way, the Egyptian military was able to extend its SAM umbrella slowly and at great cost eastward into the Canal Zone over the spring and summer of 1970. At that point Washington intervened. Fearing further escalation and the possibility of a superpower

⁸ Kirsten E. Schulze, *The Arab-Israeli Conflict* (New York: Pearson, 2008), 38–39.

⁹ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 344.

¹⁰ Smith, *Palestine and the Arab-Israeli Conflict*, 316–17.

conflict, it succeeded in compelling both sides to accept a ceasefire under the auspices of an otherwise still-born peace initiative.¹¹

The War of Attrition was at best a mixed success for Egypt. It is true that the fighting had succeeded in compelling the international community to pay attention to the Arab-Israeli conflict and had allowed Egypt to extend its now much-more-powerful SAM network into the Canal Zone—a situation that put Nasser's armed forces in a significantly better strategic position. At the same time, however, the improvement in the military balance hardly justified the destruction of canal-side cities such as Isma'iliyya in the fighting or, more importantly, the deaths of an estimated 10,000 Egyptian soldiers and civilians.¹²

Economic Problems

The cost of the War of Attrition also contributed substantially to a troubling economic slowdown. The stagnation marked a significant turnabout from the early part of the decade. Following the adoption of Arab Socialism in the early 1960s, Egypt had enjoyed an impressive economic expansion—one that had not only produced a 30 percent increase in per-capita income but that had also helped fund expanded welfare benefits for industrial workers and underwrote a new and more-aggressive program of land reform for the country's peasants. To the delight of the middle class, Nasser had even guaranteed government employment for all college graduates who could not otherwise find work.¹³

By 1965, however, serious cracks had begun to appear. The reliance on foreign borrowing and imported commodities had created significant balance-of-payments issues that left Egypt struggling to service its debt. By the mid 1960s, as a consequence, economic growth had stalled and the poverty rate had begun to climb. Unsurprisingly the Six Day War exacerbated these problems. The conflict itself had not only imposed a substantial fiscal drain on Egypt but had also deprived it of its three main sources of foreign currency—exports from its Sinai oil fields, toll revenue from the now-closed Suez Canal, and tourism. Meanwhile, though Nasser had long-since learned that the state did not have anywhere close to enough jobs to meet the employment guarantee that he had rashly given to college graduates in 1962, he had neither modified nor abandoned his promise. Instead, he had responded by having the government provide them with make-work assignments. By 1970, as a result, his

¹¹ Peter L. Hahn, *Crisis and Crossfire: The United States and the Middle East Since 1945* (Washington, D.C.: Potomac Books Inc., 2005), 55–56.

¹² Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 360–63.

¹³ Gerages, *Making the Arab World*, 288.

employment pledge had left Egypt saddled with a bloated and inefficient bureaucracy that further dragged on its economy.¹⁴

The country's worsening economic situation created a serious dilemma for Nasser. Since 1956, his regime had based its legitimacy on a combination of a rising standard of living and the international prestige that it had enjoyed as a result of his role as the leading anti-Zionist and Arab-nationalist figure. By 1970, these competing agendas had trapped him in a variation of the classic guns-versus-butter quandary. If he moved to ease his country's economic problems by abandoning his role as the leader of the pan-Arab struggle against Israel, he risked losing the support of those Egyptians who demanded justice for the Palestinians. If he instead pressed on with the struggle against Israel, he would impose new burdens on Egypt's already faltering socialist economy that would lower living standards and thus alienate the masses.

Anwar Sadat and the "Corrective Revolution"

As earlier noted, it was not Nasser who would resolve this predicament. Instead, thanks to the Egyptian leader's sudden death from a heart attack in September 1970, the task of finding a way out of the trap into which Nasser had led Egypt fell to his vice president and successor, Anwar Sadat. The son of a minor bureaucrat, Sadat was a longtime Arab nationalist and leading member of the Free Officers who, like Nasser, had joined the army after the Wafd Party had opened the military academy to members of the lower-middle class. Considered a political lightweight, he appeared to be an unlikely candidate to resolve the dilemma that Nasser's policies had created. In fact, he had managed to secure the influential Arab Socialist Union's (ASU) pivotal endorsement of his succession to the presidency only by promising the entrenched and powerful political figures who dominated it that he would continue his predecessor's policies and, critically, that he would oversee a collective system of rule in which he would be no more than a first-among equals.¹⁵

It soon became clear, however, that Sadat's political rivals had badly underestimated him. A shrewd operator, the new president deftly maneuvered them out of power in late 1970 and early 1971 by rallying to his side lower-ranking officers, mid-tier bureaucrats, leftists, the police, and others who loathed the ASU leadership for its part in the debacle in June 1967. He followed with what he called the "corrective revolution": a thorough restructuring of the government and the ruling party. Vast in scope, it included the extension of limited civil liberties, new elections for the ASU and the National Assembly designed to purge those bodies of extreme Nasserites, and—in acknowledgement of the reality that Arab unification schemes were dead—a change in the country's name from the United Arab Republic to the Arab Republic of Egypt. Thus, dismissed by many in 1970 as little more than a functionary, Sadat had succeeded in

¹⁴ Arthur Goldschmidt Jr., *Modern Egypt: The Formation of a Nation State* (New York: Routledge, 2004), 171, 180; Eugene Rogan, *The Arabs: A History – Revised and Updated Edition* (New York: Basic Books, 2012), 264.

¹⁵ Cook, *The Struggle for Egypt*, 113–18.

eliminating all of his challengers and had consolidated power in his hands by mid 1971.¹⁶

In spite of his political victory, however, Sadat was in no position to rest on his laurels. Instead, facing the same dilemma that had trapped his predecessor during his final years in power, he appeared to be caught between two equally ruinous options. On the one hand, he could seek to revive the economy by ending the expensive mobilization of the Egyptian army—a move that would effectively cede permanent control of the Occupied Territories and the Sinai to Israel and thus turn Egyptian nationalists against him. On the other, he could continue Nasser's aggressive military posturing in hopes that it would force Israel to yield the lands it had acquired in 1967—an approach that would further contribute to the country's declining economic prospects and therefore alienate the masses. It was a daunting and dangerous situation for Sadat, one that would almost assuredly cost him his position as Egypt's president if he failed to quickly find favorable resolution of it.¹⁷

Transforming the Arab-Israeli Conflict: The Yom Kippur War

Year of Decision, 1971

A shrewd statesman, Sadat opted for an innovative solution that he hoped would, by compelling Tel Aviv to engage in productive talks, free him from the binary choice of either saving the Egyptian economy or contesting Israel. It centered on improving relations with the United States. Grasping that Washington held the key to resolving both his country's domestic and foreign-policy problems, he would first, with American assistance, bring the costly contest with Israel to an end through a negotiated peace settlement that would return the Sinai and—hopefully—the Occupied Territories to Arab rule in exchange for peace and recognition. Freed from the expense of the conflict with Israel and now receiving American foreign aid thanks to its improved ties with Washington, Cairo would follow by restructuring and reviving Egypt's moribund economy through a liberalization program designed to restore growth and prosperity.¹⁸

Sadat launched his initiative in February 1971. It called for Israel to withdraw its troops forty kilometers east of the Suez Canal; in exchange, Cairo would extend the existing ceasefire with Tel Aviv, grant Israeli ships the right to transit the Suez Canal, and accept the return of peacekeepers to the Sinai. US-mediated peace talks aimed at arriving at a final settlement on the basis of Resolution 242 would follow. It was a remarkable proposal—one that constituted a dramatic break with the “three nos” of the Khartoum Resolution. Nonetheless, it went nowhere. Golda Meier was unwilling to cede any territory without a full peace treaty, while US President Richard Nixon (r. 1969-1974) was leery of applying pressure to Israel lest he alienate the Jewish vote just

¹⁶ Cook, 119–23.

¹⁷ Arthur Goldschmidt, *A Brief History of Egypt* (New York: Checkmark Books, 2008), 190; Cook, *The Struggle for Egypt*, 127.

¹⁸ Schulze, *The Arab-Israeli Conflict*, 41–42; Rogan, *The Arabs*, 364–66.

before the upcoming presidential election. As a result, Tel Aviv was able to fend off Sadat's proposal by insisting on unacceptable conditions such as Israel's retention of Sharm al-Shaykh¹⁹

Frustrated, Sadat followed with an uncharacteristically amateurish move. Calling 1971 the "year of decision," he insisted that Egypt was no longer willing to tolerate the "no war, no peace" situation that had prevailed since the end of the War of Attrition and would return to fighting if no progress was made toward a resolution of the Arab-Israeli conflict. It was an obvious bluff and a poorly executed one at that. Golda Meir's government knew that Egypt was not prepared for war and left Sadat in the lurch by refusing to respond. Embarrassed, the Egyptian leader was only able to extricate himself by lamely claiming that the Indo-Pakistani War that had broken out in December 1971 had created too much international uncertainty for Egypt to resume fighting. The episode did little to dispel doubts about Sadat's abilities. Internationally, the Israelis and Americans concluded that he was a weak and impetuous ruler and no longer took his pronouncements seriously. Domestically, his failure to back his words with actions had resulted in widespread demonstrations against his regime by frustrated nationalists. As a result, the pressure on Sadat had ratcheted up considerably. If he wanted to remain Egypt's leader, he needed to produce tangible results—and he needed to do so quickly.²⁰

The Road to War

Sadat demonstrated far greater skill in managing his next attempt to break the impasse. Having failed to lure Tel Aviv to the negotiating table through a peace proposal, he shifted gears and began preparing to go to war. The conflict he sought was an unusual one, however. Sadat did not hope to destroy Israel or even to achieve a decisive victory. Instead, he intended to wage a limited conflict aimed only at securing a modest amount of territory on the east bank of the canal. Doing so, he believed, was all that was needed to set the stage for genuine peace discussions. It would shake Israeli confidence, compel Washington to help him negotiate with Tel Aviv, and end the psychological and military imbalance that had heretofore prevented the two sides from engaging in productive talks. As National Security Advisor Henry Kissinger (1923-) later wrote, "Sadat aimed not so much for territorial gain but for a crisis that would alter the attitudes into which the parties were frozen—and thereby open the way for negotiations."²¹

He first moved to prepare Egypt for the war by breaking with the USSR. Expressing anger about the gratingly imperious attitude of Soviet troops in Egypt and about the Kremlin's foot dragging on arms deliveries, he abruptly ordered the expulsion

¹⁹ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 388–91.

²⁰ Cook, *The Struggle for Egypt*, 127–29; quotes from Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 390–91.

²¹ Cook, *The Struggle for Egypt*, 131; quote from Henry Kissinger, *Years of Upheaval* (Boston: Little Brown & Company, 1982), 460.

of 15,000 Soviet advisors in July 1972. Most observers at the time viewed this move as a rash action—a case, in Israeli Foreign Minister Abba Eban’s (1915-2002) words, of Sadat taking “an emotional satisfaction at the expense of his strategic and political power.” In fact, it was a shrewd maneuver that paid manifold dividends to the Egyptian leader. First, by suggesting that Cairo was seeking to end its close relationship with the USSR, it led the Nixon administration to reconsider its relationship with Egypt and thus set the stage for Washington to adopt a more even-handed approach to the Arab-Israeli conflict. Second, Sadat’s move had the ironic effect of compelling Moscow to change course regarding the provision of weapons to Egypt. Fearing that it might lose access to the valuable Egyptian bases from which its ships and aircraft tracked American missile submarines operating in the Mediterranean, Moscow quickly agreed to provide Sadat with sophisticated new weapons. Accordingly, during late 1972 and early 1973, the Egyptian military received a huge infusion of new arms including MiG-23 jet fighters, SU-20 attack aircraft, advanced SAM batteries, late-model tanks, and powerful wire-guided Sagger antitank missiles. Finally, the eviction of Soviet advisors helped create a sense of complacency in Tel Aviv. Coming so soon after Sadat’s failure to follow through with his threat during the “year of decision,” the rupture between Cairo and Moscow led Israeli intelligence officials to conclude that Egypt was in no position to go to war in 1973.²²

Having secured new weapons and having established the foundation for improved relations with Washington, Sadat next moved to recruit allies in the Arab world. He began with Syria. Knowing that Asad was determined to regain control of the Golan Heights as he was to retake the Sinai, Sadat struck an agreement with the Syrian leader in April 1973 to launch a coordinated, two-front attack on Israel designed to apply maximum pressure on the IDF. He followed by securing Saudi support. Meeting with King Faisal bin Abdulaziz Al Saud (r. 1964-1975) in August 1973, he asked that Saudi Arabia support his military assault with a coordinated economic campaign designed to pressure Israel’s Western backers—above all the US—by cutting the export of Saudi petroleum to those countries after the war had begun. As we have seen, the oil weapon had proven ineffective in 1967; by 1973, however, the balance between consumer and producer states had swung drastically in favor of the latter. Most significantly, Saudi Arabia had supplanted the US as the world’s swing producer and, as such, was in a position to unleash the oil weapon with devastating effect against the suddenly vulnerable West. Well aware that the balance of power had shifted in favor of the producer states, Faisal, assured Sadat of Saudi support and promised not only to deploy the oil weapon once the war began, but, in the meantime, to provide Egypt with the enormous sum of \$500 million to help it prepare for the conflict.²³

²² Eban quoted in Avi Shlaim, *The Iron Wall: Israel and the Arab World, Revised and Expanded* (New York: W. W. Norton, 2014), 318–19; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 391; Rogan, *The Arabs*, 366.

²³ Daniel Yergin, *The Prize: The Epic Quest for Oil, Money & Power* (New York: Touchstone Books, 1991), 593–98.

With Riyadh now on board, the Egyptians and Syrians hurried to complete their planning. They decided to launch the attack on Saturday, October 6, 1973, the same date as Yom Kippur, the Jewish Day of Atonement. They chose the holiday for two reasons. First, since there would be no moonlight after midnight on that date, Egyptian engineers would be able to bring bridging equipment up to the canal undetected. Second, since radio and television stations in Israel were closed on High Holy Days, the IDF would have difficulty mobilizing reservists; as a result, Egypt would be in a position to deploy substantial forces across the canal before they faced the inevitable counterattack.²⁴

Despite close coordination in the run up to the war, however, Egypt and Syria had in fact adopted very different strategies aimed at achieving decidedly different objectives. From the start, Sadat had intended to wage a cautious, limited campaign designed to break the impasse that had obtained since the end of the Six Day War. His generals had consequently developed a plan in which Egyptian troops would cross the canal and take up defensive positions ten kilometers to the east—an approach that would achieve Sadat's goal of shattering the status quo but that would also keep Egyptian ground forces safely under the cover of their powerful-but-immobile SAM umbrella. This strategy was a sound one. It was not, however, likely to entice Damascus to join Egypt in attacking Israel. As Sadat knew, Asad wanted to retake the entire Golan Heights—an objective that the Syrian army could only achieve if Egypt was committed to an all-out offensive. The Egyptian president consequently kept quiet about his true intentions, and instead told his Syrian counterpart in April 1973 that he intended to retake *all* of Egypt's lost territory and that his army would continue its offensive deep into the Sinai. Accordingly, believing that he was on the same page as Sadat, Asad ordered his generals to devise a strategy that called for the Syrian army to push forward aggressively based on the erroneous assumption that Egypt's drive into the Sinai would tie down a large portion of the Israeli army. This difference in strategy was critical, and it would have an enormous effect on the war and its outcome.²⁵

The Yom Kippur War

To the Israelis' shock, the Yom Kippur War, also known as the October War, began very differently than had any of the prior four Arab-Israeli conflicts. This time it was the Arabs rather than the Israelis who caught the enemy napping and who, at first at least, flawlessly executed a series of complex and well-planned military operations. In the north, the Syrians mounted a tactically sound offensive that scored significant gains and nearly unhinged Israel's defensive line in the Golan Heights. Egyptian troops performed even more impressively in the south. After crossing the canal in dinghies and quickly erecting pontoon bridges, engineers used 450 water cannons to blast passages for troops and vehicles in the sand berm that the Israelis had erected. As a result, the Egyptian army was able to push its way through the purportedly impregnable Bar-Lev

²⁴ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 391.

²⁵ Salim Yaqub, *Imperfect Strangers: Americans, Arabs, and U.S.–Middle East Relations in the 1970s* (Cornell: Cornell University Press, 2016), 117.

Line with little difficulty. Indeed, by the evening of October 7, 100,000 troops, 13,500 vehicles, and more than 1,000 tanks had taken up defensive positions ten kilometers east of the canal. Tel Aviv bordered on panic. Already surprised by the ease with which Sadat's troops had broken through the first line of defense, Israeli leaders were positively stunned when the Egyptians easily swatted away a concerted counterattack by two IDF armored divisions in an engagement that cost the Israelis eighty tanks—almost as many as they had lost over the course of the entire Six Day War.²⁶

As the losses mounted, Israeli leaders became increasingly concerned that the Arab attack constituted an existential threat. Emblematic of this fear was the change in the demeanor of the country's normally confident defense minister, Moshe Dayan (1915-1981). Expressing panic for Israel's survival during the early days of the conflict, he at one point histrionically declared that "the third temple [meaning Israel] is going under." Meanwhile, shocked to learn that the Arabs had knocked out 500 tanks in the first three days of fighting, Prime Minister Golda Meir begged Washington to send desperately needed equipment, supplies, and, especially, ammunition. Now serving as secretary of state, Kissinger initially demurred. Assured by Sadat that Egypt's goals were limited, he was leery of needlessly antagonizing the Arab oil states and believed that a stalemate or even a minor defeat would leave Israel more tractable and thus willing to make the concessions on which a durable peace depended.²⁷

The commencement of a huge Soviet effort to resupply Egypt and Syria on October 10 quickly changed his attitude, however. Symbolically turning the conflict into a superpower proxy war, the Soviet move impelled Nixon and Kissinger to take steps to ensure Israel's victory. Accordingly, beginning on October 14, massive US Air Force C5-A Galaxy transport planes laden with supplies, tanks, and ammunition began landing in Israel. Nixon soon went even further. On October 19, he requested that congress provide Israel with an additional \$2.2 billion of military aid—an amount greater than the total sum Israel had spent on defense in 1972.²⁸

Bolstered by the aid, the now fully mobilized IDF soon began to turn the tide. Realizing after a few days that the Egyptians were not going to advance beyond their protective SAM umbrella, the Israelis concentrated their forces for a blistering counterattack against Asad's overextended army. Driving the Syrians back nearly to Damascus, the attack not only effectively knocked Syria out of the war but, as important, also indirectly set the stage for Israeli success against Egypt. Key here was the intervention of the USSR. As the IDF counteroffensive in the Golan was unfolding, Moscow began frantically pushing Sadat to order a new advance into the Sinai that would relieve pressure on Asad's forces. The Egyptian leader initially refused, but, facing insistent Soviet demands, soon relented. Accordingly, on October 14, Egyptian

²⁶ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 411–19.

²⁷ Yaqub, *Imperfect Strangers*, 131–34; quote from Yergin, *The Prize*, 604.

²⁸ Douglas Little, "Gideon's Band: America and the Middle East since 1945," *Diplomatic History* 18, no. 4 (1994): 242–43; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 434.

forces started advancing further into the Sinai—moving well beyond not just the canal but also, critically, the SAM umbrella that protected them from Israel’s powerful air force.²⁹

The offensive proved to be a disastrous mistake. Not only was the attack costly—the IDF knocked out 250 Egyptian tanks in just a few hours—but it also completely disarranged the carefully constructed defensive position that Sadat’s soldiers had established near the canal. The Israelis were quick to take advantage. After breaking through to the east bank of the waterway on October 15, IDF troops stunned the Egyptians by securing a bridgehead on the west side of the canal. After briefly consolidating their position, they began pushing south in a drive designed to cut the Egyptian Third Army off from its supply lines.³⁰

It was at that juncture that the Arab states and Iran unleashed the oil weapon. Meeting in Kuwait City, representatives of Iran and five Arab oil-producing states agreed to raise the price of a barrel of petroleum by a jaw-dropping 70 percent. Ramping up the pressure, they announced the following day that they would immediately cut production by 5 percent and would continue to cut it by 5 percent each month “until ‘the total evacuation of Israeli forces from all Arab territory.’” They were not finished. After learning of Nixon’s request for \$2.2 billion in military aid for Israel, Riyadh and the other Arab oil-producing states announced on October 20 that they were immediately embargoing the sale of petroleum to the US.³¹

Thanks in part to the oil weapon, the war soon came to an end. On October 22, the United Nations (UN) Security Council moved to bring the fighting to a stop through the passage of Resolution 338. It called for a ceasefire, immediate implementation of Resolution 242, and talks “aimed at establishing a just and durable peace.” Nonetheless, military operations did not immediately cease. Instead, Israel’s effort to improve its negotiating position by completing the encirclement of Egypt’s Third Army extended the war for two more days. Indeed, it was only after the IDF’s maneuvers led the superpowers to engage in a brief bit of nuclear brinksmanship in support of their regional proxies that the fighting finally halted.³²

The Yom Kippur War was enormously consequential for the Middle East and greatly redressed the imbalance in power between Israel and the Arab states that had precluded productive negotiations after the Six Day War. Militarily, it is true, Israel had bounced back from its early setbacks and was in a stronger position vis-à-vis Egypt and Syria than it had been on the morning of October 6. At the same time, however, its near

²⁹ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 407–11.

³⁰ Yaqub, *Imperfect Strangers*, 133; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 407–11, 423–24.

³¹ Yergin, *The Prize*, 606–9.

³² Quote from “United Nations Resolution 338,” October 22, 1973; Daniel Sargent, *A Superpower Transformed: The Remaking of American Foreign Relations in the 1970s* (Oxford: Oxford University Press, 2015), 150–51.

defeat and its serious losses—Israel suffered 2,838 dead and 8,800 wounded, and it lost 840 tanks and 103 aircraft—had destroyed the hubris and sense of military invulnerability that had gripped the Jewish state since 1967 and had left a stunned Israeli government receptive to negotiations. The war was thus a major triumph for Sadat. Hailed thereafter as the “Hero of the Crossing,” he had succeeded in establishing a very different diplomatic climate in the region than had existed since the Six Day War and, in so doing, had—for the first time—created the proper conditions for Israel and the Arab states to finally begin to negotiate a resolution of the region’s most bitter and enduring conflict.³³

Transforming the Arab-Israeli Conflict: The PLO

Meanwhile, at the same time that the Egyptian government was groping for a way out of the costly Arab-Israeli conflict, the Palestinians were moving to once more take charge of the struggle with Israel. That effort began almost before the Six Day War had ended. Fed up with the Arab states’ inability to restore them to their land, Palestinian militant groups initiated a new campaign of terrorist and insurgent attacks beginning in 1967. Though their efforts failed to meaningfully harm Israel, they nonetheless had important consequences. They restored hope to the Palestinian people, put their national aspirations back on the front burner internationally, and won substantial diplomatic support for their cause. Above, all they proved pivotal in helping to turn what had been, since 1948, an interstate dispute back into a conflict between competing nationalisms.

The Palestine Liberation Organization and Fatah

For the Palestinians, the years between 1949 and 1967 had been dark ones. Expelled from their homes in 1948, the majority lived during that period in crowded United Nations Relief and Works Agency (UNRWA) refugee camps in Lebanon, Syria, the Egyptian-administered Gaza Strip, and the Jordanian-controlled West Bank. As we saw in chapter fourteen, the Palestinians were not treated well by the governments of those states. Denied the right to citizenship or employment everywhere but Jordan, they led a meager existence as stateless refugees—unable to meaningfully improve their circumstances let alone recover their lost homes. Too weak to challenge Israel, they placed their hopes for redemption in the hands of pan-Arabist leaders such as Nasser who, knowing that tough talk toward Israel played well with the masses, launched frequent broadsides against the Jewish state and promised to destroy it when the time was right.³⁴

The Palestinians were not completely inert between 1949 and 1967, however. In the late 1950s and early 1960s, they had established a number of competing organizations dedicated to liberating Palestine. The most important of these was the Palestine Liberation Organization (PLO). It originated out of a call Nasser had made in

³³ Schulze, *The Arab-Israeli Conflict*, 45–46; Shlaim, *The Iron Wall*, 324–26.

³⁴ Smith, *Palestine and the Arab-Israeli Conflict*, 230–31.

the early 1960s for a new body to coordinate the Palestinians' resistance to Israel. At his urgings, the Palestine National Council (PNC)—which functioned as a legislature in exile—met in May 1964 to establish the new organization and to draft its covenant, the Palestine National Charter. Uncompromising in tone and objectives, that document forthrightly called both for the destruction of Israel and for the establishment of a Palestinian state in the entirety of the former mandate.³⁵

Despite a great deal of sound and fury, however, the charter amounted at the time to little more than empty words. It was Nasser rather than the Palestinians who controlled the PLO, and he had proposed its establishment not to create an organization that could lead an armed insurgency against the Jewish state but merely to build the impression that he was actively supporting Palestinian efforts to liberate their homeland and, thus, to rebut intense criticism from his Arab rivals for his failure to back his anti-Israeli rhetoric with action. He certainly had no intention of inviting IDF reprisals by permitting the PLO to raid Israel from Egyptian territory; in fact, he deliberately arranged for a passive functionary, the Lebanese-born Palestinian diplomat Ahmad Shuqayri (1908-1980), to serve as the first chairperson of the organization, and he secured a pledge from Shuqayri that the PLO would—under no circumstances whatsoever—undertake an attack on Israel from Egyptian-controlled territory.³⁶

In contrast, a competing organization called Fatah was forthright in its commitment to armed struggle. Founded in Kuwait in 1959 by a group of Palestinians headed by Yasir Arafat (1929-2004), it eschewed a commitment to pan-Arabism or Marxism in favor of a non-ideological—albeit leftist—approach. As such, it did not concern itself with the broader pan-Arabist effort to unify the Arab world in the early 1960s and instead focused on liberating Palestine through the same violent guerrilla tactics that anticolonial organizations had successfully employed in places like Vietnam and Indonesia. Enjoying the support of the Syrian government, which at the time was seeking to displace Egypt as the leader of the anti-Zionist cause, Fatah's guerilla soldiers, or *fedayeen*, began launching raids against Israel from Syrian and Jordanian-controlled territory in January 1965. Though those operations were almost universally unsuccessful, they compared favorably with the unwillingness of Arab leaders like Nasser to match their tough rhetoric toward Israel with action; as a result, the organization quickly won the support of many Palestinians.³⁷

Ultimately, however, it was the events that followed the Six Day War rather than these ineffectual raids that vaulted Fatah to prominence. The organization initially responded to the occupation of the West Bank by undertaking insurgent operations against the occupying forces. Believing that it could spark a spontaneous “popular

³⁵ Schulze, *The Arab-Israeli Conflict*, 33.

³⁶ James L. Gelvin, *The Israel-Palestine Conflict: One Hundred Years of War* (New York: Cambridge University Press, 2005), 198–99.

³⁷ Rashid Khalidi, *The Hundred Years' War on Palestine: A History of Settler Colonialism and Resistance, 1917–2017* (New York: Metropolitan Books, 2020), 114–15.

rebellion” among the Palestinians of the Occupied Territories, it began with an “interior” campaign in the late summer and fall of 1967 in which *fedayeen* launched attacks from within the occupied West Bank. As poorly executed as it was organized, the operation soon collapsed in the face of Israel’s brutally efficient counterinsurgency tactics. Fatah was undeterred, however, and responded not by ceasing its attacks but instead by shifting to an “exterior” strategy in which guerrillas raided Israeli-controlled territory from bases in Jordan. Unfortunately for Fatah, these attacks were only marginally more effective than the “interior” campaign had been. They proved enormously consequential, however, because they led the IDF to mount a fateful retaliatory attack against the Jordanian village of Karameh on March 21, 1968—an engagement that would fundamentally change Fatah’s standing in the eyes of the Palestinian people.³⁸

Karameh, March 1968

Launched in response to Fatah incursions, the attack on Karameh aimed to destroy the organization’s forward headquarters. The Israelis anticipated that the raid would follow the pattern of earlier punitive operations: IDF soldiers would seize control of the village, kill a few guerrillas, and demolish a large number of houses before leisurely returning to Israel. This time, however, the operation did not follow the usual pattern. Instead, knowing that they enjoyed the support of Jordanian artillery and tanks dug in on the heights overlooking the village, the *fedayeen* decided to stand their ground. As a result, the IDF troops that approached Karameh found themselves in an unexpectedly serious firefight with tough, disciplined Fatah militants. Though they did manage to push their way into the village, the Israelis incurred serious losses—thirty-three dead, 161 wounded, four tanks destroyed, and a fighter-bomber shot down—and ended up retreating in disarray back across the Jordan [river](#).³⁹

The battle was a watershed moment for Fatah that benefitted the organization in a number of ways. First, thanks to its success against the IDF—and to King Hussein’s (r. 1952-1999) willingness to let the organization take full credit for the victory—Fatah was able to establish itself as the leading organization in the broader Palestinian resistance movement. Recruits thereafter flocked to its banner, and it soon enjoyed the strength to impose its control over the refugee camps near Amman. Second, Fatah was able to translate its military success into significant financial support from the increasingly wealthy Gulf principalities—funding that gave it a degree of independence from states like Egypt and Syria that had sought to control the Palestinian militant groups. Finally, Karameh helped win the diplomatic backing of scores of newly independent states in the Developing World that saw in the Palestinian struggle echoes of their own recent anticolonial efforts. As we shall see, their assistance would prove

³⁸ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 365–68.

³⁹ Morris, 368–70.

invaluable in the campaign to secure broader international recognition of the Palestinian cause in the 1970s.⁴⁰

Karamah also permitted Arafat to assume a dominant position within the PLO. By early 1968, that organization was in serious disarray. Espousing restraint at a time when Palestinians were pushing for direct action, it appeared to be out of touch and risked collapsing into irrelevance. Indeed, members of the PLO's executive board gave serious consideration to dissolving the organization in the wake of the Six Day War. Ultimately, however, they rejected that course of action. Instead, aware that the PLO's founding under the auspices of the PNC lent it a degree of legitimacy that other resistance groups lacked, executive committee members moved to remake the PLO into a genuinely independent, Palestinian-controlled umbrella organization. In short order, they forced the ineffectual Shuqayri to resign, established the PLO's autonomy from Nasser, and invited militant groups like Fatah to join. They also opted to assign seats on the revamped executive committee to the different Palestinian resistance groups on the basis of their size.⁴¹

This last move was enormously beneficial to Fatah. Enjoying broad support among Palestinians thanks to its victory at Karamah, it assumed a controlling position in the PLO and secured Arafat's election as the organization's new chair—a post he would hold continuously from February 1969 until his death in 2004. Arafat and the other Fatah leaders wasted little time in bringing the PLO's goals and methods in line with their more explicitly militaristic approach. Most notably, they successfully amended the Palestine National Charter so that it called for the liberation of Palestine through “armed struggle,” and they declared the organization's primary goal to be the establishment of a secular Palestinian state in the entirety of the former mandate.⁴²

Splinter Groups

Fatah was not the only Palestinian militant group to gain prominence following the Six Day War, however. Beginning in 1968, a number of smaller, newly created splinter groups also won followings by launching guerrilla and terrorist attacks on Israel. Motivated by the same far-left ideologies that were sweeping through the Arab states following the decline of Nasserism, organizations such as the Palestinian Popular Struggle Front (PPSF), the Iraqi-dominated Arab Liberation Front (ALF), George Habash's (1926-2008) Popular Front for the Liberation of Palestine (PFLP) and its offshoots, the Popular Democratic Front for the Liberation of Palestine (PDFLP) and the Popular Front for the Liberation of Palestine – General Command (PFLP-GC), took a far-more extreme and uncompromising line than did the more ideologically flexible Fatah. Those groups also adopted a more transnational outlook and strategy than did Arafat's organization. Most notably, committed as they were to both Marxism and pan-

⁴⁰ Chamberlin, *The Global Offensive*, 46–49; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 369–70.

⁴¹ Gelvin, *The Israel-Palestine Conflict*, 313–14.

⁴² Gelvin, 199; Smith, *Palestine and the Arab-Israeli Conflict*, 313–14.

Arabism, they believed that the radicalization and unification of the Arab states had to occur before the militants could defeat Israel and redeem Palestine.⁴³

Joining the PLO shortly after Fatah, these groups greatly complicated Arafat's ability to direct the organization. In hopes of precluding internecine conflict, he had from the start adopted an inclusive approach wherein the constituent organizations would resolve their differences through consensus, compromise, and negotiation rather than through conflict. Unfortunately for Arafat, the PFLP and PDFLP—soon renamed the Democratic Front for the Liberation of Palestine (DFLP)—were ideologically inflexible and thus disinclined to follow his consensual approach. Indeed, they not only engaged in factional infighting within the PLO, but, more importantly, used Arafat's commitment to consensus and compromise to box him in. Insisting that the organization maintain a hardline approach, for example, they prevented the PLO chair from taking advantage of passing opportunities and they delayed vital policy changes such as a shift from a commitment to Israel's destruction to a willingness to accept a Palestinian state confined to the Gaza Strip and West Bank.⁴⁴

Particularly problematic for Arafat were the radical groups' terrorist operations. Too small to mount effective attacks across the heavily defended border between Jordan and the West Bank, the PFLP and DFLP instead focused on high-profile operations against soft targets abroad. The PFLP initiated this new approach in July 1968 when it hijacked an El Al airliner and successfully exchanged the hostages it had taken for fifteen Palestinian prisoners. Wildly successful in drawing world attention to the Palestinian cause, the operation spurred the organization to follow with further hijackings in 1969. Meanwhile, in February 1970, the smaller PFLP-GC joined the campaign in spectacular fashion by detonating a bomb on a Swissair flight that killed all forty-seven people on board.⁴⁵

Arafat had been unenthusiastic about such operations through that point but could accept them as a modest price to pay for unity within the PLO. The PFLP's terror attacks in Jordan in 1970 were a different matter, however. These operations emerged out of the splinter organizations' ideological commitment to effecting radical political change in the Arab world as a first step in their campaign against Israel. Declaring "that the road to Jerusalem begins in Amman," the PFLP and DPFLP maintained that they had to first replace King Hussein's conservative government with a radical, Palestinian one before they could mount an effective challenge to the Jewish state. Accordingly, the splinter groups moved to spark a Palestinian revolution in Jordan by launching a series of terrorist attacks designed to embarrass and weaken Hussein's government. These included two assassination attempts against the king, the seizure of thirty-three Western hostages at popular tourist hotels in June 1970, and, most spectacularly, the destruction

⁴³ Smith, *Palestine and the Arab-Israeli Conflict*, 313.

⁴⁴ Gelvin, *The Israel-Palestine Conflict*, 204–8.

⁴⁵ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 376–78.

of three hijacked Western airliners at a derelict RAF airfield just outside Amman in September 1970.⁴⁶

Black September, September 1970

The hijacking incident was the final straw for Hussein. Furious both that the militant groups were challenging his rule and that Arafat had established what amounted to an autonomous Palestinian state in Jordan replete with tax collectors, courts, and uniformed military personnel, he ordered his army to move against PLO forces in the camps shortly after the PFLP's destruction of the three airliners. In a civil conflict known as Black September, Hussein's troops handily crushed Arafat's forces and seized control of the camps in and around Amman—killing an estimated 3,500 Palestinians in the process. A new offensive pushed the organization out of Jordan altogether the following year. Transiting through Syria, PLO guerrillas set up a new base of operations in Lebanon in the fall of 1971. Taking full advantage of an autonomy deal that Nasser had brokered in 1969 to ease a growing conflict between the PLO guerrillas and the Lebanese government, Arafat quickly established what amounted to a Palestinian state-within-a-state in southern Lebanon and began organizing attacks across the border with Israel.⁴⁷

The PLO's Global Offensive

Shortly after its relocation to Lebanon, however, the PLO largely abandoned its efforts to bring the fight directly to Israel. Two factors produced this change. First, at a length of only 81 kilometers, Israel's border with Lebanon is considerably shorter than the frontier that divides it and the occupied West Bank from Jordan and was thus much more difficult to cross undetected. Second, the PLO had never managed to devise a set of tactics for combating the overwhelmingly powerful IDF. As a result, while the organization continued to mount the occasional *fedayeen* raid across the border and to lob shells into Israel, it shifted in 1971 to a new approach in its effort to secure a state. Thereafter, it pursued what the historian Paul Thomas Chamberlin calls the "global offensive"—a strategy designed to draw international attention to the Palestinians' situation and to build support for their cause across the world.⁴⁸

The PLO waged this campaign on two fronts: one ideological, the other operational. First, using the existing rhetoric of anti-colonial liberation and emphasizing that it was fighting a "common imperial enemy," the organization waged a skilled public-relations campaign aimed at linking the Palestinian cause to other national liberation movements in the Global South such as the Front de libération nationale (FLN) in Algeria and the National Liberation Front (NLF) in South Vietnam, as well as to

⁴⁶ Morris, 370, 373–74; quote from Gelvin, *The Israel-Palestine Conflict*, 208.

⁴⁷ Philip Robins, *A History of Jordan*, Second Edition (Cambridge: Cambridge University Press, 2019), 138–39; Smith, *Palestine and the Arab-Israeli Conflict*, 322–24.

⁴⁸ Khalidi, *The Hundred Years' War on Palestine*, 119, 125; quote from Chamberlin, *The Global Offensive*, 5, 148.

radical terrorist organizations like Germany's Baader-Meinhof Gang and the Japanese Red Army. While the embrace of radicalism and anti-colonial ideas was more a matter of tactics than conviction for Fatah, it was central to the more-ideologically rigid splinter groups. Indeed, those organizations were unflinching in their support for radical social and political change. Most notably, the PFLP made clear its commitment to both national *and* women's liberation by having women such as Leila Khaled (1944-) and Amina Dahbour (1945-) lead several of the organization's high-profile hijacking operations.⁴⁹

Meanwhile, the more militant elements within the PLO initiated a simultaneous wave of spectacular "external operations" aimed at drawing attention to the Palestinian cause. These included the assassination of Jordan's Prime Minister Wasfi Tal (1919-1971) in November 1971 by a shadowy new group affiliated with Fatah called the Black September Organization and the murder of twenty-six people in a brutal attack by PFLP-affiliated Japanese Red Army gunmen at Israel's Lod International Airport in May 1972. Most shockingly, Black September terrorists killed eleven Israeli athletes at the Munich Olympics that September.⁵⁰

The PLO's global offensive paid handsome diplomatic dividends for the organization. The use of national-liberation rhetoric and the establishment of transnational linkages with other revolutionary non-state actors such as the NLF and the Baader-Meinhof Gang raised the PLO's stature and won it the sympathy of people living in recently decolonized African and Asian states. Likewise, its terrorist attacks ensured that the issue of the Palestinians remained at the center of global affairs. The governments of newly independent states did express regret at the loss of life in those incidents, but, having themselves only recently thrown off imperial rule, also vigorously defended the Palestinian struggle for self-determination and maintained that violence remained a legitimate tool in the pursuit of national liberation.⁵¹

By 1972, as a result, the PLO had come to win both widespread support among the people of the Developing World and the formal diplomatic backing of their governments in international bodies. The passage of a UN General Assembly Resolution in December 1972 on terrorism made clear the degree to which the PLO enjoyed the support of the world's developing states. Though it called for international efforts to control terrorism, it also upheld the right of colonized people to use violence to achieve national liberation and forthrightly condemned states that "den[ie]d] peoples their legitimate right to self determination and independence."⁵²

Perhaps unsurprisingly, the PLO's global offensive proved especially fruitful in the Arab world. It not only raised the organization's stature, but, more importantly, also helped it to gain official standing with the Arab League. That change occurred at the

⁴⁹ Rogan, *The Arabs*, 348–49; Chamberlin, *The Global Offensive*, 4–5, 52–53, 157.

⁵⁰ Chamberlin, *The Global Offensive*, 142–74, quote from page 148.

⁵¹ Chamberlin, 175–217.

⁵² Chamberlin, 175–217, quote from page 180.

Arab League Summit held in Rabat, Morocco in October 1974. There, the member states formally rejected King Hussein of Jordan's longstanding claim to represent the Palestinians. Instead, acknowledging that Arafat's organization enjoyed the support of the overwhelming majority of Palestinians and noting both the extent of its global terrorist operations and the success of its diplomatic initiatives, it voted to recognize the PLO as "the sole legitimate representative organization for the Palestinian people in any Palestinian territory that is liberated."⁵³

As its role in securing official standing for the PLO made clear, moreover, the global offensive also played a key part in accelerating the transformation of the Arab-Israeli conflict back into a contest between rival nationalisms. For years, as we have seen, the Palestinians had been passive observers in that struggle—a people who looked to leaders like Nasser to restore them to the land they had lost in 1948. Thanks to the failure of Arab armies in the Six Day War and the success of the PLO's global offensive, however, the dynamics of the Arab-Israeli contest had changed. Enjoying the recognition of nearly all Palestinians and a growing number of states, the PLO henceforth took the lead in the conflict with Israel. No longer, in other words, were the Palestinians wards of the Arab states or dependent upon them for their salvation; instead, under the PLO, they enjoyed agency in their struggle for the first time since 1948.

Progress and Frustration: Peace Talks, 1973-1977

In the meantime, the increasing strength of the oil producing states and the shock of the Yom Kippur War thoroughly scrambled the Arab-Israeli conflict. Indeed, the imbalance in power between Israel and the Arab states had been narrowed to such a degree that the two sides agreed to take part in peace talks for the first time. As a result, for a brief period after the 1973 War, hopes were high that a resolution of the Arab-Israeli conflict was at hand.

The Geneva Conference, December 1973

The peace talks began in December 1973 when the US and USSR cosponsored a meeting in Geneva under the aegis of the UN for the purpose of securing a settlement on the basis of Resolutions 242 and 338. The conference was well attended. All of the relevant parties sent representatives save for Syria and the PLO—the former because Asad insisted that Israel had to pledge to withdraw from all of the Occupied Territories as a precondition for his country's participation and the latter because Washington and Tel Aviv objected to the organization's participation. The conference opened on December 21. After a series of speeches, the attendees ended the meeting's first day by voting to put the conference on hold until the belligerent states had first negotiated a series of bilateral disengagement agreements designed to disentangle their military forces. It never restarted. Instead, thanks to Washington's desire to exclude the USSR

⁵³ "Seventh Arab League Summit Conference Resolution on Palestine," October 28, 1974, <https://www.un.org/unispal/document/auto-insert-194621/>.

from the peace process, the negotiations ended up following a bilateral path rather than the multilateral approach envisioned at Geneva.⁵⁴

The PLO's Diplomatic Push

Meanwhile, at the same time that the diplomats were preparing for the Geneva Conference, moderates in the PLO were beginning to push for a strategy oriented not on force but instead on diplomacy. They pursued this new approach for two reasons. First, they feared that the continued construction of settlements in the West Bank would eventually result in so many Jewish people living in the territory that future Israeli governments would be unable to cede it even if they wanted to; in other words, time was not on the Palestinians' side. Second, Arafat and his advisors perceived that the Yom Kippur War had altered the diplomatic situation and had made possible a comprehensive peace agreement that could potentially include provisions for a Palestinian state.⁵⁵

As Arafat was well aware, however, participation in the peace talks would require the PLO to take two important steps. First, it would need to cease the operational side of the global offensive—meaning, in other words, that it would need to end the military and terrorist attacks that had proven so effective at bringing attention to the Palestinian cause. Second, it would have to devise realistic negotiating positions that could serve as the basis for productive discussions. Accordingly, in late 1973, Arafat and his allies began to shut down the PLO's military operations. At the same time, they also started to lay the groundwork for the abandonment of the organization's longstanding support for a one-state solution, wherein the Palestinians would establish a state in the entire territory of the former mandate, in favor of a two-state solution, wherein they would have a state in the Occupied Territories that stood alongside an Israeli one that existed within its pre-1967 borders.⁵⁶

Shifting to a diplomatic approach did not promise to be easy, however. Arafat and the other moderates understood that the pursuit of a negotiated settlement and, especially, the promotion of a two-state solution would generate fierce opposition within the Palestinian community. To a degree, this was a problem of Fatah's own making. Though its leadership had been well aware that the *fedayeen* lacked the power to defeat the IDF and conquer Israel, it had nonetheless publicly maintained that the Palestinians could, through revolutionary action, achieve a complete triumph that would result in them regaining control of *all* of their ancestral lands. Primed in this way to believe that a total victory was possible, most Palestinians would strongly oppose a

⁵⁴ Yaqub, *Imperfect Strangers*, 152; Rogan, *The Arabs*, 371–72.

⁵⁵ Smith, *Palestine and the Arab-Israeli Conflict*, 334; Chamberlin, *The Global Offensive*, 218–19.

⁵⁶ Khalidi, *The Hundred Years' War on Palestine*, 125; Chamberlin, *The Global Offensive*, 218–21.

compromise solution and would instead demand the unconditional success that Fatah had earlier promised.⁵⁷

Two important groups were particularly inclined to resist any settlement short of a total victory. Drawn largely from the Palestinians who had fled Israel in 1948 and who now lived in the refugee camps, the bulk of the PLO's guerrilla soldiers were strongly against a resolution to the conflict that did not permit them to return to their homes in Israel proper. Meanwhile, the more radical splinter groups—particularly the PFLP-GC and George Habash's PFLP—were opposed to *any* compromise settlement. In fact, those organizations had earlier declared that the peace process between Israel and the Arab states was an act of treachery and had demanded that Arafat promise to refrain from taking part in the talks. The PFLP-GC had even tried to torpedo negotiations by launching a bloody terrorist attack against the Israeli village of Kirva Shimona in April 1974.⁵⁸

Arafat and his allies thus had to walk a very fine line. They needed to make clear to Israel and its Western supporters that the organization would accept a two-state solution, and they needed to do so in such a way that did not alienate the many Palestinians who remained focused on securing a total victory. Their answer was to carefully characterize the organization's willingness to accept a rump state in the West Bank and Gaza Strip not as the end product of a two-state solution—which was in fact the goal that they now sought—but instead as merely the first step in a larger effort aimed at eventually retaking all of Palestine.⁵⁹

The PLO leadership moved forward with this strategy in June 1974 at the twelfth PNC meeting when it formally endorsed a new statement of strategy: the Ten-Point Program. A formal expression of the organization's position vis-à-vis the ongoing peace process, this document reaffirmed the PLO's traditional commitment to armed struggle and to the reconquest of all of Palestine. Crucially, however, it also declared that the organization would “employ all means” toward the achievement of that goal including the establishment of a “national authority” in liberated territories as an interim step. To the untrained eye, this declaration may have appeared to be just a recapitulation of the PLO's earlier hardline position, but it in fact marked a dramatic change: with it, the organization signaled—however elliptically—that it was open to negotiations and that henceforth it was willing to accept a ministate in the occupied Gaza Strip and West Bank.⁶⁰

If the full import of the Ten-Point Program was not apparent to the casual observer, it was emphatically clear to those attuned to the subtleties of diplomatic

⁵⁷ Chamberlin, *The Global Offensive*, 221–25.

⁵⁸ Gelvin, *The Israel-Palestine Conflict*, 208; Smith, *Palestine and the Arab-Israeli Conflict*, 335–36.

⁵⁹ Yaqub, *Imperfect Strangers*, 164–65.

⁶⁰ Quote from “Ten Point Plan of the PLO,” June 8, 1974, https://ecf.org.il/media_items/995; Yaqub, *Imperfect Strangers*, 164–65.

language. US State Department officials immediately grasped that the document marked both the PLO's implicit acceptance of Israel's right to exist and its willingness to settle for a ministate in the West Bank and Gaza Strip. Palestinian hardliners in splinter groups like the PFLP and the ALF likewise understood the document's thrust. They were not happy about it. Furious that Arafat was willing to consider anything less than a total victory, they moved to counter the Ten Point Program by organizing the Rejectionist Front. Under its banner, they issued propaganda broadsides against Arafat and the moderates and sought to rally Palestinians against a settlement based on the two-state model. The existence of the front significantly constrained Arafat and his allies. The moderates understood that if they publicly pursued a diplomatic settlement and then failed to attain one, the more militant groups would be able to exploit the resulting frustration among the Palestinian people to make substantial political inroads at their expense. As a result, Arafat and his supporters remained coy about their diplomatic goals and continued to disguise their support for negotiations and their acceptance of a two-state solution in extremist rhetoric.⁶¹

Building on the organization's earlier success, the PLO's turn to diplomacy produced some exciting global victories following the Yom Kippur War. Three stand out. First, the UN invited Arafat to address the world body in November 1974. His speech was enthusiastically received, and the UN General Assembly followed by granting official observer status to the PLO. Second, thanks to a combination of rising sympathy for the Palestinian people and fear of the oil weapon, states like France extended de facto recognition to the organization. Finally, the PLO's successful lobbying and the growing power of the globe's developing states resulted in the UN General Assembly issuing a resolution in 1975 declaring that "Zionism is a form of racism and racial discrimination."⁶²

These were important achievements, and they marked a dramatic improvement in the Palestinians' position in a remarkably brief period. At the same time, however, the PLO was unable to build on them or even to use them to secure a seat at the peace talks. What accounted for this failure? Why had the PLO's diplomatic initiative flopped?

Two factors worked together to thwart the PLO's diplomatic offensive. First, Washington and Tel Aviv remained unwilling to negotiate with the organization. Arguing that the establishment of a Palestinian state in the territories "would be the beginning of the end of the State of Israel," Prime Minister Yitzhak Rabin (r. 1974-1977, 1992-1995) steadfastly refused to take part in direct talks with the PLO or to countenance concessions that could lead to the creation of a Palestinian state in the territories. Washington supported his stance. Focused largely on pushing the Soviet Union out of the Middle East, Kissinger was unwilling to apply pressure to America's most important

⁶¹ Chamberlin, *The Global Offensive*, 237–40, 252.

⁶² Quote from "UN Resolution 3379," November 10, 1975, <https://web.archive.org/web/20121206052903/http://unispal.un.org/UNISPAL.NSF/0/761C1063530766A7052566A2005B74D1>; Chamberlin, *The Global Offensive*, 246–47, 255–56.

regional ally. Second, the Palestinians' internal political differences prevented Arafat from making concessions that might force Israel to negotiate. It is conceivable that he might have leveraged the PLO's newfound international political strength to overcome Washington and Tel Aviv's opposition to direct talks if he could have openly declared the organization's willingness both to recognize Israel and to settle for a ministate in the Occupied Territories. However, the political threat posed by the splinter groups compelled Arafat to continue to call publicly for unconditional victory—a position that gave Washington and Tel Aviv the excuse they needed to refuse to negotiate. As a result, the PLO would remain outside the formal peace process during the 1970s.⁶³

Shuttle Diplomacy

In the meantime, the disengagement talks that had started in Geneva in December 1973 had begun to produce tangible progress. Taking charge of the diplomatic effort to disentangle the military forces of the combatants, Secretary of State Kissinger used shuttle diplomacy—so called because it involved him shuttling back and forth by plane between Tel Aviv, Cairo, and Damascus—to lay the groundwork for a series of promising disengagement agreements based on Resolution 242. It was a grueling effort. Traveling nearly 39,000 kilometers on forty-one separate flights and dealing with obstinate rulers like Asad and Rabin, Kissinger won plaudits in the press for his apparently single-minded determination to end the Arab-Israeli conflict once and for all.⁶⁴

Appearances, in this case, were deceiving. In reality, the secretary of state was not seeking a comprehensive settlement based on Resolution 242. Instead, he was playing a very careful diplomatic gambit designed to achieve two related American objectives: weakening the USSR's position in the Middle East and strengthening Israel by permitting it to retain much of the territory it had acquired in 1967. Grasping that Sadat was willing to break with the other Arab states to secure peace with Israel, Kissinger correctly perceived that he could achieve both of these ends by brokering a separate peace between Tel Aviv and Cairo. Such an agreement would shift Egypt from the Soviet camp to the American one and thus, in zero-sum fashion, badly weaken Moscow's position in the Middle East. It would also end the possibility that Israel would have to fight a two-front war and would thus free Tel Aviv from needing to make substantial concessions regarding the remaining Occupied Territories. Kissinger knew, however, that he could not openly pursue a bilateral agreement. A savvy diplomat, he understood all too well that Soviet and Arab pressure on Sadat would derail any straightforward effort to secure a separate peace. He consequently opted for an incremental, "step-by-step" diplomatic strategy—one designed to disguise his pursuit of a narrowly bilateral agreement between Tel Aviv and Cairo.⁶⁵

⁶³ Chamberlin, *The Global Offensive*, 242–44, 264–67.

⁶⁴ Yaqub, *Imperfect Strangers*, 160.

⁶⁵ Yaqub, 148–50.

Kissinger's efforts quickly bore fruit. As a result of his shuttle diplomacy, Israel and Egypt signed the Disengagement of Forces agreement, popularly known as Sinai I, in January 1974. It secured the withdrawal of IDF forces from the canal zone, severely limited the number of soldiers that either side could station within thirty kilometers of the waterway, and established a UN-controlled buffer area between Egyptian and Israeli troops. It was a robust agreement—one that involved meaningful concessions on Egypt and, especially, Israel's part. Indeed, it was significant enough to persuade the Arab oil states to end the embargo on the sale of petroleum to the US.⁶⁶

He next turned his attention to Syria. A deal between Tel Aviv and Damascus that resulted in Asad regaining all of the land Israel had taken in 1973 and a small part of the territory it had seized in 1967 was vital to his plan. By making it appear that he was securing agreements between Israel and *all* the Arab states from which it had taken territory, the secretary of state could provide Sadat with the diplomatic cover he needed to pursue a *de facto* separate peace. Persuading Asad to accept anything less than the return of the entire Golan Heights and overcoming Israeli doubts proved to be a serious challenge, however. Indeed, it was only after Kissinger had shuttled multiple times over a thirty-four-day period between Damascus and Tel Aviv that he was finally able to coax Israel and Syria to come to terms. Signed in May 1974, the resulting separation of forces agreement achieved Kissinger's goals. It called for Israel to withdraw from all of the territory it had won in 1973 and, of great symbolic significance to Asad, required it to return the city of Quneitra that it had acquired in 1967.⁶⁷

Kissinger found the going much tougher after Israel and Syria signed this agreement, however. The problem was the Labor Alignment's weakened domestic political situation in Israel. Punished by the electorate for failing to detect Egypt and Syria's surprise attack in 1973, the party no longer dominated Israeli politics as it had in the past. Not only did Labor Prime Minister Rabin confront a new rightist challenge in the form of the Likud Party—which future Prime Ministers Ariel Sharon (1928-2014) and Menachem Begin (1913-1978) had created in 1973 by merging a number of small rightwing parties together—but he also had to make concessions to Labor's traditional partner, the National Religious Party (NRP), which was deeply committed to retaining control of the West Bank and the Gaza Strip. As a result, Rabin found himself so boxed in politically that he was unable to make the concessions that both Sadat and Kissinger sought. He did not get much sympathy from Gerald Ford (r. 1974-1976), who had become president when Nixon resigned in August 1974 as a result of the Watergate scandal. Furious about Tel Aviv's intransigence, he signaled his displeasure by calling for a "reassessment" of America's relationship with Israel and by holding up an existing arms agreement.⁶⁸

⁶⁶ Shlaim, *The Iron Wall*, 327–29; Hahn, *Crisis and Crossfire*, 60.

⁶⁷ Yaqub, *Imperfect Strangers*, 158–62.

⁶⁸ Smith, *Palestine and the Arab-Israeli Conflict*, 331–32; Yaqub, *Imperfect Strangers*, 168–70.

The threats appeared to work. Rabin's government soon indicated that it was willing to compromise, though it also made clear that it expected the US to make concessions to Israel in return. With Tel Aviv now demonstrating flexibility, Kissinger returned to work and quickly secured a second disengagement accord between Israel and Egypt. Signed in September 1975, the Sinai Interim Agreement, commonly known as Sinai II, called for Israel to withdraw its forces in the Sinai east of the strategic Gidi and Mitla Passes. In exchange, Cairo pledged to allow ships traveling to Israeli ports to transit the Suez Canal and promised not to use force against the Jewish state.⁶⁹

Negotiated separately, the concessions that Rabin's government had demanded of Washington were far reaching. At Tel Aviv's insistence, the US agreed to provide Israel with substantial ongoing economic and military aid, promised to refrain from talking with or recognizing the PLO unless that organization first acknowledged Israel's right to exist, and pledged not to initiate any diplomatic initiatives in the region without prior consultation. These terms were enormously beneficial for Israel. It not only received a staggering \$4 billion worth of American arms over the next three years alone—twice the annual amount that the US had heretofore provided—but could now rest assured that the US would not enter into negotiations with the PLO that might compel it to cede territory to the Palestinians.⁷⁰

For Syria and, especially, the PLO, Sinai II and the promises Washington had made to Tel Aviv were a disaster. By ending the possibility of a two-front war, the agreement freed Israel from having to make any further concessions to Syria regarding the Golan heights. Meanwhile, by giving Israel what amounted to a near veto over American policy toward the PLO, Washington's pledge not to talk to the Palestinians derailed Arafat's diplomatic offensive and prevented him from turning the organization's international success into tangible gains in the struggle to establish a Palestinian state. Above all, the American concessions to Israel put the PLO in a cruel catch-22 situation that echoed the "Iron Cage" into which the British mandatory authorities had trapped the Palestinians in the 1920s. It could not talk with the US unless it first recognized Israel's right to exist, but it simultaneously could not negotiate effectively with Israel if it yielded its only real diplomatic leverage—recognition—before talks had even begun.⁷¹

Israel's Rightward Turn

Meanwhile, Israel's rightward tilt accelerated in the mid 1970s. Already under pressure from the NRP and Likud, the Labor government faced a new conservative challenge in the form of a populist movement composed of young religious people called Gush Emunim, or Bloc of the Faithful, that was dedicated to the permanent retention of the Gaza Strip, Golan Heights, and West Bank. Believing that God had

⁶⁹ Hahn, *Crisis and Crossfire*, 61.

⁷⁰ Shlaim, *The Iron Wall*, 342–46.

⁷¹ Quote from Rashid Khalidi, *The Iron Cage: The Story of the Palestinian Struggle for Statehood*, Reprint Edition (Boston: Beacon Press, 2006), x, 156; Chamberlin, *The Global Offensive*, 254–55.

promised the Occupied Territories to the Jewish people, the movement staged public protests and built illegal settlements as part of a sustained campaign to compel the government to annex the territory. The pressure on Rabin was enormous—enough, in conjunction with the Labor Alignment's worsening electoral position, to compel his cabinet to bend to public opinion and approve a substantial number of new settlements.⁷²

Demographic changes also played an important part in shifting Israel to the right. Thanks to their higher birthrates, the Mizrahim and Maghribian Jewish immigrants and their descendants had come to account for about half the country's population by the mid 1970s. Having endured discrimination, high unemployment rates, and the Labor government's failure to provide them with adequate housing in the 1950s and 1960s, they were angry with both the party and the Ashkenazi establishment. Expelled from their former homes in the Middle East and North Africa, meanwhile, they also had a special animus for the Arab states and steadfastly opposed making concessions to them.⁷³

The new Likud Party was well positioned to take advantage of the demographic and ideological changes that Israel was experiencing in the 1970s. Its pledge to take a hard line with the Arab states was sufficient to win for the party the overwhelming support of the Maghribian and Mizrahim Jewish people. Meanwhile, Likud's declaration in its party platform that "[t]he right of the Jewish people to the land of Israel is eternal and indisputable" was sufficient to secure it the backing of the settlement movement. The resulting electoral realignment proved decisive., Likud not only won a stunning electoral victory in 1977, but, in the process, displaced the Labor Alignment as Israel's dominant political party.⁷⁴

Likud's success did not bode well for the peace process. Committed to retaining the Occupied Territories, the new prime minister, Menachem Begin (r. 1977-1983), expressed little interest in continuing the talks begun under his predecessors. As a result, the already faltering negotiations between Egypt and Israel on which Sadat had staked so much unceremoniously ground to a halt.⁷⁵

The *Infitah*

Begin's victory was far from the only problem that Sadat confronted in 1977, moreover. Even more troubling for him was the failure of his economic liberalization program. Called, the *Infitah*, or opening, he had introduced it with great fanfare during the heady days immediately following the Yom Kippur War. Designed to restore growth to the Egyptian economy by effecting a rapid transition from the sclerotic state-centered,

⁷² Ari Shavit, *My Promised Land: The Triumph and Tragedy of Israel* (New York: Spiegel and Grau, 2013), 204–6; Smith, *Palestine and the Arab-Israeli Conflict*, 336.

⁷³ Schulze, *The Arab-Israeli Conflict*, 49–50.

⁷⁴ Quote from Gelvin, *The Israel-Palestine Conflict*, 188–89.

⁷⁵ Gelvin, 188–89.

Nasserist system to a more-dynamic, profit-based, free-market one, it was far reaching in terms of both scope and depth. It included the relegalization of direct foreign investment, the reintegration of Egypt into the Western economic order, and the concomitant abandonment of the Import Substitution Industrialization (ISI) model of development that Nasser had championed. As such, it promised to thoroughly remake the Egyptian economy and—Sadat promised—to dramatically raise living standards.⁷⁶

The *Infitah* succeeded in doing away with the worst excesses of Arab Socialism, but it failed to produce the promised economic growth. Instead, it brought corruption, crony capitalism, economic stagnation, and punishing inflation. The worst came in early 1977. In January of that year, the International Monetary Fund (IMF) had compelled Egypt to severely reduce subsidies on common consumer goods including flour, cooking oil, and sugar in order to qualify for a desperately needed stabilization loan. Coming on top of the recent inflation, the cutbacks hit peasants and working-class Egyptians with peculiar savagery. The result was civil insurrection. Bitter that Sadat's economic liberalization program had not only failed to live up to its promise but had rendered them even poorer, lower-class Egyptian's rose up in a series of massive, spontaneous riots. Security forces managed to reestablish order in a few days—killing 150 people in the process—and the state soon placated the lower classes by restoring the subsidies on consumer goods that it had removed in January; however, the incident had made emphatically clear that popular enthusiasm for the Egyptian president had faded precipitously since the war. For someone as politically attuned as Sadat, the lesson was clear: he needed a substantial political success if he wished to ensure his continued authority, and he needed it soon.⁷⁷

Sadat's Breakthrough, November 1977

Sadat manufactured one in the fall of 1977 with a bold gesture that jump started the faltering peace talks. As we have seen, the peace process had slowed after the signing of the Sinai II Accord before grinding to a halt with the Likud victory in the spring of 1977. It began to pick up speed in the summer of that year thanks to the efforts of the new American president, Jimmy Carter (r. 1977-1981). Committed to achieving a genuine, comprehensive peace settlement, he moved quickly to breathe life into the faltering peace process during his first year in office. However, his efforts quickly foundered over the question of whether the Arab states would negotiate with Israel bilaterally, as Begin demanded, or as a unified delegation, as the Arabs insisted. As a result, Carter's bid to restart the talks went nowhere, and the peace process remained moribund.⁷⁸

It was at that point that Sadat made his move. Speaking before the Egyptian National Assembly on November 9, he restated his commitment to a negotiated settlement before declaring that he was “prepared to go to the ends of the earth for

⁷⁶ Cook, *The Struggle for Egypt*, 137–38.

⁷⁷ Goldschmidt, *A Brief History of Egypt*, 201.

⁷⁸ Schulze, *The Arab-Israeli Conflict*, 50.

peace, even to the Knesset [Israel's parliament].” It was an audacious statement—one that not only tacitly recognized the Jewish state but also made clear that Sadat was willing to break with longstanding Arab positions to achieve peace. Begin was receptive to Sadat's overture. He may have been unalterably opposed to yielding the Golan Heights, Gaza Strip, and West Bank, but he was willing to cede territory in the Sinai Peninsula if doing so would end hostility between Israel and its most dangerous neighbor. Accordingly, on November 20, just eleven days after he had made his electrifying announcement, Sadat spoke before the Knesset at Begin's invitation. In his address, he reiterated the Arab's long-held position that any peace would have to be a comprehensive one that accorded the Palestinians self-determination. At the same time, however, he also broke with longstanding Arab League policy by calling on the other Arab governments to recognize Israel's right to enjoy peace and normalized relations.⁷⁹

It was a bold move on Sadat's part—one that broke the psychological impasse and set the stage for productive bilateral negotiations. In the short run, however, it produced more smoke than fire. Indeed, when the two sides met for substantive discussions at Sadat's winter residence in Isma'iliyya on December 25, 1977, the talks quickly deadlocked. Poor chemistry was part of the problem; Begin and Sadat never managed to establish the personal rapport that can often see difficult negotiations through to a successful conclusion. More fundamentally, the discussions failed to produce progress because the two sides were far apart on the objectives that they sought. Cairo pursued two goals: a narrow Israeli-Egyptian agreement that would secure the return of the Sinai and a broader settlement that would achieve a just resolution of the Palestinian situation. Tel Aviv, in contrast, remained unwilling to discuss anything beyond a bilateral deal with Egypt. Unsurprisingly, as a result, the talks quickly collapsed.⁸⁰

The Ba'athist States

Sadat's aggressive pursuit of peace with Israel and his effort to liberalize Egypt's economy put him decidedly out of step with the radicalism that many Arab states and movements had embraced following the Six Day War. This was particularly true with regard to the Ba'athist regimes in Syria and Iraq. Though they remained bitter rivals, the two states continued to pursue the Ba'ath Party agenda. They promoted the radical restructuring of society, sought to implement Arab-socialist economic policies, and worked to promote pan-Arabism—albeit in a watered-down form. Ironically, however, neither regime maintained control as a result of its adherence to the Ba'athist ideology that had propelled them to power. Instead, in both states, the ruling party's position ultimately rested on traditional structures such as patronage networks and the clan and religious ties that bound the ruling class together.

⁷⁹ Shlaim, *The Iron Wall*, 367–69; Yaqub, *Imperfect Strangers*, 252–54.

⁸⁰ Yaqub, *Imperfect Strangers*, 261; Rogan, *The Arabs*, 390–91.

Saddam Husayn's Iraq

The reliance on kinship networks proved especially critical in the Ba'athist leaderships' gradual consolidation of power in Iraq in the late 1960s and early 1970s. Sharing control of the government with a group of army officers, the tiny Ba'ath Party initially had only a tenuous grip on the Iraqi state. However, under the duo of al-Bakr and his relative and second-in-command, Saddam Husayn (r. 1979-2003), it quickly solidified its control of the government. The two men complemented each other well. A former prime minister, President al-Bakr lent a veneer of legitimacy and stability to the regime. Ruthless and driven, meanwhile, Saddam Husayn did the dirty work needed to undermine rival power centers in government and society. He proved peculiarly adept at this task. Using a series of imaginary coup plots and conspiracies, he purged Western businesspeople, Nasserites, independent-minded Ba'athists, and others perceived as threats to his and al-Bakr's rule. He balanced that ruthlessness with political savvy. Most notably, at the same time that he was removing potential enemies, he was also making skillful use of patronage and traditional kinship relationships to put reliable members of his Tikrit-based al-Bu Nasir clan in key positions of power. This move proved critical in his rise to power. With their help, he was able to force out the few remaining military leaders from the inner circle in 1970—thus securing the Ba'ath Party's dominance of the government.⁸¹

Importantly, Saddam Husayn had not institutionalized Ba'athist rule out of deeply held convictions. Largely non-ideological and uninterested in Ba'athist beliefs, he instead strengthened the party—and his control of it—because he saw it as a powerful instrument he could use to position himself as al-Bakr's successor. He was successful in this effort. Deftly exploiting the patronage that his position as party secretary afforded him, he gradually assumed control of the party; by the mid 1970s, as a result, he had secured his standing within the Ba'athist leadership and had emerged as the country's *de facto* ruler.⁸²

He followed by asserting total control of both the Iraqi state and society. He first secured his complete dominance of the state by insinuating loyal Ba'athists into key decision-making positions in the bureaucracy. He next moved to assert control over the Iraqi people—an end he achieved by transforming what had once been a comparatively small party into a huge apparatus that extended its tentacles to the local level. Finally, he built up a vast, interlocking web of competing security services loyal to him that kept watch on the Iraqi people—and each other—so as to ferret out any challenges to the regime before they could establish roots. Periodic show trials, purges of unreliable people, and public executions further ensured compliance.⁸³

⁸¹ Tripp, *A History of Iraq*, 194–99.

⁸² Phebe Marr, *The Modern History of Iraq*, Second Edition (Boulder: Routledge, 2004), 148–52.

⁸³ Marr, 148–52.

If the security services constituted the stick that Saddam Husayn used to control the Iraqi people, then the patronage and social-welfare benefits derived from the revenue Iraq earned on the sale of oil constituted the carrot. Having already won popular acclaim by nationalizing the Iraq Petroleum Company (IPC) in 1972, he took advantage of the growth in oil revenues from \$1 billion in 1973 to a staggering \$26 billion in 1980 to expand social-welfare benefits and to develop a modern, Arab-socialist economy. Economic development was unquestionably impressive. Under his direction, the regime constructed desperately needed infrastructure such as electrical generation plants and highways and established collective farms and state-run industrial facilities like iron works and petrochemical plants. Social reforms and welfare benefits were equally remarkable. The regime expanded educational opportunities, provided free health care to all Iraqis, and, in a move that delighted rural Iraqis, paired the revocation of the hated Law Governing the Rights and Duties of Cultivators with an aggressive land-reform program. Extending to nearly every Iraqi, the substantial welfare benefits that the state provided quickly succeeded in building popular support for the regime. As important, they drew the entire population into the patronage system that Husayn had constructed and thus rendered nearly all Iraqis beholden to the state, and—by extension—to Husayn himself.⁸⁴

Despite the remarkable economic growth and the expansion of social-welfare benefits, however, many Iraqis remained unhappy with Ba'athist rule. Dissatisfaction was particularly prominent among the country's Shi'i Arab minority. They disliked their second-class status, opposed the Ba'athist promotion of secular rule, and resented the ban on religious processions that the regime had imposed. They voiced their discontent by periodically rioting and by backing the cleric Muhammad Baqir al-Sadr's (1935-1980) al-Da'wa movement, which called for the establishment of a government based on Shi'i religious principles. Still, the Shi'a were a manageable problem for Husayn—one that his government could address through a mix of rewards and threats.⁸⁵

The Kurds were a different matter. Having long sought greater autonomy, they had viewed the Ba'athist coup in 1968 as an opportunity to extract concessions from the government. Accordingly, under Mustafa Barzani (1903-1979), the powerful head of the Kurdistan Democratic Party (KDP) and leader of the powerful *peshmerga* militia, the Kurds issued a series of sweeping demands shortly after al-Bakr and Husayn took power. They included recognition of Kurdish as an official language, Kurdish control of a share of the revenue generated from the Kirkuk oil fields in northern Iraq, and, most importantly, immediate autonomy. For a time, the regime was able to mollify the Kurds by indicating that it would meet their key demands. By 1974, however, it was no longer able to disguise the fact that it had no intention of ceding oil revenue or giving the Kurds any control over finances. In response, Barzani once again ordered the *peshmerga* to rise up in revolt. This time, the Kurds enjoyed foreign support. Seeking to pressure Baghdad into agreeing to move the border between Iran and Iraq from the Iranian shore of the strategic Shatt al-Arab Waterway to the *thalweg*—the midpoint of its deepest

⁸⁴ William R. Polk, *Understanding Iraq* (New York: Harper Perennial, 2005), 124–28.

⁸⁵ Tripp, *A History of Iraq*, 202–4.

channel— Shah Muhammad Reza's (r. 1941-1979) government began to provide heavy weapons, supplies, and ammunition to Barzani's forces. The new arms proved decisive. With them, the *peshmerga* fighters had little difficulty defeating the Iraqi army and assuming control of most of the country's mountainous north. The Kurdish uprising was an embarrassing setback for Saddam Husayn—one, he grasped, that threatened to damage his political standing if he were unable to quickly bring it to heel.⁸⁶

Accordingly, he moved with haste to find a way to defeat the insurgency. It did not take him long to arrive at one. Concluding that the military could not defeat the Kurds so long as they were receiving arms from Iran, he swallowed his pride and pursued a diplomatic settlement with the shah. The cost was substantial. In the Algiers Agreement signed in March 1975, Saddam Husayn secured an Iranian promise to immediately end its support of the Kurds in exchange for Iraq assenting to the border adjustment along the Shatt al-Arab Waterway that the shah had been seeking. While it was an unpopular agreement in Iraq, it did tip the balance in the war with the Kurdish insurgents in the regime's favor. Stripped of the supplies that had sustained their campaign, the *peshmerga* were no longer able to stand up to the Iraqi army and quickly went down to defeat. In the aftermath, the already faction-ridden Kurds split between the KDP—led by Barzani's sons after his death in 1979—and a new, leftist party, the Patriotic Union of Kurdistan (PUK), headed by future Iraqi President Jalal Talabani (1933-2017). The Algiers Accord was thus a mixed bag for Saddam had certainly tarnished his reputation, but by making the victory over the Kurds possible, the agreement had also permitted him to restore his standing within the regime.⁸⁷

Asad's Syria

Hafiz al-Asad's rule in Syria in the 1970s bore more than a superficial resemblance to Husayn's regime in contemporary Iraq. Like his counterpart and rival, Asad lacked a deep commitment to Ba'athism and established control of the party and the state far more in pursuit of his personal ambitions than in service to any larger ideological agenda. Similarly, much as Saddam Husayn had done in Iraq, Asad maintained his grip on power through traditional patronage arrangements and a reliance on kinship networks—in his case, among Syria's Alawite minority. Finally, Asad shared Husayn's reliance on internal security services to maintain order. Ruthless and corrupt, the Syrian *mukhabarat* used fear and intimidation to ensure loyalty and to find and eradicate challenges to the regime.⁸⁸

Asad did not secure acquiescence to his rule strictly through repression, however. As in Iraq, the regime also used patronage, economic development, and social welfare benefits to win the backing of the people. Here, he made effective use of Ba'athist ideology and its socialist economic program. He substantially expanded the

⁸⁶ Tripp, 199–202, 211–14.

⁸⁷ Marr, *The Modern History of Iraq*, 152–57.

⁸⁸ John McHugo, *Syria: A History of the Last Hundred Years* (New York: The New Press, 2014), 183–85, 188–89.

state-run component of the economy and strengthened its links with the traditional, capitalist sector in order to increase the state's share of economic activity and to insinuate the regime's patronage network deep into the private sphere—a move that gave him the power to reward those loyal to him and to punish those who sought to disobey his wishes. He also followed through with the Ba'athist Party's promise to build a significant amount of desperately needed modern infrastructure including the expansion of Syria's network of railways and hard-surface roads and the construction of the Tabqa Dam on the Euphrates River. Improving living standards, these projects secured the support of many poor, rural Syrians. Finally, in keeping with Ba'athism's emphasis on radical social change, he raised literacy rates and ended the landlord class's tradition of lending money to peasants at usurious rates—moves that won him the backing of those most committed to the Ba'athist call for an equitable society.⁸⁹

Ultimately, though, the economic development and social-welfare programs that Asad pursued in the 1970s were far-less successful than were Husayn's. The reason was simple: oil, or, to be more precise, the lack thereof. Absent the enormous revenue that Iraq enjoyed thanks to its vast petroleum production, Syria did not have the same easy access that Iraq did to the imported goods and capital equipment on which rapid economic development depended. The results were predictable. Without the windfall that oil revenue provided to Iraq, the Syrian economy could not overcome the corruption, the extensive use of patronage for political purposes, and, especially, the sclerotic, crony-socialist development scheme that Asad had pursued. As a result, the country endured a grim combination of weak growth and high inflation over the course of the 1970s.⁹⁰

Ironically, relations between Syria and Iraq remained frosty following the Six Day War despite both states having Ba'athist governments. The rival regimes fought over which of them was the real leader of the Ba'ath movement and worked to weaken each other through hostile propaganda. Here, too, Husayn enjoyed the upper hand. The gradual flow of prominent Syrian Ba'athists into exile in Iraq after Asad took power boosted his claims to leadership in the movement, particularly after one of the party's founders, Michele Aflaq (1910-1989), fled to Baghdad in 1970. In addition, Damascus's willingness to sign the disengagement accord with Tel Aviv bolstered Iraq's claim that it was the more militantly anti-Israel Ba'athist state. Finally, Husayn scored points with Ba'athists throughout the Middle East at Asad's expense by criticizing Syria's intervention on the conservative side in the Lebanese civil war in 1975—an action, the Iraqi president argued, that made clear that the Syrian president lacked a sincere commitment to Ba'athist ideology.⁹¹

⁸⁹ McHugo, 183–87, 195.

⁹⁰ McHugo, 183–87.

⁹¹ Tripp, *A History of Iraq*, 209–10.

The Lebanese Civil War

Despite coming dangerously close on a few occasions, Lebanon had largely managed to avoid the instability that had gripped nearby Arab states such as Iraq and Syria after 1948. Following the Six Day War, however, it could no longer insulate itself from either the effects of the Arab-Israeli conflict or the radical ideologies then coursing through the Arab states. As a result, Lebanon descended into a brutal civil war in 1975 that would last, with times of tense peace, until 1990.

Two issues set the stage for the conflict. First, the National Pact's division of political power along confessional lines—recall that it had allocated seats in the parliament according to the 1932 census—had begun to produce growing frustration among non-Christian Lebanese who increasingly chafed under Maronite dominance. The disparity in political power was bad enough; what was worse was the Maronites' steadfast unwillingness to conduct a new census that might alter the political balance—despite the fact that divergent birth rates had long since resulted in the country having a majority Muslim population.⁹²

Influenced by the radical political ideas then sweeping through the Arab World, the leftist Druze leader Kamal Jumblatt (1917-1977) responded to continued Maronite intransigence by founding the populist Lebanese National Movement (LNM) in the early 1970s. It demanded the replacement of the National Pact that had governed the country since 1943 with a new political structure based on majority rule, and it quickly gained a wide following among Shi'a, Druze, and even some Christians. Unsurprisingly, Jumblatt's campaign and the enthusiastic response it received worried the country's establishment. Concerned that they would lose power under his proposal, most Maronite politicians—and even some Sunni leaders—started to organize against the LNM. Ominously, as part of this effort, they began to form religiously defined militias.⁹³

Second, the conflict between the Israelis and Palestinians had begun to spill over into Lebanon with disastrous consequences for the country's internal cohesion. The presence of 350,000 dispossessed Palestinians had already weakened the bonds that held the country together even before the PLO relocated to Lebanon in 1971 and started using it as a base from which to attack Israel. Thereafter, *fedayeen* raids and cross-border shelling invited IDF reprisals, while Arafat's establishment of a state-within-a-state in Lebanon deepened the country's division. The results were stark. No longer able to contain either the conflict over the internal distribution of political power or the destabilizing impact of the PLO's presence, Lebanon in the mid 1970s had become a tinderbox ready to go up in flames.⁹⁴

⁹² Yaqub, *Imperfect Strangers*, 210–11.

⁹³ Rogan, *The Arabs*, 381.

⁹⁴ Smith, *Palestine and the Arab-Israeli Conflict*, 355–57; Hourani, *A History of the Arab Peoples*, 430.

It finally ignited in early 1975. That March, a bloody government crackdown on a Muslim protest over fishing rights led to a series of skirmishes in Sidon between the Lebanese army and leftist militias. The conflict gradually intensified until, on April 13, the powerful Phalange militia killed twenty-eight Palestinians on a bus in retaliation for a failed assassination attempt on Pierre Gemayel (1905-1984), the leader of the far-right Maronite Phalange Party. By then, Lebanon had reached the point of no return. The country thereafter steadily fractured along confessional lines, devolving into a complex civil war marked by sporadic firefights, brutal massacres, savage bombardments, and vicious terrorist attacks.⁹⁵

Inevitably, the fighting drew the PLO into the civil war. While the organization had tried to carefully navigate Lebanon's internal conflict up to that point, rightist attacks on Palestinian refugee camps in the winter of 1975-1976 finally compelled Arafat to join the conflict on the side of the LNM. Coupled with the sudden dissolution of the army in March 1976, the entry of the PLO into the civil war tipped the balance decisively in favor of the leftists. By the spring of 1976, as a consequence, the PLO and the LNM stood on the cusp of a decisive victory.⁹⁶

It was at that point that Hafiz al-Asad intervened in the civil war. Ironically, he did not do so to help his ostensible allies in the LNP and PLO, whose commitment to Arab nationalism and radicalism he ostensibly shared. Instead, fearing that a leftist victory would draw Israel into Lebanon—with disastrous consequences for Syria in terms of both prestige and strategic position—he opted to promote a negotiated settlement known as the Constitutional Document. It proposed the retention of the confessional system albeit with a slight increase in Muslim political representation, and it called for the Lebanese government to exercise limited authority over the heretofore autonomous Palestinian refugee camps. Unsurprisingly, while the Maronites were enthusiastic about Asad's plan, neither Jumblatt nor Arafat agreed to its terms. On the contrary, close to defeating their enemies, they rejected the Constitutional Document in favor of a final offensive designed to secure a total victory over the Maronites. Asad was furious that his nominal allies would not follow his lead. Accordingly, he responded by sending a substantial military force into Lebanon with instructions to rescue the Maronites from defeat and to restore order to the war-torn country.⁹⁷

Making allies out of longtime rivals and enemies out of previous friends, Asad's invasion thoroughly scrambled the region's fault lines. Radical Syria was now supporting conservative forces while the PLO was providing security for the US embassy; more remarkably, Tel Aviv and Damascus found themselves in the odd position of backing the same faction. The two sworn enemies even came to a *modus vivendi* in which Israel accepted Asad's military deployment based on the assumption that the presence of Syrian troops in Lebanon would curtail PLO attacks. While Tel Aviv was happy to have the Syrian army impose order in northern and central Lebanon,

⁹⁵ Rogan, *The Arabs*, 381–83.

⁹⁶ Smith, *Palestine and the Arab-Israeli Conflict*, 355–57.

⁹⁷ Yaqub, *Imperfect Strangers*, 222–26.

however, it had no intention of permitting Asad to deploy troops to the south. Instead, the Israelis moved to control that area with a proxy force called the South Lebanese Army (SLA). Unsurprisingly, the marriage of convenience between Israel and Syria left many seasoned observers bemused. As one longtime American veteran of Middle Eastern diplomacy succinctly declared, “[i]t’s weird. It’s truly weird.” It may have been strange, but it also managed to bring to a conclusion the first phase of the Lebanese Civil War—a conflict that had already claimed more than 30,000 lives.⁹⁸

The Rise of Islamism

Meanwhile, in the wake of the Six Day War, a new, insurgent ideology—Islamism—began to quietly emerge in the Arab states. With roots stretching back to Hassan al-Banna (1906-1949), Islamism held that the problems of the Muslim world stemmed not from its technological or scientific inferiority to the West, but instead from the fact that Muslims had adopted secular governments that—in direct contravention of longstanding Islamic practice—separated religion and state. Fortunately, the Islamists argued, there was a straightforward solution. By replacing its Western-style governments with Islamic ones that enforced shari‘a law, the Muslim world would again enjoy God’s grace and thus once more assume its rightful place as the globe’s dominant civilization.⁹⁹

Moderate Islamism

Most Islamists adhered to a moderate version of the ideology. First taking root in the 1970s among younger, educated Egyptians, moderate Islamism emphasized social justice and called for Muslims to abandon Western consumerism and social mores in favor of a return to traditional Islamist values and practices. The moderate Islamists had their first tangible success in Egypt’s universities. Over the course of the decade, Islamist students wrested control of the politically important student unions from Nasserists and demanded—with some success—that the country restructure higher education along Islamic lines through curricular changes and the adoption of gender-segregated classrooms. Later, many joined the Muslim Brotherhood, which, having sworn off violence, grew rapidly after Sadat released its members from prison.¹⁰⁰

Sayyid Qutb and the Emergence of Extreme Islamism

While most Islamists were moderates who hoped to persuade people to adopt their views, a small subset known as jihadis instead sought to effect change through violence. They drew inspiration from the works of Sayyid Qutb (1906-1966), the Egyptian intellectual whose writings constituted the ideological foundation of the

⁹⁸ Khalidi, *The Hundred Years’ War on Palestine*, 133; Quote from Yaqub, *Imperfect Strangers*, 124–29, 232–35.

⁹⁹ Lawrence Wright, *The Looming Tower: Al Qaeda and the Road to 9/11* (New York: Vintage Books, 2006), 28.

¹⁰⁰ Gerges, *Making the Arab World*, 321–24.

movement. Qutb's path to extreme Islamism was an indirect one. Though always a religious man, he only embraced Islamist ideas as a result of an extended trip he took to the United States in the late 1940s to study educational administration. Shocked by American materialism and by what he perceived to be the country's spiritual emptiness, he concluded that only Islamic laws and values could serve as the foundation for a just society. Accordingly, shortly after returning to Egypt in 1950, he joined the Muslim Brotherhood. Respected as an intellectual, he soon became the movement's ideological guide—a role he would retain even after his death in 1966.¹⁰¹

Despite his rapid rise in the organization, Qutb's time in the Muslim Brotherhood was anything but easy. Arrested in the roundup that followed the failed assassination attempt on Nasser in 1954, he languished in prison until the Egyptian president—having assumed incorrectly that the organization had been completely dismantled—amnestied Qutb and his associates in 1964. The movement's shockingly quick revival soon led Nasser to conclude that his decision to release the brothers had been a dangerous mistake. Accordingly, he had Qutb and the rest of the organization's leadership rearrested and tried for treason. Qutb's trial was more an exercise in government public relations than a legitimate judicial proceeding. Using torture to extract incriminating testimony from his associates, the prosecution secured a pro forma guilty verdict in July 1966. The following month, the regime had him executed.¹⁰²

If Nasser had hoped that death would end Qutb's influence, he was sadly mistaken. On the contrary, thanks to the books he had written while in prison, Qutb would enjoy a vastly greater influence in death than he ever had in life. By far the most important of those works was a slim volume called *Milestones* that he had written specifically for the most committed jihadis. Composed in small snippets that his associates smuggled out of jail, it reflected the radicalization he had experienced while confined to Nasser's brutal prison system. He argued in the book that the secular, materialistic states that predominated in the Muslim world were based on human-derived ideas rather than on God's laws; as such, they had produced not progress or enlightenment but only ignorance or *jahiliyya*—the same term Muhammad had used to describe pre-Islamic Arabia. To free themselves from *jahiliyya*, true Muslims needed to bring about governments that were rooted in the divinely provided shari'a code. But how were they to do so? That is, how could they replace the existing *jahili* regimes with Islamist ones?¹⁰³

It was in answering this question that Qutb provided the justification for terrorism and an Islamic revolution. Drawing on the ideas of the fourteenth-century legal scholar ibn Taymiyyah (1263-1328), he argued that rulers who had failed to establish shari'a law in their states were *kafirs*, or apostates, and thus subject to *takfir*—

¹⁰¹ John L. Esposito, *Islam: The Straight Path*, Fourth Edition (New York: Oxford University Press, 2011), 237–38.

¹⁰² Gerges, *Making the Arab World*, 138–39; Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 17–28.

¹⁰³ Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 34.

excommunication—and death. As such, the Islamist vanguard was not merely justified in using “physical power and jihad for abolishing the organizations and authorities of the *jahili* system” but in fact enjoined to do so. In other words, he maintained that violence, *even against rulers and others who claimed to be Muslims*, was an appropriate tool that the vanguard would need to use to return society to the justice and peace that he believed were inherent to a shari‘a-based state.¹⁰⁴

Milestones had an enormous influence on subsequent generations of jihadis. Its call for a dedicated vanguard to lead a revolution aimed at replacing the existing secular Arab states with new, shari‘a-based Islamic ones captivated a growing number of younger Muslims and resulted in the emergence of a number of new, underground militant organizations in the 1970s. As important, his promotion of jihad and call for *takfir* ensured that those groups would make violence a central component of their revolutionary effort.¹⁰⁵

Qutb’s vanguard began to take shape in Egypt in the 1970s. In a hint of things to come, a jihadi group called the Youth of Muhammad tried unsuccessfully to storm the armory at the Military Technical College in Cairo in 1974 in a failed effort to seize weapons with which it planned to overthrow the government. While the organization collapsed following the execution of its leader, other jihadi groups soon emerged to replace it. By the middle of the decade, for example, the cells that would later combine to form Jamaat al-Jihad, or Egyptian Islamic Jihad, had taken shape—including one led by Osama bin Laden’s (1957-2011) successor as head of al-Qaeda, Ayman al-Zawahiri (1951-2022). Likewise, a second organization, *Gamaat Islamiya*, or the Islamic Group, grew rapidly in the mid-to-late 1970s—particularly among university students.¹⁰⁶

Qutb-inspired groups did not remain confined to Egypt, moreover. Instead, influenced by *Milestones*, jihadi organizations began to quietly take shape throughout the Arab world over the course of the 1970s—most notably in Syria. There, the heretofore peaceful Syrian Muslim Brotherhood began to organize against the Alawite-dominated Ba’ath Party and its secular agenda. The Syrian jihadis first made their presence known in 1973 when they led demonstrations in Hama to express their opposition to the regime’s promulgation of a new constitution that did not require the president to be a Muslim. Later, in 1976, they used violence for the first time when they responded to Asad’s intervention on the Maronite side in the Lebanese Civil War with a campaign of terror bombings that ended only after a sharp government crackdown.¹⁰⁷

¹⁰⁴ Quote from Rogan, *The Arabs*, 401–2; Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 34.

¹⁰⁵ Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 35.

¹⁰⁶ Wright, 48–50, 141–42; Goldschmidt, *A Brief History of Egypt*, 208.

¹⁰⁷ McHugo, *Syria*, 191; Rogan, *The Arabs*, 405–6.

Historical Disagreements: Islamism and the Six Day War

Since the Iranian Revolution in 1979, historians and other observers have drawn a direct connection between the emergence of Islamism in the 1970s and the Arab defeat in the Six Day War in 1967. According to the standard account, Israel's triumph over the secular Arab states had so totally discredited those regimes that it produced a sudden, populist ideological sea change in the Middle East in which the masses rapidly abandoned radical Arab nationalism in favor of Islamism. Gaining broad currency after 1979, this interpretation remains to the present day the conventional wisdom regarding the growth of Islamism. Emblematic of its continued dominance are the views of the journalist Lawrence Wright, who argues in his Pulitzer Prize-winning study of al-Qaeda that "the appeal of Islamic fundamentalism in Egypt and elsewhere was born in [the] shocking debacle" of 1967.¹⁰⁸

In recent years, however, the scholar Fawaz Gerges has persuasively challenged this interpretation. He raises two important objections. First, he contends that contemporary Islamism did not originate as a populist movement but was instead a top-down phenomenon that benefited from substantial active state assistance. Sadat was the key figure here. Seeking support against the Nasserists who continued to dominate the government during the early part of his tenure, the Egyptian leader presented himself as a great champion of Islam and its values. Referring to himself as "the believer president," for example, he altered the constitution in 1971 to assert that "the principles of Shari'a are the main sources of legislation." More importantly, he provided direct state assistance to Egypt's Islamist groups. Most notably, he freed those members of the Muslim Brotherhood who remained in prison and supplied arms to *Gamaat Islamiya* in support of its effort to displace the Nasserists in the influential student unions. Sadat was not alone in nurturing Islamism, moreover. The government of Saudi Arabia, with which the Egyptian president had become close, also began to support the Islamists in the 1970s—most notably by providing huge amounts of petrodollars to finance the spread of Wahhabi religious ideas throughout the Muslim world. Thus, Gerges concludes, Islamism was not a bottom-up phenomenon, but instead a top-down one that depended heavily during its early years on state support.¹⁰⁹

Second and more importantly, he takes issue with the idea that Islamism swiftly gained in popularity at radical Arab nationalism's expense in the wake of the Six Day War. In his view, the conventional interpretation has it backwards. That is, rather than losing popularity, Arab nationalism instead *gained* in influence in the decade following the war even as Islamism remained little more than a fringe movement. As he notes, *all* of the regime changes that occurred in the Arab world in the aftermath of the conflict put Arab-nationalists rather than Islamists in power. Likewise, the splinter groups that rose

¹⁰⁸ Quote from Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 45; Fawaz A. Gerges, "The Transformation of Arab Politics: Disentangling Myth from Reality," in *The 1967 Arab Israeli War: Origins and Consequences*, ed. William Roger Louis and Avi Shlaim (Cambridge: Cambridge University Press, 2012), 287, 312–13.

¹⁰⁹ Gerges, *Making the Arab World*, 314–42.

to prominence in the PLO after the Six Day War were secular adherents of Marxism and not disciples of Qutb. Even Egypt—the home of the Muslim Brotherhood—was not immune to the shift to radical Arab nationalism. When members of the state-affiliated Organization for Socialist Youth found themselves disillusioned with Nasser following the war, for example, they did not abandon Nasserism for Islamism; instead, they embraced Marxism. Thus, as Gerges demonstrates, Islamism remained an underground movement during the 1970s while radical pan-Arabism retained its position as the dominant ideology in the Arab world.¹¹⁰

1979

Indeed, Islamism would only be in a position to begin flexing its muscles during another of the twentieth-century Middle East's watershed years: 1979. It was a long and enormously consequential year for the region. Bracketed by the shah's departure from Iran—discussed in chapter fifteen—and the Soviet Union's invasion of Afghanistan, the events of those fateful twelve months would reflect the trends, movements, and conflicts of the post-1967 Middle East and would, collectively, restructure the region socially, politically, and diplomatically.

The Egypt-Israel Peace Treaty

The signing of the Egyptian-Israeli Peace Treaty in March 1979 certainly reshaped the Arab-Israeli conflict. That Sadat and Begin were able to find common ground at that time was remarkable in light of the fraught state of bilateral talks in the early summer of 1978. As we have seen, the two sides had been unable to build on the disengagement agreements and Sadat's dramatic speech before the Knesset. Instead, the subsequent negotiations had revealed that Sadat and Begin held sharply divergent positions on key issues such as the disposition of the Occupied Territories. As a result, by the summer of 1978, the talks were at an impasse.¹¹¹

It was at that point that President Carter stepped in. Summoning the two leaders to the presidential retreat at Camp David in Maryland in September 1978, he oversaw thirteen days of difficult negotiations in an effort to bring the two sides together so that they could conclude an agreement. The differences between them were significant. Sadat offered Israel full recognition and peace in exchange for the return of the Sinai Peninsula but insisted that the agreement also include a general statement of principle that bound Israel both to return all the land it had taken in 1967 and to recognize the Palestinians' right to self-determination. Begin countered by expressing his willingness to return much of the Sinai in exchange for peace with Egypt; however, he adamantly

¹¹⁰ Gerges, 284–313.

¹¹¹ Yaqub, *Imperfect Strangers*, 261.

refused to discuss yielding the West Bank and Gaza Strip or to consider giving the Palestinians any more than limited, local autonomy under Israeli rule.¹¹²

Carter eventually arrived at a way of bridging the differences that was acceptable to both sides. Aware that the main stumbling block was the status of the West Bank and The Gaza Strip, the president devised an agreement regarding those territories that was sufficiently vague that each party could accept it. He also secured an oral commitment from Begin to impose a moratorium on the construction of settlements for five years. Thanks to this approach, the negotiations were able to produce two unlinked, interim agreements collectively known as the Camp David Accords. The first, "A Framework for Peace in the Middle East," called for the establishment of a Palestinian "self-governing authority" in the West Bank and Gaza Strip followed by a five-year "transitional period" that would conclude with "final status" negotiations. The second, "A Framework for the Conclusion of a Peace Treaty between Egypt and Israel," outlined a bilateral agreement based on Resolution 242's call for an exchange of land for peace and recognition. Sadat and Begin signed the accords on September 17, 1978.¹¹³

After further negotiations to refine the Camp David Accords, the two leaders signed the formal Egypt-Israel Peace Treaty on the White House lawn on March 26, 1979. The agreement committed Egypt to end the state of war that existed between it and Israel, to establish full, normalized relations with Tel Aviv, to acknowledge the right of Israeli ships to freely transit the Suez Canal and the Strait of Tiran, and to accept the stationing of a permanent UN peacekeeping force on its side of the two countries' common border. In exchange, Israel promised to complete a staged withdrawal from the Sinai. Despite intense criticism in the Arab world, the treaty's provisions went into effect as scheduled. The two countries established normalized relations in 1980, and Tel Aviv gradually withdrew its forces from the Sinai culminating in the evacuation of the remaining Israeli-controlled territory in 1982.¹¹⁴

Thanks to Israeli stonewalling, however, the unlinked "Framework for Peace in the Middle East" that accompanied the bilateral agreement went nowhere. As early as the fall of 1978, Begin had indicated that he interpreted the agreement in the loosest possible terms and that he did not view its provisions as binding on his country. He went even further after signing the formal peace treaty. He bluntly refused to recognize that the Palestinians enjoyed any collective rights as a people and claimed that he had agreed at Camp David merely to grant autonomy to individual Palestinians—a position that rendered the possibility of negotiations with the PLO dead on arrival. Begin also embarrassed Carter by pausing the construction of settlements for a bare three months rather than the five years he had promised, and he demonstrated his determination to establish permanent control of the territories by expropriating private Palestinian land on

¹¹² David W. Lesch, *1979: The Year That Shaped The Modern Middle East* (Boulder, Colo: Westview Press, 2001), 43–44.

¹¹³ Yaqub, *Imperfect Strangers*, 264–66; Quotes from "A Framework for Peace in the Middle East," September 17, 1978.

¹¹⁴ Yaqub, *Imperfect Strangers*, 266–70.

the West Bank. Indeed, the prime minister could not have made his position on the territories any clearer. He had no intention of yielding them or of permitting the creation of a Palestinian state, and—having secured the separate peace with Egypt that he had sought all along—he faced no pressure whatsoever to do so.¹¹⁵

Peace between Israel and Egypt had three enormous consequences. First, in the short run at least, the normalization of relations made Cairo a pariah in the Arab world. The Arab states immediately broke diplomatic relations with Egypt, imposed an economic boycott on it, and expelled it from the Arab League. Second, the peace agreement permanently tipped the balance in the Arab-Israeli conflict in Tel Aviv's favor. By removing the most powerful Arab state from the equation, it restored Israel to a position of unquestioned military preponderance and ensured that it would take part in any future negotiations from a position of overwhelming strength. Third, the new relationship between Cairo and Tel Aviv marked the point at which the Arab-Israeli conflict completed the transformation from a state-to-state contest back into a struggle between Palestinian and Israeli nationalisms that had begun with the Six Day War. Thus, just as critics had predicted, peace between Israel and Egypt gave Tel Aviv the free hand it had sought in its relations with the other Arab states and the Palestinians—a situation, as we shall see in chapter seventeen, that all-but guaranteed that Tel Aviv would take aggressive actions following the agreement's full implementation.¹¹⁶

Saddam Husayn's Coup

Meanwhile, just a few months after the signing of the peace treaty, Saddam Husayn would formally assume control of Iraq. His rise to the top had been a long time coming. Nominally al-Bakr's subordinate, he had long since established himself not only as the regime's strongman but as Iraq's de facto ruler. By 1979, however, he had determined that the time had come for him to finally rule in his own right. He made his move that July. Bloodless and quick, the coup began with al-Bakr resigning and ended with the Ba'ath Party naming Husayn as his successor.¹¹⁷

Not content merely to assume the top post, the new president followed with a series of bold and brutal actions designed to permanently eliminate any independent power bases in the Ba'ath Party that could conceivably threaten his authority. He initiated this campaign at a carefully orchestrated, televised party meeting held on July 22. Addressing the assembled leaders, he announced the discovery of a plot against the state organized by top-ranking Ba'athists who, he claimed, were in league with Syria; chillingly, he had the alleged conspirators, many of whom were in attendance, immediately arrested on trumped up charges. In the dubious judicial proceedings that followed, he arranged for twenty-two of them to be condemned to death and for thirty-three others to be sentenced to long prison terms. He followed by ordering the execution of another five-hundred high-ranking Ba'athists whose loyalty he

¹¹⁵ Shlaim, *The Iron Wall*, 371–72, 382–93.

¹¹⁶ Lesch, 1979, 82–93.

¹¹⁷ Tripp, *A History of Iraq*, 222.

doubted. Brutally effective, Husayn's coup seamlessly transformed Iraq from a one-party state into a personal autocracy—one that no longer rested on a shared ideological vision but instead on clan ties, patronage, and, above all, personal loyalty to Saddam Husayn.¹¹⁸

His seizure of power would have a decisive effect on the Middle East over the next twenty-five years. No longer bound by any internal checks or challenges, he was free to follow an aggressive foreign policy designed to achieve his dream of turning Iraq into the dominant power in the region and himself into its leading political figure. Ultimately, however, Husayn was unable to realize his aspirations. As we shall see in the next two chapters, rather than making Iraq the region's leading state, his aggressive approach to foreign relations would produce a series of disastrous conflicts that beggared his country, destabilized the Middle East, and drew the ire of the US.

The Islamist Onslaught

Meanwhile, the Sunni Islamist vanguard that Qutb had hoped to inspire announced its emergence in 1979 with a pair of spectacular terrorist attacks. The first occurred in Syria. Supported by wealthy families angry about the regime's program of land reform, the Syrian Muslim Brotherhood launched a bold assault on a military academy in Aleppo that resulted in the deaths of eighty-three Alawite cadets. The incident was the opening salvo in what would be a peculiarly savage war between the brothers and the government—one that would produce steadily rising casualties over the next few years. The second incident that year was far-more shocking. That November, more than four-hundred militant followers of Juhayman al-Otaybi (1936-1980), shocked Saudi Arabia and, indeed, the entire Muslim world, by seizing control of the Grand Mosque in Mecca and demanding the replacement of the Saudi regime with an Islamist state. It took two weeks of brutal fighting for the government to finally regain control of the mosque. Officially, 127 soldiers, a dozen hostages, and 117 jihadis died in the fighting; however, unofficial estimates put the death toll at more than 4,000.¹¹⁹

The spectacular emergence of Sunni Islamism in 1979 helped to spread jihadi thought throughout the Middle East. Indeed, in conjunction with the Iranian Revolution, it was the attacks in Syria and, especially, Saudi Arabia, that gave the jihadi groups the legitimacy needed to rapidly win new recruits to the cause of overthrowing the secular *jahili* regimes. As we shall see, those organizations would ultimately be unsuccessful in their attempts to establish Islamic governments in the Arab countries; nonetheless, their activities—and, especially, the efforts of states like Saudi Arabia to divert their zeal outward—would powerfully shape the region's development in the 1980s and 1990s.

¹¹⁸ Peter L. Hahn, *Missions Accomplished?: The United States and Iraq Since World War I* (New York: Oxford University Press, 2012), 60; Marr, *The Modern History of Iraq*, 178–81.

¹¹⁹ Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 101–9; Rogan, *The Arabs*, 405–7.

The Soviet Invasion of Afghanistan

The final event that shook the region in 1979 occurred just before the end of the year. In December, the USSR invaded Afghanistan in an effort to prop up the government of an important satellite state that was attempting to put down a Sunni insurgency. While Moscow committed 80,000 well-equipped soldiers to assist the Afghan government, it soon found itself bogged down in a costly quagmire with the *mujahidin*, the Sunni guerrillas who operated in the country's vast hinterlands. The Soviets had a hard go of it from the start thanks in part to the substantial amounts of military aid that the US and Saudi governments supplied them. Somewhat later, they also began to receive funds and recruits from newly formed transnational jihadi groups operating in the Arab states.¹²⁰

While it occurred just outside of the boundaries of the Middle East, the invasion of Afghanistan was yet another event in 1979 that would have enormous consequences for the region. Two stand out. First, the invasion dramatically boosted the position and organizational capacity of the jihadi groups. While the militarily inexperienced recruits they sent to Afghanistan had almost no impact on the course of the war, the process of recruiting, training, and supporting fighters gave the movement invaluable experience as well as the opportunity to expose sympathetic people to their ideas. Second, the Soviet invasion created fears in Washington that the move was the first step in a plan to seize the region's oil fields. It thus led President Carter to enunciate the Carter Doctrine in January 1980. Committing the US to defend the Persian Gulf from Soviet attack, this critical statement of policy would set the stage for deeper American involvement in the region in the 1980s and 1990s.¹²¹

Conclusion

The period from 1967 to 1979 was an enormously transformational one for the Middle East. Like the Ernest Hemingway character's description of his fall into bankruptcy, the change that the region experienced during this time happened "gradually and then suddenly," with 1979 marking the pivotal before-and-after moment. The momentous events of that year—above all, the Iranian Revolution—would, collectively, establish a new set of fault lines, conflicts, and issues that would reshape the region's history in ensuing decades. Indeed, as the historian David Lesch succinctly notes, "1979 constituted a major watershed, if not *the* major watershed, in modern Middle Eastern history." It is to this new, and very different Middle East that we will now turn.¹²²

¹²⁰ Lesch, 1979, 52–54, 101–2.

¹²¹ Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 121–23, 127–29; George C. Herring, *From Colony to Superpower: U.S. Foreign Relations since 1776* (New York: Oxford University Press, 2008), 854–55.

¹²² Hemingway quote from Ernest Hemingway, *The Sun Also Rises* (New York: Charles Scribner and Sons, 1926), 136; Lesch quote from Lesch, 1979, 2.

Chapter Seventeen: Transition and Stalemate, 1980-1990

Introduction

Nineteen seventy-nine came as a profound shock to the people of the Middle East. The transformative events of that pivotal year—the Iranian Revolution, the Soviet invasion of Afghanistan, Saddam Husayn's (r. 1979-2003) consolidation of power, the sudden emergence of jihadi groups, and the signing of the Egypt-Israel Peace Treaty—shook the region to its very foundations and led many observers to predict new and trying times ahead. They were not wrong. During the next decade, the Middle East experienced a brutal, eight-year-long war between Islamic Iran and Ba'athist Iraq, a wave of Islamist insurgencies, an Israeli invasion of Lebanon, and a popular uprising in the West Bank and Gaza Strip. It was, in sum, a critical decade—one of important transitions and bitter stalemates.

The Islamist Challenge

In many ways, the most fundamental change that occurred in the region after 1979 was the rapid growth of Islamism. As we saw in chapter sixteen, Islamism had emerged in the 1970s thanks to the support of the Saudi government and, especially, Egyptian President Anwar Sadat (r. 1970-1981). By the early 1980s, however, it had transformed into a decidedly grassroots movement—one that was neither beholden to nor controlled by its former patrons. Indeed, by that point, it had grown so strong that it had not only begun to supplant Arab nationalism as the most powerful insurgent ideology in the Middle East but had also—in both its mainstream and jihadi variants—become the primary challenge to pan-Arabist and conservative regimes alike.

Arab Presidents for Life

That Islamism had displaced secular Arab-nationalism as the primary revolutionary ideology in the Middle East is hardly surprising in light of the changes that had gradually occurred in the pan-Arabist states since the Six Day War. By the early 1980s, regimes such as Egypt, Syria, and Iraq that had once promoted sweeping political change in the region and that had fought to drastically reduce inequality at home had lost much of their radical drive. They continued to justify their rule by claiming that they were seeking to advance the general will and they still reliably mouthed revolutionary platitudes, but the increasingly bloated and sclerotic pan-Arabist regimes no longer invested those slogans with any real meaning. On the contrary, the leaders of those states had long since abandoned their earlier commitment to revolutionary change in favor of a more practical and cynical goal: retaining their grip on power.¹

¹ Roger Owen, *The Rise and Fall of Arab Presidents for Life* (Cambridge: Harvard University Press, 2012), 1–11.

The particulars differed in important ways, but the transformation of those regimes largely followed the same pattern. First, their rulers constructed intricate administrative structures designed to collect power in their hands and to prevent lower-ranking members of the ruling class from challenging them. For instance, Hafez al-Asad (r. 1970-2000) carefully fostered tension among the governing elite in Damascus and required all high-ranking officials to report only to him—a tactic that prevented his subordinates from creating a unified front powerful enough to insinuate itself into the decision-making process. Second, while leaders like Asad continued to rule through nominally autonomous political parties, they increasingly manipulated and controlled those organizations through traditional patronage networks that were often based on clan and ethnoreligious ties. As we saw in chapter sixteen, for example, Saddam Husayn ran Iraq and its ostensibly independent Ba‘th Party through a small, reliable group of Sunni Arabs linked together not through shared ideological convictions but instead through their common membership in the al-Bu Nasir clan to which he belonged. Similarly, Asad controlled Syria through a patronage network based in the Alawite community from which he came.²

Finally, with the partial exception of oil-rich Iraq, each of the pan-Arabist states abandoned Arab-socialist economic policies when it became clear that those approaches could no longer generate the tax revenue required to provide the subsidized goods that kept the masses satisfied or to fund the security services that held them in check when they were not. This shift did not represent an ideological change of heart or an endorsement of a free-enterprise system, however. Instead, it amounted to a cynical, calculated embrace of crony capitalism for the purpose of buttressing the regime. Compelled to auction off state assets as part of a debt settlement deal, for instance, Egyptian President Husni Mubarak (r. 1981-2011) arranged for a group of well-connected capitalists to purchase state monopolies at prices that were well-below market value—thereby ensuring that his regime continued to have access to the wealth that those assets generated.³

No subject was more emblematic of the radical regimes’ shift from being agents of revolution to defenders of the kleptocratic status quo than the issue of the succession. All of the pan-Arabist governments had originally established succession procedures that, while undemocratic, reserved to high-ranking party officials the power to select the next ruler. By the 1980s and 1990s, however, the reigning leaders of the Arab-nationalist states were dismantling those systems in favor of an informal new arrangement wherein the incumbent rulers anointed their sons as their successors. Thus, Saddam Husayn designated his son, Qusay (1966-2003), as heir apparent in 2001, Asad named his son, Bashir al-Asad (r. 2000-), to succeed him in 1994, and Mubarak began grooming his son, Gamal (1963-), for the presidency during the last years of his tenure. It is true that of the three only Bashir al-Asad went on to succeed his

² Patrick Seale, *Asad: The Struggle for the Middle East* (Berkeley: University of California Press, 1989), 340–41; Charles Tripp, *A History of Iraq* (Cambridge: Cambridge University Press, 2002), 224–26.

³ Owen, *The Rise and Fall of Arab Presidents for Life*, 40–53, 67–68, 81–82.

father, but that was only because exigent circumstances—the American invasion of Iraq in 2003 and the Arab Spring in 2011—prevented Qusay and Gamal from following suit. Thus, the radical republics that had once challenged the imperial order abroad, attacked privilege at home, and claimed to represent the general will had, by the late-twentieth century, been transformed into de facto dynastic monarchies headed by “presidents for life” dedicated above all else to keeping power in the hands of the ruling family.⁴

People might have tolerated the ruling classes’ corruption and even the autocratic nature of their regimes had those governments proven able to deliver the material improvements that they had promised. They could not. Instead, as the passage of time had so painfully revealed, the pan-Arabist regimes had proven to be as incompetent as they were corrupt. This reality was particularly obvious in Egypt. Neither Gamal Abdel Nasser’s (r. 1952-1970) much ballyhooed Arab Socialism nor Sadat’s program of economic liberalization, the *Infitah*, had succeeded in generating anything even remotely close to the living standards that the regime had assured would be forthcoming.⁵

For many, the pan-Arabist regimes’ inability to achieve the economic growth that they had promised was the last straw. Coupled with the undemocratic nature of the Arab-nationalist states and their endemic corruption, the failure to achieve sustained economic growth produced a precipitous decline in enthusiasm for the once-radical regimes beginning in the late 1970s and, more broadly, for the ideology of pan-Arabism that they claimed to champion. As a result, more and more people in the Arab world began to search for a political alternative.⁶

Increasingly, they found it in Islamism. Beginning in Egypt in the 1970s, the Islamists’ rejection of both secular government and the Western model of development in favor of a return to a social and political system based on Muslim values and laws began to resonate with people more and more disillusioned with the shortcomings and corruption of the officer states. By the early 1980s as a consequence, Islamism began to emerge as a potent political and ideological challenge to the status quo.⁷

⁴ Tripp, *A History of Iraq*, 286; Owen, *The Rise and Fall of Arab Presidents for Life*, 70–72, 83.

⁵ Arthur Goldschmidt Jr., *Modern Egypt: The Formation of a Nation State* (New York: Routledge, 2004), 170–71, 180.

⁶ Eugene Rogan, *The Arabs: A History – Revised and Updated Edition* (New York: Basic Books, 2012), 394–95.

⁷ Steven A. Cook, *The Struggle for Egypt: From Nasser to Tahrir Square*, Reprint edition (Oxford University Press, 2011), 125; Fawaz A. Gerges, *Making the Arab World: Nasser, Qutb, and the Clash That Shaped the Middle East* (Princeton: Princeton University Press, 2018), 310–11; Rogan, *The Arabs*, 394–95.

Mainstream Islamists

Moderate Islamist groups that sought to reform society from within were the prime beneficiaries of the ideology's growing appeal. This was certainly the case in Egypt. There, the revived—and now peaceful—Muslim Brotherhood experienced a surge in membership over the course of the 1980s as people who had lost faith in the secular state flocked to its banner. The organization worked tirelessly and successfully to translate its growing popularity into political power, winning an impressive thirty-six seats in the Egyptian parliament in 1987. It also moved to increase its social influence by gaining control of important civic bodies including the professional societies for doctors, lawyers, and scientists. By the late 1980s, as a result, the Muslim Brotherhood had not merely reestablished itself but had become a major social and political force.⁸

A shift in values and mores in Egypt over the course of the 1980s and 1990s reflected Islamism's surging popularity. Led by Egyptians who had absorbed Wahhabist ideas while working in Saudi Arabia, a steadily rising share of the population began to embrace traditional Islamic practices. Qur'anic recitations replaced music on the radio, alcohol consumption fell, and family photographs came down from their perches on walls and atop dressers. This change was most evidence in a return to longstanding Islamic gender norms. Men took to wearing traditional clothing and to sporting beards, while rising numbers of women embraced practices that had fallen out of favor among middle-class Egyptians during the mid-twentieth century such as gender segregation and, especially, veiling. This last change was impossible to miss. While only 30 percent of Egyptian women had covered their heads in the 1970s, fully 65 percent did so in the 1990s.⁹

Egypt was hardly alone in experiencing the Islamist revolution. Over the course of the 1980s and 1990s, a growing share of Arab Muslims abandoned secular mores in favor of Islamic practices and values. Islamic books grew in popularity, women who once favored Western modes of dress began to cover their heads and wear traditional Muslim attire, and, thanks to lavish Saudi funding, the number of mosques and the rate of mosque attendance also rose steadily throughout the region. Meanwhile, everywhere they were legal, Islamist political parties began to make big gains. Even Turkey, with its long secular tradition, witnessed the emergence of a potent Islamist political movement in the 1980s and 1990s. Its growth was so rapid, in fact, that the Islamist Welfare Party

⁸ Owen, *The Rise and Fall of Arab Presidents for Life*, 67; Gerges, *Making the Arab World*, 370.

⁹ Albert Hourani, *A History of the Arab Peoples* (Cambridge, Mass: Belknap Press of Harvard University Press, 2002), 442; Rogan, *The Arabs*, 437–38; Kim Ghattas, *Black Wave: Saudi Arabia, Iran, and the Forty-Year Rivalry That Unraveled Culture, Religion, and Collective Memory in the Middle East* (New York: Henry Holt and Co., 2020), 192–95.

managed to briefly win control of Turkey's government in 1996. It was a taste of things to come.¹⁰

Jihadis

Meanwhile, inspired by the ideas of the Egyptian writer Sayyid Qutb (1906-1966) and, after 1979, by the success of the Iranian Revolution, a small minority eschewed moderate groups like the Muslim Brotherhood and instead swelled the ranks of the militant Islamist organizations that had begun to emerge in the 1970s. The jihadis shared many of the same goals as the moderates but differed sharply with them over the best way to secure those ends. While mainstream Islamists hoped to achieve an Islamic government peacefully from within the constituted political system, the jihadis instead sought to impose Islamic rule through violence. In fact, these groups believed that individual Muslims were not merely permitted but were in fact *obligated* to overthrow the secular state—which the Egyptian theorist of radical Islam Muhammad Faraj (1954-1982) had dubbed the “near enemy”—as the first step toward establishing an Islamic polity governed by the shari‘a code. Only then, Faraj maintained, could the Muslim Middle East end its subordination to the “far enemy”—meaning the West—and regain its rightful place as the globe’s leading civilization.¹¹

Egypt once more led the way. Long the center of jihadism, the country already had a number of well-established extremist organizations at the start of the 1980s including Islamic Jihad and the “blind sheikh” Omar Abdel Rahman’s (1938-2017) *Gamaat Islamiya*. Dedicated to establishing an Islamist state through force, these groups had become bitterly disillusioned with Sadat’s government for its unwillingness to impose Islamism on the country. They seethed at his failure to implement the shari‘a code as he had promised and derided him as a hypocrite whose claim to be the “pious president” reflected little more than a cynical ploy to win popular support. They even railed at his willingness to permit his wife, Jihan (1933-2021), to engage in “immoral conduct” such as her televised dance with President Jimmy Carter (r. 1977-1981) at a White House function.¹²

The peace agreement with Israel was the final straw for the jihadis. Sadat’s de facto abandonment of the Palestinians and acceptance of Israeli rule of Jerusalem was simply too much for them. Accordingly, shortly after he signed the agreement with Israel, they began plotting his assassination. It did not take them long to succeed. During a televised military parade in 1981 celebrating the anniversary of the successful assault across the Suez Canal during the 1973 war, four soldiers affiliated with Islamic Jihad leapt from a troop transport and poured a hail of gunfire into the presidential

¹⁰ Erik J. Zürcher, *Turkey: A Modern History*, Third Edition (London: I.B. Tauris, 2004), 288–301; Hourani, *A History of the Arab Peoples*, 442.

¹¹ Fawaz A. Gerges, *The Far Enemy: Why Jihad Went Global* (Cambridge: Cambridge University Press, 2005), 9–11, 44.

¹² Gerges, 45–47.

reviewing stand. Riddled with bullets, Sadat died at the scene; lightly injured in the attack, his vice president, Husni Mubarak, succeeded him.¹³

The strength of the Egyptian jihadi organizations waxed and waned over the next few years. In the immediate aftermath of Sadat's assassination, the regime arrested so many jihadis that their organizations nearly collapsed. Thanks to the government's inexplicable decision to release most of them from prison in the mid 1980s, however, both *Gamaat Islamiya* and Islamic Jihad—now under the leadership of Ayman Zawahiri (1951-2022)—were able to reconstitute themselves in the late 1980s. They followed by initiating a brutal terror campaign in 1992 that targeted security personnel and foreign tourists, whom they blamed for bringing Western corruption to their country. To the shock of many Egyptians, they also attacked intellectuals including the prominent defender of secularism, Farag Foda (1946-1992), whom they killed in 1992, and the Pulitzer Prize-winning novelist, Naguib Mahfouz (1911-2006), who survived a stabbing in 1994. All told, more than 1,300 people died in the jihadis' wave of terror.¹⁴

Meanwhile, in Syria, the violence between the Asad regime and the jihadist Syrian Muslim Brotherhood that had begun in the late 1970s escalated dramatically in the early part of the 1980s. As we saw in chapter sixteen, its roots lay in Sunni dissatisfaction with the Ba'athist state. Utterly opposed to the idea of an Alawite-dominated secular government, the jihadis mounted a failed attempt to assassinate Asad in June 1980. The Ba'athist regime's response was brutal. Shortly after the attack, state security forces entered a prison in Palmyra and ruthlessly executed hundreds of Islamists detained there. This action did not intimidate the jihadis, however. Instead, dramatically upping the ante, they seized the city of Hama in February 1982 in a bid to precipitate a broad revolt against Asad. Though well executed, the plan ultimately failed. Thanks to the quick mobilization of the army, the regime was able to prevent the nationwide uprising that the brotherhood hoped to provoke. Under the command of his brother, Rifaat al-Asad (1937-), its forces mercilessly bombarded Hama with tanks, artillery, and missiles. To Hafiz al-Asad's satisfaction, the siege not only succeeded in retaking the city but also in breaking the Syrian Muslim Brotherhood and, thus, in ending the Islamist challenge to the government. The cost was enormous, however: the fighting flattened most of Hama and resulted in the deaths of between 5,000 and 25,000 people.¹⁵

Saudi Arabia

While the relationship between the jihadis and the government was clear in Syria, it was considerably more complex in Saudi Arabia and reflected the contradictions inherent in the Kingdom. On the one hand, the ruling house had long rooted its

¹³ Ghattas, *Black Wave*, 101–9.

¹⁴ Lawrence Wright, *The Looming Tower: Al Qaeda and the Road to 9/11* (New York: Vintage Books, 2006), 209–10.

¹⁵ John McHugo, *Syria: A History of the Last Hundred Years* (New York: The New Press, 2014), 189–94; quote from Seale, *Asad*, 322–23.

legitimacy in its embrace of the austere, Wahhabi version of Islam that had emerged in central Saudi Arabia in the eighteenth century. The people of Arabia thus lived under an Islamist system well before it would become fashionable in the rest of the Middle East. On the other, the thousands of members of the royal family had come to view the state as a vast ATM machine that they could use to fund a lifestyle of conspicuous consumption that frequently included *haram* activities such as gambling, drinking, and the intermixing of the genders. In other words, the often-very public behavior of the dynasty directly contradicted the Islamic values on which it claimed to base its rule. To obscure this inconsistency, the Saudi family had emphasized its role as the protector of the Holy Places; more importantly, it had made a point of courting the Wahhabi religious scholars and their supporters. It did so by providing patronage to the 'ulama' in the form of salaries and other perquisites, by using the state's enormous oil wealth to underwrite the spread of Wahhabism throughout the Muslim world, and by permitting the *mutawaa*—the religious police from the Committee for the Promotion of Virtue and Prevention of Vice—to enforce public morality. For decades, this arrangement had seemed to work. By promoting the Saudi family as the champion of Islam and by providing patronage to the 'ulama', it had ensured domestic stability and upheld the legitimacy of the royal family.¹⁶

The events of 1979 thus came as a great shock to the ruling house. The seizure of the Grand Mosque by extremist Muslims not only indicated that the Saudi dynasty's control of the kingdom was far-less secure than previously assumed, but, worse, also undermined its legitimacy by suggesting that it could not fulfill its role as the protector of the Holy Cities. Smelling blood, the regime's opponents were quick to pounce. At home, conservative 'ulama' and the *mutawaa* began to openly criticize the government and to demand greater autonomy and increased authority over morality and social laws. Internationally, meanwhile, the kingdom faced biting criticism from the newly empowered Ayatollah Khomeini (1902-1989). Publicly castigating the Saudis for their questionable morality, he inspired a wave of civil unrest among the Shi'i minority that lived in the kingdom's strategic, oil-rich Eastern Province. Worse, he mounted a frontal assault on the Saudi family's legitimacy by arguing that the protection of Mecca and Medina should henceforth become the collective responsibility of a consortium of Muslim states including Iran.¹⁷

The Saudis were in trouble, and they knew it. Faced with what they perceived to be an existential challenge, they responded with a vigorous three-pronged campaign aimed at blunting the regime's internal and external threats and at restoring its legitimacy. First, they moved to end domestic criticism by ceding to the 'ulama' and the *mutawaa* vastly greater control over social life. Newly empowered, the religious scholars promptly banned music, censored art, drove women from employment, and compelled the closure of the small number of cinemas that existed in the kingdom. Meanwhile, the

¹⁶ Ghattas, *Black Wave*, 52–53, 73–75.

¹⁷ Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 169; Vali Nasr, *The Shia Revival: How Conflicts within Islam Will Shape the Future* (New York: W. W. Norton, 2006), 150–52.

now-unrestrained *mutawaa* made liberal use of their swagger sticks to tighten enforcement of dress codes for women and to compel people to go to mosques at prayer times.¹⁸

Second, the Saudis sought to counter Iranian criticism among Sunnis who lived outside the kingdom by doubling down on their support for the spread of Wahhabi ideas. They did so in two ways. First, they paid for thousands of religious students to study in the kingdom under Wahhabi clerics. Second, the Saudis devoted a substantial share of oil revenue to the construction of a vast network of Wahhabi-run mosques, madrasas, and seminaries throughout the Muslim world. By the early 1990s, as a result, the Wahhabi religious scholars and their allies enjoyed near-unchecked power over social life in the kingdom.¹⁹

This exercise in soft power would have enormous consequences. In conjunction with Iran's efforts to promote Khomeini's extreme construction of Shi'ism internationally, it helped to precipitate a bitter religious cold war in which Wahhabi Saudi Arabia and Shi'i Iran competed for the hearts and minds of the globe's Muslim population. At first, that contest simmered quietly in the background. It boiled over with great suddenness in 1987, however, when clashes between Saudi security forces and Iranian pilgrims visiting Mecca during the hajj resulted in the deaths of more than 400 people. Blaming Khomeini, Riyadh broke relations with Tehran, increased markedly its financial support for Baghdad in the Iran-Iraq War (a topic we shall explore presently), and stepped up its efforts to spread Wahhabism. Iran countered by calling on Muslims to go to war with Saudi Arabia and by threatening to attack the Arab states that were supporting Iraq.²⁰

Finally, the third component of the Saudi campaign centered on Afghanistan. Like Washington, Riyadh viewed Moscow's action in that country in 1979 with alarm. Perceiving the invasion as the first step toward a potential move on the Strait of Hormuz, the Saudis were determined to prevent the USSR from solidifying its control of Afghanistan. Accordingly, in conjunction with the US, the dynasty provided billions of dollars' worth of military aid to the *mujahidin*, the Afghan guerrillas who were fighting Soviet troops with growing efficacy. With this aid—particularly with the introduction of portable American Stinger anti-aircraft missiles in 1986—the *mujahidin* were able to grind the Red Army down until, in 1989, Moscow had become so exhausted that it pulled its troops out of Afghanistan.²¹

While the Saudis financed the insurgency in Afghanistan primarily to achieve geostrategic goals, they were quick to grasp that doing so could also help them to achieve their ideological objectives. Indeed, supporting the anti-Soviet cause in Afghanistan appeared to provide them with a golden opportunity not only to prove their

¹⁸ Ghattas, *Black Wave*, 73–75; Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 167–70; Nasr, *The Shia Revival*, 155–57.

¹⁹ Ghattas, *Black Wave*, 163.

²⁰ Ghattas, 163.

²¹ Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 119–27.

Islamist bona fides and to restore their reputation as the protectors of the Holy Cities but also to divert jihadi rage away from the royal family and toward a new, external enemy. They moved to achieve this last end by providing financial backing to a new Islamist organization dedicated to fighting the Soviets called the Service Bureau. Operating out of Peshawar, Pakistan by the 'ulama' Abdullah Azzam (1941-1988) and his protégé, a young Saudi man named Osama bin Laden (1957-2011), it recruited and trained Arab Islamists to fight alongside the Afghans.²²

From a narrowly military perspective, the return on this investment was quite poor. Despite receiving substantial funds from both the Saudi government and the kingdom's private citizens, the Service Bureau and its "Arab Afghan" recruits—as the fighters it trained were known—were of negligible consequence in the war with the Soviets. For the Saudi regime, however, the military efficacy of the Arab Afghans was beside the point. It had not lavished funds on the Service Bureau for the purpose of defeating the USSR—the *mujahidin* could take care of the Soviets—but had instead done so in order to divert the ire of militant Islamists away from the royal family and toward a new, external enemy. Judged according to that yardstick, support for Azzam and Bin Laden was a smashing success. It helped to restore the government's Islamist reputation at an affordable cost, and it removed a substantial number of potentially dangerous people from the kingdom during a sensitive period. To the Saudis, in sum, it had been money very well spent.²³

There was, however, a potential downside to Riyadh's support for the Service Bureau. Diverting jihadis to Afghanistan via the organization had certainly redirected the focus of potential troublemakers away from criticizing the Saudi government and had thus served the dynasty's immediate interests. At the same time, however, support for the Service Bureau had helped to create a cadre of well-organized, battle-hardened extremists who, having helped defeat the Soviets, might next turn their ire on Saudi Arabia and the other governments of the Middle East. That was a problem for the future, however. In the meantime, the dynasty could celebrate the success of its Afghan policy.²⁴

The Lebanese Civil War, 1980-1985

Though the war in Afghanistan remained a major issue during the 1980s, renewed fighting in Lebanon overshadowed it for a period during the early part of the decade. Remarkably, in fact, Lebanon would experience even more difficulties during the early-to-mid 1980s than it had during the late 1970s. Not only was the civil war still smoldering, but Israel would invade the country in 1982 as the centerpiece of an audacious and reckless play to take advantage of the peace treaty with Egypt to permanently transform the strategic map of the Middle East in its favor. The brainchild

²² Gerges, *The Far Enemy*, 67–68, 75–76.

²³ Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 121; Gerges, *The Far Enemy*, 67–68, 75–76.

²⁴ Gerges, *The Far Enemy*, 69–70.

of Prime Minister Menachem Begin (r. 1977-1983) and Defense Minister Ariel Sharon (1928-2014), the Israeli gambit appeared for a time to be a huge success. It proved ultimately to be a spectacular failure, however, one that would dramatically—if unintentionally—alter the course of the Arab-Israeli conflict.

By the end of the 1970s, the civil war in Lebanon had become enormously complex. Replete with ambushes, massacres, assassinations, and double dealing, it had devolved into a byzantine struggle in which rival confessional factions entered into and then abandoned alliances with dizzying regularity as each struggled to carve out an autonomous position within the country. Dominating the situation like a spider was Asad. Seeking to exploit the shifting balance in order to render Syria's dominance of Lebanon permanent, he had skillfully manipulated events in the country to his advantage. Most notably, while he had intervened in 1976 to prevent a coalition of leftist Muslims and the Palestine Liberation Organization (PLO) from defeating the Christians, he had—in Machiavellian fashion—promptly shifted sides when the balance of power had begun to tip back toward the Maronites. Now once more in danger of being overwhelmed, the leader of the Christian Phalange Party, Pierre Gemayel (1905-1984), responded by entering into a tacit alliance with Israel in which Tel Aviv provided his forces with a substantial infusion of much-needed arms.²⁵

Perhaps unsurprisingly, Israel's involvement in the Lebanese Civil War went considerably further than merely providing weapons to the Phalange. With the PLO taking advantage of the instability in the country to launch terrorist raids and to periodically pummel the Galilee region with artillery and rockets, Tel Aviv had sponsored the creation of a proxy force called the South Lebanese Army (SLA) in 1976 to protect northern Israel. Equipped with arms provided by the Israeli Defense Force (IDF), the SLA maintained a security zone along the border designed to buffer Israel from PLO infiltration and attacks. It was only partially successful in achieving that end, however. It could neither prevent all incursions nor inhibit bombardment by the Palestinians' long-range weapons. Accordingly, in keeping with the ideology of the Iron Wall, the IDF periodically undertook operations in Lebanon designed to weaken or destroy the PLO. Most notably, it used a deadly terrorist attack on a bus near Haifa in March 1978 as a pretext for a large-scale invasion aimed at eliminating the organization's military infrastructure in southern Lebanon. To Tel Aviv's frustration, however, the incursion failed to degrade PLO capabilities.²⁶

Even worse from the perspective of Israel's Likud government were the results of a new round of fighting between Israeli and PLO forces along the border with Lebanon in the summer of 1981. What began as a series of artillery duels quickly escalated into IDF air attacks on the PLO's headquarters in Beirut and massive Palestinian retaliatory bombardments of Israeli settlements in the Galilee region. Eager to prevent the violence from spinning further out of control, Washington sent the veteran envoy Philip

²⁵ Benny Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920* (New York: Vintage Books, 2001), 501–5.

²⁶ Kirsten E. Schulze, *The Arab-Israeli Conflict* (New York: Pearson, 2008), 56–57.

Habib (1920-1992) to the Middle East to broker a truce. Negotiating with the Palestinians through intermediaries, he secured an agreement in July 1981 that called for the two sides to refrain from further attacks across the Lebanese border. The ceasefire thus succeeded in ending the bombardment of Israeli settlements in the Galilee; to the frustration of Begin and his political allies, however, the deal also raised the PLO's international standing and, worse, added fuel to Saudi and European peace initiatives that sought the establishment of a Palestinian state on the West Bank and Gaza Strip.²⁷

This situation was deeply troubling to Begin and Sharon. Like the rest of the Likud Party, the two men firmly backed a maximalist interpretation of Israel's territory and insisted on retaining the West Bank. Indeed, after the party's electoral victory in 1981, Begin's cabinet had, in contravention of Resolution 242 and a bevy of other agreements, formally annexed the Golan Heights and declared that the claim "of the Jewish people to the Land of Israel [is] an eternal unassailable right"—meaning, in other words, that Israel had a divinely endorsed claim to the West Bank. Unsurprisingly, given these sentiments, Begin and Sharon viewed the peace initiatives and the PLO's adherence to the ceasefire with grave concern. So long as Yasir Arafat (1929-2004) complied with the truce, pressure for a settlement in which the Palestinians would secure a state in the Occupied Territories would continue to build while, in inverse proportion, the prospects for permanent Israeli control of the West Bank would diminish.²⁸

Operation Peace for Galilee, 1982

If Begin and Sharon saw danger in Lebanon, they also sensed opportunity. Indeed, by the fall of 1981, the two men had come to perceive that country as a venue for a bold military campaign in Lebanon that they believed had the potential to permanently remake the strategic and diplomatic landscape of the region in Israel's favor. Grand in scope, their design would not only permanently end the PLO threat in Lebanon, but, more importantly, would also permit them to achieve all of the Likud Party's expansive territorial and security goals.

Their complex plan involved four steps. First, it called for the IDF to launch a fast-moving offensive into Lebanon that would maul Syrian forces in the strategic Beqaa Valley and, after surrounding Beirut, isolate the PLO's leadership and the bulk of the organization's fighters in Muslim West Beirut. Second, in exchange for the Phalangist militia entering West Beirut and destroying the PLO, Israel would use its power and influence to ensure that the Lebanese parliament selected Pierre Gemayel's son, Bashir (1947-1982), as the country's new president during elections scheduled for August

²⁷ Charles D. Smith, *Palestine and the Arab-Israeli Conflict: A History with Documents*, Sixth Edition (Boston: Bedford/St. Martin's, 2007), 378–79.

²⁸ Avi Shlaim, *The Iron Wall: Israel and the Arab World, Revised and Expanded* (New York: W. W. Norton, 2014), 400–405; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 508–9.

1982. Fourth, beholden to Israel, Lebanon's new president would sign a peace treaty that would establish normalized diplomatic relations with Israel. Bolstered by the IDF, finally, Bashir Gemayel would restore his government's authority over Lebanon and, in so doing, inflict a serious diplomatic defeat on Asad—one that would bring a permanent end to Syria's dominance of his country.²⁹

The result, Begin and Sharon hoped, would be a thoroughly reconfigured Middle East—one that would permit the Likud Party both to permanently ensure Israel's security needs and to complete the Zionist project begun a century earlier. Coupled with the peace treaty with Egypt, Syria's military and diplomatic defeat in Lebanon would mean that Israel would never again have to worry about its security. More importantly, the destruction of the PLO would so thoroughly demoralize the Palestinians as a people that it would permanently extinguish their sense of nationalism. As a result, Israel would be able to annex the West Bank—territory that the Likud Party viewed as an integral part of the Jewish state.³⁰

Breathtaking in scope, the plan faced several substantial obstacles. First, more moderate members of the cabinet had already made clear when Begin and Sharon had raised the idea of another invasion that they flatly opposed military operations that extended as far north as Beirut and would only consider authorizing an incursion in response to a major Palestinian attack. Second, though broadly supportive of Israel, the administration of President Ronald Reagan (r. 1981-1989) had shown itself to be squeamish about the kind of civilian casualties that would inevitably result from an operation along the lines that Sharon and Begin were proposing and had already indicated to Tel Aviv that it would tolerate an Israeli violation of the ceasefire only if it came in response to a serious PLO attack and only if it extended no further than forty kilometers from the border.³¹

Securing cabinet and American approval for an invasion of Lebanon thus depended on the PLO first launching an attack on Israel. Normally, Sharon and Begin could count on the organization obliging with a pretext in the form of sporadic shelling or a terrorist incursion. To their dismay, however, the PLO was pursuing a peace offensive at that time and was scrupulously adhering to the ceasefire. As a result, during the winter and spring of 1982, Begin and Sharon found themselves thoroughly boxed in.³²

To their relief, a Palestinian splinter group freed them to act by mounting a high-profile attack on Israeli interests in the spring of 1982. On June 3, the radical, Iraqi-backed Abu Nidal Group, a bitter opponent of Fatah, critically wounded the Israeli

²⁹ Shlaim, *The Iron Wall*, 406–7.

³⁰ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 509.

³¹ Peter L. Hahn, *Missions Accomplished?: The United States and Iraq Since World War I* (New York: Oxford University Press, 2012), 74–75; Smith, *Palestine and the Arab-Israeli Conflict*, 379–80; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 514.

³² Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 509–10.

ambassador to Britain in a botched assassination attempt outside London's posh Dorchester Hotel in what was likely an attempt to undermine Arafat by derailing the ceasefire. Seizing the opportunity that the incident presented, Sharon and Begin asked the cabinet to authorize a massive, 120,000-troop invasion that they disingenuously named Peace for Galilee. Their colleagues approved the operation; however, they did so only after securing promises from Begin and Sharon that the incursion would be a limited one.³³

Perhaps surprisingly, this stipulation did not faze the two men. Instead, confident that they would be able to use military exigencies to manipulate the cabinet into expanding the war, they accepted the terms without complaint. They were correct in this assumption. Taking advantage of the inevitable clashes between IDF and Syrian troops, they had little difficulty over the next few days in persuading the cabinet to approve the steady widening of the conflict.³⁴

As a result, Begin and Sharon's plan proceeded as intended. IDF forces smashed Asad's divisions in the Beqaa Valley, brushed aside the PLO's main-force battalions in the south and raced northward to seize the Beirut-Damascus Highway—an objective that lay well beyond the forty-kilometer limit they had discussed with the US. Furious with the Israelis for ignoring the conditions that he had set, Reagan sent Philip Habib back to the region. Able once again to quickly impose a ceasefire, he soon brought the invasion to a halt. Before the truce could go into effect, however, IDF troops were able to link up with Gemayel's forces and complete the encirclement of Beirut. Thus, despite significant domestic and international obstacles, Begin and Sharon's bid to remake the region was unfolding precisely as they had intended.³⁵

Evacuation of the PLO

It was at that point, however, that their plan ran into its first problem. Aware that an assault on PLO and Sunni militia positions in West Beirut would not only incur massive casualties but would also undermine the credibility of his presidency before it even began, Bashir Gemayel reneged on an earlier promise and informed the Israeli government that the Phalange militia would not lead the assault on the PLO. This decision put Sharon and Begin in a bind. Like Gemayel, they too understood that assaulting West Beirut would result in a politically unsustainable level of casualties. Accordingly, they shifted from seeking to destroy the PLO outright to attempting to force it out of Lebanon. To achieve that new goal, they ordered the IDF to impose a blockade

³³ Shlaim, *The Iron Wall*, 414–21.

³⁴ Rashid Khalidi, *The Hundred Years' War on Palestine: A History of Settler Colonialism and Resistance, 1917–2017* (New York: Metropolitan Books, 2020), 143; Shlaim, *The Iron Wall*, 414–21.

³⁵ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 515–31.

on Muslim West Beirut and to periodically shell it—a move that resulted in thousands of civilian casualties over the course of a siege that ended up lasting seven weeks.³⁶

The mounting death toll was too much for the Reagan administration. Keenly aware of the damage that the casualties were causing to Israel and the United States's international reputation, the president once again sent Philip Habib back to the region—this time with instructions to broker a settlement that would safely relocate the PLO from West Beirut. The subsequent talks were characteristically fraught and drawn out, but, on August 12, Habib finally secured an agreement. It called for the PLO to leave Lebanon by ship under the protection of a Multinational Force (MNF) composed of US, French, and Italian troops.³⁷

The operation went off without a hitch. Beginning on August 21, 11,000 Palestinian officials and soldiers left Beirut for Tunisia aboard Greek and Italian ships escorted by the US Sixth Fleet. A few days after the last contingent of PLO fighters had departed, the soldiers of the MNF followed suit. Returning to their ships under a banner that read "Mission Accomplished," they left Lebanese waters in early September.³⁸

Sharon and Begin were now ready to put the rest of their grand design in motion. The next step was to ensure that Bashir Gemayel became Lebanon's new president. To do so, they let Lebanese members of parliament know that peace would return to their country only if it had a government that enjoyed Israel's approval, and they instructed the IDF to help friendly deputies travel to Beirut to take part in the vote and to prevent hostile ones from making it to the capital. This assistance tipped the balance and allowed Bashir Gemayel to handily clear the two-thirds majority needed to become Lebanon's next president. As a result, by early September 1982, Begin and Sharon's plan appeared to be back on track. Lebanon was in friendly hands, the PLO was now incapable of attacking Israel thanks to its relocation to distant Tunisia, and Syria was isolated and weakened.³⁹

Sabra and Shatila

At that point, however, the wheels came off their plan to remake the region. Its demise began on September 14 with the assassination of Bashir Gemayel by a Syrian intelligence agent just a few days before he was to be sworn in as president. In his place, the Lebanese parliament elected his brother, Amin Gemayel (1942-) as the country's new head of state. To Begin and Sharon's chagrin, Amin intended to follow a very different approach than had his brother. Most notably, despite Israel's support for Amin's candidacy, the new president made clear once he was sworn in that he was

³⁶ Shlaim, *The Iron Wall*, 421–23.

³⁷ Schulze, *The Arab-Israeli Conflict*, 59; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 536–38.

³⁸ Andrew J. Bacevich, *America's War for the Greater Middle East: A Military History* (New York: Random House, 2016), 66.

³⁹ Shlaim, *The Iron Wall*, 425–26.

committed to a neutral, if not pro-Syrian line, and that he opposed closer ties with Tel Aviv. All at once, in other words, the ground on which Begin and Sharon's plans for Lebanon rested—a peace treaty, a tacit alliance, and diplomatic recognition—had collapsed.⁴⁰

Begin and Sharon took desperate action to recover the situation. Fearing that a force of 2,000 PLO fighters that he believed had secretly remained behind in the Sabra and Shatila refugee camps in West Beirut might further destabilize Israel's position in Lebanon, Sharon immediately ordered troops into the city. Rapidly occupying West Beirut, IDF units surrounded but did not occupy the camps. Instead, on September 16, they permitted a force of 150 Phalange militia members to enter the camps in order to clear out fighters alleged to have “gone civilian” when the PLO had evacuated the city. Once inside, however, the Phalangists made no effort whatsoever to uncover a clandestine military apparatus or to ferret out guerrillas. Instead, blaming the Palestinians for Bashir Gemayel's death, they set about methodically massacring civilian residents of the camps. Over a roughly thirty-six-hour period, they raped countless Palestinian women and murdered between 700 and 2,000 residents of the camps. The vast majority of the victims were women and children.⁴¹

News of the massacres destroyed what little credibility the invasion of Lebanon had managed to retain. Internationally, the incident greatly damaged Israel's reputation and spurred mounting pressure on Tel Aviv to withdraw from Lebanon. Domestically, IDF involvement in the killings sparked sustained opposition to a military operation for the first time in the country's history. Most notably, a rally that the antiwar group Peace Now organized to protest both the massacre and the ongoing occupation of Lebanon drew 300,000 Israelis—equivalent to 10 percent of the country's population—and succeeded in forcing the government to appoint a panel headed by the President of the Israeli Supreme Court, Yitzhak Kahan (1913-1985), to investigate the killings. Released in February 1983, the Kahan Commission report concluded that the IDF should have known that the Phalangists would seek revenge for Bashir Gemayel's assassination and that it thus bore indirect responsibility for the killings. It assigned special responsibility to Ariel Sharon, who, though he remained minister without portfolio in the cabinet, was compelled to step down as defense minister.⁴²

American Intervention, 1982-1984

Meanwhile, the revelation of the massacre resulted in international calls for the MNF to return to Lebanon and to help restore order in the capital. In response, the recently departed MNF soldiers disembarked from their ships on September 29 and took up new positions around the airport and in West Beirut. Reagan was enthusiastic about the redeployment. Eager to bring peace to Lebanon, he was convinced that their

⁴⁰ Shlaim, 427–28; Schulze, *The Arab-Israeli Conflict*, 60.

⁴¹ Schulze, *The Arab-Israeli Conflict*, 59–60.

⁴² Rogan, *The Arabs*, 417–18; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 548.

presence would both return security to the city and help the Lebanese government restore its authority over the capital and the territory immediately surrounding it. Many were dubious of his optimism, however. Indeed, critics both inside and outside the administration were skeptical that a force of only a few-thousand soldiers could achieve such a broad objective and cautioned that involvement in the Lebanese Civil War would likely lead the combatants to begin targeting MNF soldiers. Unheeded, those warnings would prove prescient.⁴³

Reagan followed by pursuing a diplomatic settlement aimed at ending Israel's occupation of Lebanon. Here, the administration abandoned any pretense of neutrality. Viewing the world through a Cold War lens, it tacitly took Israel's side in the talks and worked to produce an agreement that centered primarily on curbing Syria's position in Lebanon. While it did not include formal diplomatic recognition, the resulting treaty was nonetheless very favorable to Tel Aviv and seemed to salvage key parts of Begin and Sharon's grand design. Signed in May 1983, it ended the technical state of war that had existed between the two countries since 1948 and required Lebanon to prevent attacks against Israel from its territory. It even permitted Tel Aviv to maintain its ties with the SLA and thus to retain its grip on south Lebanon. In exchange, Israel agreed to pull its troops out—though, importantly, only after Syria had completely withdrawn its forces.⁴⁴

Predictably, the treaty met with an icy reception in both Lebanon and Syria. The Lebanese Druze, Shi'a, and Sunni factions resented what they saw as an Israeli effort to impose a one-sided agreement and began to perceive the US not as a neutral party but instead as a hostile combatant. For his part, Asad interpreted the deal as an effort by Washington and Tel Aviv to isolate his country and to compel it to abandon the dominant position in Lebanon that he had so carefully constructed over the prior seven years. Determined to retain his gains, he responded by unleashing his Lebanese proxies against the IDF and the MNF. Accordingly, in the summer of 1983, insurgents began to step up their use of hit-and-run raids, ambushes, and roadside bombings against Israeli troops, and, ominously, began to harass the marines and soldiers of the MNF with sniper and artillery fire. The Reagan administration found these attacks intolerable and responded by ordering Sixth Fleet warships to bombard Druze emplacements and by deploying the intimidating World War II-era battleship USS *New Jersey* to Lebanese waters as a show of strength. To the delight of the president, the use of force appeared to work: shortly after the navy turned its five-inch guns on them, the Druze called for a ceasefire.⁴⁵

In reality, the bombardment had not cowed the Lebanese factions in the least. Instead, it merely caused them to shift tactics. The result was a pair of horrifying, near-simultaneous truck-bombing attacks by a shadowy new group on October 23: one that killed fifty-eight French paratroopers and destroyed the nine-story building that served as their headquarters and an even-bigger one that killed 241 marines and obliterated

⁴³ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 549.

⁴⁴ Smith, *Palestine and the Arab-Israeli Conflict*, 383–84.

⁴⁵ Rogan, *The Arabs*, 419–20; Smith, *Palestine and the Arab-Israeli Conflict*, 384.

their four-story barracks in what remains the single deadliest day the Marine Corps has experienced since the Battle of Iwo Jima. Washington was in shock. Reagan initially favored keeping the marines in Lebanon and claimed that leaving would damage American credibility. Faced with staunch opposition from within his administration, however, he relented and agreed to pull US forces out in February 1984—though not before demonstrating American resolve by ordering the USS *New Jersey* to pummel Syrian and Druze positions with barrages of 1,900-pound, sixteen-inch shells.⁴⁶

Israel was not far behind. With IDF troops facing an average of more than one-hundred Shi'i attacks per month by the middle of 1984, Lebanon had become a quagmire that was growing more challenging and dispiriting by the day. Indeed, the situation had become so dangerous in early 1985 that infantry units had to have tanks accompany them when they went out on even routine patrol. The end finally came when the Labor Party's Shimon Peres (r. 1977, 1984-1986, 1995-1996), became prime minister in 1984. Peres was determined to end Israel's occupation of Lebanon, but, as the head of a national unity government, was compelled to move slowly. As a result, it was only in early 1985 that he managed to win enough votes in his cabinet for a withdrawal to the security zone just north of the Israeli border. Completed in June 1985, the move formally ended the war that Begin and Sharon had started with such anticipation three years earlier.⁴⁷

Consequences

The war in Lebanon ultimately proved to be a costly disaster for Israel. Sharon and Begin had launched it in order to restructure the strategic and diplomatic landscape of the Middle East in Israel's favor and to set the stage for the annexation of the West Bank. It had failed to achieve either goal. Instead, their gambit had not only left their country less secure than it had been before the invasion but had also greatly complicated their effort to annex the West Bank.

Ironies abounded. First, the Israeli invasion ended up solidifying Asad's control over Lebanon—an outcome that would have beggared belief in the summer of 1982. As we have seen, Syria had suffered a humiliating military setback at the start of Operation Peace for Galilee that had made emphatically clear that its army could not compete with the mighty IDF. Its defeat had been so total, in fact, that most observers had concluded that its influence over Lebanese affairs was at an end. Asad was unfazed, however. Rather than abandoning his position in Lebanon in the face of superior Israeli power, he shifted to a new strategy that relied on Syria's Lebanese proxies to achieve his ends. As we have seen, that approach proved to be wildly successful. Not only did it force the United States and Israel out of Lebanon, but it also compelled Amin Gemayel to abrogate the May 1983 agreement with Tel Aviv and to accept Syrian dominance of his

⁴⁶ Bacevich, *America's War for the Greater Middle East*, 72–73; Ghattas, *Black Wave*, 138.

⁴⁷ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 552–58.

country. Thus, ironically, a conflict designed in part to push Asad out of Lebanon had instead ended with him tightening his grip over it.⁴⁸

Second, the war ended with Israel confronting a new enemy in Lebanon: the Shi'i Islamist group Hizbullah. Prior to 1982, the Shi'a had resented the PLO's domination of Lebanon's south; as a result, they had viewed the IDF soldiers who arrived that June as liberators and even showered them with rose petals and rice—a traditional Lebanese greeting. That sentiment did not last, however. Instead, a combination of heavy-handed Israeli rule and the arrival of Iranian agents quickly turned the Lebanese Shi'i against the occupying army. Tehran's efforts were particularly important in sparking this shift. Determined to expand its influence, it built schools, subsidized goods such as gasoline, spread Khomeini's Islamist ideology, and provided training in military and terrorist operations to a growing number of recruits eager to strike back at Israel.⁴⁹

The result was the establishment in mid-1982 of a new terrorist organization, Islamic Jihad. Despite the traditional Muslim ban on killing oneself, the group initially focused on attacking Israeli and Western forces with suicide attacks. It began in November 1982, when a suicide bomber—the first in the history of the modern Middle East—killed seventy-five IDF soldiers by driving a bomb-laden truck into their command post in Tyre. It followed with the bombing of the US embassy in April 1983 and then shocking attacks on the Multinational Force the following October. Islamic Jihad also kidnapped and killed a number of Westerners in Lebanon including the president of the American University in Beirut, Malcolm Kerr (1931-1984).⁵⁰

Later, under Iranian guidance, Islamic Jihad would evolve into Hizbullah, or "Party of God," a political party with a powerful militant wing. Highly effective, its operations played a key part in impelling the IDF to retreat to the south in 1985. Hizbullah fighters did not stop their attacks following the Israeli withdrawal to the security zone, moreover. Instead, dedicated to pushing Israel out of all of Lebanon, they continued to harass IDF forces in the south and periodically fired rockets at settlements in the Galilee. Thus, in a bitter irony for Israel, the invasion that Begin and Sharon had designed in part to end PLO attacks across the Lebanese border had merely served to replace Arafat's guerrillas with far-better disciplined and equipped Hizbullah fighters who were able to launch similar strikes with far-greater efficacy.⁵¹

Third, the PLO's military defeat in Lebanon had failed to produce the collapse in morale among the Palestinians of the Occupied Territories that Begin and Sharon had anticipated. On the contrary, frustrated by the inability of the Arab states and the PLO to deliver them from the Israeli occupation, the people of the Gaza Strip and the West Bank had come to adopt a new resolve in the wake of the war in Lebanon that would soon set the stage for the most significant progress that the Palestinians had achieved

⁴⁸ Rogan, *The Arabs*, 419–20; Shlaim, *The Iron Wall*, 433–34.

⁴⁹ Rogan, *The Arabs*, 420–23.

⁵⁰ Khalidi, *The Hundred Years' War on Palestine*, 166–67; Rogan, *The Arabs*, 421–22.

⁵¹ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 556–59.

since before the *Nakba* in 1948. Thus, ironically, Operation Peace for Galilee had rendered Begin and Sharon's cherished goal of annexing the West Bank more rather than less difficult.⁵²

Finally, the invasion altered dramatically the nature of the Palestinian-Israeli struggle. It is true that Operation Peace for Galilee had succeeded in destroying the military infrastructure with which the PLO had threatened northern Israel. However, the organization had always been so feeble compared to the IDF that it could *never* have secured its objectives on the battlefield. In fact, as we have seen, nearly all the PLO's success had come through diplomacy rather than force. Thus, by compelling the organization to relocate to distant Tunisia, Operation Peace for Galilee weakened the PLO hardliners who supported military operations while simultaneously strengthening the moderates who backed negotiations—in the process setting the organization on a course that was far-more likely to succeed.⁵³

The Palestinian Revival

Initially, Arafat's efforts to pursue diplomacy following the PLO's relocation to Tunisia proved unsuccessful. Two obstacles stood in his way. First, enjoying strong US support, Israel remained implacably opposed not merely to making territorial or political concessions, but even to taking part in negotiations with the PLO. Second, Arafat's diplomatic initiatives faced concerted hostility from the Palestinian splinter groups. Indeed, those organizations were so determined to regain *all* of the territory that had composed mandatory Palestine that they deliberately acted to poison the waters anytime Arafat's diplomatic push seemed to be making headway. For example, when a joint Palestinian-Jordanian initiative appeared to show promise in 1985, the People's Front for the Liberation of Palestine (PFLP) moved to scotch it by hijacking an Italian cruise ship, the *MS Achille Lauro*, and murdering one of its passengers, a disabled Jewish-American man named Leon Klinghoffer (1916-1985).⁵⁴

By 1986, even the Arab states were distancing themselves from Arafat. Egypt's rehabilitation in the Arab world made this shift clear. Concluding that Saddam Husayn desperately needed Egypt's material and diplomatic support in his war with Iran, the Arab League states renewed diplomatic relations with Mubarak in 1987 over vigorous PLO objections. Demoralizingly, when Arafat traveled to Jordan to object, King Husayn (r. 1952-1990) did not have the PLO chair greeted with the pomp and circumstance he had enjoyed in the past. Instead, in a calculated snub that reflected

⁵² Smith, *Palestine and the Arab-Israeli Conflict*, 406.

⁵³ Khalidi, *The Hundred Years' War on Palestine*, 165; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 558.

⁵⁴ Smith, *Palestine and the Arab-Israeli Conflict*, 410.

Arafat's declining influence, the king sent a mere subcabinet-level official to pick him up at the airport.⁵⁵

The circumstances in which the Palestinians in the Occupied Territories found themselves paralleled the PLO's waning fortunes. Initially, Israeli rule had been unwelcomed but tolerable and even—in some ways—beneficial for the residents of the Gaza Strip and West Bank. Now free to work in Israel, for example, the Palestinians were able to access substantially higher-paying jobs than they had under Jordanian or Egyptian administration. Meanwhile, though the Israelis were building settlements, they were doing so at a low rate and had made a point of situating them well away from Arab towns.⁵⁶

Life under Israeli rule changed dramatically with the Likud victory in 1977, however. With the goal of preventing future cabinets from ceding land or agreeing to a Palestinian state in the Occupied Territories, the Begin government dramatically stepped up the pace of settlement construction; as a consequence, the 5,000 Israeli settlers who lived in the West Bank and Gaza Strip in 1977 ballooned to 60,000 by 1987. The Likud government also imposed increasingly obnoxious and humiliating regulations on the Palestinians. They affected travel, land use, and employment, and included water and agriculture policies so restrictive that they resulted in the Palestinians farming fewer acres in 1987 than they had in 1947. Meanwhile, the onset of a deep recession in the early 1980s produced soaring unemployment and a sharply declining standard of living in the territories that further immiserated the Palestinians. By 1987, as a consequence, the Occupied Territories were a place of barely contained frustration and fury—a tinderbox ready to burst into flames.⁵⁷

The Intifada

The match that ignited it came in the form of a traffic accident at a checkpoint between the Gaza Strip and Israel. On December 8, 1987, an IDF driver killed four Palestinians when his transport slammed into a line of cars carrying workers back from their jobs in Israel. With frustration already at a fever pitch, rumors that the crash was intentional led thousands of mostly young Palestinians to explode in a spontaneous uprising in the Gaza Strip immediately after the victims' funerals. Despite their anger, the protestors neither engaged in deadly attacks nor targeted settlements. Instead, seeking to put maximum moral pressure on Israel, they agreed to confine themselves to non-lethal acts including demonstrations, rock throwing, and a refusal to pay the taxes that supported the occupation. Coming at a time of enormous anger in the territories, the uprising—soon known as the *Intifada* or “shaking off”—rapidly spread to the West

⁵⁵ James L. Gelvin, *The Israel-Palestine Conflict: One Hundred Years of War* (New York: Cambridge University Press, 2005), 214.

⁵⁶ Gelvin, 185.

⁵⁷ Schulze, *The Arab-Israeli Conflict*, 69; Gelvin, *The Israel-Palestine Conflict*, 189–91, 215.

Bank and East Jerusalem. As a result, within just a few short weeks, the entirety of the Occupied Territories was in open revolt against the Israelis.⁵⁸

Sensing that the *Intifada* could produce tangible gains, the Palestinians in the Occupied Territories quickly moved to organize themselves to ensure its effectiveness. This effort involved a division of labor by age. While the huge number of young people in the Occupied Territories threw rocks and baited Israeli security forces, more moderate, middle-aged community leaders established a new body called the Unified National Leadership (UNL) to coordinate the uprising. Convinced that the Palestinians would enjoy greater success if they avoided causing bloodshed, the UNL took steps to ensure that the protestors continued to eschew outright violence. It also drafted and released a manifesto—the Fourteen Points—that formally articulated the *Intifada's* political goals. Moderate in tone and realistic in objective, it called for the establishment of an independent Palestinian state under PLO administration in the Gaza Strip and West Bank that would live in peace alongside Israel.⁵⁹

The Israelis badly mishandled the protests. The unity government headed by Prime Minister Yitzhak Shamir (r. 1986-1992) believed that Israel needed not only to restore order in the territories but, more importantly, to crush any hope among the Palestinians that the *Intifada* might produce a favorable political settlement for them. In keeping with this approach, Defense Minister Yitzhak Rabin (1922-1995) instructed the security forces to employ “‘might, force, and beatings’” and even to break the arms of children who threw rocks. Accordingly, the IDF enforced curfews, arrested thousands, and fired rubber bullets at unarmed teenagers. It also closed schools and universities in the territories and demolished the homes of those suspected of throwing stones. It was an over-the-top response, and it imposed enormous costs on the Palestinians. Indeed, during its first year alone, Israeli security forces killed 626 Palestinians, wounded 37,000, and jailed 35,000 more.⁶⁰

Israel's response may have imposed steep costs on the Palestinians, but it failed to cow them. Indeed, the use of overwhelming force backfired badly during the *Intifada's* early years both inside and outside the territories. In the Gaza Strip and West Bank, the resort to massive repression and reliance on collective punishment not only failed to break the resistance but also served to create a strong sense of solidarity that inspired less-committed Palestinians to join the protests. Internationally, televised footage of IDF troops firing on and beating unarmed Palestinian teenagers rapidly transformed Israel's image from that of a plucky David into that of an overbearing Goliath. The violent response even went over poorly within the Jewish state. The use of overwhelming force against unarmed Palestinians reinvigorated opposition groups such as Peace Now and demoralized soldiers who had joined the IDF under the expectation that they would be protecting Israel from foreign attack rather than violently policing civilians. In fact, during

⁵⁸ Smith, *Palestine and the Arab-Israeli Conflict*, 412, 419.

⁵⁹ Smith, 420–22; Gelvin, *The Israel-Palestine Conflict*, 217–19.

⁶⁰ Rogan, *The Arabs*, 433–37.

the *Intifada*'s first seven months alone, over 600 soldiers claimed conscientious objector status.⁶¹

The Palestinian Declaration of Independence

At the same time that the *Intifada* was posing serious challenges to the Israeli government, it was throwing a lifeline to the beleaguered PLO. Arafat was quick to take advantage. He first used the UNL's call for a Palestinian state in the territories as a way to marginalize the inflexible militants in the splinter groups. With the extremists safely contained, he followed by orchestrating a pair of bold moves designed to advance the prospects of a two-state solution in which the Palestinians would have a state in the West Bank and Gaza Strip that stood alongside a Zionist one within Israel's pre-1967 borders. First, in November 1988, he arranged for the Palestine National Council (PNC) to issue the Palestinian Declaration of Independence and to express its willingness to negotiate on the basis of United Nations (UN) Resolutions 242 and 338. Second, at American urgings, he formally and publicly renounced terrorism in December 1988. These announcements marked the culmination of the diplomatic approach that he had cautiously initiated in the 1970s, and they quickly bore fruit. Indeed, eighty-four countries including the People's Republic of China and the USSR promptly extended diplomatic recognition to the new Palestinian state. Meanwhile, with Arafat having satisfied the conditions Washington had set for direct contact, PLO and US representatives began holding regular discussions. It was a heady moment—a time when the prospects for a resolution of the Israeli-Palestinian conflict had never been brighter.⁶²

Hamas

Arafat's bold initiative did not enjoy universal support among the Palestinians, however. In addition to the now-marginalized splinter groups, a new and dynamic Islamist organization called Hamas expressed unbending opposition to his diplomatic program. The group's origins were intimately tied to the *Intifada*. Just five days after the rebellion began, the Palestinian Muslim Brotherhood established Hamas in the Gaza Strip to help coordinate the uprising. Influential from the start, the organization began laying out its goals in a formal charter that it issued in August 1988. Taking direct aim at both the PLO's secular nationalism and the two-state solution, it announced that Palestine was a divinely provided "Islamic trust" that no ruler or people could alienate or divide. Accordingly, the charter declared forthrightly that the organization's twin goals were the destruction of Israel through jihad and the establishment of a shari'a-based Palestinian state that would occupy the entirety of the former mandate. It was a strident program that ran directly counter to the UNL's more conciliatory vision. It was also one that was working fertile ground. Thanks to Saudi Arabia's generous funding of Wahhabi-influenced charities and mosques in the Occupied Territories, many Palestinians were

⁶¹ Schulze, *The Arab-Israeli Conflict*, 71; Gelvin, *The Israel-Palestine Conflict*, 216–21.

⁶² Shlaim, *The Iron Wall*, 481–82; Rogan, *The Arabs*, 435–36.

predisposed to the group's explicitly religious message. As a result, Hamas soon won a substantial following—particularly in the Gaza Strip.⁶³

Impasse

The failure of Arafat's diplomatic push to produce results accelerated Hamas's growth in 1989 and 1990. Initially, the PLO's bold initiative appeared to be successful. Thanks to Arafat's willingness to formally embrace the two-state solution, the administration of President George H. W. Bush (r. 1989-1993) had begun pressuring Tel Aviv to respond to the PLO's peace initiative. Unable to stonewall Washington, Shamir's government issued an Israeli plan for the territories in May 1989. Thin on details, it proposed elections in the Occupied Territories that would lead to a vaguely defined form of autonomy for the Palestinians under an arrangement in which Israel would continue to oversee key issues such as security, the distribution of land, and the construction of settlements.⁶⁴

While Shamir's proposal may have amounted to something of a step forward on the part of the Israeli government, it was not, at the end of the day, a substantive plan aimed at achieving peace. On the contrary, as Arafat and his allies well understood, it was an effort to delay genuine talks while Shamir's government constructed further settlements in the Occupied Territories—partly to house a sudden influx of Jewish immigrants from the USSR but, more broadly, to create further “facts on the ground” that could ensure Israel's perpetual control of the West Bank and Gaza Strip. Indeed, an addendum to the plan made explicit the limits of Shamir's proposal. It declared unequivocally that Israel would neither negotiate with the PLO nor permit the Palestinians to create a state in the occupied territories.⁶⁵

Shamir's response put the PLO chair and his moderate allies in a difficult spot. Arafat's inability to secure any meaningful gains in exchange for his renunciation of terrorism had, in zero-sum fashion, weakened the PLO and the UNL to the benefit of Hamas. In fact, Hamas felt strong enough in early 1990 to begin mounting a serious challenge to the PLO-aligned UNL for control of the *Intifada* and, thus, for a dominant position in the Occupied Territories. Arafat did not take the Islamists' success lightly. He understood that he faced a dangerous threat and that he would need to find a way to induce Israel to make concessions if he wished to avoid seeing the PLO eclipsed by Hamas. Accordingly, in the spring of 1990, he moved to restore his control over the Palestinian movement and to breathe life into his diplomatic effort by developing closer

⁶³ Gelvin, *The Israel-Palestine Conflict*, 221–26.

⁶⁴ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 609–10.

⁶⁵ Smith, *Palestine and the Arab-Israeli Conflict*, 424–25; quote from Rashid Khalidi, *Palestinian Identity: The Construction of Modern National Consciousness* (New York: Columbia University Press, 1997), xxv.

ties with the only Arab leader who possessed the military power to threaten Israel and thus to compel it to negotiate: Saddam Husayn.⁶⁶

The Iran-Iraq War, 1980-1988

At that point, Saddam Husayn's country was less-than two years removed from a devastating, near-decade long conflict: the Iran-Iraq War. By far the deadliest military contest in the Middle East since World War II, the conflict between Ba'hist Iraq and Shi'i Iran proved to be ruinously expensive for both states. Neither country intended for it have been such a grueling and deadly affair. Indeed, as we shall explore in more detail in a moment, Saddam Husayn had started the war in the expectation that it would produce a quick and inexpensive victory that would further solidify his position in the Iraqi government.

The Early War

The war began in September 1980. Unilaterally declaring the revocation of the Algiers Accord, Saddam Husayn sent the Iraqi army across the Shatt al-Arab Waterway in a limited offensive aimed at extorting concessions from Tehran. Despite ominously heavy resistance, his troops quickly seized control of 26,000 square kilometers of Iranian territory including the strategic city of Khorramshahr. Having taken what he believed was sufficient land to achieve his diplomatic ends, Husayn then ordered his army to assume defensive positions and awaited Tehran's appeal for a negotiated settlement.⁶⁷

It did not come. Saddam Husayn had been correct about Iran's internal divisions, but he had badly miscalculated the impact of his attack on its domestic political situation. Rather than undermining the revolutionary government by stoking further discord, the Iraqi invasion had instead presented Iranians with a common enemy that they could unify against. Khomeini was quick to take advantage. Exploiting the surge in nationalist and Shi'i sentiment that the invasion engendered, he rallied support for his regime and recruited soldiers for his armies. As a result, his government was able to quickly expand the Revolutionary Guard into a 120,000-strong military force and to secure hundreds of thousands of volunteers for the *Baj-e Mostazefan*, or the "Mobilization of the Oppressed": a militia composed of ideologically motivated but militarily untrained youth who were prepared to sacrifice their lives in defense of the revolution.⁶⁸

Many of them did. In a series of huge, human-wave attacks in early 1982, Revolutionary Guard, *Baj-e Mostazefan*, and regular army forces furiously struck Iraqi positions. These assaults resulted in huge numbers of Iranian casualties, but they also

⁶⁶ Smith, *Palestine and the Arab-Israeli Conflict*, 425–28.

⁶⁷ Tripp, *A History of Iraq*, 232–33.

⁶⁸ Ervand Abrahamian, *A History of Modern Iran* (Cambridge: Cambridge University Press, 2008), 171, 175.

succeeded in forcing the invaders to retreat back across the Shatt-al-Arab Waterway in June 1982. It was an impressive victory for the Iranian military. Khomeini refused to let his generals rest on their laurels, however. On the contrary, determined to finish off Saddam Husayn and his Ba' thist regime, he instructed them to press forward and to continue the offensive across the border.⁶⁹

War of Attrition

The next few years were trying ones for Iraq. To a substantial degree, its problems were fiscal in nature. The war had quickly become incredibly expensive—compelling Husayn's government to spend a staggering \$1 billion per month on its direct costs alone. Worse, thanks to the destruction of much of Iraq's port infrastructure in the fighting and to Iranian ally Hafez al-Asad's closure of the trans-Syrian pipeline, oil revenue had collapsed from \$26 billion in 1979 to a paltry \$11 billion in 1987—an amount that was insufficient to pay for the war let alone the substantial social-welfare benefits that Saddam Husayn had introduced in the 1970s.⁷⁰

Still, however bleak Iraq's fiscal position may have been, its military situation was even grimmer. Between 1983 and 1985, Iranian forces launched a new series of human-wave offensives against heavily fortified Iraqi defensive positions. Though unsuccessful, these assaults inflicted enormous casualties and left the Iraqi army demoralized and badly overstretched. The situation worsened dramatically in 1986 when Iranian-supported Kurdish guerrillas opened a second front in the north that diverted scarce Iraqi troops from the main theater. Taking advantage, Iranian forces launched a huge offensive in the south that captured the strategic al-Faw peninsula and imperiled Iraq's only port: the city of Basra.⁷¹

Ultimately, it was international concerns about an Iranian victory that rescued Saddam Husayn's regime from defeat. Determined to avoid a situation in which Khomeini's government dominated the Persian Gulf, Western countries and the Gulf States loaned Iraq the enormous sums needed to continue the war. Baghdad used those funds to procure billions of dollars' worth of high-tech equipment including Soviet T-72 tanks and Exocet-missile-equipped French Mirage jets as well as dual-use goods such as pesticides that could be converted into chemical weapons. The Iraqi government even received assistance from the US. Despite holding serious misgivings about Husayn, the Reagan administration gradually tilted toward Baghdad over the course of the mid 1980s. It took Iraq off the list of states that sponsor terrorism in 1982,

⁶⁹ Nikki Keddie, *Modern Iran: Roots and Results of Revolution* (New Haven: Yale University Press, 2006), 251; Michael Axworthy, *A History of Iran: Empire of the Mind*, Second Edition (New York: Basic Books, 2016), 268.

⁷⁰ Phebe Marr, *The Modern History of Iraq*, Second Edition (Boulder: Routledge, 2004), 187, 203–5.

⁷¹ Tripp, *A History of Iraq*, 238–42.

extended it agricultural credits in 1983, and, most importantly, began providing it with vital satellite intelligence in 1984.⁷²

What led Iraq and the US to cooperate? Why did two countries so hostile to one another that they had not had formal diplomatic relations since 1967 put aside their differences? They did so for reasons of realpolitik. That is, while their specific concerns were very different, they collaborated because each feared the prospects of an Iranian victory more than they opposed each other. Indeed, their shared concern about Iran was enough that they were able to look past not just older differences but new ones as well in pursuit of their mutual goal. Thus, Iraq did not dwell on the administration's amateurish, Iran-Contra efforts to secure the release of Western hostages in Lebanon by selling antitank missiles and spare parts to Tehran, while the US chose to look the other way when Iraqi forces used chemical weapons against Iranian troops.⁷³

Buttressed by Western and Gulf State aid, Iraq's reequipped army began to turn the tide in 1987. In brutal fighting early that year, it managed to decisively defeat a huge Iranian drive aimed at taking control of the city of Basra. Building on this success, the Iraqis followed by returning to the offensive for the first time since 1980. In the north, Ali Hasan al-Majid (1941-2010), better known as Chemical Ali, subdued the rebellious Kurds through the savage *al-Anfal*, "spoils of war" campaign that began in the spring of 1988. It was a ruthless effort even by the standards of Saddam Husayn's Iraq. Frequently employing chemical weapons—most notably in an attack on the town of Halabja that killed between 3,500 and 5,000 people—Iraqi soldiers destroyed five-hundred villages, executed countless men, and confined women and children to prison-like camps. All told an estimated fifty thousand to one-hundred thousand Kurds perished in the *al-Anfal* campaign. Meanwhile, in the south, the massively expanded Iraqi army made liberal use of chemical weapons in an offensive that regained control of the al-Faw peninsula and thus ended the Iranian threat to Basra. To support that push, Iraq fired hundreds of long-range SCUD surface-to-surface missiles at Iranian cities.⁷⁴

Tanker Wars

The Iraqi offensive also benefited from the outbreak in 1987 of an undeclared naval conflict between Iran and the US known as the Tanker War. Its origins stretched back to a campaign that Saddam Husayn had launched in 1984 to pressure Iran by firing on ships bound for its ports. Taking advantage of the fact that Iraq was equally vulnerable to such operations, Khomeini had countered by striking tankers bound for Iraq and its Gulf State backers. The impact of these attacks was substantial. Between 1984 and 1986, Iranian mines and other weapons damaged sixty-seven ships including eight Kuwaiti-registered tankers in a series of attacks that raised doubts about whether the Gulf States could continue to export oil. By 1987, such worries pushed an increasingly nervous Kuwaiti government to ask Washington to reflag eleven of its

⁷² Hahn, *Missions Accomplished?*, 77–80.

⁷³ Hahn, 77–80; Marr, *The Modern History of Iraq*, 187–88.

⁷⁴ Marr, *The Modern History of Iraq*, 188–91, 200–202.

supertankers so that they would enjoy US naval protection as they made the long and vulnerable transit of the Persian Gulf. Despite congressional concerns that complying might draw the US into a war with Iran—a fear made all-too real by an accidental Iraqi attack on the frigate USS *Stark* in May 1987 that killed thirty-seven sailors—the Reagan administration agreed to the request. Two months later, US Navy ships began escorting reflagged tankers to and from Kuwait.⁷⁵

As critics had predicted, it was not long before the operation turned into a shooting war. In a series of clashes that took place between September 1987 and April 1988, US forces sank most of Iran's navy and destroyed many of the offshore oil platforms from which Revolutionary Guard speedboats had mounted hit-and-run attacks. The operation was thus successful in ensuring the free flow of petroleum from the Gulf. It ended in tragedy, however. On July 3, 1988, the cruiser USS *Vincennes* accidentally shot down Flight 655, an Iranian Airbus 300 jetliner, it misidentified as an attacking F-14 fighter jet—killing all 290 people on board.⁷⁶

The incident played a small-but-significant part in spurring Tehran to agree to a proposed UN ceasefire in the summer of 1988. By that point, Iran was in dire straits. It was not only suffering serious battlefield defeats, but, more importantly, was witnessing the slow implosion of its economy. In response, an overwhelming consensus of Iranian officials had come to conclude that the fighting needed to come to an end and that it needed to do so as soon as possible. Khomeini was not one of them, however. Still firmly committed to the war, he adamantly rejected his advisors' insistent claims that Iran desperately needed to accept the ceasefire. As a result, despite its increasingly grim circumstances, Iran continued to fight on.⁷⁷

It was the muted response to the downing of Flight 655 that finally changed his mind. The ayatollah could ignore the bleak economic and military situation that Iran confronted, but he could no longer disregard either the growing war weariness at home or, especially, the country's worsening diplomatic isolation abroad that the subdued reaction to the downing of the airliner had revealed. As a result, though he derided it as a "poisoned chalice," Khomeini finally relented and agreed to accept the UN ceasefire proposal in July 1988.⁷⁸

Where Historians Disagree: The Iran-Iraq War's Origins

Since 1980, a lively debate has taken place regarding the origins of the Iran-Iraq War. Scholars largely agree that the Iranian Revolution set the stage for the conflict and that Iraq fired the first shot; however, they continue to argue about Saddam Husayn's motivations. Some maintain that he was behaving opportunistically. Historians like

⁷⁵ Douglas Little, *American Orientalism: The United States and the Middle East since 1945*, second edition (Chapel Hill: University of North Carolina Press, 2004), 250–52.

⁷⁶ Hahn, *Missions Accomplished?*, 82.

⁷⁷ Ghattas, *Black Wave*, 177–78.

⁷⁸ Ghattas, 177–78.

Charles Tripp contend that the Iraqi dictator had acted in hopes of taking advantage of Iran's postrevolutionary divisions and isolation to compel it to abrogate the Algiers Accord that the Iraqi dictator had signed with Iran in 1975—an outcome that would both reverse the damage that the agreement had done to his reputation and enhance his standing as a champion of Arab nationalism.⁷⁹

Others instead argue that defensive considerations predominated in motivating the Iraqi dictator to go to war. While not denying Husayn's opportunism, for example, the writer Phebe Marr maintains that aggressive Iranian actions played a critical role in his decision making. As she notes, Khomeini had begun inciting a Shi'i uprising in Iraq almost from the moment he had returned to Tehran in 1979. Deeply concerned about this threat, Husayn sought to exploit the disarray in Iran to launch a limited invasion that would compel Khomeini to halt his call for a Shi'i revolt in Iraq and, in so doing, end what had become a serious threat to his regime.⁸⁰

Consequences

Regardless of the war's causes, the end of the fighting could not have come soon enough for Iran. Indeed, the war had been a draining and devastating one for a country fresh from a cataclysmic revolution. More than 16,000 Iranian civilians and over 190,000 soldiers had died in the fighting, and the Iraqi attacks on its cities had left thousands more homeless. The conflict had also greatly weakened the economy. Per capita income was a shocking 40 percent lower at the end of the war than it had been a decade earlier, and vast amounts of vital infrastructure and industry had been destroyed or heavily damaged in the fighting. Reconstruction was projected to take many years, moreover, meaning that depressed economic conditions would likely endure long after the fighting had stopped.⁸¹

As bad as the war had been for Iran, it was even more disastrous for Iraq. The drawn-out fighting had devastated the country's cities, killed over 250,000 Iraqis, and caused the economy to contract sharply. Equally bad was the conflict's fiscal impact. The war had saddled Iraq with a staggering \$50 billion in foreign debt while wartime damage to oil facilities and a sharp decline in the price of petroleum in the late 1980s had left nominal oil revenue 58 percent lower than it had been a decade earlier. As a result, Baghdad lacked the funds either to service its international debt or to rebuild its economy—a predicament that raised real about Husayn's ability to retain his grip on power.⁸²

⁷⁹ Tripp, *A History of Iraq*, 232–33.

⁸⁰ Marr, *The Modern History of Iraq*, 182–83.

⁸¹ Axworthy, *A History of Iran*, 269; Abrahamian, *A History of Modern Iran*, 171, 175.

⁸² Tripp, *A History of Iraq*, 248–51; Marr, *The Modern History of Iraq*, 205.

Conclusion

Ultimately, the Iraqi dictator's efforts to resolve this dilemma would plunge the region into a new conflict: the Gulf War. Coming in conjunction with the conclusion of the Cold War, that conflict would not only seem to end in a decisive victory by an American-led coalition but also to presage a new dawn for the region. Indeed, in its immediate wake, the Middle East appeared to be moving toward a resolution of the many conflicts—new and old—that had marked the region during the 1980s. It is to the crisis that precipitated the Gulf War and to the apparent easing of tension that followed its conclusion that we shall now turn.

Chapter Eighteen: The Gulf War and Its Legacy, 1990-2001

Introduction

The 1990s picked up in the Middle East where the 1980s had left off: with conflict and instability endemic to the region. Many of its disputes were longstanding ones. Shi'i Iran and Sunni Saudi Arabia persisted in the fierce, Islamic cold war that they had been waging with each other since 1979, jihadism continued to pose a serious threat to regimes everywhere in the Middle East, and, most consequentially, Israelis and Palestinians remained locked in a bitter, nationalist struggle over the disposition of the West Bank and the Gaza Strip. Not all of the Middle East's conflicts were old, however. Most notably, in August 1990, Iraqi leader Saddam Husayn (r. 1979-2003) precipitated a destabilizing new crisis that threatened both the regional balance of power and the global economy when he occupied the oil-rich Emirate of Kuwait.

Just a few years later, however, the issues and conflicts that lay behind the region's instability appeared to be resolving. Most obviously, in early 1991, a US-led coalition defeated Saddam Husayn's vaunted military and ended his efforts to extend Iraq's influence. Not long thereafter, Arab states like Egypt succeeded in crushing the jihadi threat while Saudi Arabia and Iran appeared to have negotiated an end to their religious cold war. Most startlingly, the Palestine Liberation Organization (PLO) and the Israeli government seemed poised to end the region's most bitter and longstanding struggle through a negotiated settlement. For a time, as a result, the people of the Middle East cautiously dared to hope that they might finally be able to move beyond the struggles and disputes that had beset the region since the end of World War I.

The Gulf War, 1991

In many ways, the Gulf War of 1991 was the defining crisis of the decade in the Middle East. Pitting a coalition of states led by the US against Iraq, this brief conflict ended in a decisive military victory for America and its allies and appeared, for a time, to mark a real turning point for the region. Indeed, with serious efforts coming in its wake to resolve longstanding conflicts, the war seemed to mark a watershed event—one that presaged the beginning of a new-and-brighter era in the Middle East.

Origins

The difficult situation that Saddam Husayn found himself in following the Iran-Iraq War (1980-1988) led directly to the crisis. His problems were primarily fiscal in nature, and they caught him on the horns of twin dilemmas. First, with the price of oil having fallen precipitously in the late 1980s, he had insufficient revenue to service the massive, \$50 billion foreign debt he had incurred fighting the war with Iran, and still provide the patronage and welfare benefits needed to keep the Iraqi people quiescent and bound to his regime. Meanwhile, he also lacked the resources to maintain Iraq's now-massive 1.3-million-person army or to provide the employment needed to prevent demobilized

soldiers from revolting. These were serious problems for Husayn, and they seemed utterly insoluble in light of his Arab creditors' unwillingness to consider his request to write off the loans that he had incurred to finance the war with Iran. Indeed, absent a rapid, dramatic change in Iraq's circumstance, his days as its leader were likely numbered.¹

The Iraqi leader's huge army may have contributed to the fiscal vise that he found himself in, but it also suggested a way for him to wriggle free. He realized that if he could not persuade his Arab creditors to forgive the loans they had made to Iraq, he could instead use his powerful military to intimidate them into doing so. Accordingly, claiming that he had fought the war with Iran in part to protect Kuwait and Saudi Arabia from a hostile Tehran, the Iraqi leader mounted a strident diplomatic campaign against their governments in the spring of 1990—backed by the implicit threat of force—aimed at coercing them into recharacterizing the loans they had made to Iraq into grants. He even demanded that they provide an additional \$30 billion in aid to help rebuild his country's devastated economy.²

As the price of oil plummeted from \$21 per barrel in January 1990 to just \$11 per barrel that summer, his rhetoric grew even more heated. He claimed that Kuwait owed Iraq \$2.5 billion for stealing petroleum from the Rumaila oil field by drilling near the border, and he accused it of deliberately pushing the price of oil down by exceeding its Organization of the Petroleum Exporting Countries (OPEC) quotas. He followed by demanding that Kuwait and Saudi Arabia cut production and by stepping up his call for debt relief. Ominously, he asked Egyptian President Husni Mubarak (r. 1981-2011) and Jordan's King Hussein (r. 1953-1999) to let the leaders of the Gulf States know "that if they do not give this money to me, I will know how to get it." Still, despite these threats, Kuwait and Saudi Arabia remained unmoved and refused to restructure Iraq's debt or to provide reconstruction aid.³

With his intimidation campaign having failed, Saddam Husayn began to pursue a riskier but potentially more rewarding course of action. At some point in the late spring or early summer of 1990, he started to consider using Iraq's powerful army not to intimidate Kuwait but instead to occupy it outright. His thinking here echoed the logic that had informed his decision-making at the start of the Iran-Iraq War. Confident that Washington would not act, he believed that the Gulf States would have no choice but to seek a negotiated settlement—just as he had expected Tehran to have done in 1980. He would then find himself in the enviable position of being able to choose between two appealing options: either annexing the emirate outright or having the Gulf States write

¹ Charles Tripp, *A History of Iraq* (Cambridge: Cambridge University Press, 2002), 250–52; Phebe Marr, *The Modern History of Iraq*, Second Edition (Boulder: Routledge, 2004), 219.

² Marr, *The Modern History of Iraq*, 220–21.

³ Peter L. Hahn, *Crisis and Crossfire: The United States and the Middle East Since 1945* (Washington, D.C.: Potomac Books Inc., 2005), 106–7; Quote from Marr, *The Modern History of Iraq*, 220–21.

off Iraq's debt in exchange for the withdrawal of his army. Either way, Iraq would solve its immediate fiscal crisis and Saddam Husayn would both solidify his domestic standing and, finally, assume the dominant position in the Middle East that he had long sought. It seemed to be a golden opportunity—one too good for Husayn to pass up. Accordingly, in late July, he ordered Iraqi troops to begin massing along the Kuwaiti border.⁴

By August, Saddam Husayn was ready to pull the trigger. Despite Kuwait indicating a willingness to make meaningful concessions at talks hosted by Saudi King Fahd bin Abdulaziz Al Saud (r. 1982-2005) in Jidda on July 31, he ordered his army to occupy the emirate at dawn on August 2. Other than failing to capture the royal family, the invasion was a total success. Facing no meaningful resistance, the Iraqi army crossed the frontier without difficulty and took control of Kuwait and its vital oil fields in a single day. Saddam Husayn was thrilled at the apparent success of his plan. Indeed, he was so delighted that he cast caution to the wind and announced on August 8 that Iraq was annexing the emirate. In the meantime, satisfied that his control of 20 percent of the globe's supply of oil gave him real leverage, he ordered his forces to dig in and awaited requests to negotiate.⁵

Operation Desert Shield

Those offers did not come. Saddam Husayn had badly misread the post-Cold War global environment and faced almost immediate condemnation on all fronts. The United Nations (UN) Security Council moved first. Passed on the day of the invasion, Resolution 660 called for Iraq to withdraw immediately and unconditionally from Kuwait. The Arab League followed with its own formal denunciation of the invasion a few days later. Most importantly, President George H. W. Bush (r. 1989-1993) made clear that the US would not tolerate Iraq's occupation of the emirate. Declaring that "this will not stand, this aggression against Kuwait," he acted quickly to pressure Saddam Husayn into withdrawing his forces. Taking advantage of the mood of international cooperation that accompanied the end of the Cold War, he secured passage on August 6, 1990, of a second UN Security Council Resolution, number 661, that put teeth into the earlier call for Saddam Husayn to withdraw by imposing a tough sanctions regime on Iraq. Enforced by the US navy, it resulted in Iraq's exports falling by 90 percent and its imports plummeting by 97 percent.⁶

Bush also moved swiftly to prevent Saddam Husayn from seizing any more of the region's oil fields. On the same day that the UN Security Council passed Resolution 661, he obtained King Fahd's leave to begin stationing American troops in Saudi Arabia.

⁴ Tripp, *A History of Iraq*, 251–53.

⁵ Eugene Rogan, *The Arabs: A History – Revised and Updated Edition* (New York: Basic Books, 2012), 444–45; Marr, *The Modern History of Iraq*, 227–29.

⁶ Quote from Hahn, *Crisis and Crossfire*, 108; Tripp, *A History of Iraq*, 252; Marr, *The Modern History of Iraq*, 234.

Deployed under the codename Operation Desert Shield, the first contingent of soldiers arrived in the kingdom the following day. Hundreds of thousands more would follow.⁷

Why did Bush react so forcefully to Iraq's invasion of Kuwait? Two considerations informed his actions. First, so long as Saddam Husayn retained possession of Kuwait, he would enjoy a chokehold on the globe's oil supply and would thus be in a position to hold the world economy hostage at his whim. Second, with the Cold War in the process of ending, Bush had hoped to establish a "new world order" under American leadership based on the Wilsonian principles of collective security, self-determination, and the rule of law. In the president's eyes, Iraq's occupation of Kuwait amounted to a dangerous frontal assault on that system—one that the US and its allies had to defeat if his vision for the globe was to go into effect.⁸

There followed a period of diplomatic jockeying on the part of both the US and Iraq. Using a mix of persuasion and the promise of tangible rewards, Bush built a military coalition of nearly thirty states that included Britain, France, Morocco, Egypt, Argentina, Syria, Senegal, New Zealand, and Egypt and secured the diplomatic support of traditional rivals such as the Soviet Union. While the president took a global approach, Husayn instead focused narrowly on the Arab world. Seeking to fragment the coalition by peeling off its Arab states, he shrewdly drew attention to the hypocrisy evident in the contrast between Washington's forceful approach to the Kuwait crisis and its comparatively kid-glove treatment of Israel's occupation of the Gaza Strip and West Bank. He also declared that Iraq was willing to pull out of Kuwait, but only as part of a broader regional settlement in which Israel withdrew from the Occupied Territories and Syria departed from Lebanon.⁹

These arguments had little resonance globally, but they struck a chord in the Arab world. Indeed, demonstrations against participation in Operation Desert Shield broke out in many Arab states—including, notably, in countries closely aligned with the US such as Egypt and Morocco—shortly after Saddam Husayn announced his proposal. Meanwhile, drawn to the Iraqi leader's bellicose rhetoric toward Israel and his willingness to stand up to Washington, the Palestinians demonstrated their pro-Iraq feelings by organizing rallies in the Occupied Territories and Jordan. Here, Yasir Arafat (1929-2004) was in the vanguard. Having moved closer to Saddam Husayn following the failure of the PLO's diplomatic push in the late 1980s, he expressed open and enthusiastic support for the Iraqi leader following the occupation of Kuwait—a position that restored his standing among Palestinians disappointed by his inability to turn the concessions he had made into tangible gains. Thus, Bush's skillful diplomacy may have gained the support of Arab governments for Operation Desert Shield and almost assuredly tipped global popular opinion in favor of the US, but Saddam Husayn's

⁷ Douglas Little, *American Orientalism: The United States and the Middle East since 1945*, second edition (Chapel Hill: University of North Carolina Press, 2004), 256–57.

⁸ George C. Herring, *From Colony to Superpower: U.S. Foreign Relations since 1776* (New York: Oxford University Press, 2008), 908–9.

⁹ Little, *American Orientalism*, 257; Rogan, *The Arabs*, 450–51.

successful effort to link the seizure of Kuwait with the Israeli occupation of the Gaza Strip and West Bank had clearly won the hearts and minds of the vast majority of Arabs.¹⁰

Operation Desert Storm

The success of Saddam Husayn's propaganda message deeply troubled Washington. If given time, his campaign had the potential to peel off enough Arab states to make the effort to stop Iraq politically impossible—thus permitting him to retain control of Kuwait. The Bush administration reacted in two ways. First, abandoning the defensive disposition of Desert Shield in favor of an offensive deployment Bush announced in early November that he was doubling the size of the American force in Saudi Arabia to over 500,000 personnel—nearly as many as had been deployed to Southeast Asia during the height of the Vietnam War. Second, he secured approval of a new, and more forceful UN resolution designed to ensure that Saddam Husayn did not have the time needed to fracture the coalition. Passed on November 15, Resolution 678 authorized the US-led coalition to use “all necessary means” to push Iraq out of Kuwait if it did not withdraw by January 15, 1991.¹¹

In two ways, the resolution trapped the Iraqi leader. First, by leaving insufficient time for him to divide the global community or to hive off some of the Arab states, it neutralized his diplomatic strategy. Second, and more important, it confronted the Iraqi dictator with an unpalatable choice. If he withdrew, his emboldened Ba‘th Party subordinates would likely depose him from power; if he instead remained in Kuwait, the powerful US military would likely attack the Iraqi army and force it out of the emirate.¹²

Ultimately, Saddam Husayn decided that staying put was the less-risky option. After all, he reasoned, there was hope that high American casualties might spark such intense antiwar protests in the United States that Bush would be forced to call off the attack; in contrast, a retreat coming so close on the heels of the disastrous war with Iran would surely precipitate a coup. Accordingly, he ordered his troops to hunker down for the expected assault.¹³

Code-named Operation Desert Storm, the war began on January 17, 1991, with a lengthy aerial campaign by coalition forces designed to soften up Iraq's defenses. It focused initially on key military and political targets. Coalition planes attacked air defenses, command-and-control installations, intelligence facilities, Ba‘th Party headquarters, and sites suspected of producing Weapons of Mass Destruction (WMDs). The air offensive did not remain confined to such targets for long, however. Instead, it

¹⁰ Rogan, *The Arabs*, 450–51.

¹¹ Little, *American Orientalism*, 258; Rogan, *The Arabs*, 448–49.

¹² George C. Herring, *From Colony to Superpower: U.S. Foreign Relations since 1776* (New York: Oxford University Press, 2008), 909–10.

¹³ Herring, 909–10.

soon expanded to include refineries, power stations, water treatment plants, and other civilian infrastructure.¹⁴

Saddam Husayn knew that he could not hope to compete with the awesome power of the American military. Still, he believed that he might yet be able to emerge from the war victorious if he could change the terms of the struggle. Accordingly, even before the war had commenced, he had begun devising a strategy to turn the military contest into a more-complex political one that would, he hoped, shatter Bush's carefully assembled coalition and thus allow Iraq to ride out the storm. He put his plan into effect on the second day of the war. On his orders, the Iraqi military began launching SCUD ground-to-ground missiles at Israel, a country that Bush had pointedly refrained from asking to join the coalition. Unguided, the rockets Iraq fired at the Jewish state were tactically unimportant. At the same time, however, they were freighted with potentially explosive political significance. Were Tel Aviv to reply to them with force in keeping with its Iron Wall ideology, the missile attacks would have folded the Gulf War into the broader Arab-Israeli conflict, and, in so doing, compelled the Arab states to abandon the fight with Saddam Husayn—thus both shattering the coalition that Bush had so carefully put together and permitting Iraq to retain possession of Kuwait. It was a savvy ploy that suggested a subtle grasp of regional political realities.¹⁵

Ultimately, however, Saddam Husayn's gambit did not prove to be effective. Through a combination of heavy diplomatic pressure and the deployment of powerful American Patriot anti-missile batteries to Israel, the president succeeded in persuading Tel Aviv to refrain from taking countermeasures. As a result, the coalition remained intact.¹⁶

Meanwhile, the US-led war effort in the Gulf continued. The aerial bombardment of Iraq and its military forces that had begun on January 17 continued unabated and largely uncontested for six full weeks. Finally, on February 24, the coalition initiated the second phase of its war plan: a massive ground attack in Kuwait and southern Iraq. Every bit as one sided as the aerial campaign that had preceded it, the one-hundred-hour-long ground assault first drove Husayn's formations out of Kuwait and then annihilated them when they tried to retreat back to Iraq on a stretch of road that came to be known as the "highway of death."¹⁷

Having ejected the Iraqi army from the emirate, the US then faced the tantalizing prospects of taking Baghdad and deposing Saddam Husayn. For a time, Bush was sorely tempted. Doing so would result in a tidy ending to the war and would permanently eliminate a dangerous and destabilizing regime. For two reasons, however, the

¹⁴ Little, *American Orientalism*, 261.

¹⁵ Peter L. Hahn, *Missions Accomplished?: The United States and Iraq Since World War I* (New York: Oxford University Press, 2012), 101.

¹⁶ Benny Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920* (New York: Vintage Books, 2001), 612.

¹⁷ Herring, *From Colony to Superpower*, 910–11.

president ultimately opted not to push on to the Iraqi capital. First, he was leery of tarnishing America's great victory by getting bogged down in what he feared might become a costly quagmire. Second, his administration believed that regional security required the retention of an Iraqi state sufficiently coherent that it could contain Iran. The president consequently refrained from advancing on Baghdad and instead agreed to a ceasefire on February 28.¹⁸

Bush may have been content to leave the Iraqi government in place, but he still wished to remove Saddam Husayn from power. Accordingly, in an effort to signal to top-ranking Ba' thists that the US would support a leadership change, he indicated on March 1 that he hoped the Iraqis would "take matters into their own hands" and depose the Iraqi leader. In the meantime, ignorant of the 10,000-30,000 Iraqi soldiers killed in action and the 2,280 Iraqi civilians who died in the bombing campaign, Americans celebrated a triumph that came at the astonishingly low cost of only 147 US lives.¹⁹

Postwar Uprisings

Bush's effort to incite a putsch in Baghdad would have important if unintended consequences that took much of the bloom off that victory, however. The president's call for a change in leadership in Iraq did not produce the intended Ba' thist coup against Saddam Husayn, who retained his grip on power, but instead inspired the outbreak of large, popular Kurdish and Shi'i revolts against the regime in March 1991. Those rebellions began in the south. On March 1, furious Shi'i soldiers in Basra attacked symbols of the regime like local Ba' th Party headquarters and the offices of the security services.

From these beginnings, anti-regime activities quickly spread. Within days, all of the Shi'i-dominated towns and cities in southern Iraq were in open revolt against Husayn and had forced government troops and officials to withdraw. Meanwhile, inspired by news of the Shi'i rebellion, the Kurds launched their own revolt against the Ba' thist state a few days later. Over the ensuing two weeks, they established control of Iraq's mountainous north and even occupied the strategic city of Kirkuk. Thus, by mid-March, Saddam Husayn's situation looked very grim indeed. With half the country in rebellion, his days as Iraq's leader appeared to be numbered.²⁰

Unfortunately for the Shi'a and Kurds, however, the Iraqi president was well prepared for their revolts. He had held his elite Republican Guard units in reserve in anticipation of unrest, and he wasted little time in unleashing them on the rebels. They began their counteroffensive in the south. Using indiscriminate force that left tens of thousands of civilians dead and that compelled even more to flee, the Republican Guard troops quickly crushed the poorly organized Shi'i revolt. Having pacified the

¹⁸ Hahn, *Missions Accomplished?*, 104–6.

¹⁹ Quote from Hahn, 104–9; Marr, *The Modern History of Iraq*, 235–39.

²⁰ Marr, *The Modern History of Iraq*, 243–45, 248–50.

south, Saddam Husayn then redeployed his forces to the north. Once again using overwhelming force, his soldiers quickly put down the Kurdish uprising.²¹

The return of government forces created an even bigger panic in the north than it had in the south. Fearing a repeat of the recent *al-Anfal* campaign, two million Kurds responded to the Iraqi army's counterattack by fleeing into the mountains—a mass movement that created an immediate and grave humanitarian crisis. Pressure quickly mounted on Washington to provide help. In response, President Bush agreed to supply assistance through Operation Provide Comfort and to impose a No Fly Zone north of the thirty-sixth parallel that ensured a quick end to the regime's military operations in that part of the country. Later, in August 1992, the US extended the No Fly Zone to the region south of the thirty-second parallel.²²

While the Shi'a and Kurds endured enormous privation as a result of the regime's efforts to crush their rebellions, they were not the only civilians to suffer following the Gulf War. Joining them were the Palestinians, who also experienced trying times in the wake of the conflict—in their case as punishment for the PLO's rash decision to side with Saddam Husayn. Most notably, Kuwait deported 300,000 Palestinian guest workers to Jordan in the conflict's immediate aftermath. Massively disruptive, this action not only created enormous suffering for the affected Palestinians, but, by terminating the annual, \$400 million flow of remittances that guest workers in Kuwait sent to their families in the Occupied Territories, also deprived the Palestinians of the West Bank and Gaza Strip of foreign currency and thus further depressed their perennially weak economy. To top it off, Saudi Arabia and Kuwait had already suspended their heretofore generous financial support of the PLO in response to Arafat's declaration of solidarity with Husayn. The Gulf War was thus as much a defeat for the Palestinians as it had been for Iraq. Having effectively bet everything on Saddam Husayn and lost, they had—at last—seemingly found rock bottom.²³

In contrast, thanks to his army's success in suppressing the uprisings, Saddam Husayn was able to hold on to power despite his phenomenally poor handling of the Kuwaiti crisis. Still, the war had imposed serious restraints on the Iraqi leader. Passed on April 12, 1991, UN Security Council Resolution 687 called on Iraq to destroy all WMDs and their production facilities and required it to submit to on-site visits by UN inspection teams. Until they could verify that Iraq had complied with the resolution, the embargo—with limited exemptions for food and medicine—would remain in effect. Thus, Saddam Husayn may have retained his grip on power, but, with his army substantially

²¹ Tripp, *A History of Iraq*, 255–56.

²² Andrew J. Bacevich, *America's War for the Greater Middle East: A Military History* (New York: Random House, 2016), 133; Hahn, *Missions Accomplished?*, 108; Tripp, *A History of Iraq*, 256–58.

²³ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 612–13.

weakened by the war and with his WMD program shut down, he was no longer in a position to threaten his neighbors.²⁴

A New Beginning?

Despite the failure to remove Saddam Husayn from power, the Gulf War appeared for a time to have constituted a vital inflection point for the Middle East—one that had set in motion a series of changes that appeared like they would resolve many of the region's most enduring conflicts. The progress was tangible. The Lebanese civil war came to a conclusion, Mubarak's government ended the jihadis' terror campaign in Egypt, and Saudi Arabia and Iran began to engage in talks aimed at achieving détente in the religious cold war they had been waging against each other since 1979. Most importantly, the PLO and Israel entered into peace talks that seemed to be leading to a genuine resolution of the Arab-Israeli conflict. Accordingly, hopes rose that the Middle East—a region that had been beset by so much strife since the West had bequeathed to it a modern state system just after World War I—was finally entering an era of peace and stability.

Lebanon: The Taif Accord

The conclusion of the Lebanese Civil War in the early 1990s was the first indication that the Middle East might be entering a more hopeful period. Ironically, the strife that had bedeviled Lebanon since 1975 only came to an end after the country had first slid yet further into chaos during the selection of a new president in 1988. Already a persistent source of conflict, the choice of a new head of state was even more fraught than usual that year owing to the ongoing civil war. Thanks to that conflict, the country had been unable to hold parliamentary elections since 1972 and many legislators were either dead or living abroad; unable to muster quorum as a result, parliament lacked the legal standing to select a new president. In such situations, the constitution assigned the powers of the president to the sitting prime minister until a new election could be held—meaning, in this case, that the Syrian-backed Sunni prime minister, Selim al-Hoss (1929-), stood poised to serve as Lebanon's interim chief executive.²⁵

This situation deeply alarmed the Maronite head of the army, Michel Aoun (1933-). Fearing that al-Hoss might dismantle the confessional system that guaranteed Christians a substantial share of political power, he established a rival interim government. Receiving arms shipments from Iraq, Aoun was well positioned to assert his claim to power. In fact, he even felt strong enough to start a war with Syria in March 1989 aimed at liberating the country from Asad's control. The result was six months of intense-but-inconclusive fighting that solved nothing even as it further immiserated the people of Lebanon and forced tens of thousands more to flee the country.²⁶

²⁴ Marr, *The Modern History of Iraq*, 240.

²⁵ Rogan, *The Arabs*, 458–59.

²⁶ Rogan, 458–59.

The significant civilian death toll horrified people in the Arab world and created pressure for a political settlement. Accordingly, at the behest of the Arab League, Saudi King Fahd, the Algerian president, and the Moroccan king began working in the spring of 1989 to secure a diplomatic solution to the conflict that was acceptable to all parties. Though the going was tough, they succeeded in getting the different factions to agree to a ceasefire and in arranging for the Lebanese parliament to convene in Taif, Saudi Arabia in the fall of 1989 for the purpose of arriving at a final resolution of the civil war. Sixty-two members of parliament—now a full seventeen years into their terms—took part in the talks.²⁷

The resulting Taif Accord promised to restore peace to a land riven by years of factional infighting. The agreement retained the distribution of offices along confessional lines that dated to the National Pact of 1943 and, bowing to reality, acknowledged Syria's dominant position in Lebanon. Importantly, however, it also replaced the earlier practice of reserving six parliamentary seats for the Christians for every five held by Muslims with a new arrangement that accorded each faith an equal number of deputies. It also shifted substantial power from the Maronite president to the Sunni prime minister. Carefully negotiated, the Taif Accord received nearly unanimous acceptance in Lebanon; critically, however, Aoun and his Iraqi-armed militia continued to oppose the deal. Though tantalizingly close, in other words, Lebanon was still not yet free of the civil war that was entering its fifteenth year.²⁸

It was at this point that the Gulf conflict helped to finally bring the Lebanese Civil War to an end. Bush needed Syria's participation in the coalition against Iraq in order to counter Saddam Husayn's argument that the American-led effort was a case of modern-day imperialism; deftly taking advantage, Asad made clear that his price was a free hand in Lebanon. Accordingly, compelled to prioritize the effort to eject Iraq from Kuwait, the Bush administration gave its tacit sanction to a major Syrian offensive in Lebanon that began in October 1990. The resulting campaign ended in a decisive victory for Asad's troops. With the UN embargo on Iraq preventing Aoun's forces from receiving further weapons and ammunition, the Syrian military handily defeated his Christian militia and swiftly reunited the country under a government based on the arrangement that had been negotiated in Taif. Damascus followed by compelling all the militias to disarm and disband—a move that ensured that peace would endure. The only exception was Iran's powerful protégé in Lebanon, Hizbollah, which was permitted to continue its increasingly sophisticated insurgency against Israeli forces in the security zone that the Israeli Defense Force (IDF) maintained in southern Lebanon. Thus, with the exception of the far south, Lebanon was at peace for the first time in fifteen years—albeit under Syrian dominance.²⁹

²⁷ William Harris, *Lebanon: A History, 600 - 2011* (New York, N.Y: Oxford University Press, 2012), 255.

²⁸ Rogan, *The Arabs*, 460–61.

²⁹ Charles D. Smith, *Palestine and the Arab-Israeli Conflict: A History with Documents*, Sixth Edition (Boston: Bedford/St. Martin's, 2007), 386–87.

Saudi-Iranian Rapprochement

The Gulf War also played a part in producing a thaw in the religious cold war between Saudi Arabia and Iran. Frosty since the Iranian Revolution in 1979, relations between the two countries had deteriorated considerably in the mid 1980s. Critical in worsening the climate between them had been the Bloody Friday clash between Iranian pilgrims and Saudi security forces during the hajj in 1987 that had resulted in the deaths of four-hundred people. As we have seen, Riyadh had followed that incident by breaking relations with Tehran and by increasing its financial support for the Iraqi war effort, while Iran had responded by urging Muslims to go to war with the Saudis and by threatening to attack countries that supported Iraq.³⁰

Two developments made possible the start of a rapprochement in 1990. First, Iran had recently experienced an important change in leadership. Following the Ayatollah Khomeini's death (1902-1989) in 1989, the Iranian president, Ayatollah Ali Khamenei (r. 1981-1989), became Supreme Leader while the relatively pragmatic Ayatollah Ali Akbar Rafsanjani (1934-2017) succeeded to the presidency. Though they largely shared Khomeini's values, these men were considerably more flexible and willing to compromise and were thus open to a deal with the Saudis. Second, Saddam Husayn's invasion of Kuwait had scrambled the regional balance of forces and turned Iraq, the country Riyadh had heretofore relied upon to contain Iran, into a dangerous enemy. Feeling enormously vulnerable as a result of this sudden shift, the Saudis were now willing to consider sweeping foreign-policy changes that would have been off the table even just a few months earlier.³¹

Both states saw much to gain from improved relations. Concerned about the threat that Iraq had become, the Saudi government now perceived Iran as a potential partner in the containment of Saddam Husayn. Meanwhile, Tehran wanted to increase the number of Iranians who could participate in the hajj and hoped to raise its OPEC allotment so that it could generate the funds desperately needed to reconstruct its war-shattered economy. Accordingly, the Iranian and Saudi foreign ministers began to hold meetings in New York in early 1991 designed to lay the groundwork for a lasting détente. That effort bore fruit in March 1991 when they jointly announced that their countries would be resuming diplomatic relations and that Saudi Arabia was more than doubling the number of Iranians who could participate in the hajj. Thus, with their shared fear of Iraq having cut the Gordian Knot, the two bitter rivals appeared at last to have abandoned the Islamic cold war that they had been waging since the Iranian Revolution in favor of a new relationship based on cooperation.³²

³⁰ Kim Ghattas, *Black Wave: Saudi Arabia, Iran, and the Forty-Year Rivalry That Unraveled Culture, Religion, and Collective Memory in the Middle East* (New York: Henry Holt and Co., 2020), 163.

³¹ Ghattas, 181–82.

³² Ghattas, 208–10.

The Saudi response to a terror attack in the kingdom made clear just how durable the rapprochement with Iran was. In June 1996, militants detonated a massive truck bomb in front of an eight-story building that was part of the al-Khobar Towers development, a residential complex in which American service personnel stationed in the Kingdom were housed. Shearing off the entire face of the building, the huge explosion killed nineteen Americans and wounded another four hundred. While suspicion initially fell on former Service Bureau head Osama bin Laden (1957-2011), the US soon uncovered evidence indicating that the perpetrators were in fact members of a shadowy, Iran-linked Shi'i terrorist group known as *Hizbollah al-Hejaz*. The revelation of this evidence was shocking, but it did not lead Riyadh to break relations or even to censure Iran. Instead, fearing that criticism might weaken the position of the Iranian moderates they had been courting and thus imperil the five-year-old rapprochement between Riyadh and Tehran, Saudi officials chose to keep mum.³³

Jihadism Contained

In the wake of the Gulf War, even the jihadi threat appeared to be on the wane. Indeed, the subsequent decade would prove to be a tough one for the militant Islamists. Throughout the Middle East, the organizations that had arisen in the late 1970s and early 1980s suffered significant defeats in the mid 1990s at the hands of the “near enemy”: the secular regimes in the region whose downfall they hoped to achieve. Egypt was a case in point. There, the terror campaign that Ayman Zawahiri's (1951-2022) Islamic Jihad and the blind sheikh Omar Abdel Rahman's (1938-2017) *Gamaat Islamiya* had launched in 1992 to overthrow the government had gone so badly that the jihadis felt compelled to issue a unilateral ceasefire in 1997. What explained their lack of success? Why had their effort failed so quickly and so totally?³⁴

Two issues undid the jihadis' campaign in Egypt. First, the brutal terror attacks they launched alienated ordinary Egyptians. Two incidents in particular turned public opinion against the militant Islamists. One was a failed car-bomb attack on the prime minister in 1993 that wounded twenty-one bystanders and killed a schoolgirl, Shayma Abdel-Halim. Angry at Islamic Jihad for its callousness, the huge crowd that turned out at her funeral vented their fury at the militants by chanting “[t]errorism is the enemy of God!” The other occurred four years later. On November 17, 1997, *Gamaat Islamiya* gunmen disguised as police officers entered the Temple of Hatshepsut in Luxor and systematically murdered four Egyptians and fifty-eight tourists including a five-year-old girl and four Japanese newlywed couples. The incident did not go over well in Egypt. The savagery of the attack, the resulting collapse in the country's economically critical

³³ Ghattas, 210–11.

³⁴ Fawaz A. Gerges, *The Far Enemy: Why Jihad Went Global* (Cambridge: Cambridge University Press, 2005), 9–11, 44, 151–53.

tourism industry, and the attendant job losses infuriated the Egyptian people and cost the jihadis what little support they had managed to retain.³⁵

Second, Mubarak's state security succeeded in grievously weakening Islamic Jihad and *Gamaat Islamiya* through a ruthless and highly effective crackdown in the mid 1990s that involved mass arrests and the demoralizing harassment of the jihadis' wives. The incompetence of Islamic Jihad's leadership greatly aided the government in this effort. While the organization had adopted a cell structure so that an arrested member could only disclose the names of a handful of associates—a standard practice in insurgent and terrorist groups stretching back at least to the nineteenth century—its membership director had inexplicably maintained a database with the names, addresses, aliases, and *hideouts* of the entire organization's membership; as a result, state security forces were able to arrest some eight-hundred members—nearly everyone who belonged to Islamic Jihad—in a single, coordinated operation. Kneecapped in this way, the organization was unable to recover and drifted off into irrelevance. Thus, while non-violent Islamist groups continued to make gains in Egypt, the jihadis had been reduced to irrelevance, crushed by Mubarak's effective security services.³⁶

The Peace Process, 1991-1995

Perhaps the clearest sign that the Middle East was in the process of putting the strife of the twentieth century behind it was the opening of a dialogue between Israel and the PLO aimed at finally bringing to a close the region's most enduring conflict. Like the Saudi-Iranian rapprochement and the end of the Lebanese Civil War, it had its roots in the Gulf War. Stung by Saddam Husayn's accusation of hypocrisy, President Bush opted to use some of the substantial political capital he had gained from his victory to convene a conference aimed at finally resolving the Arab-Israeli conflict on the basis of Resolution 242. Though the resulting negotiations did not follow the path that the president had laid out, they nonetheless produced tangible progress and raised hopes that the long conflict between Israelis and Palestinians was at last coming to an end.

The Madrid Conference

Getting all of the relevant parties to agree to meet was a tall order. Israel and Syria, in particular, were disinclined to participate due to fears that the conference might put pressure on them to cede territory. Bush was in a peculiarly strong position thanks to the end of the Cold War and to his victory over Iraq, however; as a result, he was well positioned to compel the holdouts to come. Taking advantage of Damascus's isolation

³⁵ Lawrence Wright, *The Looming Tower: Al Qaeda and the Road to 9/11* (New York: Vintage Books, 2006), 210–11, 290–93.

³⁶ Wright, 209–11.

now that it was no longer receiving Soviet military aid, for example, the president was able to induce Syria—with its Lebanese protégé in tow—to send a delegation.³⁷

Israel proved to be a far-tougher nut to crack, however. Prime Minister Yitzhak Shamir's (r. 1986-1992) Likud government was opposed to the very idea of engaging in talks based on Resolution 242's call for the exchange of land for peace and had no intention of taking part in the conference. At the same time, however, Israel was desperate for an American loan guarantee needed to finance the construction of housing for hundreds of thousands of Russian Jews whom the USSR, in a reversal of longstanding policy, had permitted to emigrate. Bush took advantage of Israel's predicament. Using the loan guarantee as leverage, he was able to force a reluctant Tel Aviv to send a delegation to Madrid.³⁸

For its part, the PLO did not enjoy formal representation at the conference. Instead, angry at Arafat for having backed the wrong horse in the Gulf War, Bush refused to permit it to send a delegation. The president's ban was only a partial one, however. Thanks to the PLO's loose affiliation with a group of Palestinians from the territories who were participating in the conference as members of a joint Jordanian-Palestinian delegation, the organization did have a voice in the proceedings—albeit an indirect one.³⁹

Opening in Madrid to great anticipation on October 30, 1991, the conference did not prove fruitful. Instead, under the glare of intense press coverage, the discussions quickly deadlocked. The Palestinians insisted on nothing less than full, uninhibited self-determination and the establishment of an independent state in the West Bank and Gaza Strip. Firmly opposed to ceding any of the land acquired in 1967, the Shamir government rejected this position out of hand and countered with an offer of autonomy for the Palestinian people on land that remained under Israeli control—a proposal that was utterly unacceptable to the Palestinians. As a result, the conference failed to produce the hoped-for breakthrough. Still, the talks marked an important milestone. Not only were the Palestinians and Israelis negotiating directly for the first time, but they were doing so on the basis of Resolution 242 and its call for the exchange of land for recognition.⁴⁰

Oslo I Agreement, September 1993

While the public nature of the Madrid negotiations prevented the two sides from making progress in the formal talks, a separate, secret negotiating track that began in December 1992 under Norwegian auspices soon showed promise. For different reasons, both sides had become more open to negotiations by that point. While still

³⁷ Avi Shlaim, *The Iron Wall: Israel and the Arab World, Revised and Expanded* (New York: W. W. Norton, 2014), 502.

³⁸ Smith, *Palestine and the Arab-Israeli Conflict*, 431.

³⁹ Shlaim, *The Iron Wall*, 501–4.

⁴⁰ Smith, *Palestine and the Arab-Israeli Conflict*, 432–33.

wary of the PLO, newly elected Labor Prime Minister Yitzhak Rabin (r. 1974-1977, 1992-1995) was interested in ending the costly occupation of the territories and was increasingly willing to make real concessions to do so; accordingly, he authorized two Israeli academics to take part in unofficial talks in Oslo. Meanwhile, having steadily shed supporters to Hamas, Arafat, too, was willing to compromise and agreed to send the treasurer of the PLO, Ahmad Quray' (1938-), to take part in the discussions. Conducted without pressure because of their informal nature, the talks soon produced real progress toward a final settlement of the century-old conflict over Palestine. As a result, in August 1993, a shocked and delighted world learned that the negotiations had produced a draft of a comprehensive settlement.⁴¹

Known as the Oslo I Accord, the agreement had two components. The first was an exchange of letters of mutual recognition on September 9, 1993. Signed four days later on the White House lawn and consummated by an awkward handshake between Arafat and Rabin, the second component, the Declaration of Principles on Interim Self-Government Arrangements, outlined the phased devolution of power in the Occupied Territories from Israel to a provisional Palestinian government over a period of five years on the basis of Resolution 242. It called for the PLO to first assume control of the West Bank-town of Jericho and the entirety of the Gaza Strip; in later phases, Israel would cede authority over health, welfare, taxes, education, and internal security in the rest of the West Bank to an elected Palestinian Council and would gradually withdraw its police and military forces from populated areas. Tel Aviv would retain control of external defense and foreign affairs as well as the security of the Jewish settlements until the two sides concluded a final agreement at the end of the five-year period. Importantly, both parties agreed to defer discussion of contentious topics such as borders, the right of refugees to return to their homes, settlements, Jerusalem, and, above all, the Temple Mount until final status negotiations took place in hopes that the success of the phased approach would build the trust needed to resolve those thorny issues.⁴²

The Oslo I Accord marked a dramatic change in the Israeli-Palestinian relationship—one that suddenly made possible a number of important breakthroughs. Substantially weakening the rationale for the Arab states' continued rejection of Israel's legitimacy, for example, it permitted Israel to normalize relations with Jordan and to open liaison offices in Morocco and Tunisia. More dramatically, the agreement allowed Arafat to establish a government in the Occupied Territories, and it secured for the first time Tel Aviv's acknowledgement that the Palestinians constituted a people.⁴³

At the same time, however, the accord also generated fierce opposition. Rejectionist Palestinians howled that Arafat had rashly squandered the PLO's strongest

⁴¹ Rashid Khalidi, *The Hundred Years' War on Palestine: A History of Settler Colonialism and Resistance, 1917–2017* (New York: Metropolitan Books, 2020), 196–97; Shlaim, *The Iron Wall*, 530–34.

⁴² Shlaim, *The Iron Wall*, 534–38.

⁴³ Shlaim, 546–47, 556–65, 600.

card—recognition of Israel—by failing to secure the formal acceptance of a Palestinian state in return. They also complained that the deal was impossibly vague and noted that it did not ban further settlement construction. Many Israelis expressed similar disdain for the accord. Some questioned Arafat's commitment to peace and wondered whether the deal imperiled the Jewish state's security. Others—particularly militant members of the settlement movement—protested the very notion of ceding *any* land whatsoever to the Palestinians. One of them, the American-born Baruch Goldstein (1956-1994), took matters into his own hands. Seeking to derail further progress, he murdered twenty-nine Muslims while they prayed in the Mosque of Abraham in February 1994 before some of the survivors managed to beat him to death. Likewise opposed to a compromise settlement, Hamas responded with a pair of deadly suicide bombings aimed nominally at achieving retaliation but really at undermining the trust on which future progress depended.⁴⁴

Oslo II, September 1995

While the violence succeeded in temporarily halting talks and in pushing back the timeline that Israel and the PLO had agreed to in 1993, it failed to prevent further progress toward a final settlement. Indeed, the negotiations not only continued in secret but soon produced a new agreement. Signed by Arafat and Rabin in Washington in September 1995, the Israeli-Palestinian Interim Agreement on the West Bank and the Gaza Strip—known colloquially as the Oslo II Accord—laid out the next steps in the peace process. It articulated the powers that the Palestinian government would exercise, scheduled elections for the Palestinian Council and presidency, and provided for the creation of a Palestinian seaport and for the establishment of routes connecting the Gaza Strip to the West Bank. It also permitted Tel Aviv to construct a network of bypass roads linking the West Bank settlements with each other and with Israel proper. Meanwhile, of great symbolic importance, it called for the Palestinian Council to formally strike language from the Palestinian National Charter calling for Israel's destruction.⁴⁵

Finally, it divided the West Bank into three zones—the A, B, and C Areas—for the purpose of establishing where, when, and to what extent the Palestinian Council would exercise authority. Constituting 4 percent of the territory, the A Areas included the main Palestinian population centers of Jenin, Nablus, Ramallah, Bethlehem, Qalqilya, Tulkarm, and Hebron and would, by the terms of the interim agreement, fall under immediate Palestinian control. Consisting of smaller Palestinian villages, the 25 percent of the West Bank that comprised the B Areas would occupy an interim position. They would come under immediate Palestine Council civil authority while Israel continued to retain responsibility for security. Finally, the settlements, military bases, and unoccupied land that made up the balance of the West Bank constituted the C Areas and would remain under Tel Aviv's control pending final status negotiations. Importantly, this division of the West Bank was dynamic. That is, it called for the C Areas to gradually

⁴⁴ Smith, *Palestine and the Arab-Israeli Conflict*, 451–61.

⁴⁵ Smith, 461–63.

shrink as Israel withdrew its forces and ceded responsibility for that land to Palestinian civil control.⁴⁶

The Oslo II Accord began to go into effect almost immediately after Arafat and Rabin signed it. Palestinian police replaced Israeli security forces in the main Palestinian towns in the fall of 1995 save for Hebron, for which the agreement had stipulated a later withdrawal date, and elections went off as scheduled in January 1996. Predictably, Arafat handily won the presidency, and PLO candidates swept a solid majority of seats on the council. In April 1996, the new government followed by voting to delete the language in the Palestinian National Charter that called for Israel's destruction. Thus, to many observers, the main source of instability and conflict in the Middle East seemed finally to be coming to an end—a change that suggested that the region as a whole was at last poised to enter a new era of peace and stability.⁴⁷

A Return to Conflict

Such sentiments proved to be sadly misplaced. Starting in the mid 1990s, conflicts that had appeared to be winding down instead roared back to life. In the Gulf region, American efforts to contain both Iran and Iraq produced steadily rising tensions. In the Occupied Territories, meanwhile, the effort to establish a just and lasting peace between Israel and the Palestinians collapsed amid growing violence and mutual recriminations. Finally, even as the jihadis faced defeat at the hands of the “near enemy,” a new struggle pitting transnational jihadis against the “far enemy”—the United States—emerged in the Middle East. As a result, hopes that the people of the region were finally moving beyond the conflicts that had dominated the twentieth century gradually evaporated.⁴⁸

The Failure of the Oslo Accords, 1995-2000

Perhaps the biggest setback was the collapse of the Israeli-Palestinian peace process. Begun with such high hopes, it started to come apart just a few months after the signing of the Oslo II Accord. It did so thanks to the determined opposition of extremist groups on both sides. Unwilling to accept anything less than a total victory, rightwing Israelis and Hamas militants waged parallel terrorist campaigns to prevent the achievement of a final negotiated settlement. That strategy began to gain traction in November 1995, when a young Israeli named Yigal Amir (1970-) assassinated Prime Minister Rabin just after he spoke at a peace rally in Tel Aviv. A messianic Zionist, Amir believed like many others on Israel's religious right that God, Himself had made possible the conquest of the West Bank and that, by extension, ceding any of it amounted not merely to treason but to apostasy—a crime justly punishable by death. Hamas was not to be outdone, however. It moved to undermine the faltering trust on which the peace process rested by launching a wave of suicide bombings in February

⁴⁶ Shlaim, *The Iron Wall*, 545–46.

⁴⁷ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 629.

⁴⁸ Gerges, *The Far Enemy*, 10–11.

and March 1996 that killed sixty Israelis. The attacks had the desired effect. They not only furthered frayed the mutual confidence that had made peace talks possible, but, more importantly, helped tip the Israeli election held in May 1996 in favor of the Likud candidate for prime minister, Binyamin Netanyahu (r. 1996-1999, 2009-2021), who narrowly defeated Labor's Shimon Peres (r. 1977, 1984-1986, 1995-1996).⁴⁹

In retrospect, Rabin's death marked a clear turning point in the Oslo process. While he had been a tough negotiator and had expressed doubts about the establishment of a Palestinian state, he had also approached the talks with a pragmatic outlook and a willingness to compromise. Netanyahu, in contrast, showed no such flexibility. On the contrary, he was militantly opposed to any further progress on the peace process and was absolutely dead set against the creation of a Palestinian state.⁵⁰

Indeed, Netanyahu's election amounted to putting the fox in charge of the hen house—a reality made emphatically clear in a formal statement of policy that his government released a few weeks after the election. Called the “Basic Guidelines,” it explicitly rejected the possibility of a Palestinian state, declared that Israeli control of Jerusalem was “nonnegotiable,” and called for the expansion of the settlements. Separately, Netanyahu further hindered progress by insisting on reopening negotiations over previously resolved issues, by dragging his feet in implementing many provisions in the Oslo II Accord including withdrawal from Hebron and the phased handover of C Areas, and by making clear that he had no intention of engaging in final-status negotiations. Pressure from the administration of President Bill Clinton (r. 1993-2001) did belatedly force Netanyahu to hand over Hebron and to promise to agree to the Wye River Memorandum in October 1998, which committed Israel to make further withdrawals from C Areas. Thanks to intense opposition from rightwing cabinet members, however, the prime minister reneged on that pledge just weeks after he had made it.⁵¹

Netanyahu's efforts to sabotage the negotiations contributed to a precipitous decline in the Palestinians' confidence in the peace process. They had initially greeted news of the talks between the PLO and the Israeli government with great hope. The announcement of the Oslo Accords had raised expectations not only that they soon would be citizens of a Palestinian state but that they would also experience greater freedom as the negotiations unfolded. The peace process had not meet those expectations, however. Instead, rather than enjoying improved conditions, the West Bank Palestinians confronted stepped up harassment, new restrictions, and a sharp rise in the settler population.⁵²

⁴⁹ Shlaim, *The Iron Wall*, 566–72, 571, 582–84.

⁵⁰ Smith, *Palestine and the Arab-Israeli Conflict*, 470.

⁵¹ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 640–49.

⁵² James L. Gelvin, *The Israel-Palestine Conflict: One Hundred Years of War* (New York: Cambridge University Press, 2005), 238–39.

Three issues in particular rankled the Palestinians. First, the construction of Israeli-only bypass roads that the Oslo II Accord permitted had not only resulted in the expropriation of Arab land but also in the further fragmentation of the West Bank. Second, thanks to the Israeli government's aggressive construction of new settlements, the settler population doubled between 1993 and 2000—a trend hardly suggestive of the idea that the Palestinians would soon have their own state. Finally, rather than establishing a coherent Palestinian proto state, the peace process had instead produced a series of isolated enclaves in the West Bank that Palestinians could access only by passing through Israeli checkpoints and roadblocks. This change marked a significant decline in their freedom of movement. Prior to the Oslo Accords, they had been able to move with little restriction throughout the West Bank; now, as a result of the checkpoints, they had to endure lengthy waits and searches if they wished to travel more than a short distance. Thus, by the end of the 1990s, the Palestinians had not only lost faith in the peace process but had also begun to turn against the PLO leadership that had negotiated it.⁵³

By that point, moreover, the prospects for continued progress were uncertain. The collapse of Netanyahu's cabinet and the election of a Labor government under Ehud Barak (r. 1999-2001) in May 1999 had, it is true, renewed hopes that the Oslo process could produce a durable peace. Unfortunately for the prospects of a final settlement, however, Barak's political support rapidly evaporated as a result of domestic political problems and Israel's botched withdrawal from southern Lebanon in the spring of 2000—a situation that left the Likud Party poised to retake control of the government. In response, President Clinton decided to take a bold and risky step. Looking to cement his legacy, he invited Barak and Arafat to a summit at Camp David in July 2000 in hopes that he could breathe life into the Oslo process by brokering a deal that was acceptable to both parties.⁵⁴

The summit saw real breakthroughs on some issues, but ultimately ended in failure and mutual recriminations. True to expectations, Barak showed vastly greater flexibility than had Netanyahu. At one point, he even proposed ceding 91 percent of the West Bank to the Palestinians—by far the most generous proposal that Israel had yet made. Substantial as that offer may have been, however, it nonetheless included provisions that were unacceptable to the Palestinians. It would have permitted Israel to retain permanent possession of the Jordan River Valley and to maintain final say over the disposition of the West Bank's finite water resources. Most importantly, if implemented, his proposal would have left the Palestinians with non-contiguous territories, meaning that they would have had to endure the hated Israeli checkpoints in perpetuity. The two sides might have been able to resolve the impasse had they had the opportunity to complete the necessary preparatory work of narrowing differences over

⁵³ Rashid Khalidi, *The Iron Cage: The Story of the Palestinian Struggle for Statehood*, Reprint Edition (Boston: Beacon Press, 2006), 199–203.

⁵⁴ Clinton, of course, also wanted to secure a deal before his term expired. Smith, *Palestine and the Arab-Israeli Conflict*, 501; Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 655–58.

difficult issues and of establishing common frames of reference before the meeting. Barak's tenuous political position in the summer of 2000 denied them the luxury of that time, however; accordingly, the attendees—especially Arafat—were unprepared to make snap decisions in the hothouse atmosphere of a high-level summit. Thus, though intended to achieve a breakthrough, the talks instead served only to further frustrate the Israelis and to embitter and disillusion the Palestinians—particularly after Barak and Clinton unfairly blamed Arafat for the meeting's failure.⁵⁵

Having had their expectations once more raised, the Palestinians greeted the collapse of the Camp David summit with anger and frustration. Where, they asked, was the independence that they had been promised? Where was the freedom from harassment that they had come to expect? Coming as it did after seven years of peace talks, it was an infuriating moment for the Palestinians—one that left the Occupied Territories poised to explode.⁵⁶

The spark that ignited them came in the form of a visit by Likud Party leader Ariel Sharon (r. 2001-2006) in September 2000 to what is surely the most contested piece of real estate on Earth: the hill that Jewish people refer to as the Temple Mount and that Muslims know as the Haram al-Sharif. Accompanied by a phalanx of security personnel, Sharon made his appearance to assert Israel's claim to the site and to embarrass and weaken Barak, against whom he was then running for prime minister. Deliberately provocative, his action was the final straw for the Palestinians. Already seething over the failure of the Oslo process to meet the expectations it had raised, they responded by launching a new, unplanned uprising known as the Second *Intifada*.⁵⁷

It began the day after Sharon's visit. Muslims leaving the Haram al-Sharif after Friday prayers spontaneously protested his provocative move by rioting and throwing rocks at the police. Israeli security forces escalated the situation by responding with overwhelming force including live fire that killed four Palestinians and wounded another one hundred. From that point, the riots quickly got out of hand. Drawing inspiration from the recent success of Hizbullah guerrillas against the IDF in Lebanon, the Palestinians sought to use mass, populist violence to compel the Israelis to unilaterally cede the entirety of the territories. The Second *Intifada* failed to attain that goal; instead, the rioting and the forceful Israeli response merely served to destroy the few lingering shreds of trust on which the negotiations of the prior seven years had rested—especially after Sharon won election as prime minister in February 2001. Thus,

⁵⁵ Douglas Little, *Us versus Them: The United States, Radical Islam, and the Rise of the Green Threat* (Chapel Hill: University of North Carolina Press, 2016), 115–18; Khalidi, *The Hundred Years' War on Palestine*, 211.

⁵⁶ Rashid Khalidi, *Brokers of Deceit: How the U.S. Has Undermined Peace in the Middle East* (Boston: Beacon Press, 2013), 38.

⁵⁷ Morris, *Righteous Victims: A History of the Zionist-Arab Conflict, 1881-1920*, 660–62.

weakened by the deliberate actions of hardliners on both sides, the peace process that for a time had seemed so promising came to a bitter and ignominious end.⁵⁸

Historical Disagreements: The Oslo Accords

Unsurprisingly in light of the intense feelings that have long surrounded the Israeli-Palestinian conflict, scholars began offering sharply differing interpretations of the Oslo Accords almost from the moment that Arafat and Rabin signed the first agreement in 1993. Some historians cast the negotiations in a largely positive light. While acknowledging the peace process's obvious limitations and eventual failure, they argue that it constituted a major step toward the resolution of the Israeli-Palestinian struggle. For example, the historian Kirsten Schulze contends that the Oslo I Accord "was of historic significance" because it involved each party recognizing that the other was a legitimate nation and that they could together arrive at a solution to the long-standing crisis. Avi Shlaim also takes a positive view of the first accord. He contends that the talks in Norway constituted a genuine "breakthrough" and that Israel's willingness to negotiate with the PLO amounted to "a diplomatic revolution." Above all, he maintains that the shift from the mutual denial of each other's "right to self-determination" to a shared acceptance that each side had a legitimate claim to part of the territory constituted a change of such consequence as to render the Oslo Accords worthy of the "epithet 'historic.'"⁵⁹

This view does not represent a consensus of scholars, however. On the contrary, noting that the agreements offered few real concessions to the Palestinians, many historians vehemently dispute the idea that the Oslo Accords marked a substantial step toward peace. Writing shortly after the signing of the first agreement, for example, the brilliant and influential Palestinian-American academic Edward Said denounced the deal as a poorly negotiated "'Palestinian Versailles'" that conceded control of the West Bank to Israel in exchange for what amounted to a highly constrained form of autonomy. More recently, the historian Ussama Makdisi has extended Said's critique. He maintains that the Oslo Accords' failure to achieve the removal of Israeli settlements or to explicitly endorse the establishment of a Palestinian state rendered the agreements painfully one sided. Indeed, he concludes that the accords left the Palestinians with essentially the same limited autonomy arrangement that Begin had offered in 1978—save that Arafat's pseudo-government now had the added responsibility of helping to ensure Israel's security.⁶⁰

The historian Rashid Khalidi offers perhaps the most damning indictment of the peace process. He argues that the Oslo Accords were a disaster for the Palestinians in

⁵⁸ Kirsten E. Schulze, *The Arab-Israeli Conflict* (New York: Pearson, 2008), 83–84.

⁵⁹ Schulze, 79; Shlaim, *The Iron Wall*, 530, 536–37. Note, though, that Shlaim did not perceive the Oslo Accords to be as pathbreaking in the revised edition of *The Iron Wall* published in 2014 as he did in the first edition published in 2000.

⁶⁰ Said quoted in Ussama Makdisi, *Faith Misplaced: The Broken Promise of U.S.-Arab Relations: 1820-2001* (New York: PublicAffairs, 2010), 342–44.

every regard and that they made a just settlement of the ongoing conflict more rather than less difficult. In his eyes, the agreements were completely one-sided and secured for the Palestinians no more than “a highly restricted form of self-rule in a fragment of the Occupied Territories”—an arrangement that merely served to lay the foundation for future oppression and conflict. Indeed, he concludes, the agreements had achieved so little and cost so much for the Palestinians that “a failure to reach a deal would have been better than the deal that emerged from Oslo.”⁶¹

Khalidi is particularly unstinting in his criticism of Yasir Arafat’s management of the negotiations. He locates the genesis of the PLO chair’s mishandling of the Oslo process in his shortsighted decision to back Saddam Husayn in the conflict between Iraq and the United States in 1990-1991. Desperate to restore relevance to the PLO in the wake of Husayn’s defeat, Arafat followed that disastrous move by plunging headlong into the peace talks. Signing deals that Palestinian experts had not yet vetted, he agreed to let Israel retain total authority over key issues such as borders and water. Even worse, Khalidi argues, was Arafat’s decision to recognize Israel in exchange for its recognition of the PLO. Khalidi allows that Rabin’s formal acknowledgment of the Palestinians as a people marked a significant shift in Israel’s position; however, he also notes that this concession came at the cost of the only meaningful card Arafat had to play—recognition of Israel—and asserts that the PLO chair should have kept it in his hand until Israel committed to a Palestinian state in the Occupied Territories.⁶²

Iraq: Dual Containment

The Gulf region also experienced a rise in tension during the mid-to-late 1990s. Its roots lay in the Bush and Clinton administrations’ desire to restrain both Ba’thist Iraq and revolutionary Iran. They did so through a new policy called “dual containment.” Designed to neutralize Saddam Husayn while simultaneously preventing Iran from taking advantage of his weakness to expand its influence into Iraq, it was a complex policy—one that relied on a combination of force, economic leverage, and extensive international cooperation to keep both countries weak.⁶³

Washington’s efforts to limit Tehran’s influence were fairly straightforward. Involving an escalating series of sanctions designed to weaken Iran’s economy, the campaign initially banned American oil companies from participating in the extraction of Iranian petroleum and barred American citizens from doing business in the country. More controversially, after 1996 the sanctions regime expanded to include penalties on

⁶¹ Khalidi, *The Hundred Years’ War on Palestine*, 200–201.

⁶² Khalidi, 184–85, 198–201.

⁶³ Steven Hurst, *The United States and Iraq since 1979: Hegemony, Oil and War* (Edinburgh: Edinburgh University Press, 2009), 119–21.

foreign companies that invested more than token amounts of money in Iran's oil industry.⁶⁴

Centering on the aggressive enforcement of Gulf War-era UN Security Council Resolutions, the American effort to restrain Iraq was more complex. It involved three components. First, Washington sought to defang Iraq by aggressively pursuing an intrusive arms inspection regime under the United Nations Special Commission (UNSCOM) designed to uncover and destroy Iraq's WMD programs. Second, the US vigorously enforced financial and trade sanctions on Iraq that were to remain in effect until the Security Council had determined that Baghdad was in full compliance with all UN resolutions. Finally, the US sought to limit Saddam Husayn's freedom of action by maintaining and expanding the existing No Fly Zones over northern and southern Iraq and by launching punitive air attacks when Iraqi forces resisted.⁶⁵

For the first few years, dual containment was successful. The Iraqi regime went to substantial lengths to mislead the arms inspectors, but it nonetheless grudgingly permitted them access to its weapons' production facilities. Meanwhile, the strict enforcement of sanctions on Iran prevented Tehran from taking advantage of Iraq's weakness. More encouragingly, the election of the moderate Seyyed Mohammed Khatami (r. 1997-2005) in 1997 appeared to suggest that Iran was no longer the ideological threat it had once been, and that the US could consider loosening the enforcement of the sanctions it had imposed on Tehran.⁶⁶

There was no corresponding thaw in Iraqi-American relations, however. Instead, by the mid 1990s, Saddam Husayn had become increasingly aggressive in his efforts to win international support for an end to the restrictions that the US was enforcing. He began by exploiting reports that there had been a massive increase in infant mortality in Iraq following the end of the Gulf War. More detailed investigations later revealed that these studies were based on doctored Iraqi figures and that the child-mortality rate had not meaningfully increased; in the immediate term, however, Saddam Husayn's effort proved surprisingly effective in generating pressure on the UN Security Council. Indeed, the Iraqi campaign was so successful that the Security Council responded in 1995 with a new carveout to the sanctions' regime called the Oil-for-Food Program. Designed to end the suffering of the Iraqi people while simultaneously keeping sanctions in place, it

⁶⁴ Nikki Keddie, *Modern Iran: Roots and Results of Revolution* (New Haven: Yale University Press, 2006), 565.

⁶⁵ Gary Sick, "The United States in the Persian Gulf: From Twin Pillars to Dual Containment," in *The Middle East and the United States: A Historical and Political Reassessment*, ed. David Lesch, 4th ed. (Boulder: Westview Press, 2007), 323–27.

⁶⁶ Sick, 323–27.

permitted Iraq to sell large quantities of petroleum with the stipulation that it use the proceeds to buy only food, medicine, and other necessities.⁶⁷

Pleased to have cracked the sanctions regime, Saddam Husayn sought to eliminate it altogether. To do so, he first sought to drive a wedge between the US—which pressed for the continued vigorous enforcement of sanctions—and countries like France, Russia, and China that increasingly wanted to pursue commercial opportunities in Iraq. He followed by stepping up his resistance to UNSCOM's arms control inspections. Claiming that the inspectors were intruding on Iraqi security sites, he began stonewalling them in 1997 before ceasing to comply with the inspection regime altogether in the fall of 1998.⁶⁸

Iraq's increasing defiance put the Clinton administration in a bind. On the one hand, the American public showed no stomach for another war with Iraq. On the other, dual containment had clearly failed to curb Saddam Husayn. Further complicating matters, a group of newly influential conservative policymakers known as "Neocons" began to pressure the administration to abandon containment in favor of outright regime change. Clinton rejected that course but did respond to Husayn's unilateral termination of the inspection program with a massive four-day air assault on Iraq in December 1998 codenamed Operation Desert Fox. The president also signed the Iraq Liberation Act in October of that year, which made "regime change" formal American policy. The act was largely symbolic and did not meaningfully change the US approach to Iraq, but it did signal Washington's increased willingness to pursue a more confrontational approach. It also made clear the frustration of many in the US government with the fact that nearly a decade after they had celebrated a victory believed to have neutralized Saddam Husayn, they still found themselves locked in an expensive conflict with him that showed no signs of resolution.⁶⁹

Transnational Jihad

The sanctions and American punitive attacks on Iraq played a part in fueling a new struggle that pitted the United States against a shadowy transnational terrorist organization called al-Qa'ida that Osama bin Laden had first organized in 1988 while the Afghan insurgency was winding down. The seeds of that conflict first took root during the chaotic days immediately following the Iraqi invasion of Kuwait in 1990. As we have seen, Riyadh had responded to Saddam Husayn's aggression by agreeing to the deployment of American forces to Saudi Arabia. That decision shocked Osama bin Laden. He strongly agreed that Iraq needed to be pushed out of Kuwait, but he recoiled

⁶⁷ Tim Dyson and Valeria Cetorelli, "Changing Views on Child Mortality and Economic Sanctions in Iraq: A History of Lies, Damned Lies and Statistics," *BMJ Global Health* 2, no. 2 (2017): 1–5; Marr, *The Modern History of Iraq*, 281–83, 288–90.

⁶⁸ Tripp, *A History of Iraq*, 278–79.

⁶⁹ Hahn, *Missions Accomplished?*, 127–31.

at the thought of thousands of non-Muslim Westerners being stationed on land that he viewed as sacred.⁷⁰

In response, he devised a strategy to resolve the crisis that relied not on Western forces but instead on Muslim guerrillas patterned on the “Afghan Arabs” who had fought the Soviets in the 1980s. Parlaying the influence he had gained from his work in Afghanistan, he was able to present his plan to several powerful Saudi princes. The strategy he laid out was a bold one. Claiming that he could quickly raise an army of 100,000 disciplined, well-trained fighters, he argued that his soldiers would force the Iraqis to retreat using the same tactics that had compelled the USSR to abandon Afghanistan in 1988. Exuding absolute confidence in his approach, he blithely assured one prince that “we are ready to get the Iraqi forces out of Kuwait.” As such, he was shocked when Riyadh rejected his proposal out of hand. Dubious that Bin Laden could raise the promised troops or defeat Iraq’s enormous army, King Fahd instead preferred to stick with America’s more-dependable military might.⁷¹

Bitter about Saudi Arabia’s decision to allow Westerners into the land of the Holy Places, Bin Laden began to prepare to go to war with both the US and the kingdom. He did not immediately organize attacks against them, however. Instead, facing Saudi government pressure for his criticism of the royal family, he first relocated to Sudan in 1992. There, under a friendly Islamist government that happily entertained both anti-Western terrorists such as Carlos the Jackal (1949-) and jihadi figures like Zawahiri, he prepared for the coming conflict by building up al-Qa‘ida’s infrastructure and organizational capacity. As part of that effort, he created training camps, founded a series of businesses at which he employed many jihadis, planned terrorist attacks, and cemented his ties with members of Egypt’s *Gamaat Islamiya* and Islamic Jihad.⁷² Above all, however, he spent his time in Sudan fuming about the American military deployment in Saudi Arabia. Fahd had pledged that it would last only until the end of the Gulf War; with Saudi-based US aircraft enforcing the southern No Fly Zone in Iraq, however, it certainly appeared to have become a permanent occupation. Bin Laden responded by stepping up his criticism of the Saudi dynasty—a move that led Riyadh to successfully pressure the Sudanese government to kick him out of Sudan in 1996. No longer welcome in Khartoum, he found sanctuary in Afghanistan where a group of religious students called the Taliban had recently established an austere, Wahhabi-based Islamist state.⁷³

It was during this time that Bin Laden developed a radical new construction of jihad. Heretofore, a consensus had existed among the jihadis that the movement

⁷⁰ Fawaz A. Gerges, *The Rise and Fall of Al-Qaeda* (Oxford: Oxford University Press, 2014), 48–49.

⁷¹ Gerges, *The Far Enemy*, 145–48.

⁷² Lawrence Wright, *The Terror Years: From al-Qaeda to the Islamic State*, downloadable ebook (New York: Alfred A. Knopf, 2016), 185–92.

⁷³ Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 185–92; Gerges, *The Rise and Fall of Al-Qaeda*, 56.

needed to defeat the near enemy—meaning the secular regimes that dominated the Middle East—and reestablish shari‘a law before the Muslim world would have the strength to confront the far enemy. Bin Laden turned this strategy on its head. He argued that instead of focusing on the near enemy—an approach that had nearly destroyed the jihadi organizations in the 1990s—militant Muslims should go to war with the far enemy. But how would this plan work? How would attacking the world’s wealthiest and most powerful country help the jihadis to achieve their goals? For Bin Laden, the key was Washington’s reaction. He argued that attacking America would spur the US to mount a response so devastating that it would finally rouse the globe’s Muslims to action against the West. In the process, the reawakened Muslim community would overthrow the apostate, near-enemy states and replace them with Islamist ones. Living under God’s law, the Islamic world would then defeat the far enemy and once again assume its rightful position of global leadership.⁷⁴

Bin Laden’s new idea sharply divided the militant organizations. Some, most notably Ayman Zawahiri, enthusiastically endorsed his new, transnational approach and urged jihadis to attack the “head of the snake.” In contrast, many others argued that waging war against the only remaining superpower was folly and refused to take part. Indeed, *Gamaat Islamiya* went so far as to release a statement affirming that “[w]e are not a party in any front that confronts Americans.”⁷⁵

Bin Laden and Zawahiri were unmoved by these objections, however, and instead pressed ahead with their transnational campaign. Most notably, in February 1998, al-Qa‘ida announced the establishment of a new organization, the International Islamic Front for Jihad Against Jews and Crusaders, and issued a formal declaration of war on the United States. It charged the US with three offenses: occupying the sacred land of Arabia, making war on the Iraqi people via sanctions and punitive attacks, and supporting Israel. Importantly, it also declared that all Muslims bear an “individual duty” to kill Americans in order to end its war on the Islamic world.⁷⁶

The organization did not confine itself to a war of words, moreover. Instead, it followed the declaration with its first large-scale terror operation. Timing their attack to coincide with the eighth anniversary of the first Desert Shield deployment, Al-Qa‘ida suicide bombers simultaneously detonated truck bombs outside the US embassies in Kenya and Tanzania in August 1998. Twelve Americans and 212 Africans died in those attacks, and another four thousand people were injured. Alarmed, Washington retaliated by launching cruise missiles at an al-Qa‘ida training camp in Afghanistan and at a pharmaceutical plant in Sudan it suspected—incorrectly—of producing chemical weapons. Undeterred by the American response, Bin Laden followed with a series of largely failed terror operations across the globe during the celebration of the new millennium, and then with a successful suicide-bombing attack on the guided-missile

⁷⁴ Gerges, *The Rise and Fall of Al-Qaeda*, 55, 65–66.

⁷⁵ Gerges, 65–67.

⁷⁶ Wright, *The Looming Tower: Al Qaeda and the Road to 9/11*, 294–95; Quote from Gerges, *The Rise and Fall of Al-Qaeda*, 78.

destroyer *USS Cole* in Aden harbor that killed seventeen sailors and nearly made the vessel the first American warship sunk in combat since the end of the Second World War. Meanwhile, planning continued for a major attack on the American homeland—one to be of such scale and impact that it would ignite the Islamic-Christian clash of civilizations that Bin Laden believed would finally reunite the *umma* under shari‘a law.⁷⁷

Conclusion

Thus by 2000, the earlier hope that the Middle East was poised to move beyond the conflicts that had characterized the region during the twentieth century had faded. It is true that the religious cold war between Iran and Saudi Arabia had ended, but other conflicts persisted.⁷⁸ Saddam Husayn continued to defy the UN sanctions regime, transnational jihadis were increasingly determined to incite a conflict with America, and the Palestinian-Israeli conflict that had seemed so tantalizingly close to a peaceful resolution now appeared to be as depressingly insoluble as ever.

There was also an important change to the geostrategic balance in the Middle East that seemed to further complicate events in the area. The US now had a substantial military presence in the region—one that had become, in the words of the historian Andrew Bacevich, “permanent and sustained, rather than occasional and episodic” and that had, by its very presence, stoked the conflict between the emerging transnational jihadis and the West. Still, if the optimism of the early 1990s had faded, the Middle East nonetheless seemed no worse off than it had been in recent decades. Indeed, few at the turn of the century had the least inkling about the shocking events and metastasizing conflicts that would come to characterize the Middle East during the first decade of the new millennia.⁷⁹

⁷⁷ Bacevich, *America’s War for the Greater Middle East*, 205–13.

⁷⁸ Note, though, that the conflict between Iran and Saudi Arabia would roar back to life with a vengeance in the 2000s.

⁷⁹ Bacevich, *America’s War for the Greater Middle East*, 112.