CHAPTER ONE

Introduction

WHAT CAUSES VICTORY and defeat in battle? Why do the winners win and the losers lose? What makes some campaigns bloody stalemates and others apparent cakewalks? How can states maximize their odds of winning and minimize their casualties? And are the answers changing in the information age? Will new technology or changing geopolitics transform warfare, creating new winners—and new losers—on the battlefields of the future?

These are life-and-death questions, and not just for soldiers. They affect everyone: from infantrymen on the battlefield to office workers in the World Trade Center to entire nations and peoples. German victories in Poland and France condemned millions of French and Polish Jews to the gas chambers in World War II. Soviet victories in 1944–5 consigned a generation of East Europeans to a Communist oppression that was spared those who were reached first by American and British armies. The trench stalemate of 1914–18 sentenced millions to deaths that a quicker victory would have averted and ruined economies across Europe. By exhausting some countries and embittering others, it shaped the subsequent politics of the twentieth century. Today's world would be very different if European generals had fought differently in 1914–18. Defeat in battle has meant occupation and a conqueror's rule for nations from France to Poland and from Japan to Indonesia. Today, swift victory in a war against fundamentalist terrorism could spare the lives of thousands—or millions—in the West and Islam; failure could bring consequences too horrible to contemplate.

With so much at stake, great effort has been spent on these questions. The answers, however, often fall short.

As recently as 1991, a massive effort using state-of-the-art methods and the nation's best analysts radically overestimated U.S. losses in the upcoming Gulf War. The prewar congressional debate hinged on casualty expectations; these were widely seen as the key to Congress's vote on the use of force. With so much at stake, no effort was spared to achieve realistic estimates: prominent academics, government analysts, and senior military officials gave testimony using methods ranging from computer models to historical analogies to professional military judgment. Virtually all were way off. Even the closest estimate overshot the actual casualty count by more than a factor of two. The next best missed by a factor of six. The majority were off by more than an order of magnitude; official estimates were reportedly over by at least that much; while some official projections erred by a factor of over 200.[1]

The pre-Gulf war debate was hardly unique, however. For over a century, soldiers and diplomats have analyzed military balances, yet the subsequent fighting has often surprised both sides. In 1914, for example, Europeans expected a short, decisive war of movement. None foresaw a nearly four-year trench stalemate—if they had, the war might never have happened.[2] In 1940 Allied leaders were astonished by the Germans' lightning victory over France. They had expected something closer to the trench warfare of 1914–18; even the victors were surprised, with most German planners expecting an extended war of attrition in France and the Low Countries.[3] In 1973 both Israelis and Arabs alike were stunned by the October War's staggering losses, which forced Israel to beg emergency aid from the United States and spurred a wave of postwar pronouncements that tanks were now doomed, given their apparent vulnerability to guided missiles.[4] By the mid-1970s, tanks were widely seen as dinosaurs, the next generation of obsolete weapon to follow the horse cavalry or the battleship. Yet this, too, proved wrong: by the mid-1990s, the M1 tank was widely hailed as an invincible "King of the Killing Zone" after its nearly casualty-free performance in the Gulf War.[5]

Academics have done no better. Military balance estimates are central to modern political science. Much of our current understanding of international politics rests on the assumption that state behavior is shaped by the threat of war and the pursuit of military capability. The empirical study of politics thus depends on measures of capability, which play a pivotal role in the war causation, arms racing, alliance formation, conflict duration, crisis escalation, or deterrence literatures, among others.[6] Yet the standard capability measures at the heart of all this are actually no better than coin flips at predicting real military outcomes.[7] An enormous scholarly edifice thus rests on very shaky foundations.

We must—and can—do better. But real improvement will require a new approach.

Today, most analyses are either rigorous but narrow, or broad but unrigorous. Mathematical models of combat, for example, are rigorous but typically focus on material alone: how many troops or weapons do the two sides have, and how good is their equipment?[8] By contrast, holistic assessments consider issues such as strategy, tactics, morale, combat motivation, or leadership as well as just materiel but treat these variables much less systematically. Real progress demands rigor *and* breadth: a systematic treatment of both material and nonmaterial variables, backed up with a combination of empirical evidence and careful deductive reasoning. Below I advance such an analysis for one key nonmaterial variable: force employment, or the doctrine and tactics by which armies use their materiel in the field.

I hold that a particular pattern of force employment—the *modern system*—has been pivotal in the twentieth century and is likely to remain so. I argue that since at least 1900, the dominant technological fact of the modern battlefield has been increasing lethality. Even by 1914, firepower had become so lethal that exposed mass movement in the open had become suicidal. Subsequent technological change has only increased the range over which exposure can be fatal. To perform meaningful military missions in the face of this storm of steel requires armies to reduce their exposure, and since 1918 the central means of doing so has been modern system force employment.[9]

The modern system is a tightly interrelated complex of cover, concealment, dispersion, suppression, small-unit independent maneuver, and combined arms at the tactical level, and depth, reserves, and differential concentration at the operational level of war. Taken together, these techniques sharply reduce vulnerability to even twenty-first century weapons and sensors. Where fully implemented, the modern system damps the effects of technological change and insulates its users from the full lethality of their opponents' weapons.

Not everyone can master it, however. The modern system is extremely complex and poses painful political and social tradeoffs. While some have been able to surmount these challenges and implement the modern system fully, others have not. Militaries that fail to implement the modern system have been fully exposed to the firepower of modern weapons—with increasingly severe consequences as those weapons' reach and lethality have expanded. The net result has thus been a growing gap in the real military power of states that can and cannot implement the modern system, but surprisingly little change over time in outcomes between mutually modern-system opponents.

The modern system is also essential to understanding the role of numerical preponderance in war. Many suppose that victory normally goes to the preponderant side. Without modern-system exposure reduction, however, armies cannot survive long enough to make their numbers tell. Superior numbers can be decisive or almost irrelevant depending on the two sides' force employment. This in turn means that states' relative economic, demographic, or industrial strength are poor indicators of real military power: gross resource advantages matter only if they can be exploited via modern-system force employment, and many states cannot do so.

If so, then to assess military power without taking force employment into account is to risk major error. Assessments focusing solely on materiel will radically overestimate well-equipped but poorly handled armies—such as the Iraqis in 1991—and underestimate poorly equipped but well-handled troops—such as the North Vietnamese in 1965–72. The same policy initiatives can have opposite effects depending on the two sides' force employment. Typical *ceteris paribus* cost-effectiveness analyses can thus be dangerously misleading: a new weapon can look wonderful when a non–modern system opponent is assumed but terrible if the enemy uses its forces differently, and neither assumption will hold true all the time. In fact, analyses considering materiel alone may be little better than blind guesses.

The results challenge a wide variety of standard views, ranging from current U.S. defense policy and common projections for future warfare to mainstream international relations (IR) theory and orthodox interpretations of twentieth-century military history. The defense debate is increasingly focused on technology; most assume that in the information age, superior technology wins wars, fueling growing pressure to speed modernization by spending less on training and readiness. Official analyses reinforce this trend: official models that focus on materiel and exclude force employment overvalue the former and undervalue the latter. Policy decisions informed by such models are thus likely to overspend on modernization and force structure and underspend on the readiness and training needed for sound force employment. Similarly, threat assessments based on the numbers and types of hostile weapons are likely to overestimate real capability for enemies with modern equipment but limited skills but underestimate militaries with older equipment but high skills. Ensuing intervention and use-of-force decisions could produce *over*confidence and overcommitment against ill-equipped but adept militaries but *under*confidence and unnecessary caution against enemies with state-of-the-art weapons but little skill in their use.

Projections of future warfare are now dominated by the claim that technology is creating a "revolution in military affairs" (RMA) in which the nature of military power is being transformed. In the future, it is held, long-range precision air and missile strikes will dominate warfare, ground forces will be reduced mostly to scouts, and the struggle for information supremacy will replace the breakthrough battle as the decisive issue for success. These views misunderstand the relationship between technology and force employment, however. Because RMA advocates misunderstand warfare prior to the 1990s, they misread the 1991 Gulf War as a radical departure; by projecting this mistake forward into the twenty-first century, they derive a case for a radical restructuring of U.S. defense policy that is neither necessary nor desirable.

For international relations theory, the weakness of simple material proxies poses serious problems. Military power plays a pervasive role in the study of international politics; in fact, much of modern IR theory amounts to a debate on its influence over state behavior. Yet all of this rests on very simplistic treatments of its nature and determinants. Theoretically, the literature relies on logically unsound, unitary notions of military capability that mask crucial tradeoffs. Empirically, the use of weak proxies undermines existing findings and suggests that the literature may have underestimated capability's effects relative to audience costs, signaling, or resolve. Analyses of deterrence, power distribution, and polarity rest on especially thin ice given the weakness of the measures used to represent capability. The mercantilist position in international political economy is based on the proposition that economic preponderance conduces to military power, yet the relationship between economic strength and real capability is much weaker than commonly thought. More broadly, much of the empirical and theoretical literature will need to be revisited in light of a more meaningful measure of real capability.

The received view of modern military history centered on technology for much of the postwar era. The World War I trench stalemate was seen as the product of the industrial revolution in the machine gun, new artillery, and mass munition production. World War II was seen as a war of movement brought on by the tank, the airplane, and the radio. Postwar conflict was seen as overshadowed by the atom bomb.[10] A more recent interpretation emphasizes preponderance: industrial coalition wars are held to turn on the size of the combatant economies, with victory going to the side that could outproduce the other, and with the conduct of operations taking a back seat to the battle for production.[11] I suggest, however, that force employment has played a more important role than either technology or preponderance for twentieth-century warfare. How forces are used is critical; to explain historical outcomes chiefly in terms of materiel is to misinterpret the major military events of the century.

The balance of this chapter supports these contentions in three steps. First, I define what I mean by military power, or "capability," and justify my choice. Second, I discuss how I will explain capability and present some of the limitations of my analysis. Finally, I outline the book's structure and provide a road map for what is to come.

What Is "Military Power"?

My focus is on the military dimension of power. It is important to be clear on its nature and limits. State power embodies "soft" persuasive or attractive elements as well as its "hard" or military component.[12] Nor does military power in itself guarantee success in war. Militarily weak but resolute states can prevail over militarily strong but irresolute ones. War outcomes are products of more than just military power alone.

Moreover, military power (or "capability") itself can mean different things in different contexts. Military forces, after all, do many things, ranging from defending national territory to invading other states, hunting down terrorists, coercing concessions, countering insurgencies, keeping the peace, enforcing economic sanctions, showing the flag, or maintaining domestic order. Proficiency in one or even several does not imply proficiency in them all: good defenders of national territory can make poor peacekeepers; forces that can defend national territory cannot necessarily conquer their neighbors.[13] For any one mission, moreover, "success" can be defined very differently by different actors. Defenders of national territory may all value low casualties, short wars, and complete restoration of the status quo, but these goals often conflict with one another, and different defenders value them differently at the margin. Some would trade higher casualties and a longer war for complete reconquest of lost territory; others would not.[14] Some would bomb an opponent for months to avoid losing friendly ground troops; others would invade quickly to shorten the war at the cost of heavier casualties.[15] If capability is the ability to succeed at an assigned mission, different states will thus assess capability very differently for the same forces—no single, undifferentiated concept of "military capability" can apply to all conflicts in all places and times.

Any analysis must therefore focus on a subset of the tasks militaries perform, which are in turn a subset of the elements of state power. For my purposes, I concentrate on the mission of controlling territory in mid- to high-intensity continental warfare, and I define its accomplishment via three interconnected criteria: the ability to destroy hostile forces while preserving one's own; the ability to take and hold ground; and the time required to do so.

Specifically, I define offensive military capability as *the capacity to destroy the largest possible defensive force over the largest possible territory for the smallest attacker casualties in the least time*; defensive military capability is conversely the ability to preserve the largest possible defensive force over the largest possible territory with the greatest attacker casualties for the longest time. As these criteria can be fulfilled in differing degree (and often conflict with one another), I offer a theory that explains casualties, ground gain, and duration as distinct but interconnected outcomes; I then discuss the interactions and tradeoffs among them in light of the proposed theory as a whole.[16]

My unit of analysis for this theory is the *operation*. An operation is a series of interconnected battles resulting from a single prior plan. The battles associated with the invasion of France and the German drive to the Channel in 1940, for example, comprise an operation (Operation FALL GELB), as do the battles associated with the American breakout from the Normandy beachhead beginning on July 25, 1944 (Operation COBRA).[17]

By "mid- to high-intensity conflict," I mean the middle part of a spectrum ranging from guerilla warfare at the low end to global thermonuclear war at the high end. "Mid-intensity" conflicts would include regional conventional wars such as the recent campaigns in Afghanistan, the Balkans, or Kuwait; the Arab-Israeli wars; the Sino-Vietnamese War; or the Indo-Pakistani wars. "High-intensity" conflicts are conventional world wars among the great powers. I thus exclude guerilla warfare at the low end, and mass destruction warfare involving nuclear, chemical, or biological weapons at the high end of the spectrum of conflict. By "continental warfare" I mean combat fought between military forces on or over major land masses. I thus exclude war at sea, and strategic bombing against civilian targets.

I thus seek to explain the outcomes of operations to control territory in mid- to high-intensity continental warfare. Why this focus? Is this just irrelevant "old thinking" in an era of counterterrorist warfare, ethnic conflict, coercive strategic bombing, and weapons of mass destruction (WMD)?

The answer is no. While major conventional war is only one among many important missions, it remains far more important than some now suppose, and it will be for the foreseeable future. It will also remain the most expensive mission to fulfill, it will remain the central purpose for the majority of the U.S. military, and it will continue to occur between other parties in other parts of the world.

In the emerging war on terrorism, for example, counterintelligence and police work against terrorists hiding in the shadows will be accompanied by periodic major warfare against states who harbor them. This is a central implication of the "Bush Doctrine," which holds states accountable for the actions of terrorists within their borders and uses the threat of major war as a central means of enforcing accountability.[18]

The recent campaign to destroy al Qaeda's sanctuary in Afghanistan, for instance, was precisely the kind of regional conventional war I examine here.[19] Contrary to popular perception, Afghanistan was neither a guerilla war nor simply long-range bombing. The Taliban regime sought to control territory and defend key geographic objectives, not merely to harass their enemies with hit-and-run tactics. And this struggle for territorial control involved substantial close combat on the ground between Western and allied Afghan infantry on one side, and opposing Taliban fighters who had eluded American surveillance and survived American air strikes on the other. The result was a series of surprisingly orthodox ground battles, as at Chapchall on October 23, 2001, Sayed Slim Kalay on December 2–4, Highway 4 on December 2–6, or Operation ANACONDA in March 2002.[20] The critical action in the war's northern phase, for example, involved a breakthrough battle near the village of Bai Beche on November 5, 2001, in which Taliban forces that had survived more than two days of preliminary American bombing were overrun, opening the door for the Northern Alliance's advance to Mazar-e-Sharif.[21] Until ANACONDA, the *American* role in the war was mostly bombing (and spotting targets for the bombers), but the war itself involved far more than just the American contribution. That contribution was critical, but its role was to strengthen allied ground forces to enable them to prevail in traditional close combat—the air strikes did not simply annihilate the Taliban or break their will to fight. New technology played an important role in Afghanistan—and its relative importance is a major theme below—but the war itself was precisely the kind of mid- to high-intensity struggle for territorial control on which I focus here.

Nor is it clear that direct U.S. military involvement in the war on terrorism will be mostly air power or small special forces teams. Any initially limited U.S. intervention will face powerful pressures to escalate if small-scale efforts prove insufficient.[22] And among America's most powerful escalatory threats is the ability to topple regimes by invading and taking political control of their territory—that is, by fighting and winning a major conventional theater war. Against regimes like Mullah Omar's, Saddam Hussein's, Bashar Assad's, or Kim Jong-Il's, this is the ultimate sanction. It threatens what they value most: their hold on power. And it is the single most credible threat one can direct at this value: few regimes can survive an American march on their capital. Even where this ultimate sanction is unused, its existence makes other coercive means more effective: it makes one's opponents consider one's ability to remove them by force if they ignore lesser threats. In Kosovo, the United States initially took this threat off the table only to regret it later; it would be a similar mistake now to assume that lesser threats will always suffice, and thus that America will never again have to wage a major theater war.

Major war is also the primary planning yardstick not only for U.S. forces, but for most world and regional powers. For most of the post–Cold War era, the U.S. military was sized and structured to win two, nearly simultaneous major regional conflicts; the Bush administration has modified this standard to winning one while holding the line in another, but the standard is still set in major-war terms.[23] Most of the U.S. military is oriented to this threat; by contrast, the special forces, which some now see as the vanguard of American military action, consume only about 1 percent of annual U.S. defense spending.[24] Many of the world's other major militaries are similarly oriented: India's, for example, is designed for major war with Pakistan or China; Pakistan and China must be ready for major war with India; Israel must prepare for major war with Syria or other Arab neighbors.

Nor are concerns with major warfare limited to great and regional powers, or wholly superseded by ethnic disputes, guerilla warfare, or other low-intensity conflicts elsewhere. The recent wars in Bosnia, Croatia, Eritrea, Zaire/Congo, Rwanda, Azerbaijan, and Kuwait were all mid- to high-intensity conflicts in which combatants sought to take and hold territory in conventional ways.[25] If war breaks out tomorrow in Kashmir or the Bekaa Valley, the fighting would not be low intensity. The conflicts I focus on here are hardly a thing of the past for anyone, and their centrality to U.S. interests makes them an appropriate place to start in understanding capability.[26]

I exclude countervalue strategic bombing chiefly because its dynamics are well studied elsewhere ("counterforce" violence is directed against hostile military forces; "countervalue" violence is aimed at hostile populations, economic centers, or political leadership).[27] Purely countervalue bombing is also less common than many think—and much less successful. To date, almost all strategic bombing has blended countervalue coercion with military counterforce to reduce an opponent's ability to wage war.[28] The latter is within my scope and has provided the overwhelming bulk of strategic bombing's actual historical impact. Whatever its intent, the primary result of Allied strategic bombing in World War II was to reduce Germany's war-making capacity.[29] Countervalue strategic bombing in Korea and Vietnam failed to bring political concessions directly; where it succeeded it was by reducing the target states' ability to wage a counterforce war effort.[30] In the Gulf War, strategic bombing was mostly counterforce in intent; limited countervalue bombing aimed at coercing Saddam by threatening his hold on power failed.[31] In Kosovo, NATO bombing combined countervalue strikes against Serbian leadership and economic infrastructure with counterforce missions against Serbian ground forces; Serbian concessions occurred only after NATO began preparations for a major land invasion of Kosovo.[32] In Afghanistan, early hopes that bombing Taliban leadership targets would yield concession-proved unrealistic. Strategic bombing has thus shown little ability to succeed via countervalue coercion, and recent experience in Kosovo and Afghanistan gives little reason to expect change any time soon.

Weapons of mass destruction (i.e., nuclear, chemical, and biological weapons) are a major threat and will become an increasingly common problem as technology proliferates. They are clearly an important issue for capability, but not the *only* important issue. Many states will either fail to acquire WMD or choose not to use them; conventional capability will thus remain important even as WMD proliferate.[33] Moreover, to understand WMD's military effects, one must explain conventional capability first. Regional mass destruction warfare would probably not shut down conventional operations by a great power: regional nuclear arsenals will probably be tiny for the foreseeable future, and most great powers train their troops to fight in chemical and biological environments.[34] The nature of the fighting would change, perhaps drastically, as the combatants seek to cope with damage incurred and reduce vulnerability to further attacks. But most do this by modifying their conventional-war methods for the special conditions of WMD (e.g., by spreading out troops and supporting infrastructure). To understand such measures' effects, one needs to understand the effects of dispersion, for example, on military outcomes generally. Without this, assessments rest on unnecessarily speculative ground. Understanding conventional operations alone is obviously insufficient to assess WMD, but it is a necessary precondition.

Finally, international relations theory rests on the empirical record of state behavior—its central function is to explain that behavior as we have observed it—and for modern interstate warfare, most of the empirical record concerns conventional continental counterforce. Of the forty-six twentieth-century interstate wars in the University of Michigan's Correlates of War dataset, for example, fully forty have been primarily conventional continental counterforce in nature.[35] If capability is an important contributor to state behavior, then the issues I address here are essential foundations for understanding the politics of international conflict.

Methodology

I explain capability in modern counterforce warfare using a combination of complementary methods. Social science has yet to provide a single perfect methodology; each of the major research traditions has important shortcomings taken alone. Each tradition also has important strengths, however. Taken together, a combination of contrasting approaches offers an opportunity to cover the weaknesses of each with the strengths of others, providing what Donald Polkinghorne has termed "methodological triangulation," and increasing our ability to tease knowledge from the imperfect data available to us.[36]

In particular, I combine close review of recent historiography with formal theory, case method, statistical analysis, and simulation experimentation. As for the first of these, historians are now developing a major reinterpretation of the two world wars (especially the first), focusing on the role of doctrinal adaptation for the wars' course and outcome.[37] This emerging view has important—but as yet largely unrecognized—implications, not just for the military history of the remainder of the twentieth century, but also for current defense policy and IR theory. Historical narrative, however, is an awkward medium for addressing such broader concerns. Its scope is typically limited narrowly to particular events. And natural-language narrative is ill-suited for sorting out the internal logic of complex, multivariate relationships, or for projecting trends into new situations where particulars differ from past events.

By contrast, formal theory, or the use of mathematical language to describe causal relationships, has advantages in sorting out the internal logic of complex, interconnecting claims. It also facilitates inference from observations of the past to conjectures about the future. Its specificity can strengthen policy prescription while disciplining thought. Formal theory alone, however, can be precise and specific but wrong—or, maybe worse, irrelevant where it abstracts away the real issues in the interest of mathematical clarity or tractability. I thus place the history first: formal language is used to generalize, systematize, and extend ideas drawn from serious historiography, harnessing deductive rigor to historically critical substance, and producing new ideas with important implications.

This new theory is tested using three different methods. First, I provide three detailed, archivally based case studies of actions fought under conditions chosen to provide maximum theoretical leverage. Small-*n* case method permits the depth of analysis needed to characterize variables, like force employment, that have not heretofore been included in large-*n* datasets. It also allows detailed process tracing to help distinguish real causation from mere coincidence. This depth of detail, however, makes it impossible to consider more than a handful of cases.

I thus complement the case studies with a series of large-*n* statistical analyses. The statistical findings speak to larger trends extending over a much wider body of experience. But since force employment has not been studied systematically, it is absent from available datasets and must thus be treated indirectly via enabling assumptions and proxy variables. Taken alone, the statistical results would offer at best a partial test; combined with case method, however, they enable methodological triangulation and the prospect of greater confidence than either could provide alone.

Both small-*n* case method and large-*n* statistics, however, are limited to battles already fought (and the peculiarities of surviving documentation). The future may differ from the past, and key details of past events may now be lost to history. Moreover, any ex post facto method—whether large or small *n*—faces a problem of selection on wars when testing theories of capability. The outcomes of interest here (casualties, territorial gain, combat duration) can be observed only in wartime. Yet "capability" exists as a potential in peacetime as well—in fact, many of its most important applications rest on claims about peacetime capability.[38] Deterrence, for example, is a product of peacetime capabilities.

International relations theories resting on capability must measure it in peacetime as well as wartime if its causal role is to be properly tested. But what if war is such an unusual special case that the relationship between capability and its causes is different there from what it is under normal, peacetime conditions? Perhaps, for example, states normally assess each others' capability in peacetime, determine likely winners and losers in advance, and avoid war by settling disputes in accordance with the mutually understood balance of power; if so, then war could occur only under unusual conditions wherein the "normal" determinants of capability were clouded by private information in ways that foreclosed peaceful settlement.[39] This would imply that all actual wars were outliers characterized by peculiar relationships among materiel, force employment, and capability—and thus that tests using only observations of actual warfare would be misleading for claims about capability in general (which pertain mostly to times of peace). While the severity of this problem is hard to gauge, it is a widespread conundrum in international relations scholarship on war, and one to which any ex post test of capability theories is inherently subject.

I thus complement ex post observation of real combat with a series of ex ante experiments using a Defense Department combat simulation, Janus, as a kind of laboratory, changing key features of a battle while holding all other aspects constant. This provides a unique ability to control for extraneous variation and observe whatever details of the (simulated) fighting may be of theoretical interest. And because experimentation creates new events under conditions chosen by the experimenter, it can create conditions of theoretical importance but historical rarity, allowing a theory's entire parameter space to be explored systematically, and freeing the analysis from dependency upon any special conditions thought to be unique to the wars states have actually fought. It can thus illuminate conditions that might be more characteristic of peacetime confrontations that do not escalate all the way to open warfare. The result is an unusually systematic form of counterfactual analysis. Any counterfactual, however, is simulation rather than reality. As such it lacks the verisimilitude of historical observation. In combination with multiple forms of empirical observation, however, it offers an important source of contrasting perspective.

Each method thus has strengths and weaknesses. None would be sufficient alone, yet each offers something unique and important. By combining them I thus exploit a wider range of potential insight and reduce the odds that my findings are artifacts of a given method's blind spots.

Plan of the Book

Chapter 2 critiques the state of the art in capability assessment in greater detail, providing a more sustained case for the need for improvement. It also uses the particular shortcomings identified to suggest some properties needed in potential improvements. Especially, it shows why a focus on the tactical and operational levels of war is an especially promising approach.

Chapters 3 and 4 present the new theory. Force employment is central to this, but little studied by theoretical social scientists. Given this, I devote chapter 3 to a detailed exposition of the modern system as it emerged by 1918. Chapter 4 then assesses post-1918 technological change, variations in numerical preponderance, and their interactions with modern-system force employment.

The theoretical discussion in chapters 3 and 4 is entirely qualitative. While this conveys the key arguments, the formal presentation in the appendix is needed for a complete understanding of the details. The appendix presents a dynamic model of territorial gain, casualties, and duration as a function of force employment, technology, and preponderance. It then treats the model's *comparative statics* (its predictions for how capability changes with controlled variation in causal variables) with much greater rigor than the qualitative discussion in chapters 3 and 4 can provide.

Chapters 5 through 7 test this theory via three historical case studies of actions selected to provide maximum theoretical leverage for the claims in the new and orthodox explanations of capability. Chapter 5 examines Operation MICHAEL, the first of the German 1918 Spring Offensives. Chapter 6 considers Operation GOODWOOD, the penultimate Allied attempt to break out of the Normandy beachhead in July 1944. Chapter 7 explores Operation DESERT STORM, the Coalition offensive in the Persian Gulf War of 1991. MICHAEL and GOODWOOD present conditions of nearly ideal defense- and offense-dominance, respectively, in orthodox theories' terms. By contrast, the new theory predicts offensive success in MICHAEL and defensive success in GOODWOOD. The historical results correspond much more closely with the new theory's predictions than its orthodox competitors,' and under conditions that should have offered easy, unambiguous predictive successes for orthodox theories if the latter were correct. While this neither proves the new theory nor falsifies its competitors, the unusual conditions merit a greater shift in confidence than would otherwise be warranted from such a small sample of cases. In Operation DESERT STORM, by contrast, both the new and orthodox theories predict a Coalition victory. The reasons behind the prediction are very different, however, and by process tracing I show that the *way* the Coalition won was consistent with the new theory but inconsistent with orthodox views, in a case whose prominence in the public debate makes it especially important for any policy-relevant theory of capability.

Chapter 8 presents the statistical analysis. To mitigate the shortcomings of individual datasets taken alone, I use a combination of three different databases with contrasting coverage, units of analysis, and sources of potential error. The results show a strong, consistent pattern of greater correspondence between the data and the new theory than for its orthodox competitors; for this to hold across such diverse data sources offers significantly greater confidence in the findings than could be obtained from any one dataset alone.

Chapter 9 presents the simulation experimentation using the results of the joint U.S. Army, Defense Advanced Research Projects Agency, and Institute for Defense Analyses "73 Easting" project. The Battle of 73 Easting was a representative action from Operation DESERT STORM in which elements of the U.S. 2d Armored Cavalry Regiment struck the Iraqi Tawakalna Division on a stretch of featureless desert near a map reference line called 73 Easting. Immediately after the war, researchers were dispatched to the battlefield to collect a historically unprecedented mass of detailed data on the minute-by-minute activities of each participating tank, troop carrier, truck, or infantry team, which was then represented electronically in a modern combat simulation, Janus. I use the simulation to explore a series of seven controlled variations in technology and force employment, "refighting" the historical battle for each change in conditions and observing the resulting differences in combat outcomes. The experimental findings support the new theory but contradict both orthodox material-based theories of capability in general, and orthodox explanations of the Gulf War outcome in particular.

Chapter 10 concludes the book. It provides a more detailed summary of my main arguments and findings; most of the chapter, however, develops their implications for scholarship and policy and contrasts these with the views now typically held on the basis of current understandings. I argue that these contrasts are quite sharp, and that neither scholarship nor policy can be conducted on a sound basis without a more systematic consideration of force employment and its role in military power.