The Games of July: Explaining the Great War
Frank C. Zagare
The University of Michigan Press, 2011
http://press.umich.edu/titleDetailDesc.do?id=2098928

The science of politics . . . explains the mystery of how social choice evolves out of individual values.

WILLIAM H. RIKER AND PETER C. ORDESHOOK

CHAPTER ONE

Purpose and Method

Until Germany invaded Poland in 1939, it was known simply as the Great War. The label was an understatement. The carnage, both human and political, was unprecedented. According to one estimate, more than 9,000,000 lives were lost: 1,800,000 Germans and 1,700,000 Russians died; France lost 1,350,000 of its citizens, Austria-Hungary more than 1,200,000; Great Britain lost almost 1,000,000 soldiers, and Italy more than 650,000. Lesser but still substantial losses were borne by the United States, Serbia, Belgium, Turkey, and Rumania. Japan, Bulgaria, Portugal, and Greece also sustained casualties (Singer and Small 1982: 89).

As horrific as the loss of human capital was, it tells only part of the story. In the aftermath of the war, four great empires—the Russian, the Ottoman, the German, and the Austro-Hungarian—were obliterated. Another, the British, was placed on the critical list. More significantly, the war set loose a variety of forces that would, in short order, lead to the establishment of totalitarian regimes in Germany, Italy, Russia, Spain, and several other European countries. Within a generation, history would repeat itself, and all the major powers of the world would be at war once again. At this point, sadly, the Great War assumed another name, World War I.

Although some have, it is difficult to read the diplomatic record of World War I and conclude that it was inevitable. There were so many points at which it could easily have gone the other way (Mombauer 2001: 288). So how could such an unmitigated disaster have occurred? How could a war that "no one wanted" break out? These are, of course, loaded questions that immediately raise a number of vexing philosophical, ontological, and epistemological issues. Nonetheless, one purpose of this book is to address them by explaining:

- why Germany entered into a defensive alliance with Austria in 1879, an "unlikely" alliance that seemed to offer Germany few tangible benefits;
- why Germany issued a blank check to Austria on July 5, 1914, and in so doing abrogated control of its foreign policy to another, lesser, power;
- why a local war between Austria and Serbia escalated into a continental war that included Germany, Russia, and France; and
- why British foreign policy failed to deter a German declaration of war against France on August 3, 1914, a failure that brought about a worldwide conflict when Britain declared war on Germany the next day.

Taken together, the answers to these questions form the basis of an explanation of the initiation, the escalation, and the expansion of a long-simmering dispute between a major and a minor power that would eventually include all of the great states in the international system. Explanation, therefore, rather than description is a primary objective of this book. In consequence, many of the details normally found in histories of the Great War are omitted from this analysis. The reader looking for a "thick" description of the events of July and early August 1914 will have to look elsewhere.[1] Nonetheless, in addressing these questions, I hope to provide a coherent description of the chain of events that brought about World War I.

To construct an explanation for the outbreak of the Great War, I take advantage of some recent advances in the theory of noncooperative games. More specifically, I use a set of interrelated game-theoretic models called *perfect deterrence theory* (Zagare and Kilgour 2000) to answer the four central questions.[2] My explanation therefore is unabashedly theoretical. For this I offer no apologies. As King, Keohane, and Verba (1994: 46) note, "The best scientific way to organize facts is as observable implications of some theory or hypothesis."

Perfect deterrence theory is well suited for this purpose. As discussed in detail in chapter 3, perfect deterrence theory is a logically consistent and empirically plausible theory of interstate conflict initiation, limitation, and escala-

1. Albertini's (1952) three-volume account would be a good place to start. For an overview of the recent historiography relevant to the July Crisis, see Williamson and May 2007. A useful summary can be found in Meyer 2006: 1–87. Nomikos and North's (1976) study, however, is the most judicious description of the diplomatic record. Finally, Trachtenberg's (1990/91) analysis of the run-up to the war is an exemplar of rigorous historical research that should not be overlooked.

2. Basic game-theoretic concepts are introduced as they become relevant. For expository purposes, I assume that the reader is uninterested in certain technical details, including some threshold definitions and all formal proofs. Section 1.6 provides a nontechnical overview of noncooperative game theory and its solution concepts.

tion.[3] Because it is game-theoretic, it explicitly takes into account the interconnected decision-making process that is part and parcel of the web of facts that surround the Great War. Any explanation of a complex event such as the Great War that "assumes the absence of strategic interaction and anticipated reactions" is likely to be deficient (King, Keohane, and Verba 1994: 45–46).[4]

More important, though, is the fact that the equilibria of a (noncooperative) game model provide a solid foundation on which to construct explanations of real-world events and processes, such as those that preceded the outbreak of war in 1914. In section 1.6, I provide an informal discussion of the theoretical significance of game-theoretic equilibria and the precise sense in which they contribute to the development of an explanation.

For some scholars (e.g., Rapoport 1958), game theory is strictly a normative theory that prescribes behavior to (instrumentally) rational decision makers (players) who are involved in an interactive situation (a game) whose outcome partly depends on the choices made by at least one other player.[5] Game theory has long been used in the field of security studies for strictly normative purposes.[6] Indeed, I use it in this way in chapter 7 to both analyze and assess British foreign policy in 1914.

There is little doubt that game theory is a potentially powerful prescriptive tool. But Rapoport notwithstanding, there is no good reason why it cannot also be used descriptively to describe *and* explain decision making in the real world.[7] As with many deductive models, game-theoretic models start with a small set of assumptions and rigorously defined terms that permit the derivation of statements that, when they are operationalized, are subject to empirical application and testing.

There is also a long tradition in the field of security studies of using game theory descriptively to explain interstate behavior. Riker (1962), for example,

---

3. Perfect deterrence theory is a general theory of interstate conflict; its logical and empirical domain is not restricted to great-power wars or other types of armed conflict.

4. F. Fischer's (1967, 1975) explanation of the outbreak of the Great War has been criticized on precisely these grounds. Fischer's strong assumptions about Germany's preferences eliminate the necessity to take into account the policies of other states. It is highly likely, however, that Fischer's failure to take into account strategic interaction led him to what many scholars consider questionable conclusions about the motivation of Germany's leadership group. These conclusions are discussed in detail in chapter 6.

5. *Instrumental rationality* should be distinguished from the concept of *procedural rationality,* with which it is often conflated. Instrumentally rational behavior should be understood as behavior that is no more and no less than purposeful behavior (Riker and Ordeshook 1973). The requirements of procedural rationality, by contrast, are much more demanding (Simon 1976). Procedurally rational actors are, in essence, omniscient actors who make "cool and clearheaded ends-means calculations" (Verba 1961: 95) based, inter alia, on a careful evaluation of all possible courses of action. For further discussion, see Zagare 1990.

6. See, e.g., Zagare 1983.

7. See Elster 2000 for the counterargument.

used his theory of minimal winning coalitions to explain why the three known examples of grand coalitions forming in the international system fell apart so quickly.[8] More recently, Bates et al. (1998) have used game theory to construct explanations (that they call analytic narratives) of a number of real-world processes. In so doing, their aim was "not only to bring theory to bear upon data but also to bring data to bear upon theory" (2000: 698).

Their aim is also mine. In addition to providing a theoretically consistent explanation of the Great War, I use the events of July 1914 to *evaluate* perfect deterrence theory. Theoretical assessment, therefore, constitutes the second major purpose of this work. Although it is second, the goal of assessing perfect deterrence theory is in no sense secondary. Neither of this work's two purposes should be understood to trump the other. Indeed, the development of a theory-based explanation of the Great War and the assessment of perfect deterrence theory within the context of the July Crisis are, in general, synergistic enterprises, and there are many advantages of treating them as such.

1.1. Bringing Theory to Bear on Data

One very good reason to examine specific events or processes within the confines of a well-articulated theory is the additional organizing power that is acquired by doing so. From the infinite variety of observations about an event or a process that might be made, a theory singles some out for special consideration. In the case of a game-theoretic framework such as perfect deterrence theory, these categories include, but are not limited to, the identification of the players, the choices they face, the set of possible outcomes, the players' preferences over the outcome set, the private information each player possesses, information that is common to all the players, and each player's beliefs about what is likely to take place as the game is played out. At the same time, the explicit use of a theoretical framework makes analysis more tractable by suggesting what information can or should safely be ignored.

Explorations of real-world political phenomena are also rendered less ad hoc when they are theoretically informed. Theoretical frameworks severely limit not only the number but also the cast of variables that can be called on to provide a coherent explanation. While this is true of all theories, it is especially true of formal theories such as perfect deterrence theory, since formalization

8. The three instances of overwhelming majorities in international politics are the Concert of Europe, an alliance composed of England, Prussia, Russia, France, and Austria that was institutionalized at the Congress of Vienna in 1815 following Napoleon's final defeat, and the allied coalitions that came about when Germany collapsed after the two twentieth-century world wars.

requires an explicit statement of assumptions and arguments. Formal theories, in other words, are more transparent. In consequence, they are at once subject to more intense scrutiny and less amenable to even unintended manipulation (Snidal 2002: 80).

Explanations derived from deductive methodologies such as game theory have the added benefit of clarity. The relationship between the premises and the conclusions of game-theoretic models explains why something must be the case. As Kaplan (1964: 339) pointed out long ago, "The explanation shows that, on the basis of what we know, the something cannot be otherwise. Whatever provides this element of necessity serves as an explanation. The great power of the deductive model is the clear and simple way in which necessity is accounted for."

Theoretically based explanations of real-world events are also more compelling than atheoretical accounts. When a seemingly unique event can be identified as an instance of a more general category that is part of some theory's empirical domain (Rosenau and Durfee 2000: 3) or, to use Hempel's (1965: 345) terminology, when its dynamic is placed under a covering law, a deeper understanding of an apparently singular event is achieved (Riker 1990: 168). For example, in chapter 6, I explain one aspect of the July Crisis first by identifying it as a specific representation of a wider category of events (i.e., conflicts that escalate) and then by associating it with a specific prediction of perfect deterrence theory (i.e., an equilibrium of an escalation model) and the conditions (i.e., the beliefs) that give rise to it.

To summarize briefly, one important goal of this book is to gain a deeper understanding of the outbreak of World War I by examining it from the vantage point of a well-articulated theory of conflict initiation, escalation, and resolution—perfect deterrence theory. Explanations of the Great War are plentiful, but the explanation I construct here will be more coherent, less ad hoc, less opaque, and more forceful than other attempts to clarify the events of late July and early August 1914.

1.2. Bringing Data to Bear on Theory

In addition to constructing a theoretically informed explanation of the Great War, this book subjects perfect deterrence theory to more intense empirical scrutiny.[9] A number of good reasons exist for doing so.

First, empirical examination provides an opportunity to illustrate the the-

9. For large-*n* studies, see Quackenbush 2003, 2010; Senese and Quackenbush 2003.

ory in application or at work. Most theories are complexes of assumptions, concepts, and hypotheses. All things being equal, parsimonious theories are to be preferred to those that are not, but even parsimonious theories composed of a limited number of variables may be difficult to appreciate or comprehend. At minimum, therefore, empirical exposition can enhance the understanding of a theory's purpose and its underlying logic.

The application of a theory to a specific historical event can also serve as a "plausibility probe" of a theory's explanatory and predictive (or, in this case, postdictive) power (Eckstein 1992: 147–52). A theory's ability to explain a complex event and to anticipate its dynamic is a clear indicator of the theory's continuing viability. A theory that survives its first plausibility probe should have greater standing than a theory that does not; moreover, a passing grade signals that further investment in the theory—in the form of elaboration, refinement, extension, or systematic testing—is warranted.

Of course, this statement does not necessarily imply that a theory's failure to account for all conceivable eventualities is a sign that it is degenerative in the sense of Lakatos (1970). The discovery of a lack of fit between theory and evidence may in fact be another reason for examining a theoretical structure in the context of a single event—if that term can be used to describe a case as complex as the July Crisis. The failure of a theory's structure to conform to a particular feature of the real world provides an opportunity to adjust the theory to increase its scope.

However, such adjustments should never seek simply to account for an empirical anomaly, for that kind of modification would indeed be evidence of degeneration (Vasquez 1997). Rather, one wholly defensible purpose of theory modification is to broaden the class of events to which it can be fruitfully applied. For example, when originally specified, perfect deterrence theory applied only to deterrence situations in which there were two principal players. The reason is straightforward: none of its component models were *n*-person games. In consequence, the theory's implications for those extended deterrence situations in which the choices of a third actor, a pawn or a protégé, partly determine an outcome were circumscribed.

To overcome this limitation, another game form—the Tripartite Crisis Game—was developed (Zagare and Kilgour 2003) to explore the impact of alignment patterns on the success or failure of extended deterrence relationships. Since the assumptions underlying the analysis of this model are derived from, and are consistent with, those initially used to cast the theory, the Tri-

partite Crisis Game can be considered not only an extension of, but also a critical component of, perfect deterrence theory. Indeed, it has proven indispensable to this analysis of the July Crisis. In part or in whole, in one variant or in another, it is called on to help answer three of the critical questions about the Great War.

Put in a slightly different way, well-articulated theories always have clear boundary conditions. The assumptions on which a theory rests can be considered limiting conditions that define the set of circumstances to which the theory applies. When empirical scrutiny reveals these conditions to be too restrictive, their relaxation can be a proper analytic reaction. Indeed, it is the dominant scientific response (Bates et al. 1998: 16).

1.3. A Possible Objection and a Rejoinder

Not all analysts would agree that game-theoretic models and the theories derived from them should be used to develop an explanation of a real-world event, even if that event can be considered a member of a wider category.[10] Morrow (1997: 29), for example, asserts that "individual case studies are poor tests of rational choice models." He argues that "ex post reconstructions of historical events use information that the actors could not have at the time, subtly influencing us away from the strategic problems they faced[and] cannot explain the case precisely because the game cannot fail to fit the specified facts of the case."[11]

Morrow's position is extreme. It overlooks the fact that "critical case studies" can and do serve important theoretical purposes (Eckstein 1975). Still, Morrow is more correct than not in arguing that under most circumstances, even a theoretically informed single case study would fall short of a comprehensive *test* of a game-theoretic model or, for that matter, any theoretical construct. The reason is not difficult to fathom: one cannot generalize from an *n* of one. And the July Crisis is clearly a "one."

Generalization, however, is not the motivation for this examination of the July Crisis. Explanation and theoretical illustration take pride of place, so Morrow's observation is of little moment here. But even if it were pertinent, it does not necessarily follow that ex post reconstructions of real-world interactions

10. King, Keohane, and Verba (1994: 12) make the case for developing causal explanations of "even highly unusual events that do not fall into a large class of events."
11. See also Elster 2000.

that rely solely on information that was available to the actors at the time they made their decisions cannot be developed, as Bueno de Mesquita's (1998) analysis of the end of the Cold War so well demonstrates.

Morrow is also correct in asserting that *some* game model can always be molded to fit the facts of almost any individual event. But here too his objection misses a critical point. Closeness of fit is not the only criterion by which explanatory models are judged. Like any historical description, a game-theoretic explanation of an individual case can also be evaluated in terms of the plausibility of its assumptions about the motives of the players and other critical components of the model (i.e., the choices available to the players, the set of outcomes, and so on). For example, in an analysis of the Geneva Conference of 1954, I used a game-theoretic model of strategic deception to reconcile the well-known but unexplained discrepancy between the public and the private pronouncements of U.S. decision makers both before and during the negotiations that brought the Franco-Vietminh war to a close (Zagare 1979). Using the same model, Thakur (1982) developed six seemingly plausible alternative explanations of the diplomatic process at Geneva. But these competing descriptions of the conference were either empirically dubious or logically deficient and, hence, could be dismissed (Zagare 1982).

1.4. Blueprint

This discussion brings up yet another criterion by which theories and explanations based on them can be judged: their competitive advantage or lack thereof. This criterion, which is generally associated with the philosopher of science Imre Lakatos (1970), is especially pertinent to any explanation of the July Crisis, simply because they are so plentiful. With so many to choose from, it is easy to go wrong. It is also easy to blend many of these explanations together to include every variable that can plausibly be associated with the onset of the Great War. Simplification is clearly in order. Inferior explanations must be cast aside. Stronger explanations, however, must be closely scrutinized and evaluated. Are they internally consistent? Are they empirically supported? Are they falsifiable? And are they superior to their competitors?

Chapter 2 reviews and evaluates extant explanations of major power wars in general, and of the Great War in particular. After quickly eliminating a number of candidate explanations, either because they are monocausal or because they are nonfalsifiable, attention is turned to the leading contender: realist, or balance-of-power, theory and its corollary, classical deterrence theory.

The underlying assumptions, the theoretical characteristics, the major propositions, and the policy implications of the realist research program are examined and found wanting. Chapter 3 lays out the general contours and empirical expectations of perfect deterrence theory, which should be viewed as a theoretical alternative to the standard explanatory paradigm.

Chapters 4–7 bring data to bear on theory. Component models of perfect deterrence theory are used to explain several critical determinants of the chain of events that led to the Great War. In chapter 4, the decision of Germany's first chancellor, Otto von Bismarck, to enter into a defensive alliance with Austria in 1879 is examined in the context of the Tripartite Crisis Game. Understanding Bismarck's motivation is essential for an appreciation of German foreign policy up to and including the July Crisis; it is also essential for comprehending the structural characteristics of the European state system just prior to the war.

Chapter 5 explores both the Austro-Hungarian decision to seek Germany's support, and the German decision unconditionally to provide that support, in the wake of the assassination at Sarajevo of Archduke Franz Ferdinand, the heir apparent to the Austro-Hungarian throne. The so-called blank check is found to have a natural explanation in the form of an equilibrium prediction of the Protégé-Defender Subgame of the Tripartite Crisis Game.

Another model, the Asymmetric Escalation Game, is used in chapter 6 to explain the initial escalation of the July Crisis from a local war in the Balkans between Austria-Hungary and Serbia to a continental war that included all the great European powers with the singular exception of Great Britain. Because the fit between theory and fact is unusually close, this chapter exhibits many of the characteristics of what George and Bennett (2005) call process tracing. Since much of the action of the crisis took place before Great Britain entered the war, this chapter contains deeper description than any other. The description demonstrates a close association between the choices available to the players in 1914, their beliefs about each other's intentions, and the intensification of the conflict.

Chapter 7 analyzes British foreign policy in the run-up to the Great War. Why did Britain's foreign secretary fail to make clear a threat to support France in the event of a German attack, and precisely why did his policy fail? Again, the answer to this question is developed in the context of an equilibrium prediction of the Tripartite Crisis Game. The existence conditions associated with this prediction provide an answer to the first question; the theoretical characteristics of the equilibrium offer an answer to the second.

Chapter 8 brings this analysis to a conclusion. A explanation for the outbreak of the Great War is summarized. The theoretical consequences of bringing data to bear on perfect deterrence theory are discussed. And the policy implications of this study, many of which are relevant to the world in which we currently live, are highlighted.

1.5. Scylla and Charybdis

Under normal circumstances, the two tasks I have undertaken here—historical explanation and theory illustration and application—are fraught with difficulties. The dangers are magnified when the goals are pursued simultaneously, for these twin objectives are sometimes at odds with one another, working at cross-purposes. When placed between a rock and a hard place, compromises are often required. These compromises, however, are likely to satisfy no one.

Historical analysis, for example, is generally descriptively rich and long on details. But explanation demands simplification. Details that stand in the way of explanation frequently must be overlooked or ignored. In the analysis that follows, an abstract, formal, structure does most of the work of explanation. In consequence, deep description is frequently cast aside. Many historians will understandably be unhappy. But as the great American artist Georgia O'Keeffe once said in a completely different context, "Nothing is less real than realism. Details are confusing. It is only by selection, by elimination, by emphasis that we get at the real meaning of things."

It is also inevitable that some historians will be less than pleased with the connections I make between the events leading up to the Great War and the game models I use to analyze them. Among these connections, the assumptions I make about the preferences of the players have been the most problematic. And historians (and some political scientists) have offered little help here. For example, what goal was Germany pursuing in the run-up to the war? There is great controversy and a stunning lack of consensus among diplomatic historians about the answer to this question. In the end, judgments had to be made. And while I have made every attempt to justify these coding decisions, I appreciate that reasonable objections persist. Alas, there is no sure way to avoid the ambiguities in the historical record. I accept this reality as a certain limitation of my analysis. There may indeed be more than one plausible interpretation of the "facts."

Like historians, formal theoreticians are likely to find objectionable some

of what follows. The empirical application of a game-theoretic model to diplomatic history in general, and to the July Crisis in particular, is anything but straightforward. The primary problem involves bringing the theoretical concepts of an abstract model into an isomorphic relationship with a reality characterized by, among other things, nuance, ambiguity, equivocation, and duplicity. The real world is rarely as contained or as neat as any formal representation of it. Again, compromises must be made.

Take, for instance, the issue of player aggregation. Purists (and I generally count myself among them) are likely to object not only to the fact that I treat states as players in a game, but also to the fact that in chapter 6 I consider two state-based coalitions as single players. After states have been admitted as players, there is a great temptation to reify them. I have tried to avoid that temptation by resting my conclusions about the policy objectives of Germany, Austria-Hungary, France, Russia, and Great Britain on appropriate primary and, when necessary, secondary sources rather than from some "objective" analysis of the national interest. In other words, rather than positing state goals, I have tried to uncover them from an examination of the contemporaneous diplomatic record.

Still, Elster (2000: 693) argues that "rational choice explanations divorced from methodological individualism have a dubious value." In principle, I agree. But strict adherence to principle all but eliminates the possibility of analysis, which is odd because Elster goes on to assert—wrongly, in my opinion—that "rational choice theory is the only theory in the social sciences capable of yielding sharp deductions and predictions." The disaggregation of states, even to the level of a half dozen or so leading decision makers, destroys any semblance of tractability and with it any possibility of a satisfying explanation of all but the most trivial of circumstances. Readers must judge for themselves whether the violation of principle comes at too great a price.

Some theorists, formal or otherwise, may also find fault with my decision to use generalized models to analyze certain events rather than to fine-tune these models more precisely to adhere to the specific circumstances I examine. I have felt uncomfortable with this approach, especially in chapter 6, where I specify the causal mechanism underlying an actual escalation spiral. I have nevertheless proceeded in order to maximize generality, but only when I have convinced myself that a more nuanced model would only yield the same conclusion.

King, Keohane, and Verba (1998: 43) assert that "where possible, social science research should be both general and specific: it should tell us something

about classes of events as well as about specific events at particular places. We want to be timeless and timebound at the same time." It would be difficult to state the objective of this book more succinctly.

1.6. Coda

Although game theory has been a fixture of the security studies literature since the early 1950s, it has not been used until recently to study a complex of real-world events such as the one that unfolded during the middle of the summer of 1914. Accordingly, in this section, I provide an overview of those aspects of the mathematical theory of games that are relevant to an understanding of the explanation I develop, a brief (nontechnical) discussion of the intuition that lies beneath the explanation, and a sense of some of the limitations of using game-theoretic models to organize a case study. I begin with a few definitions.[12]

The basic concept is that of a game itself. A *game* can be thought of as any situation in which the outcome depends on the choices of two or more actors—that is, when the choices of these actors (*players*) are interdependent. Players may be individuals or groups of individuals who act as coherent units. The players are assumed to be *rational.* Simply put, this means that they are purposeful, that is, that they have objectives and that they act to bring them about. It does not mean, however, that the players are necessarily intelligent in the sense that their objectives are wise, realistic, or even admirable. It also does not mean that a player will succeed. In game theory as in life, players are often misguided, shortsighted, imprudent, and unsuccessful. The lions of July were no exception.

Games can be categorized in many ways. For my purposes here, the fundamental distinction concerns the rules that govern play. When there is no legal or other kind of constraint on the players' ability to coordinate their actions, they are said to be playing a *cooperative game.* By contrast, a *noncooperative game* occurs whenever coordinated decision making is not possible. In the real world, coordinated action might be precluded for any number of reasons. For example, players might not be able to communicate with one another. Collusion might also be prohibited by statute, as would be the case in an economic market that was regulated by an antitrust law. Finally, a commitment problem

12. There are many excellent book-length expositions of the fundamentals of game theory. For the political scientist, there is none better than Morrow 1994. Condensed treatments can be found in Gibbons 1997; Zagare 2008.

will likely exist whenever "there be no Power erected, or not great enough for our security [such that] every man will and may lawfully rely on his own strength and art for caution against all other men" (Hobbes 1968 [1651]: 223–24).

For what may be obvious reasons, noncooperative games are the rule and cooperative games the exception in international politics. Since the international system is a self-help system that lacks an overarching authority with the ability to force powerful and sovereign states to honor their commitments, any agreement that a state makes to cooperate with another will always be contingent on the state's willingness to honor that agreement. Thus it should come as no surprise to learn that noncooperative game theory holds a particular attraction for theorists of international conflict, as it does here.

At the heart of the theory of noncooperative games is the concept of an *equilibrium* outcome (or strategy set). In game theory, an outcome is considered to be in equilibrium whenever no single player can induce a better outcome by selecting a different course of action, given that the choices of all the other players in the game do not change. At an equilibrium outcome, therefore, the strategy choice of each of the players is a best reply to the choices of all the other players in a game.

An equilibrium outcome, however it is defined, should be thought of a self-enforcing agreement among the players about how to play a game. The agreement is self-enforcing simply because it is in the interest of each player to carry it out. Whenever a noncooperative game is the only game in town, a self-enforcing agreement will clearly be the only meaningful type of agreement available to the players.

Put in a slightly differently way, only equilibrium choices are consistent with (individually) rational behavior by all of the players in a game. When an outcome is not in equilibrium, at least one player can do better by switching to a different course of action. Thus, the assumption that the players in a game are rational leads naturally to the expectation that they will make choices that are associated with some equilibrium outcome. Game-theoretic explanations and predictions derive from this expectation. When players in a real-world game make choices that can plausibly be associated with an equilibrium outcome, an explanation is at hand. As Riker (1990: 175) has observed, "Equilibria are . . . identified consequences of decisions that are necessary and sufficient to bring them about. An explanation is . . . the assurance that an outcome must be the way it is because of antecedent conditions. This is precisely what an equilibrium provides."

Similarly, a game-theoretic prediction about future play presumes rational choice—that is, the assumption is that an equilibrium choice will be made by each of the players. Unfortunately, since a game will often have more than one equilibrium, this assumption does not always lead to a point prediction or to a fully satisfying explanation. The technical reason is that when multiple equilibria exist, the choices that lead to them may not always be *interchangeable:* even when all the players make choices associated with an equilibrium outcome, the resulting outcome may not be in equilibrium. This vexing problem is compounded by the fact that the equilibria of a noncooperative game may not always imply the same payoff to one or more of the players—that is, the equilibria may not be *equivalent.* Thus, in a game with multiple equilibria, it may be impossible to specify before the game is actually played out which course of action the players will take, or to explain after the game has concluded, why one rational strategic possibility rather than another came about.

It is more than an understatement to suggest that the existence of multiple nonequivalent and noninterchangeable equilibria confounds both explanation and prediction in a game-theoretic model and, as Niou and Ordeshook (1999: 89) point out, in any theory of rational agency. For the most part, however, the games analyzed in this book almost always have equilibria that exist under unique (parameter) conditions. Thus, the explanations developed in this volume are, on their face, more compelling than explanations derived from models with multiple coexisting equilibria.[13]

Still, some problems remain. One difficulty is demonstrating that the conditions associated with the existence of a particular equilibrium, unique or otherwise, are satisfied empirically. This is not always an easy and straightforward task. Whenever this problem surfaces, a judgment call is necessary, and of course great care must be taken not to mold all-too-convenient facts. But this challenging problem is not confined to game-theoretic studies. Indeed, it is endemic to the social science enterprise itself. Historians, for example, must deal with it all the time.

Another theoretical predicament manifests itself whenever observed action choices are consistent with more than one equilibrium; when this happens, as it does in chapter 5, it is impossible to say for sure which equilibrium form actually came into play, although it may still be possible to speculate rig-

13. The problem is especially acute in a repeated game model. One of the so-called folk theorems of game theory states that in an infinitely repeated game, *any* combination of action choices that is individually rational can be part of some equilibrium. Thus, just about any behavioral pattern can be justified in a repeated game. There are no such models in this book.

orously. That an unambiguous conclusion may not always be possible is a fact of life I have come to accept. Even so, this limitation of game theory is tempered by the realization that even when there is more than one rational strategic possibility, some possibilities may in fact be ruled out—logically, empirically, or both. In the end, additional insights into real-world processes are obtained.

The technical requirements (i.e., the definition) of an equilibrium outcome in a noncooperative game turn largely on whether a game's strategic structure is specified in strategic- or in extensive-form. In a *strategic-form* representation, the players are assumed to choose, simultaneously, a complete course of action (i.e., a *strategy*) before a game is played out.[14] In the static environment of a strategic-form game, therefore, each player makes but a single choice. By contrast, in an *extensive-form representation,* the players are assumed to make *choices,* sequentially, at *moves* (or decision points) that are represented by nodes on a game tree.[15] Thus, extensive-form games better capture a game's ongoing dynamic.

In a static (strategic-form) game, the standard equilibrium concept is due to Nash (1951). A strategy pair is a *Nash equilibrium* if no player could achieve a better outcome by switching, unilaterally, to another strategy. In a dynamic (extensive-form) game, where the players' choices are sometimes contingent, the central equilibrium concept is *subgame-perfect* (Selten 1975). Nash equilibria exist in the dynamic context, but they may be based on incredible threats (i.e., on threats of irrational choice), whereas subgame-perfect equilibria require the players to plan to choose rationally at every node of the game tree. Nash and subgame-perfect equilibria are the accepted measures of rational behavior in games of *complete information,* in which each player is fully informed about the preferences of the other players. In games of *incomplete information,* in which, for example, at least one player is uncertain about the other's preferences, rational choices are associated with *Bayesian Nash equilibria* (in static games) and with *perfect Bayesian equilibria* (in dynamic games).[16]

The fine distinctions among the various equilibrium definitions are less important, at least for the purposes of this book, than the intuition that lies beneath all of them. An equilibrium outcome—whether Nash, subgame-per-

14. Strategic-form games are sometimes referred to as *normal-* or *matrix-form* games. See figure 2.1 for an example.
15. For this reason, the extensive-form is also termed the *game-tree* form of representation. See figure 2.2 for an example.
16. The list is not exhaustive. For a precise definition and a detailed discussion of the standard equilibrium forms mentioned in this paragraph, see Gibbons 1992.

fect, Bayesian Nash, or perfect Bayesian—should be thought of as one of perhaps several rational strategic possibilities. In the context of any noncooperative game model, an explanation requires at minimum a plausible association of action choices with both an equilibrium and the conditions required to bring it about. All things being equal, the strength of the explanation is amplified as the plausibility of the association increases and as the number of rational strategic possibilities decreases. For this reason, I have gone to great lengths in this volume, but especially in chapters 4–7, to make as compelling a case as possible for the connection between the beliefs of the players, their action choices, and the existence conditions associated with (as it turns out) some perfect Bayesian equilibrium. Since my goals in this work mandate that these connections be established, I hope that an intuitive sense of the central role played by an equilibrium outcome in a noncooperative game model, along with the brief descriptions of some fundamental concepts found in the footnotes, will enable those readers either uninterested in technical details or with a low tolerance for formal definitions to follow the argument. I also hope that game theorists and other formal modelers will understand the absence of proofs and symbolic argumentation. Of course, the mathematical particulars do exist, but since there is no apparent need to reproduce them here and I lack the space to do so, I simply note their source location whenever relevant.