

The Limits of Strategic Choice

Constrained Rationality and Incomplete Explanation

Arthur A. Stein

In *Strategic Choice and International Relations*, pp. 197–228. Edited by David A. Lake and Robert Powell. Princeton, New Jersey: Princeton University Press, 1999.

This file may contain errors from the optical character recognition performed on the scanned copy of the original.

CHAPTER SEVEN

The Limits of Strategic Choice: Constrained Rationality and Incomplete Explanation

ARTHUR A. STEIN

GAME THEORY, especially as applied in microeconomics, has been ballyhooed as the savior of the study of international politics and attacked as its destroyer.¹ It will, proponents argue, provide the firm analytic foundations on which to build a rigorous science. This volume itself reflects such hubris. But like other innovations, this one has been shunned as well as welcomed. If its supporters have the zeal of religious converts, its detractors reflect no small degree of mindless defensiveness. This essay highlights the strengths of a strategic-choice approach, which lie in its flexibility and rigor, but also elucidates its weaknesses, which lie in its excessive simplification, causal incompleteness, and post hocery.

My task in this essay is both that of defender and critic of the faith. I emphasize the strengths of the approach but also detail its core weaknesses. I argue that some typical criticisms are actually of particular modeling choices rather than of the approach itself. But I also delineate a set of more fundamental weaknesses. As would be expected of a believing agnostic asked to play the role of church ombudsperson, I conclude that though limited and incomplete, the approach is both useful and unavoidable.

A strategic-choice approach is particularly suited to the study of international relations.² Long before the development of strategic interaction

This paper was written with the financial assistance of the University of California's Institute on Global Conflict and Cooperation and UCLA's Academic Senate. The author thanks Alan Kessler for research assistance. My thanks to the volume editors and contributors, and to Jim Caporaso, Amy Davis, Jim Fearon, Steph Haggard, Alan Kessler, Lisa Martin, Paul Papayoanou, Beth Simmons, Cherie Steele, as well as anonymous referees of Princeton University Press.

¹ This is also happening in other subfields of political science and, indeed, other disciplines.

² The area of economics most devoted to strategic choices has been regularly monitored by political scientists who mine it for intellectual nuggets. Economists, who typically assume competitive markets, often ignore strategic interaction, assuming that individuals and firms simply take markets as given and do not make strategic choices. But economists do see firms

models, the field of international relations (like classical diplomatic history) focused on the interests of states and how interactions affect choice and strategy. Similarly, a strategic-choice approach begins with purposive, intentionalist, rational explanations of behavior and adds the component of actor interaction. The actors' choices reflect not only their preferences and the constraints they confront but also the existence of other actors making choices. Not surprisingly, therefore, game theory's formal tools for analyzing strategic choice quickly found application in the study of international politics.³ That it met so ready an audience and was synonymous with key assumptions in the field also meant, however, that its impact was not revolutionary.⁴

This chapter highlights the benefits and strengths of a strategic-choice approach to the analysis of international politics. First, the approach provides the benefits of mathematical modeling: rigor, deduction, and inter-

as making strategic choices when there are but few of them, and political scientists have looked to economists' work on duopoly and oligopoly for analogues to international politics. International-relations scholars were affected by the work of economists in the late 1940s and early 1950s on monopolistic competition. This is quite evident in Waltz (1959), who cites the work of Fellner and Chamberlain. It is also evident in Kaplan's (1957) characterization of different international systems, which comport with economists' characterizations of different kinds of markets (Boulding 1958).

The connection between economists and international relations has resulted in many economists being taken to the bosom of international-relations scholars (Viner, Hirschman) and to economists addressing international relations issues directly (Schelling, Boulding).

³ Economists played a central role in World War II in the military application of operations research and strategic gaming (Shubik 1992; Kaplan 1983; Bernstein 1995). Game theory in the immediate postwar period was sustained by military spending and immediately applied to strategic security issues long before it was embraced by economists (Leonard 1992; Mirowski 1991).

The relevance of game theory to international politics was noted in two of the most important books in international-relations theory written in the 1950s. Waltz (1959) and Kaplan (1957) both refer to game theory. Indeed, Kaplan's book includes an appendix on the subject. Economist Thomas Schelling published, in 1960, his classic *The Strategy of Conflict*, applying game theory to issues of international conflict. A special issue of *World Politics* in October 1961, devoted to theoretical work on the international system, includes two essays by economists, both on game theory (Schelling 1961; Quandt 1961). Two other essays in that volume also discuss game theory (Kaplan 1961; Burns 1961), making the number of contributors to the 1961 special issue discussing game theory almost half the total. By 1985 another special issue of *World Politics* devoted to international cooperation dealt almost entirely with game theoretically based work.

For a discussion of the isomorphism of game theory and international relations, see Stein 1990. For a review essay, see O'Neill 1994. One can argue that strategic choice is not an approach to, but a definition of, international politics. The focus here is on the former.

⁴ In contrast, a strategic-choice approach came much later to, and was far more revolutionary in its impact on, other subfields of political science, such as electoral politics or public law, in which sociological perspectives dominated (and still do). For the spread of rational choice and game theory into political science, see Rogowski 1978; Riker 1992; and Ordeshook 1986; also see Miller 1997.

nal consistency. In addition, its flexibility allows it to be applied very broadly. Indeed, its flexibility is such that many of the criticisms typically leveled at it are actually about specific modeling choices rather than fundamental criticisms of the approach per se. I argue, that is, that many criticisms imply alternative modeling choices rather than an alternative approach, so that criticisms characterized as fundamental are more typically squabbles within a church. Scholars must make strategic choices in applying strategic-choice models to the study of international politics. Scholars must be self-conscious about critical modeling assumptions and intermediate steps rather than simply parrot the substantively driven assumptions made in other fields. The approach is generic, but particular scholarly choices make it meaningful. There is, in other words, an art to applying the science of choice?⁵

Yet, the chapter is skeptical as well as affirming, for the approach has serious limitations, both generically and specifically as applied to international relations. I argue that the core assumptions of the model are ideal ones, originally made with a normative intent and that there are fundamental problems in the use of such models for explanation. Indeed, the use of such models for positive explanation either is viable only because the models are self-fulfilling or fails because the models make problematic simplifying assumptions about human capabilities. When the models are expanded to achieve greater verisimilitude, they become indeterminate and incomplete—consistent with a variety of outcomes and a variety of plausible models and paths to any particular outcome.

THE ART OF THE SCIENCE OF CHOICE: MODELING AND POST HOC EXPLANATION

Given the flexibility of the approach, constructing models of strategic choice is an art rather than a science, requiring scholars to make an array of choices about the nature of the actors, their preferences, their choices, their beliefs, and so on.⁶ These scholarly choices can be debated.

Purposive Explanation and Actors in International Relations

A strength of the strategic-choice approach is its applicability to any actor. Economists apply it to individuals and to firms. In the case of interna-

⁵ Aumann (1985) sees both game theory and mathematical economics as art forms. Much the same can be said of mathematics in general.

⁶ This point is more general in that science is a logic of justification rather than a logic of discovery.

tional relations, it can be applied to people, bureaucracies, nations, and others. Thus it effectively finesses the unit-of-analysis debate in international relations. Moreover, the strategic-choice approach explains the existence of collectivities in terms of the interests that brought them into being and thus provides microfoundations for aggregation and delineates the requisite assumptions for the application of the approach to collective actors.

Explanations of behavior in the social sciences typically begin with reference to intentions (Whether of individuals or larger social aggregates like firms, interest groups, or governments).⁷ Scholars treat actions as purposive and so assess actors' interests, options, and calculations in explaining them. More specifically, purposive intentionalist explanation is one standard approach to the study of international relations.⁸ Explanation of a government's action often starts by addressing that government's interests, and the "national interest" lies at the heart of classical models of explanation in the field.⁹

Concomitant with intentions is the question of whose interests: who are the actors whose intentions serve to explain their actions. In the debate among scholars of international politics over the appropriate unit of analysis, some scholars reduce international politics to human behavior and explain world politics as the result of individual choices. For them, the relevant unit of analysis is the individual, and state policy is ultimately reducible to the actions of individuals. Explaining foreign policy and international politics means explaining leaders' choices. Critical questions are reformulated to frame the question in such terms. Explaining the origins of World War I devolves to explaining the confluence of the personal choices that resulted in war: the ambivalent signaling of Lloyd George, Ferdinand's decision to attack Serbia, Bethmann-Holweg's decision to issue a blank check to Austria-Hungary, and so on.

Others find the relevant behavioral units to be aggregations of people. Economists treat firms as actors; international relations scholars treat states the same way. They see foreign policy as collective action and deem

⁷ For a discussion of rational-choice theory as a subset of intentionalist explanation, see Elster 1986.

⁸ It is this very centrality of purposive explanation that places the discipline of economics at the heart of the social sciences. For a discussion of the evolving meanings of economics and political economy, see Groenewegen 1991. The centrality of purposive explanation in the social sciences makes possible both the importation of economic ideas now so prevalent in political science and elsewhere (this volume being an excellent reflection) and the imperialism characteristic of recent economic theorizing. Economists have applied their models to all sorts of decision problems (Becker 1976; Hirshleifer 1985; Radnitzky and Bernholz 1987; Radnitzky 1992; Baron and Hannan 1994).

⁹ Graham Allison (1971) began his critique of standard international-relations theorizing by elucidating this approach and dubbing it "model I." The link between balance of power,

the collectivities to be actors. Often dubbed "the unitary-actor approach," this perspective treats international politics as the result of state policy. Explaining the origins of World War I is then reduced to a determination of the confluence of state choices, like the British failure clearly to deter and the German blank check to Austria, that resulted in war.

Not only is strategic-choice analysis silent on the unit-of-analysis issue, applications of the approach effectively expand it by disaggregating the person. Scholars working in this research tradition have given us "the rational gene" and conceptualized weakness of the will as an "intrapersonal prisoner's dilemma." Such social and natural scientists reduce the explanation of behavior to a unit of analysis more fundamental than the human being. Just as psychoanalysts are prepared to disaggregate the human being into the constituent components of ego, id, and superego, these theorists of strategic choice are prepared to see those components as engaged in a strategic game against one another.

The very flexibility of the approach means that critics often focus on modeling choices rather than challenge purposive explanation in general. Critics of the state-as-actor perspective, for example, equate strategic-choice analysis with a focus on unitary states. Since proponents of bureaucratic politics explanations do not view states as integrated entities that can be modeled as having preferences and making choices, they offer an alternative unit of analysis, the bureaucracy or organization, and then focus on how its preferences, interests, choices, and interactions determine a state's foreign policy. But bureaucratic politics explanations, although billed as a process alternative to purpose, simply shift purposive explanation to a different (albeit still aggregated) level of analysis.¹⁰ Organizational interests replace national interests, and interacting bureaucracies replace interacting states. Hence bureaucratic politics arguments can be modeled using strategic interaction as readily as the perspective they criticize (Bender and Hammond 1992).

A strategic-interaction approach also breaks down the level-of-analysis problem in international relations. Scholars have argued that scholarship in the field could focus on individuals, states, or the international system, and, more important, that the levels of analysis could not interact or be combined.¹¹ Analysts would have to choose, they argued, between undertaking a true systemic study or one that was reductionistic (focusing on individuals or states) because they could not combine domestic politics

typically thought of as the key theory in the field, and a purposive explanation is provided by Morgenthau's famous statement, "interests defined as power."

¹⁰ See Allison 1971 and the subtle changes in Allison and Halperin 1972. For critiques and discussions, see Art 1973; Ball 1974; and Perlmutter 1974. For an excellent paper applying rational actor models to Allison's formulation, see Bender and Hammond 1992.

¹¹ Some scholars add bureaucratic politics as an additional level of analysis (Jervis 1976).

and the international system. In contrast, a strategic-choice approach can be seen as a vehicle for integrating the levels of analysis, demonstrating that some levels act as constraints on choice at other levels and that micro-foundations for macro-outcomes are essential.¹² Moreover, the framework provides the ability to move across levels of analysis, to move from individuals' preferences to states' preferences and to international outcomes.¹³

Indeed, strategic-choice models can be so broadly applied that they can be, and have been, applied to an array of actors, including people, firms, and states. Moreover, the interactions analyzed can involve combinations. The models can be applied to the interactions of actors at different levels of aggregation: individuals with states, and states with international organizations, multinational organizations, and nonstate actors (the Palestinian Liberation Organization, for example).¹⁴ Using a strategic-choice approach thus allows the assessment of interactions across levels of analysis.

Although a strategic-choice approach is silent on the unit-of-analysis question, it does stipulate the necessary assumptions for its application to larger aggregates. The actors in a strategic-choice framework are presumed to have consistent utility functions and to be capable of choice, perception, and calculation. These specifications provide more precise meaning to the notion of unitariness. For states to be seen as actors in a strategic-choice sense, they must be unitary in more than that a decision by such actors implicates everyone in the collectivity. A nation that declares war on another implicates all its members in that decision. But that meaning of unitariness is insufficient for a strategic-choice approach, which requires that aggregate actors be unitary in that they have a definable utility function.¹⁵

¹² I make this argument in a more extended fashion in Stein 1990, esp. 175–84.

¹³ The thrust of this discussion is that levels of analysis can be, and typically must be, combined in an explanation of some outcome—that independent variables at different levels of analysis can be combined. But there is another sense in which a level-of-analysis problem remains: At what unit and level of analysis should one choose to explain. Put another way, at what level of analysis should one couch the dependent variable. To take a concrete example, one can choose to explain why President Kennedy responded to Soviet missiles in Cuba and selected a blockade or why the United States responded to Soviet missiles and adopted a blockade or why a bipolar system periodically generates superpower crises that are resolved short of war. Each question can be answered by combining levels of analysis, but the questions differ.

¹⁴ For a different conception of problems in game theory with the notion of a player, see Güth 1991.

¹⁵ The longstanding application of purposive explanation to states as actors in international politics means that the whole held, not just strategic interaction analyses, must confront the question of the viability of treating such aggregates as actors. Few if any scholars are troubled by discussing national interests, a few more are taken aback by discussing the utility functions of states, but many more balk when a psychohistorian asserts that “nations

Not only can the strategic-choice approach deal with actors at different levels of aggregation, such models explain the existence of aggregate entities by providing explanations both for why actors create institutions and the consequences of institutions for choice. Self-interested actors, whether individuals, firms, or states, create institutions to deal with collective-action problems and the suboptimality associated with autonomous choice.

As Rogowski demonstrates in chapter 4 of this volume, institutions are mechanisms for collective choice, and their design significantly determines the nature of that choice. The collective choices of institutions can be affected by their constituent members, and so issues of franchise and membership are central to their design. The specific mechanisms by which the preferences of constituent members are aggregated also significantly determine outcomes.¹⁶ And, as Gourevitch demonstrates in chapter 5 of this volume, that institutions and their design matter means that they are significant and important foci for political struggles. Thus, although the approach assumes the nature and utility functions of actors as given and does not address how they change (from empires and city-states to states, for example), it can be used to explain the strategic-choice bases of changes in institutional form.¹⁷

In short, a strategic-choice framework provides microfoundations for macrobehavior. It links levels of analysis by treating aggregate entities as the products of individual choice (Gourevitch) and collective choice as the consequence of different mechanisms for the aggregation of individual preferences (Rogowski).

The interest of purposive self-interested autonomous actors in mechanisms for collective action breaks down the realist/institutionalist divide in international relations. Many scholars have contrasted realism and institutionalism as two alternative approaches to the field; the former emphasize autonomous choice under anarchy, and the latter stress the role of international institutions. The strategic-choice approach developed in this volume transcends this intellectual split by explaining institutional creation, design, contestation, and collapse as the products of autono-

have psychologies, just as individuals do; they have dreams and fantasies that can be analyzed; they have urges that arise from childhood fears and traumas of their populace" (*New Yorker* 1995, 55–56). Yet analytically (in the scientific, rather than psycho, sense), extending interests and purposive calculation to states is perhaps as questionable as extending them to psyches. For a discussion of the problem of social choice and aggregate entities, see Sen 1995.

¹⁶ The centrality of rules and procedures to outcomes is the key conclusion from Arrow's (1951) Nobel-prize winning work and is a core motif of the literature on social choice.

¹⁷ More broadly, as the title of an edited volume suggests, organizations can be seen as games (Binmore and Dasgupta 1986).

mous choice. Actors create, maintain, argue about, and attack institutions. Their self-interest in different settings explains all these responses.

Taking Preferences Seriously

Strategic interaction analysis is also flexible because it is built on the foundation of subjective utility and so can encompass a variety of preferences and bases of calculation. Whether the units of analysis are individuals, firms, or states, the analysis begins with some articulation of their preferences.¹⁸ Ironically, this theoretical flexibility is not always evident in scholarly practice.

In chapter 2 of this volume, Jeffrey Frieden delineates the ways that scholars ascertain the preferences that animate their analyses. Frieden stresses an analytic preference for positing actor preferences or rooting them in some deductive theory. Further, he points out, it is better to explain change (whether across actors or for one actor over time) by reference to changing circumstances rather than changing preferences. This avoids the circularity associated with explaining behavioral change by preference change when our knowledge of preference change comes from observing behavioral change. Despite the predisposition of strategic-choice theorists for fixed preferences, Frieden shows that we can explain changes in behavior by changes in preference if we have some independent means by which to ascertain the change in preferences.

However arrived at, whether held constant or allowed to change, preferences are central to the analytic enterprise. Yet, scholars of international relations have disagreed about how to characterize actor preferences. Given the focus on subjective utility in a theory of choice, it is ironic that scholars typically posit the interests and preferences of actors rather than investigate them empirically. There is a long-standing tradition in international relations of imputing the national interest. Classical realists argue that all states have a core national interest of assuring their physical and territorial integrity and so act to maximize their power.¹⁹ This assumption was to play the role in studies of international relations that the assumption of wealth maximization played in economics. But power maximization came to be seen as either an infalsifiable assertion or a falsifiable and false one (Rosecrance 1961).

¹⁸ As already mentioned, some criticisms of strategic choice are actually criticisms of intentionalist explanation. After all, causal explanation is provided in the natural sciences without reference to intention and purpose. It is possible to conceive of a social science that similarly ignores the preferences of individuals and groups of individuals. But this line of criticism has not developed in the field.

¹⁹ The maximization of power is at the heart of Morgenthau's (1948) work.

The neorealist revolution replaced power maximization with an assumption that states minimally act to assure their own survival.²⁰ Yet, this assumption is neither self-evident nor adequate. Assuming the primacy of survival as a minimum is not enough to provide an unambiguous basis for decision. It is not self-evident because what constitutes survival must itself be defined. Survival raises issues of self-definition and identity. Some states fit the classical realist vision of focusing solely on physical and territorial integrity. But other states define their survival more broadly, including such issues as ideology and ethnicity (Stein 1995). Moreover, the presumption of a minimal concern with state survival is simply inadequate. It confines the field to be able only to address those international relations in which states' survival is on the line, and this ignores too much of international relations.²¹

Further, in international relations (and in other fields), imputing the same interest to all actors has drawn criticism and led to calls for inducing subjective national interests and not just assuming objective interests.²² This view holds that actors' formulations of their interests, the nature of their utility functions, must be investigated directly. This problem is at its most interesting among anthropologists, who must choose whether to explain the behavior of people and groups within the framework of their own worldviews or by applying concepts that have no meaning to those people whose behavior is to be explained.²³

This debate mirrors the controversy in international relations between generalists (or theorists) and regionalists in the study of international relations. During U.S. involvement in Vietnam, for example, debates raged not only about appropriate policy but about the knowledge necessary to

²⁰ The difference between positing the maximization of power and minimally assuring survival is the main, self-consciously articulated distinction between classical realism and neorealism (Waltz 1979, 1990).

²¹ This animates Krasner's (1978) inductive search for the national interest. Krasner begins by arguing that the core realist assumption about state preferences is inadequate to explain the bulk of a great power's foreign policy.

²² Rosenau (1968) distinguishes between subjective and objective assessments of the national interest; Krasner (1998) distinguishes between deductive and inductive means of assessing the national interest.

²³ See the distinction between emic and etic in Harris 1979. This anthropological problem has entered political science in discussions of peasant behavior (Scott 1975, 1976; Popkin 1979). The problem also arises for historians, who confront the issue of being constrained by what was self-consciously known by those in the past.

There is a long-standing argument about the universality of economic exchanges and the existence of premodern modes of calculation and assessment (Finley 1973; Polanyi 1944; Kindleberger 1937, 1944; Humphreys 1969). For a refutation of the argument for classical economies, see Conybeare 1987, chap. 4. The issue is hotly debated among economic anthropologists, who refer to it as the substantivist-formalist controversy. For an essay that links these strands, see Lowry 1979.

make policy. Theoretically oriented scholars were prepared to apply the lessons of game theory and deterrence to the American intervention in Vietnam. John McNaughton, assistant secretary of defense in the Johnson administration, asked his old friend, Thomas Schelling, to apply his ideas about bargaining and signaling and coercive warfare to the problem of intimidating North Vietnam through the use of U.S. air power.²⁴ In contrast, regionalists are horrified by the ahistorical and acontextual use of general arguments. In this case, they held that a knowledge of Vietnamese history, society, and culture were essential to understanding how to interact with the North Vietnamese.

In addition to the issue of how to ascertain states' utility functions is the question of what to include in them. The larger issue is what sorts of preferences about states will be assumed in the study of international relations. The analytic predisposition of adherents of strategic-choice models is to keep the utility functions as spare as possible. But a core question for the field is how spare can assumptions of state preferences be. No matter how compelling the analytic desire for a simple stipulated preference, the preference functions of states will have to be expanded in order to explain the range of observed international relations.²⁵

Much of the work in the field of international political economy, for example, simply borrows assumptions about state preferences from international trade theory and holds that states are interested in maximizing national wealth. But as recent criticisms of this work point out, foreign economic policy is driven by both security and material concerns. This is the heart of Gowa's argument that foreign economic policies have *security externalities* and that security interests as well as economic ones underlie foreign-trade policies (Gowa 1989; Gowa and Mansfield 1993). Similarly, security policies have *wealth externalities*, and these figure in the formulation of state security policies. Concerns about the effects of security policies and military spending on a state's export prospects and future economic growth have been and remain key elements in the formulation of national security policy. Scholars are focusing on how material and security interests combine in the formulation of foreign economic policies and even of security policy (Frieden 1994b; Papayoanou 1996, 1997, 1999; Skålnes 1998, 1999; Steele 1995; Weber 1997).

²⁴ Significantly, Schelling had no idea where to begin (Kaplan 1983).

²⁵ Much the same has occurred in other subfields of political science. Scholars who have run up against the limits of what they can explain in American politics simply by positing that politicians have a reference to be reelected have been on to expand politicians' utility functions to include policy preferences. Similarly, positing that voters have a simple preference for maximizing income has proved inadequate, and more recent work combines voters'

Conjoined with the issues of how to define and assess preferences is the question of whether preferences can change. Analysts often treat preferences as fixed and unchanging.²⁶ Some even hold that changing preferences pose a problem for a theory of choice. Yet, it is possible to model how preferences are formed and how they change (Kapteyn, Wansbeek, and Buyze 1980; Hansson 1995; March 1978; Schelling 1984).²⁷

Scholarly debate about the nature of preferences, the basis for inferring them and their malleability, is within the church and constitutes no challenge to strategic interaction analysis per se. Making one set of modeling choices (bets, in the language of Lake and Powell in chapter 1 of this volume), preferring to keep preferences unchanged and explaining by reference to changing constraints and opportunities (Frieden's chapter 2), does not mean that other modeling choices imply a fundamentally different view of causality.

STRATEGIC CHOICE AS MATHEMATICAL MODEL, MATHEMATICS AS THE LANGUAGE OF SCIENCE: MODELING AND INTELLECTUAL BRUSH CLEARING

The main benefits ascribed to the use of a strategic-choice approach in international relations are those of any mathematical model, greater analytical coherence and rigor. Game theory, which is the basis of the strategic-choice approach discussed in this volume, is a branch of mathematics, and its strengths are those all mathematical modeling imparts.²⁸ Mathe-

interests in their own incomes with some general preference (held in varying degrees by different voters) for equity.

²⁶ Stigler and Becker (1977) make a strong argument for this position. Frieden's essay in this volume (chapter 2) makes this case for international relations.

²⁷ The issue of preference change is related to that of the dependence of preference on belief. Chapter 1 of this volume takes the orthodox position of keeping these separate, but there are important cases in which preference depends on belief (Hausman and McPherson 1994).

The issue of preference change is related, too, to the question of whether actors themselves change and whether they are capable of taking into account anything other than present aims. It is not just that actors heavily discount future payoffs but that they have no sense of, or interest in, their future selves and simply pursue present aims (Parfit 1984; Walsh 1994). Is the IBM that made business machines in the first half of the century the same company that made computers in the second half of the century when the products, organizational structure, and leadership had all changed? Similarly, is the middle-aged corporate executive the same person as the hippie of two decades earlier? Can they be said to share the same preferences? Can the young hippie be thought of as making choices with calculated temporal consequences?

²⁸ The father of the theory of games, John von Neumann is, by at least one reckoning, one of the greatest mathematicians of all time (Paulos 1991). His important work, *The*

matics is the language of science.²⁹ It adds precision and rigor and makes assessments possible that simply are not possible in ordinary language.³⁰ Moreover, mathematical models are formally true.³¹ Hence the empirical adequacy of game theory, indeed of formal and mathematical work more generally, is not at issue. The statements formally derived within an axiomatic structure are formally true and need not be empirically assessed.³² But because the central role of empirics involves the isomorphism between analytic and modeling assumptions and the underlying reality being assessed, many of the disputes over strategic choice actually concern the specific modeling decisions scholars make.

Further, although mathematics is substance-free, specific uses fill this generic tool with meaning, and verbal formulations of substantive problems translated into formal (or numerical) terms must be subsequently translated back.³³ Quite different substantive domains may be studied using similar mathematical tools, but the use of similar tools need not imply common substantive links. Specific kinds of mathematics have been developed with particular substantive issues in mind, but the tools can be applied in other domains as long as the problems are isomorphic. Newton developed the mathematics of calculus in order to study planetary motion, but calculus is applicable to any question involving how fast something

Theory of Games and Economic Behavior, focused on economic behavior largely because of the impact of his coauthor, Oskar Morgenstern, the economist (Leonard 1992; Rellstab 1992; Schotter 1992; Mirowski 1992). But the earlier roots of game theory are all in mathematics (Dimand and Dimand 1992). Ironically, during game theory's early history at Princeton, the mathematics department encouraged, and the economics department frowned on, the new enterprise (Shubik 1992).

²⁹ Galileo expanded this point further when he wrote, "The book of nature is written in the language of mathematics; without its help it is impossible to comprehend a single word of it" (quoted in Pinker 1997).

³⁰ The renowned economic game theorist, David Kreps (1990c), offers a rather restrained and limited picture of the contributions of game theory to economics. Indeed, the contributions he lists are not specific to game theory at all but are generic ones applicable to mathematical modeling generally.

³¹ Indeed, the different appellations applied to this enterprise in political science include "formal theory" and "mathematical political theory."

³² This discussion finesses debates in the philosophy of mathematics.

³³ Not all mathematical possibilities have real-world counterparts. Many games can be analyzed as hypotheticals, but relatively few have been studied because they model important social reality. Kenneth Boulding once decried modern mathematical economics as "a quantum mechanics for an unknown universe." Mathematicians disagree about whether their constructions exist in their minds or in the world. For an introduction to such issues in the philosophy of mathematics, see Barrow 1992; Davis and Hersh 1981; and Paulos 1991.

is changing or how much a changing quantity totals.³⁴ One should be wary of assuming that two domains are comparable simply because some models are usable in both. That calculus can be applied to elections and planetary motion does not mean that the domains are not fundamentally different.

Similarly, game theory is generic in character and has also been widely applied. In addition to its extensive use in economics (Kreps 1990b, 1990c; Tirole 1988; Fudenberg and Tirole 1991; Binmore 1992; Rasmusen 1989; Osborne and Rubinstein 1994), it has been used in social psychology (Colman 1982; see also Kelley and Thibaut 1978), political science (Brams 1975; Ordeshook 1986), international relations (Brams 1985; Brams and Kilgour 1988; Nicholson 1992), and philosophy (Braithwaite 1955; Gauthier 1986; Parfit 1984; Lewis 1969).³⁵ Indeed, game theory has become widely applied in biology to interactions between animals (Maynard Smith 1976, 1982, 1984, and comments that follow 1984; Dugatkin and Reeve 1998).³⁶ In all fields, every application involves modeling choices, and most scholarly debates are about such choices.

Still, the core assumptions of game theory would appear to demand greater constraints on their use than other branches of mathematics. In arithmetic, for example, little need be assumed about objects and events in order to count, add, subtract, multiply, and divide them. In contrast, applying a strategic-choice model would seem to presuppose more stringent substantive assumptions. The actors in a strategic model must be sentient creatures capable of comparison, assessment, and choice. Yet, the components of a strategic-choice approach are not so constraining as to limit its applicability to human beings alone. That game theory can be used to model animal behavior³⁷ implies neither that inten-

³⁴ On the centrality of the substantive importance of the study of moving bodies for critical developments in mathematics, see Kline 1985; for the simplest description of differential and integral calculus, see Paulos 1991.

³⁵ For an imaginative application to literary interpretation, see Brams 1980.

³⁶ For discussions of the analogous quality of economics and biology by economists, see Hirshleifer 1977, 1978a, 1978b; and Samuelson 1978, 1985.

³⁷ Hammerstein (1989) notes that, surprisingly, a model presuming human rationality has not only found wide applicability in the study of animal behavior but seems to do a better job explaining it than it does human behavior. For a review of game theory and evolutionary biology, see Hammerstein and Selten 1994.

The Kahler essay in this volume (chapter 6) beautifully demonstrates the issues associated with the use of evolutionary models in international relations. In their hard biological form, these models dismiss intentionality, but Kahler discusses the ways that such models have been imported into the social sciences and the attendant modifications that see a role for directed variation and adaptive learning

tionality can be attributed to animals nor that intentionality need not be part of game-theoretic explanations in the social sciences.³⁸ That noncooperative games can be applied to pretrial bargaining, legislatures, and international politics does not mean that the domains are not analytically and theoretically distinctive.

The central benefit of mathematical modeling is its logic and coherence, its internal consistency. These are not trivial advantages in an intellectual domain beset by analytic sloppiness. The analytic consequences can be readily grouped into three types. First, formal modeling can demonstrate *nondeducibility*—that the deductions argued to follow from certain premises do not actually do so. As has been argued in recent years, for example, assumptions of anarchy do not by themselves lead to a deduction of the conflictual nature of international politics.³⁹ Second, formal modeling can demonstrate *multifinality*—that multiple outcomes flow from the same premises. The multiple equilibrium problem discussed below provides one example. Third, formal modeling can demonstrate *equifinality*—that there are alternative paths to the same outcome. The example often given, as in chapter one of this volume, is that a concern with relative gains can emerge from a purely self-regarding calculus without the need to change the underlying utility function.⁴⁰ These consequences of mathematical modeling have proved powerful enough to transform the scientific enterprise, both natural and social.

THE NORMATIVE-POSITIVE PARADOX

Ironically, the strategic-choice approach offered here as a retrospective explanation for behavior and outcome began as part of a normative enterprise intended to improve decisions, not to explain them. The use of such normative tools and ideal types for positive explanation is inherently problematic—they are either true or false because of their self-conscious application. Moreover, modifying these tools to make them more isomorphic with reality, and thus more useful for explanation, has generated a knowledge of the incompleteness and indeterminacy of such models.

Game theory has been used in international politics, as elsewhere, both to explain behavior and as a tool for improving the quality of decisions. The predominant focus in this volume, and thus for most of this essay, is on the use of strategic choice as an explanation for international politics.

³⁸ In a similar vein, the use of a conception of equilibrium in the natural and social sciences does not imply that economic equilibrium means the same thing as physical equilibrium (Phelps 1991).

³⁹ For my formulation, see Stein 1990.

⁴⁰ For my take on this, see Stein 1990, chap. 5.

But to its progenitors and original practitioners, game theory has been a normative tool, one devised to make decisions more rational.⁴¹

*A Normative Enterprise: Constrained Actors
in Search of Rationality*

Probability, logic, decision theory, and game theory were all developed to improve human decision making. They were not developed as accurate representations and reconstructions of what people actually do but as tools individuals should apply to achieve more rational decisions than they otherwise might.⁴² People were assumed to be purposive but limited in their capacity to be fully rational. Decision-making tools were thus desired, created, recommended, and adopted. The logician Gottlob Frege argued that logic addresses the way people “must think if they are not to miss the truth” (Lowe 1993).⁴³ Similarly, probability theory has its roots in gamblers looking for an edge. On their own, people were seen as poor intuitive statisticians prone to mistakes, and the use of probability theory could improve their decisions.⁴⁴ Psychologists’ findings about the heuristics and biases of human judgment underscore the point; in the words of Kahneman and Tversky (1973, 237): “In making predictions and judgments under uncertainty, people do not appear to follow the calculus of chance or the statistical theory of prediction. Instead, they rely on a limited number of heuristics which sometimes yield reasonable judgments and sometimes lead to severe and systematic errors” (see Kahneman, Slovic, Tversky 1982). These fields all began with a normative orientation rather than a positive one; they were intended to improve human decision making and were not intended to explain what people actually do.

⁴¹ The only value to which decision theorists and game theorists were committed was rationality: optimizing and maximizing actor preferences.

⁴² Nozick 1993, xi) characterizes the contrast between two perspectives regarding people and their rationality as expressed by two of the greats: “Descartes attempted to show why we should trust the results of reasoning, Hume questioned the rationality of our doing so.”

⁴³ The contrary position was held by Locke who argued that people have minds that do not require instruction in logic. Knowledge of the rules of logic can help, but the mind can reason in logical terms even absent a knowledge of logic. This is akin to the argument in modern macroeconomics that individuals calculated rational expectations even before the elaboration of the model and the mathematics that made it tractable.

⁴⁴ Some early probability theorists saw their enterprise as capturing intuitive reasoning, or, as Pierre Laplace put it, probability theory was “only good sense reduced to calculus” (Laplace 1951 [1814], 196). When probability theory conflicted with intuition, it was the theory that needed reformulation (Daston 1980, 1988).

Similarly, the use of game theory can be seen as a normative tool useful for decisions.⁴⁵ Just as the telescope extends what one can see beyond the range of the unaided human eye, so game theory extends what one can logically assess beyond the means of the untrained and unaided human brain. Seen as such a tool, game theory (and decision theory) may be quite useful for making decisions but cannot be used to explain decisions made by actors in the past who did not have this knowledge.⁴⁶

Ironically, though, what started as a normative enterprise has become a positive one. Tools once created to improve the quality of otherwise imperfect human decision making are now being used to explain choice. The social constructions of humans wishing to improve the quality of their cognitively constrained mental faculties, both bounded and semi-autistic, are now being used retrospectively to explain past human choices.⁴⁷

Decision- and game-theoretic explanations of choice are thus inherently problematic in that they make inappropriate assumptions about the capacity of individuals and groups to make rational individual and social choices. The models were developed as normative instruments because individuals were seen as wanting to be rational but constrained in their ability to be rational. Applying such models as positive explanations, then, has led to the assault on rational choice by cognitive psychologists. As discussed below, individuals' presumptive ability to make rational choices has been systematically attacked by work in cognitive psychology showing that individuals deviate from the requisites of individual rationality.

Finally, that formal work was developed with an avowed normative objective of improving the quality of decision making implies that it is absurd to criticize it as inherently conservative. Yet, one criticism of formal work in economics and political science is that it is conservative since it takes the world and its constraints as given and so ignores alternative possibilities.⁴⁸ In international relations, rational-choice theory is seen as

⁴⁵ Nozick (1993) argues that it is inadequate even as a normative construction and that the standard normative view of conditions that a rational decision should satisfy must be expanded to include the symbolic meaning of actions.

⁴⁶ One can push the point further. Perhaps game theory is not an experimental apparatus being used to explore the world but a tool that can itself generate structure, as a computer algorithm might. If so, game theory is creating something new. On experimental tools, see the discussion in Barrow 1992, 261.

⁴⁷ Camerer (1997, 167) points out that "it is remarkable how much game theory has been done while largely ignoring [the] question [of whether it is meant to describe actual choices by people and institutions]" and goes on to note that even when it "does aim to describe behavior, it often proceeds with a disturbingly low ratio of careful observation to theorizing."

⁴⁸ The critique is also related to the politics of the modern age. Earlier generations of economists were vociferous critics of socialism and communism. Some of those writing in the public-choice arena in the last two decades have been staunch critics of the welfare state.

blocking the prospects for envisioning and creating a world that transcends the anarchic and conflictual system of competitive nation-states.⁴⁹ But this criticism applies to current practice or practitioners rather than to the nature of the intellectual tool. Nothing in the nature of strategic choice generates a particular political, as opposed to analytic, perspective. Indeed, people developed the perspective in order to overcome their natural limitations and better fulfill their natural desires.

*Bounded Rationality, Political Autism, Knowledge,
and Self-Validating Theory*

Strategic-choice explanations will certainly be correct in explaining the choices of actors who self-consciously and correctly use strategic-choice theory in making their decisions. A socially constructed edifice of rationality created by self-interested and purposive beings will function as a self-validating or self-fulfilling theory when actors use a model of the world to make their choices and to explain the behavior of those who make choices using such a model.

But since the model is a normative ideal, its use as a positive explanation will sometimes prove wrong when it is applied to cases in which actors do not use the model self-consciously. Since the models were constructed by those who deemed themselves crippled rationalists who needed decision-making tools to make better decisions, using the models to explain the decisions of those who did not use them will prove problematic at least some of the time.

The very use of rational-choice models in positive explanation has opened the door for cognitive psychologists to demonstrate all the ways in which people are crippled rationalists who fail to achieve the normative ideal captured in rational-choice theory.⁵⁰ The psychologists' assault is not on purposive explanation per se. Abelson (1976) argues that social psychologists subscribe to a "limited subjective rationality" that qualifies standard notions of rationality by recognizing that people may have dis-

⁴⁹ Some feminists and Marxists view rational choice as hostile to their agendas and have gravitated to the seemingly more hospitable intellectual soil of constructivism or postmodernism. Yet, rational choice can as readily be used by Marxists and feminists (there are many of the former and few of the latter) as by conservative proponents of decentralized market exchange.

Note that these labels are only specific to a stylized and caricatured depiction of modern political alignments. In an earlier time, political liberals were the staunch proponents of decentralized markets against conservative statist supporters of mercantilism. For the evolution of liberal views of the state and exchange, see Stein 1993.

⁵⁰ For an accessible history of the cognitive revolution, see Gardner 1985.

torted notions of reality and may filter information. The qualifications capture the mental-processing rules that people actually use in place of formal logic.

A key implication of these criticisms is that careful attention must be paid to the human ability to process information and to how people conceptualize the world in which they function. To reflect reality, formalizations of choice must necessarily be built on accurate behavioral foundations.⁵¹ In addition, context not only matters, it may be all important.⁵² The substantive nature of actual choices may be an essential element of decision.

But the criticisms emerging from cognitive psychology do not constitute attacks on purposive explanations per se but on their assumptions about rationality. Psychologists, who have argued that human beings do not have the cognitive ability always to carry out the requisites of rational choice, focus on the heuristics and shortcuts people use to process information and make decisions.⁵³ They have also clearly demonstrated the importance of context to decision making. Unlike economists, who emphasize the generality and universality of rational choice, psychologists point to situational factors. Their experimental work clearly shows, for example, that individuals do not perceive gains and losses symmetrically. People take risks to avoid certain losses but are risk averse with regard to gains. Thus how options are framed, as losses or gains, is critically important in determining actual choice (Tversky and Kahneman 1981, 1987).⁵⁴

⁵¹ Perhaps the greatest such problem posed by psychology for modern game theory is whether people are Bayesian rational. Formal models of incomplete information games depend on Bayesian updating (Mariorri 1995), yet the empirical evidence is that most people are not naturally Bayesian.

⁵² The existence of constraints on cognitive ability (Oaksford and Chacer 1993) and the context-dependence of cognition (Stevenson 1993) are widely recognized, and psychologists do want to ground theory in cognitive process (Shafir 1993). Some economists, too, see the importance of incorporating actual reasoning (Rubinstein 1991) and stress the importance of bounded rationality and of empirics (Binmore 1988). Or, as Simon (1990) puts it, the invariants of human behavior lie in cognitive mechanisms for choice.

⁵³ Economists were originally quite skeptical of many of these psychological assaults but have now conceded many of their points. A good way to trace this is in the work of Charles Piotr (Grether and Piotr 1979; Piotr 1987). One implication has been the burgeoning field of experimental economics and the use of small group experiments to assess economic arguments (Smith 1992).

The key work is that of Tversky and Kahneman 1974. Economists have debated how troubled they should be that people do not function in the ways assumed by economic theory (Friedman 1953; McClelland 1975, 136-43; Lagueux 1994). Most defensively, Piotr (1987) accepts the problematic nature of the psychological evidence on economic assumptions but argues that the enterprise be continued in the absence of a viable alternative.

⁵⁴ Still another example of the importance of context is provided by the argument that people make decisions sequentially using some aspect of the options available to them. Econ-

Critical choices between peace and war can also reflect framing. Leaders who see a choice between war now and war later, rather than between a cooperative gesture and a military response, are much more likely to take an escalatory military step during a crisis (Snyder 1978). The decisions that led to World War were made by leaders who saw no alternative (Farrar 1972).

The constraints on rationality include human emotions, as well as how brains process information, for people have passions as well as interests (Hirschman 1977).⁵⁵ Economists have developed utilitarian explanations for the existence of our emotions and argue that self-interested beings find emotions quite helpful (Frank 1988, 1993; Hirshleifer 1987a, 1993).⁵⁶ Nevertheless, the utilitarian basis of emotions does not undercut the perverse implications of emotionalism for rational-choice explanations of behavior.⁵⁷ Again, formal theorizing with a normative focus may be driven by a desire to minimize the impact of emotions on decisions—to facilitate interests and minimize the ability of passions to get in the way—but positive theory that seeks to explain what people actually do must incorporate the impact passions, as well as interests, have on choice.⁵⁸

Not only are humans boundedly rational and emotional, they are also somewhat politically autistic. They do not make appropriate attributions, especially in social settings. A classic example comes from the actor/observer literature, which finds that people typically attribute their own choices to structural constraints but attribute others' choices to prefer-

omists would argue that each of three equally valued options would have a one-third probability of being chosen. Tversky (1972) argues that this is nonsensical when two of the options have a common feature—for example, different recordings of the same symphony—and the third is quite different (a book, for example). Tversky's argument is that a person would first choose between a book and a recording and only then select between recordings; so the odds of selecting the book are one in two, and the odds associated with each recording are one in four. Again, it is not a model of calculated choice that is being questioned but the specific presumptions of the process.

⁵⁵ Hirschman (1991a, 357) argues that turn-of-the-century economists rejected the "instinctual-intuitive, the habitual, the unconscious, the ideological and neurotically-driven"—that is, turned away from the "nonrational that characterized virtually all of the influential philosophical, psychological, and sociological thinking of the time." In so doing, they emptied their concepts, most specifically self-interest, "of their psychological origin."

⁵⁶ One philosopher argues that impulses have utility in finding salient solutions to coordination problems (Gilbert 1989b). Others argue that habitual and routine behaviors can be seen as rational (Hodgson 1993a). Simon (1978a) points out that even psychoanalytic theory contains a functional component and a sense of rationality (contrast with Cohen 1976).

⁵⁷ Human emotions need not be seen as antithetical to rational choice; indeed, they may be essential to any human ability to comprehend and evaluate gain and loss (Damasio 1994).

⁵⁸ Some game theory does include the possibility of unexplained deviations from presumed rationality (threats that leave something to chance, trembling hand equilibria).

ences, predispositions, and character. International relations are rife with such political autism. The security dilemma describes how states reduce their own security by not taking the reactions of others into account when they take steps to improve their security. The belief that one's own actions are merely reactions to others' provocations, when those others' actions are not understood as reactions themselves, is a form of political autism.

Hence using strategic choice as an explanation requires making insupportable assumptions about individual capability. People are crippled rationalists, wanting to make rational decisions but constrained and limited by their own psychology: by their emotions and cognitive processes.

Yet self-interested purposive beings also use the knowledge at their disposal to improve the quality of their decisions—to improve on the limitations of their biology. They create tools, like computers, and develop knowledge, including game theory, in order to do better.

This means that *actors' knowledge, including that of social-science theory, must be incorporated into the explanation of choice and outcome.*

Knowledge of strategic choice (and economics) is self-fulfilling and self-validating. These bodies of knowledge were created by people who wanted to do better, and they provide guidelines for doing just that. When self-interested actors apply them, their knowledge claims are true precisely because they are used. The self-conscious normative use of economics makes economics a self-fulfilling truth as a positive theory.⁵⁹

By contrast, *knowledge of psychology is self-falsifying.* Although economists and psychologists agree on the desirability and utility of purposive explanation, the former presume rational choice, the latter question it. The work of psychologists focuses on the ways that people fall short of the ideal model of rational choice. Much psychological work, therefore, including that in international relations, proffers recommendations on overcoming peoples' shortcomings that make the theories self-falsifying. Studies of crisis decision making typically find that stress adversely affects decisions, but they conclude with recommendations on how to avoid the impact of stress and achieve more rational decisions (Holsti 1989; Janis 1982). The implicit, if not explicit, message of much psychological work is that being sensitive to cognitive distortions can improve the quality and rationality of human decision making.⁶⁰

A complete model of choice and behavior must necessarily include the knowledge actors have about the world. Since people can be aware of, and

⁵⁹ In recent bidding for government cellular phone rights, all the bidders hired game theorists to advise them on bargaining strategy, and, not surprisingly, the game-theoretic prediction for the bidding worked quite well.

⁶⁰ Much the same can be said of psychoanalytic theory. Once individuals became aware of the factors driving their behavior, the analysts held, they would be able to deal with them.

indeed can make use of, social scientists' theories about their behavior, explanations of their behavior need to incorporate people's knowledge and beliefs about how the world works and the implications of such knowledge for their decisions. This means that some knowledge claims become true because of their self-conscious application by the subjects of the theory, and other knowledge claims become false because of their subjects' conscious awareness of them.

*Positive Verisimilitude and Incomplete and Indeterminate Explanation
(The Multiple-Equilibria Problem)*

One solution to the problem of using a normative model for positive purposes is to expand the formal model to make it more isomorphic with the reality being explained, that is, to make it less of an ideal for which humans strive and more reflective of constrained human beings with constrained information. In fact, scholars have modified the nature of rationality presumed in strategic-choice explanations.⁶¹ Early critics focused on the presumably unbounded nature of rational explanation, which posits that actors assess all possible options and maximize. Even reflective game theorists recognize this, or, as one major game theorist puts it, "homo rationalis [a species that acts purposefully and logically] is a mythical species, like the unicorn and the mermaid" (Aumann 1985). Economist Brian Arthur characterizes the domain of rational explanation in similarly dismissive terms: "If one were to imagine the vast collection of decision problems economic agents might conceivably deal with as a sea or an ocean, with the easier problems on top and more complicated ones at increasing depth, then deductive rationality would describe human behavior accurately within a few feet of the surface" (Arthur 1994, 406).⁶²

Troubled by the synoptic quality of rationality, Herbert Simon and others replace it with a notion of bounded rationality. Simon, for example, argues that actors satisfice rather than maximize, that they stop their process of assessment when they hit on a minimally acceptable option. At issue here is not an alternative to purposive explanation per se but to its

⁶¹ For many who accept the logic of purposive explanation, the notion of rationality still draws wrath. Although the discussion that follows treats criticisms analytically, understand that some criticism stems from the connotation of the word *rationality*. Rational-choice explanations do contain some assumptions in addition to those of intentionalist explanation (Elster 1985, 1986).

⁶² Arthur (1994, 406) goes on to suggest that deductive rationality would apply to a simple game such as tic-tac-toe but that "rational 'solutions' are not found at the depth of checkers, and certainly not at the still modest depths of chess and go."

character.⁶³ Rather than being hyperrational, individuals are boundedly rational, and bounded rationality has emerged as a growth industry within rational-choice scholarship.⁶⁴

The recognition of the inadequacy of decision theory and game theory as positive explanations has sent scholars in the direction of making these models more isomorphic with the reality being modeled. The simplistic assumptions of early models—simultaneous choice by two actors, each choosing between two options and having complete information about strategies and payoffs—were replaced by subsequent refinements that dealt with sequential choice and incomplete information.⁶⁵ These refinements expand the problem of multiple equilibria and thus explode the indeterminacy associated with the models. Further expanding the models to achieve greater verisimilitude (e.g., adding options and actors, etc.) quickly reaches the limits of mathematical tractability, but even if these problems are eventually solved, the solutions will surely expand the range of indeterminacy.

As currently developed, strategic situations of the slightest complexity are plagued by multiple equilibria and multiple solutions concepts. Many recent solution concepts have been developed in order to reduce the number of equilibria generated by simpler criteria.⁶⁶ But the mere existence of multiple equilibria (as well as multiple solution concepts) implies that knowing the players' strategies and payoffs is inadequate completely to determine a unique outcome. In other words, structure and choice are indeterminate because they are explanatorily incomplete.⁶⁷ This *strategic*

⁶³ The way Simon (1976, 1978a, 1978b) and others put it is to distinguish between different types or aspects of rationality (Evans 1993). Simon's distinction is between substantive and procedural rationality.

⁶⁴ See Conlisk (1996); and Lipman (1991, 1995). Aumann (1997, 8) argues that equilibrium refinements that have been proposed to deal with the multiple-equilibrium problem "don't really sound like bounded rationality. They sound more like super-rationality."

⁶⁵ Game theory has been extended from perfect to imperfect information, from dealing with cases in which the payoffs sum to zero to those in which they do not, from assuming that utility (payoffs) can be transferred among actors to cases in which they cannot, from assuming simultaneous choice (normal or strategic form) to sequential choice (extensive form), from dealing only with the interaction of two actors to that of n actors, and from complete to incomplete information. Aumann (1992) extends game theory to cases in which actors need not assume that others are necessarily rational. For a description of advances, see Harsanyi (1977, 1988).

⁶⁶ For a discussion of the relationship among equilibria, see Morrow (1994a).

⁶⁷ The existence of multiple equilibria implies that the enterprise is inadequate for prescription and incomplete as explanation. The reliance of solution concepts on beliefs about unreachable states is also a problem. Some scholars confronting the issue of equilibrium selection and individual deviation from strategic rationality have even abandoned the formal enterprise and moved toward the use of experiments to assess the basis of human choice

indeterminacy principle represents the explanatory boundary of strategic choice and minimally implies that in cases with multiple equilibria, complete explanation will require conjoining this approach with something else. Strategic-choice explanations thus demonstrate their own theoretical incompleteness.

COMMON KNOWLEDGE, THE CONSTRUCTION OF CHOICE, AND POST HOC EXPLANATION

Strategic-interaction models presume that actors have common knowledge of the rules of the game, of each other, their choices, and the probability distribution of one another's preferences. The limitations associated with this modeling requirement are enormous and, in all but the most structured social situations, mean that explanation is invariably post hoc.

The essence of the strategic problem is the assumptions made about the common knowledge of the actors.⁶⁸ This was unproblematic when game theorists only studied games, like chess, and could assume the players had common knowledge of the rules. Much the same is true of highly structured social settings, as when a government establishes fixed procedures to regulate bidding between firms in an auction. But in many social settings, common knowledge is far more problematic. Can we, for example, even be certain that people mean the same thing when they use the same word?⁶⁹ This form of the problem is particularly important in international relations. At one point, Yasir Arafat, the Palestinian leader, created a roadblock in Israeli-Palestinian negotiations when, in addressing an Arab audience, he used the Arabic word, *jihad* (usually translated as holy war). A debate ensued about the meaning of the word in Arabic and its typical translation; namely, had Arafat alluded simply to a political struggle or to a holy war.

(Plott 1991). Those unwilling to depart from the formal enterprise are increasingly moving toward evolutionary game theory that does not depend on strong rationality assumptions (Binmore and Samuelson 1994; Van Damme 1994; Robson 1995).

⁶⁸ For the ongoing discussion of the meaning and necessity of an assumption of common knowledge about the rules of the game, see Lismont and Mongin 1994. This focus has brought together the frontiers of game theory with that of logic in philosophy (Bacharach 1987, 1994; Stalnaker 1994). For some implications of a deviation from common knowledge, see Geanakoplos 1992, 1994.

Putting oneself in another's shoes is biologically hard-wired. The autistic suffer from a neurological disorder that makes them largely blind to the existence of other minds and makes them unable to assess others' beliefs and intentions. Simon Baron-Cohen (1995) describes this inability to infer what others are thinking as "mindblindness."

⁶⁹ This is, in fact, a critical issue in the philosophy of language.

More broadly, how states come to see themselves as engaged in a particular game is itself important. Even if we assume that states know the targets, both direct and indirect, of their actions, it is not at all clear that states accurately recognize themselves as targets of others' actions. Scholars have debated the extent to which Truman's decision to drop the atomic bomb on Japan was intended as a signal to the Soviet Union. Conversely, in certain situations states incorrectly see themselves as targets of others' actions, or sometimes fail to recognize that they are targets.⁷⁰ Even more broadly, actions have audiences as well as targets, and actions taken for domestic audiences can have international consequences, and vice versa. The general point is that in certain games, it is not always self-evident who the actors are.

The issue of what is posited on behalf of actors extends to the nature of their choices. Typically, scholarly analysis stipulates the strategic choices actors confront. Yet, crafting alternatives lies at the heart of statecraft and creative diplomacy. When diplomats search for "formulas" to resolve conflicts, they are looking for options other than those that created the deadlock in the first place (Zartman and Berman 1982). Ulysses's choice could have been framed as between traveling within hearing distance of the sirens or navigating a path that kept him out of hearing range. It could also have been put as the choice between keeping his ears plugged (not hearing) or leaving his ears unplugged but paying the consequences. But Ulysses devised an alternative, one that effectively allowed him to eat his cake and have it, too.⁷¹ This human creativity in structuring alternatives is completely outside the modeling of strategic choice. An analyst can elucidate the payoffs associated with various options and can explain choices and even detail circumstances in which actors find themselves aggrieved and would prefer to have different options than the ones they seem to have or be given. But the delineation of choice is a human activity, so one cannot know a priori whether actors will craft new alternatives and of what kind.⁷² In other words, the analyst cannot know a priori that Ulysses will devise a way that will allow him both to listen to the sirens'

⁷⁰ Jervis (1976) argues that states often mistakenly see themselves as the targets of others' actions.

⁷¹ Elster (1979) uses the story to demonstrate a case in which the possibility for rational choice exists before the fact rather than during it. Elster refers to this as "imperfect rationality."

⁷² The point can be put even more broadly: actors do not know what their preferences are until they actually begin to analyze a problem. Anderson (1983) argues that during the Cuban missile crisis, the preferences of the members of the executive committee advising President Kennedy only emerged during their deliberations through the process of assessing options. For a debate on whether the mutual determination of wants and benefits poses a problem for decision theory or is merely an attack on the implicit view of how deliberations take place, see Kusser and Spohn 1992 and Broome 1994.

call and to avoid their deadly temptation. Post hoc, it is easy to explain Ulysses's choice. Strategic-choice explanation presumes a knowledge of the complete set of options and is thus post hoc.

Constructing options and transforming payoffs plays fundamentally with the conception of choice under constraints. Some actors are pragmatists who take constraints as given; others act as revolutionaries more interested in overcoming constraints. The exchange between U.S. President Jimmy Carter and Israeli Prime Minister Menachem Begin at Camp David illustrates a dialogue between a pragmatist and a visionary. Carter focused on the extant situation and what it required from the Israeli leader. Begin's response, irritating to Carter, was to lecture the U. S. president on Jewish history. His point was implicit but clear: Had he and others given in to presumed requisites of the day, the state of Israel would not have been created in 1948. Begin preferred to await better circumstances than to compromise with current reality. Alternatively, he could try to create an alternative future and reality, what Israelis refer to as "creating facts."

Actors cannot only create options and attempt to transcend constraints, they can also structure choices for one another. The Prisoner's Dilemma is, after all, a case of three actors making choices in which one of them, the district attorney, has the power to structure the options and payoffs available to the other two (Burns and Buckley 1974).⁷³

Institutional design is also about structuring choice and creating new alternatives. Such designs, for example, can transform public goods into collective ones, and vice versa.⁷⁴ The most-favored-nation clause fundamentally alters the inherently private and divisible nature of trade and trade policy. States can and do pursue specific and differentiable trade relations with other states. Historically, trade agreements were signed by pairs of countries stipulating general terms for their bilateral trade. These illiberal tools, originally mercantilist devices states used to discriminate among their various trading partners, were completely transformed through the inclusion of the most-favored-nation clause. By guaranteeing that the signatories would extend to one another any preferences they offered third nations in subsequent treaties, this clause ensured that no state would be more favored; thus international trade agreements

⁷³ Relational power—the ability to affect another's payoffs—can be distinguished from metapower—the ability to structure another's options and payoffs (Baumgartner, Buckley, and Burns 1975; Baumgartner, Buckley, Burns, and Schuster 1976).

⁷⁴ Other elements of institutional design include the construction of private property rights in order to "marketize" some domain rather than regulate it and the construction of different oversight structures (the fire alarms/police patrols distinction).

changed from private discriminatory arrangements into accords that created a public good among a club of members.⁷⁵

The constraints by which actors are limited or which they attempt to transcend vary in character. Physical and biological limitations differ from social ones. Without external assistance, human beings cannot fly. This inability to soar constitutes a tight constraint. Social constraints are less confining. Constructivists are correct, to some extent, in stressing human creativity and construction: Social structure is, in part, a product of human agency. The strategic-choice theorist has no problem explaining and modeling the strategic constructions of human beings. The norms and social artifacts that can be modeled as constraining choice can also be explained as the products of choice. Post hoc, there is little that cannot be included in the strategic-choice enterprise.

In using the simple building blocks of choice and payoff, models of strategic choice decontextualize calculation, relationship, and the international system. Everything about the respective actors and their relationship must be contained in the structure of the game (who goes when, what choices they have), their payoffs, their beliefs, and the information conditions. Power asymmetries between the actors, for example, are either irrelevant or their impact must be contained in one of these elements (the powerful may have more options than the weak, they may have better information, they may have higher payoffs, etc.). But power is not a direct factor in these models; it operates through the components of strategic choice delineated in chapter 1 of this volume.⁷⁶

The same is true of relationships. Some states see one another as enemies; others see one another as allies. Does the existence of a prisoner's dilemma imply the same things in both cases? Because the alliance problem is a prisoner's dilemma (Snyder 1984) and the armament choice between enemies is also a prisoner's dilemma, does that mean there is no difference between alliances and rivalries in international politics?⁷⁷

Here, too, the strategic-choice theorist can make accommodations. That game-theoretic modeling effectively treats actors on a par, in much the way the notion of sovereignty does, does not mean that prisoner's dilemmas between asymmetric powers cannot be modeled differently

⁷⁵ This is discussed in Stein 1984, 1990, and picked up by Ruggie 1992.

⁷⁶ The five elements of a strategic situation are "i) the collection of players; ii) the physical order in which play proceeds; iii) the choices available whenever it is a player's turn to move; iv) the information about previous choices made by others available to a player whose turn it is to move; and v) the payoffs to each of the players resulting from any play of the game" (Reny 1992, 103). As Philip Reny (*ibid.*) points out, this is so flexible that, "remarkably, one is hard-pressed to uncover a real-life strategic situation which cannot be usefully modelled [*sic*] by a carefully chosen extensive-form game."

⁷⁷ See the discussion in Stein 1990, chap. 6.

than those between states of relatively equal power, or that such games between allies cannot be modeled differently than those between rivals.

Strategic choice can be applied to actors and interactions at any level of analysis, but, in so doing, it decontextualizes the role of domain. The distinction between anarchy in international politics and hierarchy in domestic politics disappears. This generates important insights—that there are elements of anarchy in domestic politics and elements of hierarchy in international politics and that there exist wide swatches of international cooperation and domestic conflict. But it also misses something critical, namely, that the difference between legitimate authority in domestic politics and its absence in international politics is ignored in the blithe comparison of international conflict to pretrial bargaining within societies (cf., chapter 1 of this volume).

Formal theorists not only impute actors' preferences and choices, they also assume their bases for calculation and assessment. Although a smattering of studies employ alternative decision criteria, the application of a strategic-choice approach has typically treated expected utility as the basis for choice. But there are bases of calculation and assessment other than expected utility (not to mention the different ways expected utility has itself been formalized; see Schoemaker 1982).⁷⁸ The neorealist emphasis on a preeminent concern with survival can be read to imply a lexicographic utility function in which states act to maximize their chances of survival without engaging in any trade-offs among other interests (Stein 1990). This comports with the recognition that a variety of decision criteria are available for actors making purposive calculated choices.⁷⁹ Again, these can be incorporated in the structure of models of strategic choice.

Models of strategic choice are enormously flexible, and scholars can invariably construct post hoc models with an equilibrium outcome that matches the behavior actually chosen. Yet, except where the situation is quite constraining, the models can only be constructed post hoc. Even then, they do not typically generate one unique equilibrium and also face equally plausible contending models (ones that generate an alternative set of equilibria that include the observed outcome).

We have now come full circle. An approach created to improve human decision making has created an intellectual enterprise that falls short both

⁷⁸ A variety of paradoxes reflect different ways of arriving at a rational choice. More specifically, backward induction as the basis of rational choice generates certain paradoxes (Selten 1978; Binmore 1987; Pettit and Sugden 1989; Basu 1994; and, for their linkage with issues of common knowledge, Reny 1992).

⁷⁹ Seemingly perverse decisions by calculating actors are not new to political life. Radicals who oppose reform and support reactionaries in the belief that this will bring revolution sooner provide an intriguing example of the contorted calculations that underlie political decisions.

as normative recommendation and as positive theory, and for the same reasons. The limitations of human cognition in conjunction with the inherent constraints of uncertainty in a strategic setting result in models that, while still excessively simple, depend on strong common knowledge assumptions and decisive situational constraints and nevertheless result in multiple equilibria. Our knowledge of strategic reasoning gives us models that neither prescribe a unique strategy nor identify a complete explanation for any choice (even retrospectively, except in the simplest cases).⁸⁰

The strategic-choice approach discussed in this volume has generic problems and limitations with international-relations manifestations.⁸¹ Although the papers here demonstrate the utility of the approach, they also implicitly demonstrate that it does not resolve some core issues plaguing the study of international politics. At the end of the day, despite greater rigor, tighter conceptualization, and more precise specification, unresolved questions remain.

ALL BETS ARE OFF: WHITHER STRATEGIC CHOICE

This volume is written with a note of triumphalism.⁸² The use of strategic choice is growing. The lag in importing an idea from an economics article into political science grows ever shorter. Every new development is assessed for its potential application in political science.

Yet, this is not the first coming of strategic choice.⁸³ Game theory made its way into international relations in the late 1950s, was accepted and

⁸⁰ The limitation of formal models beyond some level of complexity can be seen as related to Godel's incompleteness theorem about mathematics more generally: that beyond some level of complexity, all logical systems are incomplete in that they contain propositions that cannot be proven true or false within the rules of that system.

⁸¹ Sutton (1990, 507) argues that game-theoretic models in industrial economics, the very area transformed by game theory in the 1980s and which has been the inspiration for international-relations scholars are similarly beset by problems of indeterminacy. He wonders "whether the old taunt is true that 'with oligopoly anything can happen,'" and continues to ask "in 'explaining' everything, have we explained nothing?" To the extent that international relations, too, is a domain dominated by a small set of great powers engaged in strategic behavior, then it may also be a domain in which many outcomes are possible and in which many equally viable models can be developed. Sutton's description of game-theoretic models of industrial economics, that the "richness of possible formulations leads to an often embarrassingly wide range of outcomes supportable as equilibria within some 'reasonable' specification," could as easily apply to international relations.

⁸² The combination of triumphalism and proselytizing exhortation suggests either a religious movement or a Ponzi scheme.

⁸³ It should be noted that the initial manifesto that organized this volume was written as if Schelling had had no impact in the late 1950s.

absorbed, and the field moved on and elsewhere. Two explanations can be proffered for the eclipse of game theory in the 1960s.⁸⁴ First, the formal tools seemed to have been fully developed. The apparent end of technical development, in conjunction with the incorporation of the extant body of findings, led international-relations scholars to other pursuits. More directly, the deterrence problem on which early game theory focused had been worked through, there seemed little more that could be said about deterrence in formal terms, and no other interesting applications for the method appeared to exist. Second, the assumptions that underlay the formal work on deterrence could be directly assessed empirically. Thus one response to game-theoretic studies of deterrence was to undertake empirical assessments of the nature of decision making. Scholars interested in deterrence questioned the assumption of unitary actors (directly leading to the development of the bureaucratic politics literature) and the presumed rationality of decision making during crises.⁸⁵ Empirical work on decision making in the late 1960s and early 1970s attempted to assess directly some of the assumptions of formal deterrence theory.

Formal game-theoretic work returned to the forefront of international relations in the late 1970s and early 1980s. It came after the quantitative work of the mid- and late 1960s had either been discredited or was no longer generative, when empirical work seemed excessively inductive and too little informed by theory, and when extant theorizing was flighty and undisciplined. By then, it also reflected new developments in game theory: the use of simulation (Axelrod 1984) and games with sequential moves (Selten 1975, 1978, on perfect equilibrium; Kreps and Wilson 1982, on sequential equilibrium).⁸⁶

Two elements seem to drive developments in the evolution of international relations theory. One is the generative force of a research tradition. Approaches with numerous followers tend to have research agendas that drive intellectual effort. Eventually, though, these run out of steam—research avenues get fully spun out or reach the limits of the technologically

⁸⁴ A striking indication of this is the relative attention paid to game theory in two special issues of *World Politics* in the 1960s. As noted above, the 1961 volume paid considerable attention to game theory. The 1969 issue paid scant attention to that subject, however, concentrating largely on the wave of quantitative empirical work that had become the focus of the field's attention in the mid-1960s.

⁸⁵ For the evolution of work on deterrence, see Jervis 1979. As a historical note, the direction of Graham Allison's work was affected by Andrew Marshall, director of research for the defense department and a key figure at the RAND Corporation during the mid-1950s, when game theory and economic modeling were applied to strategic questions.

⁸⁶ Snyder and Diesing's 1977 publication was also an important event, although their work reflected the interest in decision making and empirical work that was a hallmark of the 1970s.

tractable and feasible. Absent new technical developments, scholarly interests and efforts shift elsewhere. Something like this happened to content analysis, simulation, and events data modeling, among other approaches. A similar occurrence seems evident more recently in the shift toward experimentation and evolutionary game theory among economists and some political scientists.⁸⁷

The other important force changing the focus of scholarship in international politics is the real world. World War II and the events of the late 1930s dealt a death blow to the field's attention to international law. Changes in the perceived efficacy and utility of the United Nations have regularly fueled and dampened interest in international organizations. The perceived successes and failures of European integration have affected the study of regional integration more broadly. The Cuban missile crisis generated the literature on crises and crisis decision making and crisis management. The collapse of Bretton Woods and the oil crisis brought the return of international political economy. The rhetorical stridency, increased defense budgets, and collapse of détente in the early Reagan years brought a rebirth of security studies, a subfield that had atrophied in the mid- and late 1970s.⁸⁸ The end of the cold war is already having profound effects on the research interests of those in international relations. Whether the widespread use of game theory survives this shift will depend on the ability of strategic-choice theorists to address substantive issues of current concern.⁸⁹ Given that today's game theorists recognize the importance of behavioral approaches, and given the current state of formal theorizing, another cycle is not out of the question.⁹⁰

Yet, there remain both new methodological developments not fully tapped and numerous potential agendas for the continued application of strategic choice in international relations. Many components of strategic interaction deemed important by common sense have been profitably illuminated by technical developments in game theory. Reputation and signaling are two that have been studied formally and applied to interna-

⁸⁷ The economists' turn toward evolutionary game theory can be seen in Binmore and Samuelson 1994 and Van Damme 1994. For a characterization of experimentation as behavioral game theory, see Camerer 1997. Also see the adaptive-learning approach discussed in Honkapohja 1993.

⁸⁸ Jervis's (1978) work on military doctrine helped drive the appearance of this new intellectual agenda, as did Robert Powell's revisiting of Schelling with the new developments in game theory.

⁸⁹ For example, will the formal work of Fearon, Cetinyan, and others, illuminate more about the impact of ethnicity in international politics than alternative approaches?

⁹⁰ The phenomenon of cycling between formal and behavioral work can be seen in the history of economics as well (Seligman 1971; Latsis 1972). Formal theorizing generates great insights, but pressures build both for empirical assessments of assumptions and for information about what formalists ignore.

tional politics. But even here, much work can yet be done. Concepts such as resolution, conviction, stubbornness, stamina, and probity are all elements of reputation that cannot yet be addressed by game theory (Shubik 1993, 220). Some of the continued vitality of the approach in international relations depends on such developments.⁹¹

What Alternative?

Explaining behavior by reference to purpose is a standard practice in the social sciences. Not surprisingly, therefore, many debates are within the approach of purposive explanation. Indeed, the formal tools discussed here are constructions of purposive human beings who wanted to improve the rationality of their decisions. Is there, then, any alternative outside, and different from, a purposive approach?

It is not clear what the alternative to purposive explanation is and whether any of the suggested possibilities lie outside purposive explanation. Harsanyi (1969) contrasts rational choice with functionalist and conformist explanation but then argues that rational choice provides an explanation for institutions and social values. Elster (1986) points to structuralism and social norms as alternative possibilities, although he finds the former implausible and the latter incorrect. Then, too, his potential alternatives have also been the focus of rational-choice explanation.⁹²

Even the constructivist or postmodern alternative, which I cannot even sketch here, is hardly at odds with the rational-choice approach. It merely takes the primitives of strategic choice, the nature of the system and of preferences, as malleable and in need of explanation. But strategic choice can address these matters as well. The common law and norms, along with such other factors as private property rights and institutions, can be seen as constraining and framing choice. But they can also be studied as the products of human agency and choice.⁹³

Further, parallels may be drawn between constructivism and postmodernism, on the one hand, and game theory and economics, on the other:

⁹¹ The injunction "grow or stagnate" may not apply to political entities, but it does apply to intellectual enterprises.

⁹² One can easily generate models of choice in which actors are, in effect, strategic clummies without choice (the view of structural models). One can also focus on the choices underlying social norms. Plott (1987) simply says there is no alternative to rational choice.

⁹³ A rational-choice basis can be established for tradition and norms, which are the alternatives often sketched in sociological thought. For discussions of rationality by some of the classic sociologists of the twentieth century, see Cohen 1976 and Lane 1974. For one assessment of when norms do or do not undercut rational choice, see Mortimore 1976. For my discussion, see Stein 1996.

Both emphasize doubt and uncertainty, and the open-endedness emphasized by postmodernism is also found in game theory (Heap 1993). Ironically, constructivists characterize a strategic-choice approach, despite its emphasis on choice and its normative roots, as structural and contrast it with one that emphasizes agency. Yet, a strategic-choice approach can elucidate both the situations in which actors have an interest in constructing political alternatives and the requisites for institutional design and construction in the modern world. One can even argue that there is a new-age quality to modern game theory: Discussions of the equilibria attainable in incomplete information games under some sets of assumptions and expectations come close to the argument that peace will come when we all think peace. In short, strategic choice emphasizes agency and can deal with social constructions such as norms and institutions.

CONCLUSION

Strategic choice is more than a language and a set of tools, but it is far less than a theory. It is more than a language because it has content and entails substantive assumptions. As a result, it is richer than calculus (a branch of mathematics specifically developed to study planetary motion but applicable to many phenomena that share nothing of a substantive nature). Yet, as an approach, strategic choice remains a largely empty vessel—not shapeless but in need of content if it is to have something to contribute. It is therefore specific applications of strategic choice, rather than strategic choice *per se*, that often draw criticism. The continued success of strategic choice lies in its indigenous development in confronting concerns specific to international politics. Yet, we know that complexifying the model to make it more isomorphic with reality typically makes it indeterminate and incomplete as an explanation. So our response, if asked if we should believe in game theory and strategic choice, should be that of the Zen master when asked if he believed in God. Sounding as if caught in a strategic game himself, he answered, “If you do, I don’t; and if you don’t, I do.”⁹⁴ And our response to the question of whether a strategic-choice approach should constitute the foundation of explanation in international politics is “yes, one cannot do without it; and no, it is simply not enough by itself.”

⁹⁴ Related by Abraham Kaplan (1967) to explain that he did not lack faith in models of rationality but wanted to say “however” to the faithful.