

# Rigor or Rigor Mortis? Stephen M. Walt

## Rational Choice and Security Studies

The past decade has witnessed a growing controversy over the status of formal approaches in political science, and especially the growing prominence of formal rational choice theory. Rational choice models have been an accepted part of the academic study of politics since the 1950s, but their popularity has grown significantly in recent years.<sup>1</sup> Elite academic departments are now expected to include game theorists and other formal modelers in order to be regarded as “up to date,” graduate students increasingly view the use of formal rational choice models as a prerequisite for professional advancement, and research employing rational choice methods is becoming more widespread throughout the discipline.<sup>2</sup>

Is the increased prominence of formal rational choice theory necessary, inevitable, and desirable? Advocates of formal rational choice approaches assert that these techniques are inherently more scientific than other analytic

---

*Stephen M. Walt is Professor of Political Science and Master of the Social Science Collegiate Division at the University of Chicago. He will join the faculty of the John F. Kennedy School of Government at Harvard University in July 1999.*

---

I thank the following individuals for their comments on earlier drafts of this article: Graham Allison, Robert Art, Michael Desch, George Downs, Erik Gartzke, Charles Glaser, Joseph Grieco, Robert Jervis, John Mearsheimer, Barry Posen, Jack Snyder, Stephen Van Evera, and the participants at seminars at Princeton, Rutgers, Duke, and the John M. Olin Institute for Strategic Studies at Harvard. Barry O'Neill provided an exceptionally detailed and useful set of criticisms, for which I am especially grateful, and Ann Ducharme, David Edelstein, and Seth Jones were able research assistants and made valuable suggestions as well.

---

1. Seminal early applications of rational choice theory include Anthony Downs, *An Economic Theory of Democracy* (New York: Harper and Row, 1957); Gordon Tullock and James Buchanan, eds., *The Calculus of Consent* (Ann Arbor: University of Michigan Press, 1962); Thomas C. Schelling, *The Strategy of Conflict* (Cambridge, Mass.: Harvard University Press, 1960); and Mancur Olson, *The Logic of Collective Action: Public Goods and the Theory of Groups* (Cambridge, Mass.: Harvard University Press, 1965). Surveys of the basic literature include Dennis Mueller, ed., *Public Choice II* (Cambridge, U.K.: Cambridge University Press, 1989); James Alt and Kenneth Shepsle, eds., *Perspectives on Positive Political Economy* (Cambridge, U.K.: Cambridge University Press, 1990); and Peter C. Ordeshook, ed., *Models of Strategic Choice in Politics* (Ann Arbor: University of Michigan Press, 1989). An excellent introductory textbook is James D. Morrow, *Game Theory for Political Scientists* (Princeton, N.J.: Princeton University Press, 1995).

2. According to one estimate, rational choice scholarship now comprises nearly 40 percent of the published articles in the *American Political Science Review*, and another scholar reports that 22 percent of the *APSR* articles published between 1980 and 1993 were rational choice in orientation. Similarly, the annual report of the *APSR*'s editor suggests that 15–20 percent of all *ASPR* submissions and published articles were rational choice in orientation. See Donald Green and Ian Shapiro,

approaches, and argue that the use of more sophisticated models has produced major theoretical advances.<sup>3</sup> They are also prone to portray skeptics as methodological Luddites whose opposition rests largely on ignorance. Thus Robert Bates draws a distinction between “social scientists” and “area specialists” (a distinction that implies the latter are not scientific), and suggests that the discipline is finally “becoming equipped to handle area knowledge in rigorous ways.” According to Bates, these “rigorous ways” are rational choice models, and he chastises area experts for raising “principled objection to innovations . . . while lacking the training fully to understand them.”<sup>4</sup>

Not surprisingly, other scholars have greeted such claims with considerable skepticism, and argue that rational choice theory has yet to produce a substantial number of important new hypotheses or well-verified empirical predictions.<sup>5</sup> Indeed, some critics of rational choice methods question whether formal techniques are of any value whatsoever, and regard the modeling community as a group of narrow-minded imperialists seeking to impose its preferred method on the entire discipline.<sup>6</sup>

---

*Pathologies of Rational Choice Theory: A Critique of Applications in Political Science* (New Haven, Conn.: Yale University Press, 1995), p. 2; Norman Schofield, “Rational Choice and Political Economy,” *Critical Review*, Vol. 9, Nos. 1–2 (Winter–Spring 1995), p. 210, n. 10–12; and Ada W. Finifter, “Report of the Editor of the *American Political Science Review*, 1996–97,” *PS: Political Science and Politics*, Vol. 30, No. 4 (December 1997), pp. 783–791.

3. For example, the late William Riker once argued that social science laws “must be encased in a deductive theory,” and suggested that rational choice models were the basis for the only successful social science theories. Similarly, Bruce Bueno de Mesquita argues that “formal, explicit theorizing takes intellectual precedence over empiricism,” and Peter Ordeshook once claimed that “understanding politics requires sophisticated tools of deduction. . . . If mathematics is a necessary part of that analysis, then such mathematics is necessarily a part of political theory.” See Riker, “Political Science and Rational Choice,” in Alt and Shepsle, *Perspectives on Positive Political Economy*, pp. 168, 175–177; Bueno de Mesquita, “Toward a Scientific Understanding of International Conflict: A Personal View,” *International Studies Quarterly*, Vol. 29, No. 2 (June 1985), pp. 121–136; and Peter C. Ordeshook, “Introduction,” in Ordeshook, *Models of Strategic Choice in Politics*, p. 2.

4. See Robert Bates, “Letter from the President: Area Studies and the Discipline,” *APSA-CP: Newsletter of the APSA Organized Section on Comparative Politics*, Vol. 7, No. 1 (Winter 1996), pp. 1–2; Bates, “Area Studies and the Discipline: A Useful Controversy,” *PS: Political Science and Politics*, Vol. 30, No. 2 (June 1997), pp. 166–169; and Bernard Grofman, “The Gentle Art of Rational Choice Bashing,” in Grofman, ed., *Information, Participation, and Choice* (Ann Arbor: University of Michigan Press, 1993).

5. See Green and Shapiro, *Pathologies of Rational Choice Theory*; Raymond Wolfinger, “The Rational Citizen Faces Election Day, or What Rational Choice Theories Don’t Tell You about American Elections,” in M. Kent Jennings and Thomas E. Mann, eds., *Elections at Home and Abroad: Essays in Honor of Warren E. Miller* (Ann Arbor: University of Michigan Press, 1993); and Lars Udehn, *The Limits of Public Choice: A Sociological Critique of the Economic Theory of Politics* (New York: Routledge, 1996).

6. See, for example, Chalmers Johnson and E.B. Keehn, “A Disaster in the Making: Rational Choice and Asian Studies,” *National Interest*, No. 36 (Summer 1994), pp. 14–22; and Johnson, “Preconcep-

The stakes in this dispute are considerable. Because technical proficiency is often used as a surrogate for professional competence—and even to define what constitutes “legitimate” scholarship in a particular field—the outcome of this debate will have a powerful impact on the basic nature of the social sciences and on the allocation of finite academic resources. To put it bluntly, if reliance on formal methods becomes the *sine qua non* of “scientific” inquiry, then scholars who do not use them will eventually be marginalized within their respective fields. Like most methodological debates, therefore, the struggle has been quite contentious.<sup>7</sup>

One unfortunate result of this polarization has been the stifling of genuine debate on these issues. Instead of debating and acknowledging the actual strengths and weaknesses of competing research traditions, scholars are increasingly reluctant to criticize one another openly for fear of being seen as intolerant. Because such a reputation can have chilling effects on one’s professional prospects (particularly for younger scholars), the result is a narrowing of intellectual exchange. This is antithetical to scientific progress, which is furthered by an unfettered clash of ideas.<sup>8</sup>

What is at stake goes beyond the evolution of particular academic departments or the career prospects of individual scholars, however. Much more important, the outcome of this debate will also guide the nature of scholarly discourse on important political topics and shape the intellectual capital of the scholarly community. Subfields that are dominated by rational choice theorists will inevitably emphasize certain types of work over others, will privilege certain questions at the expense of others, and will prize certain analytical talents rather than others.<sup>9</sup> Thus the debate over the role of formal rational choice theory will have a powerful effect on what we think we know about

---

tion vs. Observation, or the Contributions of Rational Choice Theory and Area Studies to Contemporary Political Science,” *PS: Political Science and Politics*, Vol. 30, No. 2 (June 1997), pp. 170–174.

7. Past examples include the *Methodenstreit* among economic theorists in Germany in the nineteenth century, the struggle between institutional and neoclassical economists before and after World War II, and the “great debate” between so-called behavioralists and traditionalists in American political science (including international relations) in the 1960s. See Klaus Knorr and James N. Rosenau, *Contending Approaches to International Politics* (Princeton, N.J.: Princeton University Press, 1967); and Yuval P. Yonay, *The Struggle over the Soul of Economics: Institutional and Neoclassical Economics in America between the Wars* (Princeton, N.J.: Princeton University Press, 1998).

8. An exception to this observation is the special issue of *Critical Review*, Vol. 9, Nos. 1–2 (Winter–Spring 1995), entitled “Rational Choice Theory and Politics.”

9. The same is true for fields that employ nonformal approaches, of course. The central point is simply that the content of a field of inquiry is inevitably shaped by the techniques and procedures that are used to study it.

politics, and thus on what the academic community will be able to contribute to the wider public debate on important social issues.

This article seeks to advance this debate by evaluating the contribution of recent formal work in the field of security studies. Formal rational choice theory has been part of security studies for several decades, but recent formal scholarship is quite different from the seminal early work of scholars like Thomas Schelling, Daniel Ellsberg, or Mancur Olson. “First-wave” theorists like Schelling used simple formal illustrations and did not place much emphasis on mathematical rigor. Indeed, Schelling explicitly warned against the tendency for social scientists “to treat the subject of strategy as though it were, or should be, solely a branch of mathematics.”<sup>10</sup> The “second wave” of formal theorizing has largely ignored Schelling’s warning and placed far more emphasis on formal proofs and mathematical derivations. The question, therefore, is whether this latest wave of formal theorizing has contributed significantly to our understanding of international security.

My argument is straightforward. The central aim of social science is to develop knowledge that is relevant to understanding important social problems. Among other things, this task requires theories that are precise, logically consistent, original, and empirically valid. Formal techniques facilitate the construction of precise and deductively sound arguments, but recent efforts in security studies have generated comparatively few new hypotheses and have for the most part not been tested in a careful and systematic way. The growing technical complexity of recent formal work has not been matched by a corresponding increase in insight, and as a result, recent formal work has relatively little to say about contemporary security issues.

Two caveats should be noted before proceeding. First, this article does not offer a comprehensive analysis of recent formal work in security studies. Space does not permit me to discuss every application of formal theory to a security studies topic; instead, I have concentrated on the work of a number of prominent figures in the rational choice field and on scholarship that has been regarded by members of that subfield as especially sophisticated. By focusing on some of the best and most widely cited work, therefore, my sample is if anything biased in favor of formal approaches.

Second, this article does not compare the relative merits of formal theory with other methodological approaches. Such a comparison would be extremely valuable, but a proper assessment would require far more space than is avail-

---

10. Schelling, *Strategy of Conflict*, pp. 10–11, n. 4.

able here. Moreover, even if one were able to show that a particular approach had been especially productive, that would hardly mean that alternative research techniques should be entirely discarded.

The remainder of this article is organized as follows. The first section summarizes the basic principles of rational choice theory and describes the literature under scrutiny in the rest of the article. The second section describes the essential aims of social science and discusses several criteria for judging a body of scholarship. I argue that three criteria are especially important: precision and consistency, originality, and empirical validity. The third, fourth, and fifth sections apply these criteria to recent formal work in security studies, focusing on a number of especially important or prominent examples. The last section summarizes my assessment and offers some concluding remarks about the place of formal theory in this field.

### *What Is Formal Theory?*

Formal rational choice theory is defined “more by the *method* of theory construction than by the *content* of its theories.”<sup>11</sup> It refers to the use of mathematical models to derive propositions from a set of basic premises. The use of mathematics helps ensure logical consistency among the propositions, especially when dealing with complex relationships where the use of ordinary language might lead to logical errors or vague predictions.

Formal rational choice theory is more than just the assumption of purposive behavior on the part of social actors (as in the familiar “rational actor” assumption). Similarly, it does not refer to any scholarship that uses simple game theory concepts like the “prisoner’s dilemma” or “mixed strategies” primarily for heuristic purposes or as an illustrative analogy. Strictly speaking, formal theory involves the construction of specific mathematical models intended to represent particular real-world situations and the use of mathematics to identify the specific solutions (“equilibria”) for the model(s).<sup>12</sup>

---

11. See David Lalman, Joe Oppenheimer, and Piotr Swistak, “Formal Rational Choice Theory: A Cumulative Science of Politics,” in Ada W. Finifter, ed., *Political Science: State of the Discipline II* (Washington, D.C.: American Political Science Association, 1993), p. 78 (emphasis in original).

12. Barry O’Neill makes a useful distinction between (1) “proto-game theory” (where formal ideas provide a convenient vocabulary or offer useful analogies), (2) “low game theory” (where solutions to specific games are used to analyze particular social interactions), and (3) “high game theory” (where scholars construct general proofs for whole classes of games). See O’Neill, “Game Theory and the Study of Deterrence in War,” in Paul C. Stern, Robert Axelrod, Robert Jervis, and Roy Radner, eds., *Perspectives on Deterrence* (London: Oxford University Press, 1989), p. 135. Prominent

In security studies, formal rational choice theory usually means the use of game theory. Game theory is a set of techniques for analyzing individual decisions, in situations where each player's payoff depends in part on what the other players are expected to do. Game theory thus differs from decision-theoretic approaches, which analyze individual utility maximization against an exogenous, noncalculating environment. Because security studies generally focuses on situations where actors frequently try to anticipate what others will do, and where the outcome for each actor will be affected by the choices that others make, the attractiveness of game theory is not surprising.<sup>13</sup>

Formal rational choice theorists do not agree on everything, of course, and there are important epistemological and methodological differences within the modeling community. Nonetheless, most applications in the field of international relations or security studies employ the following basic assumptions and techniques.

1. Rational choice theory is *individualistic*: social and political outcomes are viewed as the collective product of *individual* choices (or as the product of choices made by unitary actors).

2. Rational choice theory assumes that each actor seeks to maximize its "subjective expected utility." Given a particular set of preferences and a fixed array of possible choices, actors will select the outcome that brings the greatest expected benefits.

3. The specification of actors' preferences is subject to certain constraints: (a) an actor's preferences must be complete (meaning we can rank order their preference for different outcomes); and (b) preferences must be transitive (if *A* is preferred to *B* and *B* to *C*, then *A* is preferred to *C*).<sup>14</sup>

4. Constructing a formal theory requires the analyst to specify the *structure* of the game. This typically means identifying the set of players, the likelihoods

---

examples of proto-game theory include Robert Jervis, "Cooperation under the Security Dilemma," *World Politics*, Vol. 30, No. 2 (January 1978), pp. 167–214; and Robert O. Keohane, *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton, N.J.: Princeton University Press, 1984). This article focuses primarily on low game theory, which accurately describes the bulk of recent formal work in security studies.

13. I have not examined the extensive use of operations research and other decision-theoretic techniques in the analysis of military policy, largely because this work has had less impact in the social sciences. An encyclopedic survey is Barry O'Neill, "Game Theory Models of War and Peace," in Robert Aumann and Sergiu Hart, eds., *Handbook of Game Theory with Economic Applications*, Volume 2 (Amsterdam: Elsevier Science, 1994).

14. A more demanding condition is that the actors' utility functions are consistent with the Von Neumann-Morgenstern expected utility theorem. This theorem imposes additional constraints (such as the exclusion of infinite utilities), but is not necessary in simple contexts. See Morrow, *Game Theory for Political Scientists*, chap. 2.

of each player's pattern of preferences, each player's information at every choice point, and how they see their moves as connected to the possible outcomes.

5. Once the game is fully specified, the analyst usually looks for its *equilibrium*. An equilibrium is an assignment of strategies to the players, such that each player's strategy maximizes his or her expected utility, given that the others use their assigned strategies. Thus an equilibrium is a strategy from which a rational actor would have no incentive to deviate unilaterally.<sup>15</sup>

Within a formal rational choice model, therefore, an equilibrium is a prediction. If the game structure is an accurate representation of the phenomenon in question, and if there are no mathematical mistakes, the equilibria of the game identify the only outcomes that are logically possible. These equilibria form the basis for any subsequent empirical testing.<sup>16</sup>

As noted above, there are important differences among formal theorists regarding the epistemological status of these models. For example, formal theorists are divided between those who endorse a "thin" conception of rationality (which assumes only that the actors choose rationally to achieve whatever goals they may have) and those who rely on stronger assumptions ("thick rationality") about each actor's preferences. In the latter case, the analyst assumes that preferences are consistently ordered and also specifies what those preferences are (e.g., that the actors seek to maximize power, or wealth, or whatever). There are also disputes over whether rational choice theories must merely be consistent with the observed outcome, or whether they must also be consistent with the actual process by which decisions are made.<sup>17</sup> This

---

15. More precisely, this is the definition of a "Nash equilibrium," first established by John Nash. See *ibid.*, pp. 91–98. For a formal discussion, see David M. Kreps, "Nash Equilibrium," in John Eatwell, Murray Milgate, and Peter Newman, eds., *The New Palgrave: A Dictionary of Economics: Game Theory* (New York: W.W. Norton, 1989), pp. 167–177.

16. As William Riker puts it, "Equilibria are valuable, indeed essential in theory in social science because they are identified consequences of decisions that are necessary and sufficient to bring them about. An explanation is . . . the assurance that an outcome must be the way it is because of antecedent conditions. This is precisely what an equilibrium provides." See Riker, "Political Science and Rational Choice," p. 175.

17. The question is whether the actors must *consciously* select the course of action that will maximize their expected utility through a process of reasoning that is at least roughly consistent with the logic of the model. According to Jon Elster, a proper rational choice explanation requires that "the action must not only be rationalized by the desire and the belief, but it must also be caused by them and, moreover, caused 'in the right way.'" See his "Introduction," in Elster, ed., *Rational Choice* (New York: New York University Press, 1986), p. 16; and Terry Moe, "On the Scientific Status of Rational Choice Theory," *American Journal of Political Science*, Vol. 23, No. 1 (February 1979), pp. 215–243.

latter issue does not affect the construction of a formal model, but it is critical to any effort to test its implications.<sup>18</sup>

### *How to Judge Social Science Theories*

The fundamental aim of social science is to develop *useful knowledge* about human social behavior. Such knowledge may take the form of a deeper and more accurate understanding of the past, or the elaboration of a new theory that explains some important aspect of human conduct, or a largely descriptive account of a particular social group or event. Whatever its precise form, the essence of the enterprise is the discovery of powerful, well-founded claims about human behavior. Social science should not be merely an intellectual exercise undertaken for the benefit of its practitioners. Given that what we know (or think we know) about human nature and social institutions can have powerful effects on the fates of whole societies, social science should always strive to produce accurate and relevant knowledge about the human condition.

Given this basic objective, there are three main criteria for evaluating social science theories.

First, a theory should be *logically consistent* and *precise*. Other things being equal, theories that are stated precisely and that are internally consistent are preferable to theories that are vague or partly contradictory. An inconsistent theory is problematic because (some of) the conclusions or predictions may not follow logically from the initial premises. In this sense, an inconsistent theory creates a false picture of the world. Inconsistent theories are also more difficult to test because it is harder to know if the available evidence supports the theory. Similar problems arise when a theory is vague, because a wider range of empirical outcomes will be consistent with the theory as stated. Precision also means identifying underlying assumptions and boundary conditions, which helps us guard against applying the theory in circumstances for which it is not suited.

The second criterion is *degree of originality*. Although the level of originality can be difficult to measure and subject to dispute, it is still one of the most prized features of any scientific theory. We prefer creative and original theories because they tell us things that we did not already know and help us see familiar phenomena in a new way. A novel theory imposes order upon phe-

---

18. A good summary of these issues is found in Green and Shapiro, *Pathologies of Rational Choice Theory*.



nomena that were previously hard to understand, and solves conceptual or empirical puzzles that earlier theories could not adequately explain. Not surprisingly, therefore, both natural and social scientists place a premium on the creation of new ideas.<sup>19</sup>

The final criterion is *empirical validity*. The justification for this criterion should be obvious as well: the only way to determine if a theory is truly useful is to compare its predictions against an appropriate body of evidence. Theories may be tested either by examining the correlation between independent and dependent variables (i.e., do they covary in the manner predicted by the theory?) or by testing the causal logic directly through detailed process-tracing.<sup>20</sup>

When evaluating a particular research tradition, therefore, we want to know if its propositions receive adequate empirical support. Are efforts to test key propositions carefully done, and are the results consistent with the theory? Other things being equal, a research tradition that ignores or discounts this requirement is being too easy on itself.

These criteria provide a set of hurdles that any social science approach must try to overcome. Although all three are important, the latter two criteria—originality and empirical validity—are especially prized. A consistent, precise yet trivial argument is of less value than a bold new conjecture that helps us understand some important real-world problem, even if certain ambiguities and tensions remain. Similarly, a logically consistent but empirically false theory is of little value, whereas a roughly accurate but somewhat imprecise theory may be extremely useful even though it is still subject to further refinement.<sup>21</sup> We do not expect every article or book to receive a high score on

---

19. Famous examples of especially original and fruitful theories include Darwin's theory of natural selection, Newton's mechanics, and Einstein's theory of relativity. In the social sciences, one might point to Keynesian economic theory, collective goods theory, deterrence theory, the theory of bureaucratic politics, and the application of cognitive and social psychology to international conflict. Although some of these theories later fell from favor, each was properly regarded as a creative and potentially valuable conceptual vision, and each spawned a large and influential literature.

20. On these methods of theory testing, see Stephen Van Evera, *Guide to Methods for Students of Political Science* (Ithaca, N.Y.: Cornell University Press, 1997).

21. Thus Christopher Achen and Duncan Snidal argue that the first wave of rational deterrence theory was "astonishingly fecund, both for theory and for policy," and of "immense practical importance." Yet the theory was not developed through formal modeling and contains many features that Achen and Snidal judge to be "woefully underconceptualized." See Achen and Snidal, "Rational Deterrence Theory and Comparative Case Studies," *World Politics*, Vol. 41, No. 2 (January 1989), pp. 153, 159.

all three criteria, of course, but we have reason to be skeptical if a particular research tradition consistently slights one or more of them.

Let us now consider how well recent formal work meets these standards.

### *Logical Consistency and Precision*

This section examines whether formal rational choice methods contribute to the development of logical and precise theories. After describing the virtues of formalization, I explain why formalization is neither necessary nor sufficient for scientific progress and consider some of the costs that it imposes. I conclude that although rational choice models can help increase the precision of our theories, this contribution does not justify privileging them over other social science approaches.

#### WHY USE FORMAL MODELS?

In social science, the main virtue of formalization is its contribution to logical consistency and precision. According to James Morrow, the primary advantage of formal modeling is “the *rigor and precision of argument* that it requires.”<sup>22</sup> Not surprisingly, therefore, scholars who use these methods place a very high value on this criterion. Thus Bruce Bueno de Mesquita has declared that “logical consistency is a fundamental requirement of all [scientific] hypotheses,” and he further suggests that “our main problem [in the study of international conflict] is not a lack of facts . . . but a lack of rigorously derived hypotheses that can render our facts informative.”<sup>23</sup> Properly employed, the formal language of mathematics can impart greater precision to an argument, and helps guard against inconsistencies arising either from a failure to spell out the causal logic in detail or from the ambiguities of normal language.<sup>24</sup>

---

22. See Morrow, *Game Theory for Political Scientists*, p. 6 (emphasis added).

23. Bueno de Mesquita, “Toward a Scientific Understanding of International Conflict,” p. 128. See also Steven Brams and D. Marc Kilgour, *Game Theory and National Security* (New York: Basil Blackwell, 1988), p. 2. The mathematical economist Gerard Debreu offers a similar defense of abstract modeling, arguing that “economic theory has had to adhere to the rules of logical discourse and must renounce the facility of internal inconsistency. A deductive structure that tolerates a contradiction does so under the penalty of being useless, since any statement can be derived flawlessly and immediately from that contradiction.” See Debreu, “The Mathematization of Economic Theory,” *American Economic Review*, Vol. 81, No. 1 (March 1991), pp. 2–3. For a typically witty rebuttal of Debreu’s claims, see Deidre N. McCloskey, *The Vices of Economists—The Virtues of the Bourgeoisie* (Amsterdam: Amsterdam University Press, 1996), pp. 78–81.

24. Kenneth Arrow offered a similar assessment more than four decades ago, writing that “mathematics . . . is distinguished from the other languages habitually used by the social scientists chiefly by its superior clarity and consistency.” Arrow, “Mathematical Models in the Social Sciences,” in Daniel Lerner and Harold D. Lasswell, eds., *The Policy Sciences* (Stanford, Calif.: Stanford University

In this sense, expressing an existing theory in formal language provides one type of test: Do the predicted consequences follow logically from the stated premises? In this sense, formalization can give us greater confidence in theories that were originally stated in verbal form. Formalization can also make the assumptions that drive a conclusion more apparent, thereby spurring further investigation and discouraging any tendency to overgeneralize.

These virtues should not be dismissed lightly. In the field of security studies, for example, formal analysis has shown that certain widely accepted propositions were not strictly deducible from the standard premises, as in Robert Powell's formal demonstration that the state with higher resolve may not always prevail in a nuclear crisis.<sup>25</sup> Similarly, a formal model can suggest new ways to interpret a body of empirical data, as in James Fearon's use of a simple bargaining model to show how selection effects can alter how one interprets historical cases of extended deterrence.<sup>26</sup> Finally, formal theory can also provide the tools to analyze especially complex interactions, where one might not be able to work through the causal processes in purely verbal form. Thus, even if formal techniques made no other contribution, their capacity to verify and maintain logical consistency is an undeniable asset.

#### WHY FORMALIZATION IS NEITHER NECESSARY NOR SUFFICIENT FOR SCIENTIFIC PROGRESS

This endorsement of formalization needs to be qualified in several important respects, however. First, the use of formal techniques is not a prerequisite for logical consistency. Many nonformal works of natural and social science contain precise, logically consistent theories, which casts doubt on the claim that formal methods are a necessary part of science.<sup>27</sup> Darwin's theory of natural

---

Press, 1951), p. 129. See also David M. Kreps, *Game Theory and Economic Modelling* (Oxford, U.K.: Oxford University Press, 1990), pp. 6–7.

25. See Robert Powell, *Nuclear Deterrence Theory: The Search for Credibility* (Cambridge, U.K.: Cambridge University Press, 1990).

26. James D. Fearon, "Signaling versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model," *Journal of Conflict Resolution*, Vol. 38, No. 2 (June 1994), pp. 236–269. Strictly speaking, one does not need a formal model to recognize that selection effects should be taken into account when interpreting the success or failure of deterrence. For a nonformal discussion of the same point, see Jack Levy, "Quantitative Studies of Deterrence Success and Failure," in Stern et al., *Perspectives on Deterrence*, pp. 117–120.

27. Examples from the social sciences include Benjamin Cohen's restatements of the various strands of Marxist theory in Cohen, *The Question of Imperialism: The Political Economy of Dominance and Dependence* (New York: Basic Books, 1973); Andrew J.R. Mack, "Why Big Nations Lose Small Wars: The Politics of Asymmetric Conflict," *World Politics*, Vol. 27, No. 2 (January 1975), pp. 175–200; Stanislaw Andreski, *Military Organization and Society* (Berkeley: University of California Press, 1971); Ronald Rogowski, *Commerce and Coalitions: How Trade Affects Domestic Political Alignments*

selection, for example, was not formulated in axiomatic terms, and Darwin could not even specify the full causal mechanism by which favorable traits were passed on to successive generations. Yet the theory was clearly a stunning achievement.<sup>28</sup>

Second, complete logical consistency—and in particular, the ability to deduce testable propositions from a set of general assumptions—is neither necessary nor sufficient for scientific progress. Larry Laudan notes that “inconsistent theories have often been detected in almost all . . . branches of science,” and argues that efforts to resolve such inconsistencies often form an important part of specific research traditions.<sup>29</sup> Some evidence even suggests that working scientists routinely ignore the strict canons of logic in their daily work, and are more prone to inferential “errors” than are ordinary citizens.<sup>30</sup>

The social sciences are replete with inconsistent or incomplete but nonetheless highly useful theories. John Maynard Keynes’s *General Theory* contains a number of important gaps and inconsistencies (whose exploration dominated the macroeconomic research agenda for several decades), but it was nonetheless a major watershed in economic thought.<sup>31</sup> Similarly, Mancur Olson’s *The*

---

(Princeton, N.J.: Princeton University Press, 1989); John J. Mearsheimer, *Conventional Deterrence* (Ithaca, N.Y.: Cornell University Press, 1983); Robert D. Putnam with Robert Leonardi and Raffaella Y. Nanetti, *Making Democracy Work: Civic Traditions in Modern Italy* (Princeton, N.J.: Princeton University Press, 1993); Robert A. Pape, *Bombing to Win: Air Power and Coercion in War* (Ithaca, N.Y.: Cornell University Press, 1996); Norman H. Nie, Jane Junn, and Ken Stehlik-Berry, *Education and Democratic Citizenship in America* (Chicago: University of Chicago Press, 1997); and Chaim D. Kaufmann, “Out of the Lab and Into the Archives: A Method for Testing Psychological Explanations of Political Decision Making,” *International Studies Quarterly*, Vol. 38, No. 4 (December 1994), pp. 557–586.

28. Moreover, Darwinians remained deeply divided over a number of central features of natural selection, such as the inheritability of acquired characteristics. See Robert J. Richards, *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior* (Chicago: University of Chicago Press, 1987).

29. Laudan goes on to say that “the only conceivable response to a conceptual problem of this kind is to refuse to accept the offending theory until the inconsistency is corrected,” but he cautions that “the refusal to accept an inconsistent theory need not require that one cease working on such a theory.” See Laudan, *Progress and Its Problems: Towards a Theory of Scientific Growth* (Berkeley: University of California Press, 1977), pp. 49–50, 230 n. 5 (emphasis in original).

30. Commenting on a study of the psychological traits of scientists, the biologist and philosopher of science David L. Hull points out that “in deductive logic reasoning from ‘if *p* then *q*’ and ‘*q*’ to ‘*p*’ is fallacious. . . . When scientists and nonscientists were tested, scientists tended to commit this error more frequently than ordinary people. They also tended to reason quite rapidly from minimal data to possible explanations. . . . After all, scientists are involved primarily in nondemonstrative forms of inference, and by definition nondemonstrative inferences fail the canons of deductive logic. As fallacious as affirming the consequent may be in deductive logic, it is central to science.” See David L. Hull, *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science* (Chicago: University of Chicago Press, 1988), p. 301 (emphasis added).

31. See Marc Trachtenberg, “Keynes Triumphant: A Study of the Social History of Economic Ideas,” in Robert Alun Jones and Henrika Kuklick, eds., *Knowledge and Society: Studies in the Sociology of Culture Past and Present* (Greenwich, Conn.: JAI Press, 1983).

*Logic of Collective Action* contains a number of logical ambiguities, yet is properly regarded as a seminal contribution.<sup>32</sup> In international relations, Kenneth Waltz's *Theory of International Politics* argues that bipolar worlds are stable in part because the two leading powers would compete everywhere and thus reduce the danger of miscalculation. But Waltz also argues that peripheral areas are of little or no strategic consequence, which raises the question of why a rational superpower would compete there in the first place. Despite this contradiction, Waltz's theory has probably been the most influential work in the field over the past two decades, and deservedly so.

These examples suggest that although logical consistency is highly desirable and efforts to achieve it are a central aim of science, it is not the only thing we look for in a theory. Put differently, an incomplete but highly suggestive theory may be an important advance, even if it requires additional work to clarify its deductive logic and identify critical assumptions and boundary conditions.

Logical consistency is also insufficient when a theory's core assumptions are subject to question. Formal rational choice models derive logical conclusions from a set of initial premises about how human beings (or states) make decisions.<sup>33</sup> If human decisions in the real world are not made in the way that rational choice theorists assume, however, then the models may be both deductively consistent and empirically wrong.<sup>34</sup>

For example, many formal models relax the assumption of full information by making additional assumptions about the way each player will revise his or her beliefs.<sup>35</sup> Such models typically assume that actors with incomplete

---

32. For example, Olson argued that small groups are more likely to provide a collective good than are larger groups, based on the claim that each member's share of the good will decline as group size increases, thereby decreasing the individual incentive to contribute. This argument assumes that the collective good is not in "joint supply" (meaning that consumption by one actor does not reduce the amount available to others), and ignores the possibility that certain collective goods (such as the ability to lobby legislators on behalf of some position) may require a large membership to be effective. See Russell Hardin, *Collective Action* (Baltimore, Md.: Johns Hopkins University Press, 1982), chap. 3.

33. A good critique of the utility assumptions common to rational choice theory is Amitai Etzioni, *The Moral Dimension: Toward a New Economics* (New York: Free Press, 1988). I am indebted to Robert Jervis for bringing this source to my attention.

34. Of course, such theories may do a better job of explaining and predicting than do rival theories. This is an empirical issue, however, which is one of the main reasons why it is necessary to subject rational choice models to careful empirical testing.

35. More sophisticated models also make heroic assumptions about the ability of actors to perform complex calculations to determine what course of action to take. This implicit assumption reveals a modest irony: formal modelers are admired because they are able to devise elaborate games and work out increasingly complicated solutions, yet the games themselves are supposed to describe the behavior of ordinary human beings (or collectivities) who have never had a course in game theory and may not even understand simple algebra.

information will revise their initial beliefs according to Bayes's rule, which states how probability estimates are optimally revised in light of new information.<sup>36</sup> Bayes's rule is a principle of probability theory, however, not an empirical law of human decisionmaking. Quite the contrary, in fact, for there is abundant experimental evidence confirming that human beings do not revise their beliefs in this manner.<sup>37</sup> This result means that a formal model in which actors revise their beliefs according to Bayes's rule can be logically consistent but empirically false, because the predictions it generates have been calculated with an empirically inaccurate algorithm.

Another reason why logical consistency is not enough is the well-known problem of "multiple equilibria." Over the past three decades, game theorists have devised ways to build more realistic models by relaxing certain key assumptions (such as the belief that the players have full information). Unfortunately, these more complicated games often contain several equilibrium solutions (i.e., solutions a rational actor would not depart from unilaterally), which means that logical deduction alone cannot tell you which outcome is going to occur.<sup>38</sup> This problem is compounded by the so-called folk theorem,

---

36. Bayes's rule states that the probability that a particular state of the world is true given the occurrence of a particular event is the probability that both the state and the event will occur, divided by the probability that the event will occur independent of the actual state of the world. Formally, if there are two possible states of the world ( $X$  and  $Y$ ), and an event  $A$ , Bayes's rule can be written as

$$p(X/A) = \frac{p(X)p(A/X)}{p(X)p(A/X) + p(Y)p(A/Y)}.$$

37. See Daniel Kahneman and Amos Tversky, "On the Psychology of Prediction," *Psychological Review*, Vol. 80, No. 4 (July 1973), pp. 237–251; David Grether, "Bayes Rule as a Descriptive Model: The Representativeness Heuristic," *Quarterly Journal of Economics*, Vol. 95, No. 3 (November 1980), pp. 537–557; and Maya Bar-Hillel, "Similarity and Probability," *Organizational Behavior and Human Performance*, Vol. 11, No. 2 (April 1974), pp. 277–282. For general discussions of the rationality assumption in economics, see Robin M. Hogarth and Melvin W. Reder, eds., *Rational Choice: The Contrast between Economics and Psychology* (Chicago: University of Chicago Press, 1987); and Karen Schweers Cook and Margaret Levi, eds., *The Limits of Rationality* (Chicago: University of Chicago Press, 1990). A critical assessment of the use of Bayesian assumptions in game theory is Ken Binmore, "DeBayesing Game Theory," in Binmore, Alan Kirman, and Piero Tani, eds., *Frontiers of Game Theory* (Cambridge, Mass.: MIT Press, 1993).

38. For example, Harrison Wagner's 1992 article on rationality and misperception presents a two-stage game in which a potential challenger is uncertain about the willingness of the defender to retaliate. Wagner shows that there is an equilibrium in which a strong defender always retaliates in the first stage (so as to deter a challenge in stage 2), while a weak deterrer retaliates only occasionally, and the challenger is more likely to be deterred in stage 2. But as Barry O'Neill has pointed out, the model contains another equilibrium in which neither strong nor weak deterrers retaliate in stage 1, and the challenger always challenges in the second. See R. Harrison Wagner, "Rationality and Misperceptions in Deterrence Theory," *Journal of Theoretical Politics*, Vol. 4, No. 2

which says that in repeated games with incomplete information and an appropriate discount for the future payoffs, there are always multiple Nash equilibria.<sup>39</sup> Although it is sometimes possible to identify which equilibria will be preferred—Schelling’s famous discussion of “focal points” was an important effort in this area—“formal mathematical game theory has said little or nothing about where these expectations come from.”<sup>40</sup>

Within the field of game theory, the main response to the problem of multiple equilibria was to develop more restrictive “solution concepts.”<sup>41</sup> A solution concept (such as “Nash equilibrium,” “subgame perfect equilibrium,” or “perfect Bayesian equilibrium”) is a set of restrictions on what a “rational” actor would do. In a formal model of bargaining, for example, a more refined solution concept will eliminate certain equilibria by forbidding players from making logically permissible but otherwise incredible threats, or by placing certain limits on the inferences that players may draw from one another’s behavior. By formally restricting certain choices, more refined solution concepts eliminate some of the equilibria and thereby permit more determinate predictions. The problem is that the empirical predictions one draws may depend on the particular solution concept that is employed.<sup>42</sup> Thus more

---

(April 1992), pp. 115–142; and Barry O’Neill, “Are Game Models of Deterrence Biased towards Arms-Building?: Wagner on Rationality and Misperception,” *Journal of Theoretical Politics*, Vol. 4, No. 4 (October 1992), pp. 472–473.

39. See Drew Fudenberg and Eric Maskin, “The Folk Theorem in Repeated Games with Discounting or Incomplete Information,” *Econometrica*, Vol. 54, No. 3 (May 1986), pp. 533–554.

40. Kreps, *Game Theory and Economic Modelling*, p. 101. The inability of some formal models to identify clear predictions has led some formal theorists to invoke ideas, institutions, or culture as exogenous, ad hoc variables in order to explain which of the various equilibria is chosen. In these accounts, however, the exogenous variables are doing the real work of explaining the outcomes; the game itself is essentially a descriptive framework in which to embed the narrative. See, for example, David Kreps, “Game Theory and Corporate Culture,” in Alt and Shepsle, *Perspectives on Positive Political Economy*; Geoffrey Garrett and Barry R. Weingast, “Ideas, Interests, and Institutions: Constructing the European Community’s Internal Market,” in Judith Goldstein and Robert O. Keohane, eds., *Ideas and Foreign Policy: Beliefs, Institutions, and Political Change* (Ithaca, N.Y.: Cornell University Press, 1993); and Mark Blyth, “Any More Bright Ideas?: The Ideational Turn in Comparative Political Economy,” *Comparative Politics*, Vol. 29, No. 2 (January 1997), pp. 229–250.

41. The search for increasingly refined solution concepts may be going out of fashion among pure game theorists, in part because there is no good way to determine which concepts should be preferred. I am indebted to Michael Chwe for discussion on this point.

42. This point is nicely made in Barry Nalebuff’s 1991 article on reputation building. Like the Wagner article just discussed, Nalebuff’s model depicts situations where a state is contemplating intervention in one area to deter subsequent interventions elsewhere. The problem is that it is not clear how other actors will interpret different responses. Intervening could be seen as a sign of weakness (“I will intervene now because I cannot meet a later challenge and am therefore trying to bluff”), or it could be seen as a sign of strength (“I will intervene now to demonstrate how strong I am”). The model generates multiple equilibria—for example, there is an equilibrium in

realistic game models can be both logically consistent and indeterminate in the absence of subjective judgments about the particular equilibria the actors are going to prefer.

These criticisms of rational choice theory are not new, and sophisticated game theorists are fully aware of the limitations of the method as well as the strengths. Nor do these difficulties discredit the use of formal models, including such efforts in the field of security studies. But these concerns ought to sound a cautionary note. In game theory, as in life, one rarely gets something for nothing. One can relax the unrealistic assumption of full information, for example, but only at the cost of unrealistic assumptions about the way that actors update beliefs and the ability of real-world decisionmakers to perform complex calculations. Although logical consistency and precision are desirable and formal techniques can help us achieve them, this capacity does not ensure accurate or useful results by itself. By themselves, in short, the potential gains in precision and logical consistency do not demonstrate the superiority of formal techniques over other approaches.

#### THE COSTS OF FORMALIZATION

Moreover, the potential increase in precision and consistency is bought at a price. Unlike the first wave of formal theorizing, which relied on simple models largely for illustrative purposes, recent formal work has become less and less “user-friendly.” Some of the inaccessibility arises from the use of more sophisticated mathematics, but an equally serious barrier is the tendency for many formal theorists to present their ideas in an overly complex and impenetrable manner. In general, formal theorists rely heavily on a specialized jargon and what Donald McCloskey has termed a “scientistic” style, in which formal proofs, lemmas, and propositions are deployed to lend a quasi-scientific patina to otherwise simple ideas.<sup>43</sup> Formal methods also make it easier to bury key assumptions within the model, thereby forcing readers to invest considerable time and effort to unearth the basic logic of the argument.

---

which all states intervene whether they are strong or weak, and another where no states intervene—and Nalebuff goes on to show that the question of whether intervention is “rational” depends on the solution concept that is employed. See Nalebuff, “Rational Deterrence in an Imperfect World,” *World Politics*, Vol. 43, No. 3 (April 1991), p. 329.

43. See Donald N. McCloskey, *The Rhetoric of Economics* (Madison: University of Wisconsin Press, 1985); and McCloskey, *Knowledge and Persuasion in Economics* (Cambridge, U.K.: Cambridge University Press, 1994), especially chap. 9.



The obvious defense of increased formalization is simply that this is the price that must be paid for theoretical progress. We do not expect physicists to dumb down superstring theory; by the same logic, one should not expect formal theorists to simplify their work to help other scholars understand it. This argument has some merit. Sophisticated mathematical tools have been of considerable value throughout the social sciences, and one would never want to rule out such techniques a priori. But the increased emphasis on mathematization is not an unalloyed good, particularly if it does not yield a substantial increase in explanatory power.

First, other things being equal, a theory that is easy to grasp and understand is inherently easier to evaluate than one that is impenetrable or obscure. Accessibility increases the number of potential critics, thereby increasing the number of challenges that a theory is likely to face. Facilitating potential challenges contributes to rigor, because the larger the audience that can understand and evaluate a theory, the more likely it is that errors will be exposed and corrected and the better a theory has to be in order to retain approval. By contrast, an incorrect theory that is presented in an opaque or impenetrable way may survive simply because potential critics cannot figure out what the argument is.<sup>44</sup>

Indeed, when a research tradition prizes mathematical rigor above all else, incorrect or trivial ideas may survive because they are presented in a technically impressive way. As Thomas Mayer has written, "With modeling held in such high regard, there is the danger that a trivial idea, if it is accompanied by a large enough bodyguard of equations, will succeed in surmounting the refereeing process. Many published models merely 'algebraize' the obvious." Even scholars who have mastered the requisite techniques may be forced "to plough through an elaborate set of equations to get at what could have been said much more briefly."<sup>45</sup>

Second, the time invested learning formal techniques is time that cannot be spent learning a foreign language, mastering the relevant details of an important policy issue, immersing oneself in a new body of theoretical literature, or compiling an accurate body of historical data. Similarly, the time required to

---

44. This principle is not limited to scholarship using mathematics or other technical tools. If qualitative work is written in obscure and inaccessible jargon, or it is based on source materials that are not available to other scholars, this will inhibit critical evaluation and make it easier for dubious work to evade challenge.

45. See Thomas Mayer, *Truth versus Precision in Economics* (Aldershot, U.K.: Edward Elgar, 1993), pp. 123–127.

understand an elaborate formal demonstration (or the time spent perfecting the mathematical details of one's own models) is time that cannot be spent questioning underlying assumptions or testing the empirical validity of the argument. My point is not that these other skills are more valuable than the use of formal techniques, merely that there are opportunity costs involved in relying on any one particular analytic approach.

Finally, a logically consistent and mathematically rigorous theory is of little value if it does not illuminate some important aspect of the real world. As the economist George Stigler (who was hardly opposed to rational choice theory) once commented: "At leading centers of economic theory . . . it has been the practice to ask: Is the new theory logically correct? That is a good question but not as good as the second question: Does the new theory help us to understand observable economic life? . . . Until the second question is answered, a theory has no standing and therefore should not be used as a guide to policy."<sup>46</sup> And even if a formal theory does contribute to scholarly understanding, a forbidding level of technical complexity will make it more difficult for policymakers to use, thereby reducing its practical value.

Once again, these arguments do not imply that formal modeling is not a useful part of the social science toolkit. Rather, they suggest that this research tradition has both strengths and limitations; it imposes costs as well as conferring benefits. The technical complexity of recent formal work might be justified if these techniques led to lots of useful new hypotheses, and if these hypotheses were well supported by careful empirical tests. In other words, if formalization was more likely than other approaches to produce important policy-relevant knowledge, then we might (rationally) disregard these costs. As I shall now show, however, this does not seem to be the case.

### *Creativity and Originality*

Despite the confident claims of some of its practitioners, recent rational choice work in security studies has not produced a noteworthy number of important new theories or hypotheses. Formal rational choice theorists have refined or qualified a number of existing ideas, and they have provided formal treatments of a number of familiar issues. When compared to other research traditions, however, their production of powerful new theories is not very impressive.

---

46. Quoted in *ibid.*, pp. 27–28.

The lack of originality takes two closely related forms. The first form I term “methodological overkill”; the second might be called the problem of “old wine in new bottles.” Let us consider each phenomenon in turn, along with some prominent illustrative examples.

#### METHODOLOGICAL OVERKILL

Methodological overkill refers to the tendency of some elaborate formal models to yield rather trivial theoretical results. Here the problem is not that the arguments are incorrect; rather, the problem is that the elaborate formal machinery does not produce very interesting findings.

EXAMPLE NO. 1. James D. Morrow, “Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining.”<sup>47</sup> This article presents a complex model of crisis bargaining, which assumes that states are uncertain about the balance of power, specific military advantages, and the opponent’s resolve. The model is a fairly realistic depiction of some of the factors that influence crisis bargaining, and Morrow also offers some informal tests of the model’s predictions.

Unfortunately, Morrow’s sophisticated model yields rather trivial results. The central finding is that crises and wars do not arise in the absence of some form of uncertainty, a proposition that has been advanced by a number of other scholars in the past. Morrow also finds that “war is most likely when the initiator’s forces are superior to the defender’s forces, although war becomes unlikely when the initiator is grossly superior to the defender.” The model also reveals that “militarily weak nations are willing to initiate crises when they hold advantages that compensate for their objective military inferiority,” and “as the status quo becomes more favorable to the initiator . . . crises and wars become less likely because deterrence is more likely to hold.” The model also shows that “the costs of war do discourage the sides from fighting,” and suggests that “crises do not occur when the initiator holds strong beliefs that the defender has an advantage, regardless of the true state of affairs” (pp. 956–957, 959). In other words, states do not begin a crisis when they think the other side has a big advantage. There is nothing obviously wrong with these conjectures, but nothing very earth-shattering about them either.

---

47. James D. Morrow, “Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining,” *American Journal of Political Science*, Vol. 33, No. 4 (November 1989), pp. 941–972. Subsequent references to works discussed appear in parentheses in the text.

EXAMPLE NO. 2. Jeffrey S. Banks, "Equilibrium Behavior in Crisis Bargaining Games."<sup>48</sup> This article develops a formal solution to certain bargaining games of incomplete information that is robust with respect to the actual specification of the game. In other words, Banks shows that a certain class of formal results do not depend on the individual features of the game (e.g., the specific number of moves, the order in which the players choose, etc.).

What new hypotheses does the analysis yield? After elaborating a simple two-actor model in which one player possesses private information about the benefits and costs of war, Banks demonstrates that "in any equilibrium of any game with the above format, the probability of war is an increasing function of the expected benefits from war of the informed player." He elaborates: "In any equilibrium of a crisis bargaining game . . . 'stronger' countries (i.e., those with greater expected benefits from war) are more likely to end up in a war; yet if the bargaining negotiations are successful and war is averted, stronger countries receive a better settlement as well" (p. 601). In other words, states that know they will reap greater benefits from war are more likely to enter one; and states with greater power (and greater incentives for war) can strike a better deal when bargaining short of war. Few international relations scholars will find these results surprising, even if one accepts that the model is an accurate representation of real-world crises.<sup>49</sup>

EXAMPLE NO. 3. D. Marc Kilgour and Frank C. Zagare, "Credibility, Uncertainty, and Deterrence."<sup>50</sup> Kilgour and Zagare construct a three-stage formal model of deterrence that is explicitly designed to incorporate uncertainty about each player's willingness to retaliate. The model also allows the players to revise their behavior in light of the opponent's prior conduct. The model and the analysis are fairly sophisticated, but the theoretical and practical results are for the most part affirmations of the conventional wisdom.

Kilgour and Zagare suggest that the "signal contribution" of the model is to provide a "measure of the circumstances in which deterrence can emerge in an uncertain world" (p. 326). The actual results are not very illuminating, however. For example, the authors also find that (1) "the higher each player's evaluation of the . . . status quo, the more likely the sure-thing deterrence equilibrium exists" (p. 321); (2) "deterrence stability is enhanced by increasing

---

48. Jeffrey S. Banks, "Equilibrium Behavior in Crisis Bargaining Games," *American Journal of Political Science*, Vol. 34, No. 3 (August 1990), pp. 599–614.

49. It is worth noting that this article does not offer any empirical support for its claims.

50. D. Marc Kilgour and Frank C. Zagare, "Credibility, Uncertainty, and Deterrence," *American Journal of Political Science*, Vol. 35, No. 2 (May 1991), pp. 305–334.

the costs associated with mutual punishment" (p. 321); and (3) "if at least one player is willing to endure the costs of mutual punishment, deterrence can, but need not, fail" (p. 323). They conclude that "when the credibility of each player's threat is sufficiently high, deterrence is very likely," and further observe that "in core areas, where both players have inherently credible threats, increasing the costs of mutual punishment past a certain point does little to enhance deterrence stability." Given these familiar results, it is not surprising that they recommend "policies of deterrence that are sufficient to inflict unacceptable damage on an opponent yet are survivable enough to be available for a *retaliatory* attack," and that they endorse arms control, single-warhead ICBMs, hardened silos, and other familiar elements of nuclear strategy (pp. 326–327, emphasis in original). In short, Kilgour and Zagare have reinvented the central elements of deterrence theory without improving on it, despite the elaborate formal exercise they perform.<sup>51</sup>

EXAMPLE NO. 4. David Lalman and David Newman, "Alliance Formation and National Security."<sup>52</sup> This article develops an expected utility model of alliance formation and tests it against a body of quantitative data. The analysis is straightforward and clearly presented, but the conclusions are prosaic. For example, the authors find that "nations generally enter into alliances in the expectation of improving their security position," adding that "the pattern of alliance formation through time is related to the opportunity to enhance security . . . *realpolitik* considerations of security are crucial to alliance formation decisions" (p. 251). Although the analysis itself is careful and straightforward, it is not clear what has been gained from formalization.

EXAMPLE NO. 5. James D. Morrow, "Alliances, Credibility, and Peacetime Costs."<sup>53</sup> This article presents a sophisticated game-theoretic model in which alliances are a means of signaling interests in the presence of uncertainty. Although Morrow's formulation challenges the idea that alliance credibility is largely driven by concerns about reputation, the conclusions for the most part

---

51. Interestingly, Kilgour and Zagare's model produces results different from Morrow's model described above. Morrow's central finding was the impossibility of war in the absence of some form of uncertainty, whereas Kilgour and Zagare find that "misperception is neither necessary nor sufficient for the failure of mutual deterrence." *Ibid.*, p. 317. Among other things, this shows that modeling alone does not ensure truth, as one can create a model to produce any particular conclusion that one might want.

52. David Lalman and David Newman, "Alliance Formation and National Security," *International Interactions*, Vol. 16, No. 4 (1991), pp. 239–254.

53. James D. Morrow, "Alliances, Credibility, and Peacetime Costs," *Journal of Conflict Resolution*, Vol. 38, No. 2 (June 1994), pp. 270–297.

echo the conventional wisdom. In particular, the model implies that (1) “tighter alliances improve the ability of allies to fight together while imposing higher peacetime costs,” and (2) “tighter alliances tend to produce greater deterrence and a higher probability of intervention (on behalf of one’s ally).” This is probably correct but neither surprising nor counterintuitive, and Morrow himself notes that “the implications of the model appear to be consistent with stylized facts about alliances” (p. 294).

In each of these examples, in short, technical sophistication and logical consistency did not yield particularly creative or original results.

#### OLD WINE IN NEW BOTTLES

In addition to producing rather trivial results, formal models sometimes use new concepts or labels for familiar ideas, so what at first glance seems like a wholly original contribution turns out to be an old argument in a slightly different guise.<sup>54</sup> Consider the following examples.

EXAMPLE NO. 6. Robert Powell, “Absolute and Relative Gains in International Relations Theory.”<sup>55</sup> This article presents a simple formal model showing how the willingness of states to cooperate is affected by the distribution of benefits. The argument is simple and consistent with familiar realist logic: states worry more about the distribution of benefits when they fear that others might use their share of the gains to increase their military power and attack. As this fear increases, incentives to cooperate will decline.

Although Powell’s specification of the problem is an improvement over earlier treatments (including the seminal work of Joseph Grieco), the basic argument is an old one.<sup>56</sup> In Powell’s model, the critical variable that determines the prospects for cooperation is “the technology of warfare.” In his words, “If the use of force is at issue because the cost of fighting is sufficiently low, cooperation collapses. . . . But if the use of force is no longer at issue,

---

54. Needless to say, formal theorists are not the only social scientists who engage in this practice.

55. Robert Powell, “Absolute and Relative Gains in International Relations Theory,” *American Political Science Review*, Vol. 85, No. 4 (December 1991), pp. 1303–1320.

56. Grieco distinguished between two types of states with different preferences: those that cared about “relative” gains (which he called “defensive positionalists”) and those that did not (or “rational egoists”). By contrast, Powell assumes that all states have similar preferences, but their behavior varies with the external constraints they face. This formulation is simpler and more consistent with realist premises than Grieco’s. See Joseph Grieco, “Anarchy and the Limits of Cooperation,” in David Baldwin, ed., *Neorealism and Neoliberalism: The Contemporary Debate* (New York: Columbia University Press, 1993); Grieco, *Cooperation among Nations: Europe, America, and the Non-Tariff Barriers to Trade* (Ithaca, N.Y.: Cornell University Press, 1990); and Robert Powell, “The Neorealist-Neoliberal Debate,” *International Organization*, Vol. 48, No. 2 (Spring 1994), especially pp. 334–338.

cooperation again becomes feasible" (p. 1311). What Powell calls the "technology of warfare," however, is essentially identical to the concept of the offense-defense balance identified by George Quester, Robert Jervis, and Stephen Van Evera. As Jervis put it back in 1978: "If the defense has enough of an advantage . . . , the security dilemma [will] cease to inhibit status quo states from cooperating."<sup>57</sup> This is identical to Powell's claim that cooperation becomes more likely as the technology of warfare makes using force more costly (which implies that defenders can inflict high costs on attackers). Powell's article is a useful contribution to the absolute/relative gains debate, but it does not make a fundamentally new argument.

EXAMPLE NO. 7. James D. Fearon, "Rationalist Theories of War."<sup>58</sup> This article presents a rationalist framework for understanding the outbreak of war, using a simple formal bargaining model. Given the plausible assumption that fighting is always costly, Fearon argues that a satisfactory rationalist theory of war has to explain why the parties involved could not reach the same outcome via negotiation, thereby avoiding the costs of war.<sup>59</sup>

According to Fearon, the theoretical existence of outcomes that rational states should prefer to war implies that war can arise in only one of two ways. First, war arises because states have "private information" about power and resolve *and* powerful incentives to lie about it. They may misrepresent their strength or resolve to try to gain a better deal in a given confrontation, but this tactic may also lead them to overlook a negotiated solution that would have been preferable to war. Second, war can result from what he calls the "commitment problem." Even if both sides may know that a satisfactory bargain exists, they cannot accept the deal because they cannot be sure that it will be kept.

This article is useful because it suggests that rationalist theories of war are really of only two kinds, and it identifies how rational states can end up fighting even when there are negotiated solutions that each prefers to war.<sup>60</sup> But the central argument—that wars arise either from the "commitment prob-

---

57. See Jervis, "Cooperation under the Security Dilemma," p. 187.

58. James D. Fearon, "Rationalist Theories of War," *International Organization*, Vol. 49, No. 3 (Summer 1995), pp. 379–414.

59. The logic of this argument is similar to the theory of industrial disputes advanced by John Hicks in the 1930s. In Hicks's case, the question is why labor and management cannot reach agreement on a contract solely via negotiation, thereby avoiding the costs of a strike. See Hicks, *The Theory of Wages* (London: Macmillan, 1932), chap. 7.

60. Fearon's argument applies only to the final decision to wage war, once there is a concrete dispute between two states. It does not address the other conditions that might operate to make war more likely, such as ideological differences, shifting balances of power, the perceived weakness of a particular regime, or the domestic incentives that might drive a particular regime to seek war for its own sake.

lem" or "private information"—is not new. What Fearon calls the "commitment problem" (a term borrowed from recent formal work in economics) has long been recognized as a central feature of international anarchy. As Robert Art and Robert Jervis put it in 1976: "International politics takes place in an arena that has no central authority. . . . *States can make commitments and treaties, but no sovereign power ensures compliance.*" Similarly, Kenneth Oye noted in 1986 that because "states cannot cede ultimate control over their conduct to a supranational sovereign, *they cannot guarantee they will adhere to their promises.*"<sup>61</sup> Thus, to say that war arises from the "commitment problem" is merely to give a new label to a well-established idea.

In addition, although the concept of "private information" is broader than the more familiar idea of "secrecy," its effects on crisis bargaining are essentially the same.<sup>62</sup> It is not a new idea to claim that states are more likely to miscalculate when their opponents conceal information from them, although it is important to distinguish this source of miscalculation from errors arising from cognitive or organizational sources of misperception (a task Fearon accomplishes very well). Thus, although Fearon's analysis clarifies these issues in an insightful and intelligent way, the formalization does not yield a new theoretical claim.<sup>63</sup>

EXAMPLE NO. 8. "Costly Signals" and Reputation Building. A similar conclusion emerges when we examine the formal literature on reputation, and especially its reliance on the idea of "costly signaling."<sup>64</sup> The basic claim of these

---

61. See Robert Art and Robert Jervis, "Introduction," in Art and Jervis, *International Politics*, 1st ed. (Boston: Little, Brown, 1976), p. 2 (emphasis added); and Kenneth A. Oye, ed., *Cooperation under Anarchy* (Princeton, N.J.: Princeton University Press, 1985), p. 1 (emphasis added).

62. Unlike some forms of secrecy (such as number of weapons, for example), "private information" includes information (such as a player's level of resolve) that could not be reliably revealed to the other side even if one wanted to.

63. Fearon's discussion of "rationalist" theories does not explain when war will or will not occur. As he notes at the end of the article, both the problem of "private information" and the "commitment problem" created by anarchy are constant features of international politics and thus cannot explain why war occurs in some circumstances but not in others. See Fearon, "Rationalist Theories of War," p. 410.

64. Of this large and growing literature, see especially Reinhard Selten, "The Chain-Store Paradox," *Theory and Decision*, Vol. 9, No. 2 (April 1978), pp. 127–159; David M. Kreps, Paul Milgrom, John Roberts, and Robert Wilson, "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," pp. 245–252; Kreps and Wilson, "Reputation and Imperfect Information," pp. 253–279; and Milgrom and Roberts, "Predation, Reputation, and Entry Deterrence," pp. 280–312. All appear in *Journal of Economic Theory*, Vol. 27, No. 2 (August 1982). See also Wilson, "Deterrence in Oligopolistic Competition," in Stern et al., *Perspectives on Deterrence*; and Wilson, "Reputations in Games and Markets," in Alvin E. Roth, *Game-Theoretic Models of Bargaining* (Cambridge, U.K.: Cambridge University Press, 1985).



models is straightforward: because an actor's true preferences are unobservable and talk is cheap, an actor can signal its true preferences only by employing a "costly signal." Such signals are actions that impose higher costs on an actor with low resolve, and thus are more likely to be made only by actors who are more resolute. A number of formal theorists have used this concept in interesting ways, but the basic idea is virtually identical to Robert Jervis's distinction between "signals" and "indices," which he laid out more than twenty-five years ago.<sup>65</sup> As Jervis puts it, "Signals can be as easily issued by a deceiver as by an honest actor . . . they do not contain inherent credibility." By contrast, indices (which is Jervis's term for costly signals) "are statements or actions that carry some inherent evidence that the image projected is correct." Specifically, "behavior that is felt to be too important or costly in its own right to be used for other ends is an index."<sup>66</sup>

Again, the point is not that the formal literature on costly signaling has added nothing to our understanding of international politics. Rather, my point is that the idea that reputations could rest on what are now called "costly signals" did not emerge from a formal analysis. Accordingly, this literature cannot be used as evidence that formal theory is a superior source of new concepts or hypotheses. Formalization has refined our understanding, perhaps, but even that claim has not gone unchallenged.<sup>67</sup>

Taken together, these examples reveal that even sophisticated formal analyses often lead to familiar conclusions about the behavior of states. Does this mean that formal theory has added nothing new? Of course not. Schelling's early work was extremely influential, as was the application of collective goods theory to the question of alliance burden-sharing.<sup>68</sup> Robert Axelrod's analysis of the logic of cooperation in an iterated prisoner's dilemma has had a far-

---

65. Examples include Nalebuff, "Rational Deterrence in an Imperfect World"; Wagner, "Rationality and Misperception in Deterrence"; D. Marc Kilgour, "Domestic Political Structure and War Behavior: A Game-Theoretic Approach," *Journal of Conflict Resolution*, Vol. 35, No. 2 (June 1991), pp. 266–284; and Andrew Kydd, "Game Theory and the Spiral Model," *World Politics*, Vol. 49, No. 3 (April 1997), pp. 371–400.

66. See Robert Jervis, *The Logic of Images in International Relations* (Princeton, N.J.: Princeton University Press, 1970), pp. 21, 26. For a discussion of the relationship between his early work and recent economic work on reputation, see Jervis, "Drawing Inferences and Projecting Images," paper presented at the annual meeting of the International Studies Association, San Diego, California, April 1996.

67. For a fair-minded critique of the formal literature on reputation, see Jonathan Mercer, *Reputation and International Politics* (Ithaca, N.Y.: Cornell University Press, 1996), pp. 28–42.

68. The seminal piece is Mancur Olson and Richard Zeckhauser, "An Economic Theory of Alliances," *Review of Economics and Statistics*, Vol. 48, No. 2 (August 1966), pp. 266–279.

reaching impact in many areas, including the study of strategy.<sup>69</sup> Recent formal work has also shown how certain phenomena (such as the stability of an arms control agreement) can be sensitive to the level of information available to the actors, and formal analysis can increase our confidence in an existing theory by confirming its logical soundness.<sup>70</sup> So the criticisms noted above should not be interpreted as a blanket condemnation either of formalization in general or of its recent manifestations in security studies.<sup>71</sup>

Yet it should also be clear that formal theory enjoys no particular advantage as a source of theoretical creativity.<sup>72</sup> In addition to confining the analysis to an individualistic, rational actor framework, the technical requirements of modern game theory tend to shape both the topics that are chosen and the ways they are addressed. It is not surprising, for example, that much of the formal work in security studies focuses on two-party interactions (and especially on crisis behavior), because these situations are mathematically tractable. This tendency makes good methodological sense, but it may also help explain why other approaches have been more theoretically fruitful, and have made more useful contributions to other security problems.

---

69. Strictly speaking, Axelrod's results emerged from a computer simulation rather than from a formal model. His argument rests on the logic of the iterated prisoner's dilemma, however, and helps highlight the importance of these games for the formal analysis of cooperation. Given that some parts of Axelrod's argument did not stand up to careful formal scrutiny, this example also supports the claim that creativity and theoretical fertility are more important than strict logical consistency. See Robert Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984).

70. See George Downs and David Rocke, *Tacit Bargaining and Arms Control* (Ann Arbor: University of Michigan Press, 1989); and Downs and Rocke, *Optimal Imperfection?: Domestic Uncertainty and Institutions in International Relations* (Princeton, N.J.: Princeton University Press, 1995).

71. For a nuanced appreciation of the contributions and limitations of rational choice theorizing, see James B. Rule, *Theory and Progress in Social Science* (Cambridge, U.K.: Cambridge University Press, 1997), chap. 3.

72. It is worth noting that some of the most interesting and important theoretical innovations in security studies over the past two decades have come from nonformal theorists. For example, John Steinbruner and Bruce Blair made a major contribution to deterrence theory by highlighting the importance of organizational and operational considerations; Robert Jervis offered a comprehensive inventory of the ways that psychological biases could affect foreign policy decisionmakers; John Mearsheimer developed and tested a simple theory of conventional deterrence; Barry Posen showed how external conditions and organization theory could explain key elements of great power military doctrine; and Robert Pape constructed and tested a theory of military coercion. More recently, scholars like Peter Katzenstein, Elizabeth Kier, and Alastair Iain Johnston have applied cultural and constructivist approaches to security studies, all of them based on extensive empirical work. One need not be persuaded by all of these works to recognize that they were important efforts to bring social science to bear on important security problems. For a useful survey, see Richard K. Betts, "Should Strategic Studies Survive?" *World Politics*, Vol. 50, No. 1 (October 1997), pp. 7–33.

Furthermore, the history of both natural and social science suggests that theoretical innovations emerge not from abstract modeling exercises, but primarily from efforts to solve concrete empirical puzzles. Indeed, simple observation and largely atheoretical experimentation can be as important as subsequent efforts to devise a deductive structure to explain the observations.<sup>73</sup> In other words, induction and deduction are equally valid avenues for creating a theory, and the former may in fact be more fruitful.

Deterrence theory offers an obvious example. Contrary to the traditional view, formal theory played no role in the creation of deterrence theory.<sup>74</sup> Rather, it emerged largely from Bernard Brodie's historically informed speculations and from Albert Wohlstetter's famous RAND study on the optimal location for U.S. bomber bases. A similar example is the initial development of the bureaucratic politics paradigm by Andrew Marshall and Joseph Loftus, who were trying to account for apparently irrational Soviet military allocations.<sup>75</sup> Among other things, these episodes confirm that case studies can be an extremely fertile source of new theories as well as a useful way to test both formal and nonformal theories.<sup>76</sup>

The bottom line is that although formal approaches to security affairs have produced a number of interesting refinements, the overall level of theoretical innovation is not superior to other social scientific methods. Formalization can impart greater precision and help identify inconsistencies or qualifications, but it enjoys no particular advantage as a source of new hypotheses.

### *Empirical Validity*

The ultimate measure of a theory is its ability to explain real events in the real world. As Maurice Allais warned in his address accepting the Nobel Prize for economics, "Mere logical, even mathematical deduction remains worthless in

---

73. See Ian Hacking, *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science* (New York: Cambridge University Press, 1983), pp. 154–158, 248–249; and Stephen Toulmin, *Human Understanding*, Volume 1: *The Collective Use and Evolution of Concepts* (Princeton, N.J.: Princeton University Press, 1972), pp. 189–190.

74. On this point, see O'Neill, "Game Theory Models of War and Peace," pp. 1010–1013.

75. See Bernard Brodie, ed., *The Absolute Weapon: Atomic Power and World Order* (New York: Harcourt, Brace, 1946); A.J. Wohlstetter, F.S. Hoffman, R.J. Lutz, and H.S. Rowen, *Selection and Use of Strategic Air Bases*, RAND Report P-266 (Santa Monica, Calif.: RAND, 1962); and Marc Trachtenberg, *History and Strategy* (Princeton, N.J.: Princeton University Press, 1991), pp. 29–31.

76. On the latter point, see Bruce Russett, "International Behavior Research: Case Studies and Cumulation," in Russett, *Power and Community in World Politics* (San Francisco, Calif.: W.H. Freeman, 1974), pp. 16–17.

terms of an understanding of reality if it is not closely linked to that reality. . . . Any theory whatever, if it is not verified by empirical evidence, has no scientific value and should be rejected."<sup>77</sup> A research tradition that insists on careful and systematic empirical testing is setting a higher standard for itself than one that places relatively little value on the provision of empirical support. Mere logical consistency is not sufficient.

Does recent rational choice scholarship in security studies pay sufficient attention to this criterion? The answer is no. With a few notable exceptions, the bulk of formal work in security affairs does not engage in any empirical testing at all. Anecdotes and "stylized facts" are sometimes used to explicate a point and to enhance the plausibility of the argument, but relatively little effort is devoted to rigorous empirical evaluation.<sup>78</sup> Other formalists have used mathematical simulations or referred to supportive quantitative evidence, but

---

77. Quoted in Mayer, *Truth versus Precision in Economics*, p. 27. Albert Einstein shared this view. He praised Johannes Kepler for recognizing that "even the most lucidly logical mathematical theory was of itself no guarantee of truth, becoming meaningless unless it was checked against the most exacting observations in natural science." And Einstein called Galileo Galilei "the father of modern physics" because Galileo realized that "pure logical thinking cannot yield us any knowledge of the empirical world. . . . Propositions arrived at by purely logical means are completely empty as regards reality." Quoted in Timothy Ferris, *The Whole Shebang: A State of the Universe(s) Report* (New York: Touchstone, 1997), p. 28. Gary King, Robert Keohane, and Sidney Verba offer a similar appraisal, writing that "formal models do not constitute verified explanations without empirical evaluation of their predictions." See King, Keohane, and Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton, N.J.: Princeton University Press, 1994), pp. 105–106.

78. For example, the following formal works contain no empirical evidence, and rely entirely on formal deduction: Powell, *Nuclear Deterrence Theory*; Powell, "Absolute and Relative Gains in International Relations Theory"; Powell, "Guns, Butter, and Anarchy," *American Political Science Review*, Vol. 87, No. 1 (March 1993), pp. 115–132; Powell, "Stability and the Distribution of Power," *World Politics*, Vol. 48, No. 2 (January 1996), pp. 239–267; Fearon, "Rationalist Theories of War"; Banks, "Equilibrium Behavior in Crisis Bargaining Games"; James D. Morrow, "A Spatial Model of International Conflict," *American Political Science Review*, Vol. 80, No. 4 (December 1986), pp. 1131–1150; R. Harrison Wagner, "Deterrence and Bargaining," *Journal of Conflict Resolution*, Vol. 26, No. 2 (June 1982), pp. 329–358; Wagner, "Rationality and Misperception in Deterrence"; Wagner, "Nuclear Deterrence, Counterforce Strategies, and the Incentive to Strike First," *American Political Science Review*, Vol. 85, No. 3 (September 1991), pp. 727–749; Kilgour and Zagare, "Credibility, Uncertainty, and Deterrence"; Kilgour, "Domestic Political Structure and War Behavior"; Morrow, "Alliances, Credibility, and Peacetime Costs"; Nalebuff, "Rational Deterrence in an Imperfect World"; Nalebuff, "Minimal Nuclear Deterrence," *Journal of Conflict Resolution*, Vol. 32, No. 3 (September 1988); Nalebuff, "Brinkmanship and Nuclear Deterrence: The Neutrality of Escalation," *Conflict Management and Peace Science*, Vol. 9, No. 2 (Spring 1986), pp. 19–30; Alasdair Smith, "Diversionary Foreign Policy in Democratic Systems," *International Studies Quarterly*, Vol. 40, No. 1 (March 1996), pp. 133–153; Smith, "Alliance Formation and War," *International Studies Quarterly*, Vol. 39, No. 4 (December 1995), pp. 405–425; Smith, "International Crises and Domestic Politics," *American Political Science Review*, Vol. 92, No. 3 (September 1998), pp. 623–638; and Kydd, "Game Theory and the Spiral Model."

these approaches fall short of a careful empirical test.<sup>79</sup> Empirical testing is not a central part of the formal theory enterprise—at least, not in the subfield of security studies—and probably constitutes its most serious limitation.<sup>80</sup>

EXAMPLE NO. 9. James D. Fearon, “Domestic Audience Costs and the Escalation of International Disputes.”<sup>81</sup> The limitations that arise from the low priority placed upon empirical testing are nicely revealed in James Fearon’s formal analysis of domestic audience costs and crisis bargaining. Fearon’s argument is intuitively plausible, technically sophisticated, and informed by his knowledge of international history. Fearon goes to some lengths to identify the real-world implications of his analysis, and the article is in many ways an exemplary use of the formal approach.

Fearon defines audience costs as the domestic political costs that leaders incur when they back down in a crisis. Using a simple bargaining model in which leaders of different states face different audience costs, he finds that once a crisis is under way, “the side with a stronger domestic audience (e.g., a democracy) is always less likely to back down than the side less able to generate audience costs (a nondemocracy).” He also suggests that the constraints imposed by audience costs may explain why democratic states are less prone to conflict with each other; specifically, the presence of high audience costs allows democratic leaders to signal their intentions more credibly, thereby minimizing the miscalculations that can lead to war (pp. 577, 586).<sup>82</sup>

---

79. Examples of these approaches to testing include Jean-Pierre Langlois, “Rational Deterrence and Crisis Stability,” *American Journal of Political Science*, Vol. 35, No. 4 (November 1991), pp. 801–832; Downs and Rocke, *Tacit Bargaining and Arms Control*; Downs and Rocke, *Optimal Imperfection?*; and Morrow, “Capabilities, Uncertainty, and Resolve.”

80. Between 1989 and 1998, for example, *World Politics* published twelve articles that contained a formal model. Of these, only five contained systematic empirical evidence. The *Journal of Conflict Resolution* published thirty-seven formal articles in the same period (excluding articles dealing solely with technical aspects of game theory), of which thirteen contained empirical support for the model. *International Studies Quarterly* contained twenty-six formal articles in this period, but only ten contained empirical evidence, and *International Organization* published fourteen, seven of which provided empirical support for the model. Overall, roughly 60 percent of these articles relied solely on formal deduction and anecdotal illustration, rather than systematic empirical testing. Similarly, a total of ninety-four formal theory manuscripts were submitted to the *American Political Science Review* between August 1996 and August 1997, but only twenty-five of them (26 percent) contained systematic empirical evidence. See Finifter, “Report of the Editor of the APSR,” p. 784.

81. James D. Fearon, “Domestic Audience Costs and the Escalation of International Disputes,” *American Political Science Review*, Vol. 88, No. 3 (September 1994), pp. 577–592.

82. The idea that relative audience costs will affect bargaining power is not new. As far as I know, the idea was first articulated by Thomas Schelling, who also suggested that democratic and nondemocratic states might differ on this dimension. See Schelling, *Strategy of Conflict*, pp. 27–29.

What is not clear is whether the argument is in fact correct. To begin with, Fearon's model assumes that democracies typically face greater domestic audience costs than nondemocracies, and that both democratic and nondemocratic leaders recognize that this is the case. If an authoritarian regime *believed* that its own audience costs were higher (e.g., due to the fear of a coup or because the regime thought that democratic publics were easily manipulated), the model's predictions would not hold.<sup>83</sup> Yet Fearon offers only anecdotal evidence that authoritarian states actually face lower audience costs, or that this belief is widely shared by democratic and nondemocratic leaders.<sup>84</sup>

Second, the model also assumes that leaders and publics hold similar preferences about the proper course of action. Domestic audiences will punish leaders who back down, but they may also reward a leader who overreaches at first and then manages to retreat short of war. Thus the British and French governments did not suffer domestic audience costs when they backed down during the Rhineland crisis of 1936 or the Munich crisis in 1938, because public opinion did not support going to war.<sup>85</sup>

Third, although Fearon does present some illustrative anecdotes and refers to several quantitative studies that are consistent with his argument, he does not test the logic of the model directly.<sup>86</sup> And it is not difficult to think of possible exceptions: (1) the United States gave in to North Korea's demands following the seizure of the *Pueblo* in 1968 and also granted many of Iran's demands following the seizure of the U.S. embassy in 1980, even though it

---

83. The personal cost to a deposed tyrant could be higher on average than the cost to a democratic incumbent who risks losing the next election, and authoritarian leaders often face other domestic pressures that limit their ability to back down once a crisis is under way. For example, the belief that the Hapsburg monarchy faced internal revolt unless it eliminated the threat from Serbia drove Austro-Hungarian decisions in the July crisis that led to World War I, and the three authoritarian states involved in the July crisis (Germany, Austria-Hungary, and Russia) were clearly less willing to compromise than were democratic Britain and France.

84. As Fearon notes, "The idea that democratic leaders on average have an easier time generating audience costs is advanced here as a plausible working hypothesis." See Fearon, "Domestic Audience Costs," p. 582.

85. A further complication arises if neither leaders nor publics know how resolved they are until *after* the crisis is under way. As a result, a particular leader may adopt a hard-line position at first, based on the belief that it is in the "objective" interest of the state and that this position has popular backing. As the crisis continues, however, the citizenry may become alarmed by the danger of war and eager for a peaceful resolution. A democratic leader who backed down at this point might be rewarded rather than penalized, whereas a leader who continued to escalate might be punished for adventurism.

86. A recent quantitative study that supports some of Fearon's predictions is Joe Eyerman and Robert A. Hart, Jr., "An Empirical Test of the Audience Cost Proposition: Democracy Speaks Louder Than Words," *Journal of Conflict Resolution*, Vol. 40, No. 4 (December 1996), pp. 597-616.

probably faced higher audience costs than did North Korea or Iran; (2) Britain, France, and Israel backed down to U.S. (and possibly Soviet) pressure following the Suez War in 1956, even though their “audience costs” were much higher than those of their opponents;<sup>87</sup> (3) higher “audience costs” did not enable the United States to prevail against England in the *Trent* affair in 1861, and public opinion had virtually no impact in the Venezuelan crisis in 1895–96;<sup>88</sup> and (4) domestic audience costs were dwarfed by other considerations prior to the Six-Day War in 1967 and the War of Attrition in 1969–70.<sup>89</sup>

Fourth, the absence of empirical testing is also important because the model omits another potentially important determinant of crisis behavior. For some states (including some democracies), the principal cost of backing down in a crisis may not be domestic censure but the fear that allies may defect and that adversaries will be emboldened. If there is a sharp difference in the *external* audience costs that each state faces, then a difference in *domestic* audience costs may fade into insignificance.<sup>90</sup> This question is ultimately an empirical one, of course, and cannot be resolved by a purely formal analysis.

In short, this article offers an interesting and intuitively plausible conjecture about crisis bargaining, one well worth further exploration. Until it is rigorously tested, however, there is no way of knowing how significant the actual contribution really is.

---

87. Robert A. Pape argues that Soviet nuclear threats played a critical role in convincing British leaders to withdraw from Egypt after the Suez invasion, but other scholars have reached different conclusions. See Pape, “Why Economic Sanctions Do Not Work,” *International Security*, Vol. 22, No. 2 (Fall 1997), pp. 115–117; and James G. Richter, *Khrushchev’s Double Bind: International Pressures and Domestic Coalition Politics* (Baltimore, Md.: Johns Hopkins University Press, 1994), p. 93.

88. See Christopher Layne, “Kant or Cant: The Myth of the Democratic Peace,” *International Security*, Vol. 19, No. 2 (Fall 1994), pp. 5–49.

89. Israel’s decision to preempt in 1967 is consistent with Fearon’s model, insofar as the model suggests that democracies will escalate more readily once a crisis is under way. But domestic audience costs did not play a key role in either state’s decisions to escalate the crisis or in the final decision to go to war. Both Egypt and Israel seem to have been equally resolute, in part because both believed they were ready for war. Nasser’s reluctance to back down was based in part on his concerns about *external* audience costs (and especially the loss of prestige in the Arab world), which underscores the unimportance of relative *domestic* costs in this case. Similarly, nondemocratic Egypt gained a tactical victory over democratic Israel during the 1969–70 War of Attrition, in part because Israeli resolve waned as the conflict continued and in part because the fear of escalation led the superpowers to impose terms that were favorable to Egypt. See Jonathan Shimshoni, *Israel and Conventional Deterrence: Border Warfare from 1953 to 1970* (Ithaca, N.Y.: Cornell University Press, 1988), pp. 169–171.

90. These concerns are not unrelated, of course, because failure to preserve one’s external reputation is one reason why a domestic audience might seek to remove a particular leader. Nonetheless, the two concerns are conceptually and empirically distinct.

## TAKING TESTING SERIOUSLY?

When formal theorists do engage in extensive empirical testing, moreover, the tests themselves are not as “rigorous” as they might initially appear. To illustrate this point, let us consider two ambitious attempts to combine formal analysis with extensive empirical testing: (1) Bruce Bueno de Mesquita and David Lalman, *War and Reason: Domestic and International Imperatives*, and (2) Emerson M.S. Niou, Peter C. Ordeshook, and Gregory F. Rose, *The Balance of Power: Stability in International Systems*. I have chosen these works because their authors emphasize the importance of testing theories empirically, and because both have been seen as salient demonstrations of the power of formal theory.<sup>91</sup> The question is how rigorous are the tests and how well do the theories perform?

EXAMPLE NO. 10. Bruce Bueno de Mesquita and David Lalman, *War and Reason: Domestic and International Imperatives*.<sup>92</sup> The centerpiece of *War and Reason* is a formal model with two players (state *A* and state *B*) and eight possible outcomes: the status quo (SQ), negotiation (N), capitulation by *A* (Cap<sub>A</sub>), capitulation by *B* (Cap<sub>B</sub>), war begun by *A* (War<sub>A</sub>), war begun by *B* (War<sub>B</sub>), acquiescence by *A* (Acq<sub>A</sub>), and acquiescence by *B* (Acq<sub>B</sub>). They make a number of general assumptions about state preferences and assume that all states act to maximize their expected utility. The bulk of the subsequent analysis explores the additional restrictions (e.g., on the domestic costs of using force, the cost of giving in after being attacked, etc.) that would make each outcome the “equilibrium” outcome. These additional assumptions are then interpreted as the underlying conditions that yield each outcome in the game.<sup>93</sup>

91. For example, Frank Zagare’s review of *War and Reason* in the *American Political Science Review*, Vol. 87, No. 3 (September 1993), p. 811, praised it as “the most significant application to date of game theory to the question of war and peace,” and Glenn Snyder called *The Balance of Power* “a valuable, ground-breaking effort” that “blazes a useful trail.” See Snyder, “Alliances, Balance, and Stability,” *International Organization*, Vol. 43, No. 1 (Winter 1991), pp. 140, 142.

92. Bruce Bueno de Mesquita and David Lalman, *War and Reason: Domestic and International Imperatives* (New Haven, Conn.: Yale University Press, 1992).

93. As with some of the other formal work discussed above, many of the results derived from the model are rather trivial. After deriving and testing more than twenty hypotheses, for example, Bueno de Mesquita and Lalman offer the following general conclusions: “To state it crudely: national leaders wage war when the expected gains minus the expected costs of doing so outweigh the net expected consequences of alternative choices. War can be stumbled into when one nation judges the intentions of a rival too optimistically. War can begin even with full information if it is motivated by a fear of ceding any advantage that is attached to the first use of force. The anticipated net gains from war may be real and tangible acquisitions, or they may be the avoidance of a future expected to be worse than the one anticipated through warfare.” They also find that war will not occur if two states prefer negotiation to using force, and if both sides know this with 100 percent confidence. In other words, if both sides would rather talk than fight and if both sides know this, they do not fight. See Bueno de Mesquita and Lalman, *War and Reason*, p. 250.



The predictions derived from the model are tested with a series of large-*N* analyses and a number of briefer case studies. At first glance, these tests appear to show overwhelming success for the model. Upon closer examination, however, the results are not compelling and do not achieve a high level of rigor.

Bueno de Mesquita and Lalman offer numerous statistical tests of their various hypotheses, based on a data set of 707 dyadic relationships between European states from 1815 to 1970.<sup>94</sup> Although they do not describe their statistical procedures in much detail (which makes it difficult to evaluate their conclusions), the results appear to provide convincing support for the model.<sup>95</sup> Unfortunately, there are at least three noteworthy problems with their statistical tests.

First, the quantitative indicators they employ face severe problems of internal validity, given the intrinsic difficulty of obtaining valid quantitative indicators for concepts like “risk propensity,” “utility,” and “uncertainty,” and then applying them to 707 dyads going back to 1815. To their credit, Bueno de Mesquita and Lalman recognize the difficulty of the task and admit that their indicators are quite crude (p. 280). Unfortunately, this also means that the quantitative tests are not very rigorous, because the indicators on which they are based do not adequately capture the theoretical relationships set forth in the formal model.

Second, the statistical tests are compromised by the lack of precise measures for key variables in the model. As they admit, Bueno de Mesquita and Lalman were unable to devise measures for some of the critical variables in the equilibrium conditions that yield different outcomes. The practical result of these missing conditions is that events they code as consistent with a particular outcome may be equally consistent with several other outcomes, which means that the tests blend successful and unsuccessful predictions. Thus we do not know how many successful predictions the model actually makes.

Even if this problem were corrected, Bueno de Mesquita and Lalman’s specific testing procedure exaggerates their model’s performance. In particular,

---

94. Specifically, their data set consists of 469 events in which two states engaged in a dispute with each other, plus another 238 observations on randomly paired dyads, included to represent the “nonevents” that are often excluded from quantitative studies. Thus the analysis is based on 707 dyad observations.

95. To note one example, on p. 84 they describe a dummy variable labeled BACQ, which is meant to satisfy the theoretical conditions of Proposition 3.5 (the “acquiescence by B theorem”). The reader is told that “the details of the operationalization are in appendix 1,” but there is in fact no mention of this dummy variable anywhere in the appendix and only a very general discussion of the actual measures they employed. The reader is also referred to a number of earlier articles and books for explanations of key elements of the methodology, thereby making it even more difficult to figure out what they have done.

although their model contains eight possible outcomes, most of the hypotheses are tested by constructing a  $2 \times 2$  table of observed and predicted outcomes, asking simply whether or not the predicted outcome occurred. Unfortunately, collapsing eight categories into two lumps together cases where a particular outcome was predicted and actually occurred and any cases where a particular outcome was predicted but did not occur, thereby generating inflated chi-squared and goodness of fit statistics.<sup>96</sup> Taken together, these flaws undermine their otherwise laudable effort to test the model through a detailed quantitative analysis.

One response to these problems would be to supplement the quantitative analyses with case studies, where one could hope to obtain more valid and reliable measures of key variables.<sup>97</sup> Ideally, a detailed process-tracing of an appropriate set of case studies would have allowed the authors to determine if the participants in interstate crises made choices in the manner depicted by the model, thereby providing a more convincing demonstration of its explanatory power. Unfortunately, the case studies contained in *War and Reason* do not provide this sort of evidence and do not achieve a high standard of rigor. Consider the following three examples.

*The Fashoda Crisis.* In chapter 3 of *War and Reason*, Bueno de Mesquita and Lalman deduce Proposition 3.5, the “acquiescence by B theorem.” It reads as follows:

With full information conditions, assumption 2.A.7b, . . . and strict preferences,  $Acq_B$  is a full-information equilibrium outcome of the international interaction game if and only if the equilibrium outcome of the crisis subgame at node 5 is either  $Cap_B$  or  $War_A$ , and for State B,  $Acq_B > War_A$ . (p. 81)

---

96. For example, when testing whether or not the model successfully predicts “acquiescence by state B,” Bueno de Mesquita and Lalman present a  $2 \times 2$  table comparing observed and predicted outcomes. Entries on the main diagonal of this table (Yes/Yes or No/No) appear to be successful predictions, but the 442 entries in the No/No cell contain both cases where the model successfully predicted a specific outcome different from  $Acq_B$  (such as “capitulation”) and cases where it predicted an outcome other than  $Acq_B$  but where some other outcome (different from both  $Acq_B$  and the predicted outcome) actually occurred. Collapsing categories in this way thus masks the unsuccessful predictions. See Bueno de Mesquita and Lalman, *War and Reason*, pp. 81–85; and Curtis S. Signorino, “Estimation and Strategic Interaction in Discrete Choice Models of International Conflict,” Occasional Paper No. 98–4 (Cambridge, Mass.: Weatherhead Center for International Affairs, Harvard University, 1998), pp. 20–23.

97. Bueno de Mesquita has previously emphasized the problem of internal validity in large-*N* research, noting that “the close scrutiny of individual decisions yields better estimates of utilities than do gross applications of general evaluative criteria.” See Bueno de Mesquita, “Toward a Scientific Understanding,” p. 133.

As discussed earlier, the manner of presentation is not very transparent, and one must refer back to the original model and assumptions to figure out what is actually being said. Once translated, however, this proposition in effect predicts that state *B* will acquiesce to state *A*'s demands if it prefers acquiescing to beginning a war itself *or* to letting the opponent (state *A*) begin the war.

Bueno de Mesquita and Lalman illustrate this hypothesis with a brief case study of the 1898 Fashoda crisis between Great Britain and France. To do this properly, one would want to obtain independent evidence about British and French preferences, and then show that each side acted as the model predicts. Specifically, one would have to show that the leaders of each state held the preferences identified in Proposition 3.5, and that France backed down because it preferred that outcome to a war launched by Britain. Ideally, one would also seek evidence showing that the key elites made the choices they did via a process of reasoning at least roughly similar to the mechanism implied by the model. Yet the only evidence that Bueno de Mesquita and Lalman provide about French preferences is a quotation by French Foreign Minister Théophile Delcassé, stating that "war is preferable to national dishonor" (quoted on p. 84).

In short, the *model* says France will acquiesce only if it prefers this outcome to war, yet the French foreign minister apparently believed exactly the opposite. It is possible (even likely) that Delcassé was bluffing and his statement was not a true reflection of French preferences. Nonetheless, given that this is the only independent evidence Bueno de Mesquita and Lalman provide about French preferences, the case (as they portray it) actually contradicts their theoretical argument.

*The Greco-Turkish Confrontation in Cyprus.* A second example follows from their analysis of the democratic peace literature. To explain why democracies do not fight each other whereas democracies and nondemocracies do, Bueno de Mesquita and Lalman argue that democracies face higher domestic constraints to using force and that this is common knowledge. As a result, all democracies know that they prefer peace to war, and each knows that other democracies know this, so war between them is not a rational outcome.

When a democracy and a nondemocracy face each other, however, their model identifies two main paths to war. In the first path, the nondemocracy assumes that the democracy is reluctant to use force and attacks, mistakenly believing that the democracy will capitulate. In the second path, which is the one emphasized by Bueno de Mesquita and Lalman, the democracy fears that the nondemocracy will try to exploit its reluctance to use force and chooses to

preempt, thereby obtaining the first-strike advantage. In their words, "The high domestic political constraint faced by democracies makes them vulnerable to threats of war or exploitation and liable to launch preemptive attacks against presumed aggressors" (p. 159). This result seems quite counterintuitive: for democracies, their reluctance to use force actually makes them more likely to employ it!<sup>98</sup>

Unfortunately, this surprising result is not well supported by the empirical record, including the evidence contained in *War and Reason* itself. Bueno de Mesquita and Lalman suggest that democracies are prone to preempt in a crisis, but more extensive empirical studies have shown that preemptive wars are very rare and that democracies almost never fight preventive wars.<sup>99</sup> The outbreaks of World Wars I and II contradict their model as well, insofar as none of the threatened democracies tried to launch a preemptive attack on their nondemocratic adversaries. Furthermore, the model suggests that domestic constraints are the key to the democratic war puzzle, whereas other empirical studies have suggested that normative factors or alliance commitments are more important.<sup>100</sup>

Finally, and most important for our purposes, the evidence presented by Bueno de Mesquita and Lalman often undermines their own argument. First, as already discussed, the 1898 Fashoda crisis was a confrontation between two democracies, but the dovish nature of democracies and their ability to signal peaceful intentions played little or no role in its outcome (and is not even mentioned in their own account).<sup>101</sup> Second, to show how democracies and nondemocracies interact in ways that lead the former to use force preemptively, Bueno de Mesquita and Lalman offer a brief case study comparing the 1967 and 1974 confrontations between Greece and Turkey over Cyprus. In 1967 the use of force was averted, but democratic Turkey occupied the disputed

---

98. As they elaborate: "If the first-strike advantage is large enough, *A* will prefer to initiate the use of force rather than risk being compelled to capitulate or to fight under the most adverse conditions. Thus, *A*'s democratic institutions make it susceptible to exploitation and incline it toward preemption." See Bueno de Mesquita and Lalman, *War and Reason*, pp. 159–160.

99. See Randall L. Schweller, "Domestic Structure and Preventive War: Are the Democratic States More Pacific?" *World Politics*, Vol. 44, No. 1 (January 1992), pp. 235–269; and Dan Reiter, "Exploding the Powder Keg Myth: Preemptive Wars Almost Never Happen," *International Security*, Vol. 20, No. 2 (Fall 1995), pp. 5–34.

100. See, in particular, Bruce Russett, *Grasping the Democratic Peace* (Princeton, N.J.: Princeton University Press, 1995), pp. 86–93, 119–120; and Joanne Gowa, *Ballots and Bullets* (Princeton, N.J.: Princeton University Press, forthcoming).

101. On Fashoda, see Christopher Layne, "Kant or Cant?"; and Susan Peterson, "The Lessons of the Fashoda Crisis for Democratic Peace Theory," *Security Studies*, Vol. 5, No. 1 (Autumn 1995), pp. 3–37.

island in 1974. They argue that the absence of force in 1967 and its employment in 1974 are both consistent with their theoretical predictions.

Unfortunately, this case neither constitutes a rigorous test nor offers persuasive support for their argument. The model predicts that Turkey will preempt to prevent Greece from taking advantage of the domestic constraints on Turkey's use of force, yet they offer no evidence demonstrating that this is in fact the reason Turkey chose to act in 1974 but not in 1967. More important, this case is a poor choice for testing this proposition because the democratic Greek government was overthrown by a military coup at the beginning of the 1967 crisis. Thus Greece was not a democracy in either 1967 or 1974, yet Turkey did not use force in the first confrontation but did use force in the second. Faced with this clear challenge to the model, they argue that the Greek military dictatorship faced domestic constraints that were "more typical of a democracy" (p. 162). This sort of flexible coding is the antithesis of scholarly rigor, and casts further doubt on the empirical validity of the model.

*The Sino-Indian Border War.* Bueno de Mesquita and Lalman also include a case study of the Sino-Indian border war of 1962. Although this case is included to illustrate a proposition about the impact of shifts in the balance of power, their explanation for the war is the same as the causal mechanism they depict for a war between a democracy and a nondemocracy. In their words: "All the conditions for war were there. India believed China preferred to capitulate rather than fight back. China knew India held this belief. China sought negotiation and offered concessions, whereas India sought capitulation or acquiescence. The Chinese were prepared to fight back, but India, a low probability of success in war notwithstanding, pursued the use of force through its forward policy. China ultimately met force with force" (p. 202). This is precisely the causal pattern suggested for a preemptive war begun by a democracy against a nondemocratic challenger; the only problem is that the regime types are exactly the opposite of the ones depicted by the model! In this case, China (a nondemocracy) is acting the way that their model says a democracy should behave (i.e., it is reluctant to use force and prefers negotiation, but eventually preempts when pressed). India (a democracy) is acting the way the model says that authoritarian challengers will behave (i.e., it is trying to take advantage of the other side's reluctance to use force). Yet the contradiction is never explained.

Thus the three case studies they provide of democratic-democratic and democratic-nondemocratic interactions either do not support or actually contradict the predictions of the model. The empirical record is not being used to test the theory; it is being tailored to fit it.

In sum, neither the quantitative analysis nor the case studies contained in *War and Reason* provides compelling empirical support for the theoretical model developed in the book. Significantly, neither type of test is performed in an especially careful or rigorous fashion—among other things, the case studies themselves appear to be based on a cursory number of historical sources—and little effort is made to test the model directly. Although Bueno de Mesquita and Lalman are to be commended for stressing the importance of empirical testing, their effort does not achieve a high standard of scientific rigor.

EXAMPLE NO. 11. Emerson M.S. Niou, Peter C. Ordeshook, and Gregory F. Rose, *The Balance of Power: Stability in International Systems*.<sup>102</sup> This book is an ambitious effort to formalize balance-of-power theory, and to test the resulting model through an in-depth study of great power diplomacy. Unlike most of the recent formal work on security topics, which uses two-person, noncooperative game theory, *The Balance of Power* relies on *n*-person, cooperative game theory.<sup>103</sup> Although in many ways an exemplary study (the presentation is reasonably clear and accessible, and the authors frequently acknowledge the limits of their model), the empirical results are not convincing.

The central focus of the book is the concept of *stability*, which takes two forms. System stability refers to any distribution of resources in which none of the “essential” members can be eliminated by the others. Resource stability, by contrast, refers to situations where there is no incentive or capacity to alter the existing distribution of resources. The central question, therefore, is under what conditions will a given international system exhibit either form of stability.<sup>104</sup>

To answer this question, Niou, Ordeshook, and Rose construct an *n*-person model of the international system. The model assumes that states seek to maximize their share of the system’s resources while preserving their own independence. It further assumes that (1) all states have perfect information, (2) resources are infinitely divisible and readily transferable, (3) states prefer to gain additional resources through negotiation rather than war, and (4) all

---

102. Emerson M.S. Niou, Peter C. Ordeshook, and Gregory F. Rose, *The Balance of Power: Stability in International Systems* (Cambridge, U.K.: Cambridge University Press, 1989).

103. In cooperative game theory, players can communicate and make binding agreements; in noncooperative game theory, binding agreements are forbidden and communication may or may not be permitted. A central question for cooperative game theory is what types of coalitions are likely to form among the players, as each seeks an arrangement that will maximize their own utility.

104. A system that is “resource stable” is also “system stable” (after all, eliminating an essential actor would by definition alter the distribution of resources), but it may be possible to alter the distribution of resources without eliminating an essential member.

states grow at an equal rate.<sup>105</sup> The model yields a healthy number of unsurprising results as well as a number of more interesting and counterintuitive predictions. In particular, Niou, Ordeshook, and Rose demonstrate that system and resource stability can be achieved simultaneously only when one state controls exactly 50 percent of the resources in the system.

The basic logic is straightforward: when one state has exactly 50 percent of the resources in the system, the others must cooperate with one another and isolate it, because any further increase in the strongest state's resources would allow it to absorb the others. Paradoxically, this result implies that any state (or coalition) that is facing the threat of elimination can avoid it by voluntarily transferring resources to the strongest state until the other state controls exactly 50 percent. Because all states know that this tactic is possible, certain distributions of power may be resource stable if all members realize that they could not unilaterally improve their share of resources (taking into account how the other members will respond). The model also implies that wars will never occur, because rational states would prefer to readjust resources through negotiation and voluntary transfers rather than through the use of force.

Niou, Ordeshook, and Rose test the model through an empirical analysis of European great power diplomacy in the period 1870–1914. In contrast to the statistical procedures used in *War and Reason*, their statistical procedures are explicated clearly and are generally convincing, and the historical narrative is based on an array of primary and secondary sources.<sup>106</sup> What is lacking, unfortunately, is a strong correspondence between the theoretical model and the empirical results.<sup>107</sup>

First, the bulk of the empirical testing involves a comparison of the gains from different alliance combinations, measured by an index of material capabilities. The authors predict that states will prefer coalitions that are just large enough to win (meaning they are larger than any combination of the remaining actors), while maximizing their own share of the overall alliance resources.<sup>108</sup>

---

105. Some of these assumptions are subsequently relaxed in order to analyze specific issues. In chapter 5, for example, they relax the assumption of equal growth in order to investigate the logic of preventive war.

106. In terms of pages, more than 25 percent of *The Balance of Power* is devoted to a discussion of measurement procedures, empirical tests, and historical narratives.

107. For a judicious but telling critique of the Niou, Ordeshook, and Rose results, see Snyder, "Alliances, Balance, and Stability."

108. Formally, they show that if  $g(c)$  equals a state's gains from alliance  $C$  and  $r(c)$  equals the total resources in  $C$  (i.e., the sum of each member's capabilities), then states in the system will choose allies in order to maximize  $g(c)/r(c)$ . They do not necessarily assume proportionality, however,

They present a series of tables summarizing the gains from different alliance combinations, and some of the results are clearly consistent with the model. Unfortunately, these results depend on a series of ad hoc restrictions that are wholly separate from the earlier theoretical analyses, and alliances that do not fit are explained by invoking other exogenous factors. Thus they arbitrarily exclude the possibility of a Franco-German alliance because of the conflict over Alsace-Lorraine, and they invoke German “mistrust” of Russia to explain why Germany chose a less profitable alliance with Austria. Similarly, Britain is excluded in some cases because its interests conflicted with several potential partners, even when the model predicts that these coalitions would have brought it greater resources. Instead of vindicating the formal model, in short, the empirical analysis ultimately relies on ad hoc factors like interest, revisionism, or ideology.

Second, the central mechanism contained in the model—the voluntary transfer of resources from declining state(s) to strong states (up to the level of 50 percent of total resources)—is largely absent from the empirical discussion. And when it does appear, the authors recognize that this mechanism is not a realistic possibility. Thus, although they mention that France and Germany might have been brought together by a German decision to relinquish Alsace-Lorraine, they add that this step “would almost certainly have led to the demise of the nascent German state” (p. 262). The absence of such transfers from the empirical account is not surprising, of course. Real states in the real world are notoriously reluctant to transfer resources (voluntarily) to more powerful rivals, and certainly not on the scale that is implied in the model.<sup>109</sup>

More generally, although the narrative in the empirical chapters is couched in formal terms, there is little direct evidence showing that policymakers made choices for the reasons depicted in the model. Such evidence may not be entirely necessary to prove the worth of a formal explanation, but a rigorous effort to test the theory would have sought at least some evidence indicating that states made alliance choices or war decisions in roughly the manner they imply. Instead of using history to test the model, in short, the model is used to organize the historical narrative.

---

which means that a state might in some circumstances receive gains larger than its initial contribution to the coalition. See Niou, Ordeshook, and Rose, *Balance of Power*, p. 220.

109. It is also worth noting that territorial cessions have been declining steadily over the past two centuries, probably as a consequence of the rise of nationalism and political participation in most great powers.



Finally, the historical analysis of World War I is unconvincing. Because the model views war as inherently irrational, the outbreak of fighting in July 1914 can be explained only by domestic politics, misperception, asymmetrical growth, or some other exogenous factor. Using the model and their statistical analysis as a guide, Niu, Ordeshook, and Rose reject the view that World War I was caused by German fears of rising Russian power. Instead, they argue that it was in effect a preventive war begun by Russia, intended to check the rise of German power. As Germany approached preponderance, the model says that it would be countered by a coalition of the rest. But because this response was not possible (for various exogenous reasons), they argue that Russia was forced to choose the second-best alternative of preventive war. Yet they offer no evidence that key Russian elites saw the July crisis as an opportunity for preventive war; on the contrary, there is abundant historical evidence that Germany was the driving force throughout the crisis. At best, the analysis of World War I is not a rigorous test of the model; at worst, it underscores its limitations.<sup>110</sup>

In sum, although the model is logically consistent and the authors make an admirable attempt to demonstrate its empirical value, *The Balance of Power* is not a persuasive demonstration of the power of formal theory. Formalization does not clarify the argument and does not lead to new, well-confirmed hypotheses, and the empirical evidence does not support the main theoretical claims. Although the authors deserve praise for their ambitious effort to combine rigorous formal analysis with careful historical research, the results cast additional doubt on the claim that formal theory is an intrinsically superior approach to the study of international politics in general and security affairs in particular.

### Conclusion

Several conclusions may be drawn from this survey of formal rational choice approaches to security studies. First, formal theory is most useful for enhancing the precision of a theory, and for verifying and refining its deductive logic.

---

110. In contrast to their behavior in the 1909 Bosnian crisis, German officials repeatedly pressed Austria to inflict harsh measures on Serbia in 1914. Key German officials were also obsessed with the specter of rising Russian power and the growing cohesion of the Triple Entente. As Chancellor Theobald von Bethmann Hollweg put it in 1909, "The future belongs to Russia, as it grows and weighs upon us like an ever-deepening nightmare." See David G. Hermann, *The Arming of Europe and the Making of the First World War* (Princeton, N.J.: Princeton University Press, 1996), p. 214.

This can be a valuable contribution, and provides ample justification for the continued use of formal techniques.

Second, formalization has not led to powerful new explanations of important real-world phenomena. For the most part, recent formal work has tended to take arguments derived from other scholars and place them in mathematical form. Such efforts have helped qualify and refine these existing theories, but the initial creative insights have generally come from scholars employing other approaches.

Third, recent formal work generally lacks rigorous empirical support. Formal theorists have devoted relatively little effort to testing their propositions, and the tests they have provided are often unconvincing. Although there are good reasons to value formal theory, in short, it should not be seen as inherently more valuable or “scientific” than other well-established research traditions.

Taken together, these characteristics help explain why recent formal work has had relatively little to say about important real-world security issues. Although formal techniques produce precise, logically consistent arguments, they often rest on unrealistic assumptions and the results are rarely translated into clear and accessible conclusions. And because many formal conjectures are often untested, policymakers and concerned citizens have no way of knowing if the arguments are valid.

In this sense, much of the recent formal work in security studies reflects the “cult of irrelevance” that pervades much of contemporary social science. Instead of using their expertise to address important real-world problems, academics often focus on narrow and trivial problems that may impress their colleagues but are of little practical value. If formal theory were to dominate security studies as it has other areas of political science, much of the scholarship in the field would likely be produced by people with impressive technical skills but little or no substantive knowledge of history, politics, or strategy.<sup>111</sup> Such fields are prone to become “method-driven” rather than “problem-driven,” as research topics are chosen not because they are important but

---

111. The field of economics offers a cautionary tale. A 1990 survey of elite economics graduate programs reported that 68 percent of the students believed a “thorough knowledge of the economy” was unimportant for professional success; indeed, only 3.4 percent thought such knowledge was “very important.” Similarly, the American Economic Association’s commission on graduate education warned in 1988 of “the extent to which graduate education in economics may have become too removed from real economic problems. . . . graduate programs may be turning out a generation with too many *idiots savants*, skilled in technique but innocent of real economic issues.” See Arjo Klammer and David Colander, *The Making of Economists* (Boulder, Colo.: Westview, 1990), p. 18; and Mayer, *Truth versus Precision in Economics*, p. 159.

because they are amenable to analysis by the reigning *méthode du jour*.<sup>112</sup> Instead of being a source of independent criticism and creative, socially useful ideas, the academic world becomes an isolated community engaged solely in dialogue with itself.<sup>113</sup>

Throughout most of the postwar period, the field of security studies managed to avoid this danger. It has been theoretically and methodologically diverse, but its agenda has been shaped more by real-world problems than by methodological fads. New theoretical or methodological innovations have been brought to bear on particular research puzzles, but the field as a whole has retained considerable real-world relevance.

By contrast, recent formal work in security studies has little to say about contemporary security issues. Formal rational choice theorists have been largely absent from the major international security debates of the past decade (such as the nature of the post-Cold War world; the character, causes, and strength of the democratic peace; the potential contribution of security institutions; the causes of ethnic conflict; the future role of nuclear weapons; or the impact of ideas and culture on strategy and conflict). These debates have been launched and driven primarily by scholars using nonformal methods, and formal theorists have joined in only after the central parameters were established by others.<sup>114</sup> Thus one of the main strengths of the subfield of security studies—namely, its close connection to real-world issues—could be lost if the narrow tendencies of the modeling community took control of its research agenda.

The solution should be obvious, however. Instead of embracing formal rational choice theory as the only true way to do science or seeking to banish

---

112. Even the mathematical economist Gerard Debreu, in a speech extolling the virtues of formalization, has warned that “the values imprinted on an economist by his study of mathematics . . . do not play a silent role: they may play a decisive role. The very choice of the question to which [a mathematical economist] tries to find answers is influenced by his mathematical background. Thus the danger is ever present that the part of economics will become secondary, if not marginal to that judgement.” See Debreu, “Mathematization of Economic Theory,” pp. 4–5 (emphasis added).

113. For a practitioner’s views, see David Newsom, “Foreign Policy and Academia,” *Foreign Policy*, No. 101 (Winter 1995–96), pp. 52–67.

114. An interesting example of this tendency is Rui J.P. de Figueiredo, Jr., and Barry R. Weingast, “The Rationality of Fear: Political Opportunism and Ethnic Conflict,” in Barbara F. Walter and Jack L. Snyder, eds., *Civil War, Insecurity, and Intervention* (New York: Columbia University Press, 1999), which formalizes several earlier writings on the origins of ethnic conflict. Their model implies that ethnic conflict is more likely when (1) endangered elites “gamble for resurrection” by stirring up ethnic hatred, (2) when the opposing groups act in ways that appear to confirm the elites’ claims that they are a threat, and (3) when the perceived cost of failing to heed these warnings is high. In other words, ethnic conflict is more likely when one ethnic group has reason to believe that the another group is hostile and that ignoring the potential threat might be fatal.

it from the field, members of the security studies profession should actively strive to retain the intellectual and methodological diversity of our field. Just as natural sciences profit from the fruitful collaboration of theoreticians and experimentalists, security studies should welcome contributions from formal theory, large-*N* statistical analysis, historical case studies, and even the more rigorous forms of interpretive or constructivist analysis.<sup>115</sup> Although individual scholars will emphasize different techniques in their own work and place different values on the contributions made by each approach, the field as a whole will be far richer if such diversity is retained *and esteemed*.<sup>116</sup> Given the continued relevance of security issues and the tragic consequences that accompany ignorance, it would be irresponsible to accept anything less.

---

115. Hacking once again provides the appropriate caution: "What is scientific method? Is it the experimental method? The question is wrongly posed. Why should there be *the* method of science? There is not just one way to build a house, or even to grow tomatoes. We should not expect something as motley as the growth of knowledge to be strapped to one methodology." See Hacking, *Representing and Intervening*, p. 152.

116. As Peter C. Ordeshook has observed, "Regardless of the mathematical rigor of our models, we need to drop the view of science as an enterprise directed by academics armed with theorems and lemmas or by experimentalists scurrying about in white smocks. Science proceeds less coherently, through induction and deduction informed by attempts to be practical and to manipulate real things, where those manipulations rely as much on experience, intuition, and creative insight as on theory." See Ordeshook, "Engineering or Science: What Is the Study of Politics?" in *Critical Review*, Vol. 9, Nos. 1–2 (Winter–Spring 1995), p. 180. See also the defense of methodological pluralism offered by Gabriel A. Almond in "Separate Tables: Schools and Sects in Political Science," *PS: Political Science and Politics*, Vol. 23, No. 4 (Fall 1988), pp. 828–842.