In this response to Stephen Walt’s critique\(^1\) of the application of formal analysis to international security studies, we take strong issue with a number of Walt’s arguments and assertions, and we try to clarify what we believe are his misconceptions about the nature and mechanisms of progress in scientific research. We begin, however, by identifying some of the issues we do not dispute with Walt. First, it is true that formal analysis, especially in the area of security studies, is only infrequently motivated by the attempt to explain some well-documented empirical regularity or universally recognized empirical anomaly. If there is room for disagreement here, it is the extent to which regularities or anomalies can be found in the security studies literature that are sufficiently precise to allow careful analysis. Second, there is little disagreement that some formalism exists for its own sake, although we need to be cautious here because much of this rigor seeks to understand the very definition of rationality in complex strategic environments. Third, despite the proliferation of competing models of deterrence, bargaining, coalitions, threats, and so on, those models are rarely set against each other for competitive empirical assessment. Finally, we cannot ignore the fact that very little of what researchers label “theory” is theory in any true sense, but instead is often best described as a demonstration of one’s ability to cobble together assumptions and derive something that can be labeled “lemma” or “theorem.”

**Rational Choice, Game Theory, or Formalism?**

Despite these agreements, we believe that much of Walt’s discussion is wrong-headed and counterproductive to his objective of sustaining a policy-relevant subfield of security studies. To begin, Walt’s article is not a dispassionate attempt at “evaluating the contribution of recent formal work in the field” (p. 8); rather, it is an unconstructive critique. But what is it a critique of—rational choice, game theory, or formalism? His article begins by bemoaning the limitations and increasing predominance of the rational choice paradigm

---


© 1999 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology.
(citing such “experts” as Chalmers Johnson). Then, via a superficial review of Bayesian analysis, Walt’s focus detours to game theory, but soon is directed at three things—the paradigm, game theory, and formalism—after essentially equating the paradigm with game theory and formalism. This blurred focus is occasioned, doubtlessly, by the fact that the paradigm, game theory, and formalism, though intimately related, are not equivalent: nonmathematical scholars such as V.O. Key certainly embraced rational choice perspectives; William Riker, arguably the father (or at least midwife) of rational choice thought in political science, rarely, if ever, proved a theorem and instead relied on the formal insights of others; and the use of mathematics often falls outside the domain of what anyone might argue is rational choice theorizing.

Walt’s true target, though, appears to be formalism. He gives only passing reference to the usual shopworn critiques of the rationalist paradigm (although he cannot refrain in footnote 35 on page 17 from swallowing the misconception that the paradigm presupposes people who are mathematical geniuses), and he seems only mildly discomfited by the folk theorems of game theory, which place the as yet unmodeled and poorly understood mechanisms of equilibrium selection at the center of any complete theory of social processes. The reason for Walt’s redirected focus, we suspect, is that the study of international relations and security is a subfield of political science that has long accepted the rationalist premise of self-interested action and depended, albeit imprecisely, on the strategic imperatives of game theory (recall that much of game theory’s early development was motivated by the inherent inadequacies of other modes of analysis into strategic matters). Here, of course, we need not refer only to contemporary scholars, but to our classical predecessors (e.g., Taylor, Morgenthau, Claude, etc.). Notions of rationality, self-interest, and strategy permeate these earlier writings, while debates among their heirs (e.g., Waltz, Keohane, Jervis, Gilpin, etc.) are more likely to concern individual motives, perceptions and beliefs, the role of institutions in constraining individual action, and the specification of strategic environments.

Consistency

It would seem, then, that there should be considerable room for agreement among us. Unfortunately, that room is circumscribed by Walt’s failure to understand at least five aspects of scientific study: (1) the value of consistency; (2) the creativity inherent in a formal model’s development; (3) the necessity of proceeding with clearly explicated null models; (4) the practical nature of
assessing the empirical content of a model; and (5) the different ways in which science advances in other disciplines.

Our difficulties begin with Walt's discussion of the relative value of logical consistency and precision. Put simply, Walt stacks the deck against formal analysis by asserting that originality and empirical validity are more highly valued than are logical consistency and precision: “Although all . . . are important, the latter two criteria—originality and empirical validity—are especially prized” (p. 13). But while an argument can easily be original if it is incoherent, illogical, or imprecise, we do not see how any idea, hypothesis, or argument can be empirically valid if it is any of these things. Punch lines sustained by obscure or disconnected reasoning may be true, but only accidentally so. And how do we know they are true if their opposite cannot be refuted because we do not know what that opposite is owing to vague conceptualizations? To have content, all arguments must possess domain constraints in the form of initial assumptions, a definition of terms, and evident logical connections, and we cannot assess empirical validity if, as is the case with much of what is labeled “theory” in security studies, one or more of these things is nonexistent. To impose only the criteria of originality and the appearance of empirical validity establishes a field upon which the rational choice theorist quite properly rejects playing.

In the same vein, Walt laments that formal essays are difficult to read, because doing so requires wading through pages of dense notation and mathematical argument to learn the hidden assumptions and restrictions of the analysis. This lament, however, is different than Walt’s assertion that “formal methods . . . 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” (p. 20). That assumptions are sometimes less than self-evident is true. But, and this is the big “but,” it is hardly an established fact that formalism makes it easier (than what?) to camouflage the logic of an argument. On the contrary, a careful reading will either uncover all assumptions or reveal their absence if the analysis is logically flawed. Walt may bemoan his difficulty with discovering or understanding assumptions that are formally stated, but to assert that they cannot be discovered by those who make the effort is patently false.

Having opted for the artful values of originality and the appearance of empirical validity, Walt fails to acknowledge fully that logical consistency, precision, and the attendant discoverability of assumptions are not uniform characteristics of much of anything else. If consistency, precision, and discoverability did characterize other less formal approaches, formalism and even the
rational choice paradigm would not have gained the foothold in our discipline that so concerns Walt. It is silly to deny that a reasoned and informal contemplation of events and processes cannot yield insights that often move the focus of scientific investigation. But the rational choice paradigm and formalism are not mushrooms that sprung up in an unattended intellectual forest. They are reactions to a discipline mired in imprecision, vagueness, obscure logic, ill-defined constructs, nontestable hypotheses, and ad hoc argument. They are a reaction to a discipline that in the 1920s proclaimed the Weimar constitution the greatest political-intellectual achievement of its age; a discipline that in the 1960s substituted correlation for cause; a discipline submerged in such conveniently vague and ill-defined ideas as “power,” “leadership,” “authority,” “group,” “alliance,” “function,” “ideology,” “culture,” “regime,” “stability,” and “balance.” They are reactions to a discipline that substituted the well-turned phrase for concrete constructs, operational measures for theoretical primitives, and the gloss of methodological sophistication for true theory. They are, in short, a reaction to a discipline that substituted the well-turned phrase for concrete constructs, operational measures for theoretical primitives, and the gloss of methodological sophistication for true theory. They are reactions to a discipline that does and does precisely what Walt critiques the formal analyst of doing—burying key assumptions in an indecipherable format, although generally that format was a language more to the liking of those who studied French and Plato in college rather than calculus.

As part of his critique of the weight that formal analysts place on consistency and precision, Walt acknowledges that the limitations of the rationalist paradigm and game theory “do not discredit the use of formal models.” But he then attempts to advance the argument that “the potential gains in precision and logical consistency do not demonstrate the superiority of formal techniques over other approaches” (p. 20). We might agree if we knew the identities of these other approaches. If Walt’s argument is to be constructive, it is incumbent that he establish a better treatment of the specific difficulties noted by “other approaches,” for as we are all taught in elementary philosophy of science courses, we reject a theory only when a better one becomes available. The questions that immediately come to mind here, of course, are: Is there a theory of risk that we can substitute for Bayesian analysis? Where in the study of international relations do we find a more coherent or empirically valid theory of strategic complexity? What other part of social science treats the indeterminacies of social processes that game theory’s folk theorems uncover?

Creativity

Perhaps our sharpest disagreement with any specific point in Walt’s argument, however, is his assertion that “technical sophistication and logical consistency
did not yield particularly creative or original results” (p. 26). Again, this is less a coherent argument than a simple expression of prejudice, because logical consistency is itself a profoundly important creative contribution. How many trees have been cut to publish attempts at explicating Robert Keohane and Joseph Nye’s synthesis of realist and liberal perspectives, Waltz’s alternating conceptualizations of realism, the true content of neorealism, or, more to the point of Walt’s substantive interests, the preconditions for a viable deterrent strategy, the avoidance of trade wars, and stable alliances? The weaving together of a complex argument that appears to be consistent with some real-world process is valuable, just as it is useful to secure qualitative insights into the likely meaning of events and their preconditions. But that is something different from logical consistency. This is not to say that such consistency is always absent from verbal or purely statistical examinations of political events and processes, but it is not their uniform characteristic. It may be true, moreover, that science often advances without first establishing the logical foundations of an idea, but ultimately, those foundations must be established before the word “theory” can be uttered with any meaning and the true value of an insight established. To rank raw conclusions above logical consistency in the overall scheme of what we demand from ourselves as scientists, as Walt does, or to judge the construction of a mathematically precise argument that establishes a sufficient condition for something to be true as less original than the addition of a suggestive, yet vague or incomplete conceptual scheme is not a constructive comparative assessment of alternative intellectual approaches.

The preceding rejoinder, however, gives too much legitimacy to Walt’s critique. Consider, for example, the issue of the causes of war. Kenneth Waltz, in Man, the State, and War, offers a puzzle, but not one that he solves. Geoffrey Blainey, in The Causes of War, helps resolve this puzzle by suggesting that the problem concerns the extent to which, using the game theorist’s jargon, the assumption of common knowledge is not satisfied. Finally, Bruce Bueno de Mesquita and David Lalman, in War and Reason—one of Walt’s whipping boys—dissect Blainey’s hypothesis and begin the search for the conditions under which incomplete information rather than common knowledge is the critical parameter in determining the likelihood of conflict. Walt may quibble about their empirical methods, but this sequence of intellectual developments is precisely what any scientist would want to see—paradox, hypothesis, and a logical refinement that clearly differentiates the alternative possibilities. Because understanding requires each step in this process, to judge one step more or less original than the others is nonsense.
Originality in formal analysis does not reside, moreover, in the mere derivation of some result. Given that Walt apparently values models and modeling so little, he cannot see the level of creativity that often goes into a model’s design. A precise specification of the problem and the attendant assumptions and constructs do not appear out of thin air. Indeed, the process of model construction (which is often separate from analysis) is much like the informal contemplative processes and freewheeling imagination that Walt so admires. But even this assessment understates the originality of ideas to be found in formal analysis. Consider again the folk theorems of game theory. At first glance, those theorems appear to be mere statements of mathematical logic pertaining to abstract notions of equilibria. Moreover, Walt refers to them as if they were limitations of game theory. We would suggest, however, that barring a demonstration of the reliance of those theorems on constructs and assumptions that other approaches can avoid while treating the same issues, those constructs and assumptions are no more limitations of game theory than they are of any theory. More to the point, however, even if it is true that formal analysts must appeal to such ideas as culture and norms as a way of refining a game-theoretic prediction, we at least have learned the role of such ideas in specific social processes and the theoretical constructs required to study their genesis and evolution. Indeed, it is only the game-theoretic perspective and its attendant formalism that have brought the problems of indeterminacy, coordination, and equilibrium selection to light. And as a consequence, they can now be used to explore such issues as the sources of stability in constitutional design; the possibility that the disagreements over paradigms in international relations concern only an assessment of the likelihood that one equilibrium versus another will be feasible; and, in that context, the functions performed by international organizations, trade, and parallel political structures. If Walt does not see these contributions as original, then his definition of originality must encompass only the generation of incomplete, atheoretical insights.


Null Models

Walt stacks the deck against formalism in another way. Much of his discussion is framed by the ostensible purpose of “evaluating the contribution of recent formal work” (p. 8). Any scientific assessment, however, requires a clear and reasoned null hypothesis that is sustained when the evidence fails to support its alternative. In Walt’s case, that null appears to be a rather imprecise “inquiry in any other form.” Time and again, he tells us of the profound insights gained by other approaches that, we presume, are either inherently beyond the reach or somehow beyond the capacity of formal analysts—profound insights that, if we are to judge by what formal analysis lacks in Walt’s view, are “well-verified empirical predictions” (p. 6) that have “been tested in a careful and systematic way” (p. 8). Unfortunately, with the exception of a single footnote citing a few empirical studies of which he approves (see p. 30 n. 72), we are at a loss to learn the identities of those systematically tested hypotheses. Walt cites approvingly Waltz’s influential book (p. 17), Theory of International Politics, but the editions we see are missing the chapters that offer a critical empirical test of any specific hypothesis. Walt’s own The Origins of Alliances is an admirable effort at dissecting the processes whereby alliances form and dissolve—one we have found useful in our own research—but eighteen tables (of which nine merely summarize the historical record discussed in the text, four summarize the author’s interpretation of specific events, four offer macrodata that can be gleaned from standard sources, and one offers an ad hoc index of “capabilities”) hardly qualify as a rigorous empirical test of anything.4 Walt cites the less formal work of Thomas Schelling, Daniel Ellsberg, and Mancur Olson, although he fails to note the more formal contributions of that age by Kenneth Arrow and Duncan Black or John von Neumann and Oskar Morgenstern’s seminal volume, Theory of Games and Economic Behavior, upon whose mathematical shoulders all the above stood (though the shoulders upon which Olson stood were those of economists who formalized the preconditions for market failure in the presence of externalities). But with respect to strategic studies, we are unaware of much that meets Walt’s demand for systematic testing. Certainly, Walt cannot be referring to the demarcation of ideas among realists, neorealists, liberals, and neoliberals; to any conclusions pertaining to the relative importance of domestic politics; or to insights into the operation of strategic deterrence. And we doubt he is referring to the mainstay “fact”

that democracies never war among themselves, and the dependence of this conclusion on ad hoc operationalizations of the concept of democracy and measures of the severity of conflict.

Walt’s failure to contrast clearly the accomplishments of formal analysis with “other approaches” is all the more frustrating because it is true that our less formal intellectual predecessors offer a plethora of valuable insights and ideas that frame the research most students of international affairs, regardless of persuasion, pursue. But in most cases, that research takes the form of attempts to resolve some ambiguity or inconsistency in arguments, for, as we believe any true comparative assessment would show, those predecessors offer a full plate of competing, contradictory, imprecise, and incomplete arguments, hypotheses, and perspectives. It may be true that much of what the formal analyst offers as “substantive conclusion” is well understood or at least consistent with some prior argument (we know colleagues who claim that everything can be found in Aristotle). But showing that a prior conclusion follows logically from some set of initial assumptions is a form of reproducibility that science demands—it tells us that the models in question are not mere fantasy and may not even be fundamentally flawed. Indeed, such redundancy is a form of empirical test that Walt demands to the extent that the initial ideas or conjectures arose from empirical observation. But perhaps more important, we learn something else from the “proof” of an otherwise known conclusion or assertion—we learn the hidden assumptions or incomplete logical connections of those earlier arguments, the sufficient conditions for their validity, and, if we are lucky or sufficiently imaginative, their necessary conditions. Indeed, if these assumptions and logical connections were not hidden or incomplete, the prior arguments leading us to them could not have been any less formal than the essays Walt singles out for criticism.

**Testing**

The fundamental problem with Walt’s critique, however, is not the prejudices it reveals or its failure to do what is demanded of the formal analyst—comparative empirical assessment. Rather, it lies in Walt’s failure to understand the methods whereby formal and empirical analyses complement each other in any real ongoing scientific enterprise. Walt requires direct empirical application and assessment. Yet the researchers he cites approvingly—Schelling and Olson—as well as Riker, for example, apply formalism differently. Schelling’s seminal contribution, *The Strategy of Conflict*, is neither strict formalism nor
strict empiricism, but interpretation. He takes a few formal lessons of game
theory (e.g., the possibility of multiple equilibria and the importance of se-
quencing in a game’s extensive form) and transforms those ideas into useful
applications (to, in the cases cited, the importance of mechanisms of coordi-
nation and the nature of viable threats). Olson nowhere offers a test of the
theorems about market failure in the presence of externalities upon which his
analysis relies (and we are not aware of any prior tests of those theorems in
the literature), but by expanding the domain of those rigorously derived
results, he opens the door to additional theorizing about substantive matters
(e.g., political entrepreneurship and the applied subfield Elinor Ostrom terms
“common pool resource” issues). Riker offers perhaps a clearer example of
sophisticated “application.” Rather than suppose that Arrow’s Impossibility
Theorem (an exercise in pure deductive formalism) or McKelvey’s wholly
abstract investigation of the properties of the majority preference relation in
generalized spatial preference structures offer results that require direct testing,
in Liberalism against Populism, Riker uses these ideas to compel a confronta-
tion between two seminormative paradigms of democratic theory, each with exten-
sive policy-relevant implications.6

Schelling, Olson, and Riker do not propose to test anything directly, but
instead their discussions are skillfully woven and interpreted elements (theo-
rems, structures, and propositions) of formal theory. Walt may lament the small
number of such volumes (just as we all do and just as we all should lament
the limited number of truly seminal and clear presentations of alternative
paradigms in international relations), but we need to understand that the
formal essays Walt critiques for having failed to be either sufficiently original
or for not offering a direct empirical assessment of their conclusions are only
small pieces of a larger puzzle—bricks in an incomplete wall. That the wall is
incomplete, however, is merely a call to use those bricks and replicate the
talents of Schelling, Olson, and Riker. Riker, in fact, knew only the rudiments
of game theory, and barely concerned himself with the proof of such things as
McKelvey’s instability theorem. But rather than decry the impenetrability of

---

5. Elinor Ostrom, Governing the Commons: The Evolution of Institutions for Collective Action (New
the Theory of Social Choice (Prospect Heights, Ill.: Waveland Press, 1982). The essays we cite here as
forming the basis of Riker’s analysis are Kenneth J. Arrow, Social Choice and Individual Values (New
Haven, Conn.: Yale University Press, 1951); and Richard D. McKelvey, “General Conditions for
that proof (and its reliance on such concepts as continuous and convex preference sets, measurable spaces, upper-semicontinuous functions, etc.), or bemoan the fact that the proof was accompanied by a minimal substantive assessment of its implications or assumptions, Riker succeeds in reshaping our thinking about democratic elections and referendums, the potential for majority tyranny in a democracy, and the strategic options of otherwise disadvantaged political actors. If Walt complains that he cannot find the development of equivalent applications in the formal literature of strategic studies, then one reasonable response is to ask the critic to engage in the interpretive discourse he believes is essential and valuable and set his ideas before us.

The Process of Scientific Development

To comprehend fully our disagreements with Walt, however, it needs to be understood that the development of formal models and their interaction with the empirical world often proceed differently than do the methods of inductive reasoning and “informed opinion.” First, by the interchangeable use of words, Walt suffers a confusion between “model” and “theory” when he writes as if they are equivalent. They are not. Theory in our paradigm encompasses a single entity—game theory and its underlying constructs—whereas the research erected on this structure is best described as modeling. It may be true that if models are sufficiently “connected” and if they concern a broad enough range of substantive matters (e.g., microeconomics), then there is little lost in labeling the package “a theory.” But nothing approaches this packaging in strategic studies. Of course, a distinction between model and theory does not by itself detract from Walt’s core critique—that formal analysis lacks a coherent body of empirically tested or testable propositions. But the distinction does explain that absence as well as the presence of “testing” that Walt deems unsatisfactory. Here let us turn to Walt’s discussion of our own work, The Balance of Power, which he admittedly treats more kindly than even we might (owing in part to our research subsequent to that effort). Two aspects of our analysis bother Walt. First, he argues that our “empirical analysis ultimately relies on ad hoc factors.” Second, he states that “instead of using history to test the model, the model is used to organize the historical narrative” (p. 44). Although we fail to see how this differs from what passes as empirical analysis in strategic studies (or how it differs much from Walt’s own research into alliances), we nevertheless take these “criticisms” as compliments. To suppose that a formal model can wholly encompass a complex process that stretches
over some forty-five years without resorting to some ad hoc discussion is ludicrous. By definition, that reality must be orders of magnitude more complex than any model, in which case the question becomes: Can the events and patterns we perceive as reality and our understanding of it be organized by anything at all? If a few pages of notation does this in even a modest way—and if, in particular, we can identify those events and patterns that are not yet accommodated by a model and gain a sense of what assumptions are not satisfied in each deviant case—then certainly the analyst has accomplished a great deal.

More specifically, however, it is Walt's first comment that is the more bothersome, for it is here that he reveals his faulty image of empirical scientific research—something akin to researchers scurrying about in white smocks, conducting critical experiments. In fact, most of science consists of the ill-defined and even sometimes random "play" between models, theories, and a complex reality. Rarely does science consist of definitive hypothesis testing, and then only in very restrictive and controlled environments. The reason for this is simple: unless we are concerned with some utterly basic scientific issue, reality is far too complicated to be accommodated in any straightforward way by any simple tractable model. And one implication is that, in general, scientific testing is an imprecise, often informal process. Put simply, productive empirical research, whether in political science or any other field, proceeds differently than what Walt demands of the formal analyst.

The study of politics is, as we argue elsewhere, a field more akin to engineering than to science.7 Of necessity, our discipline must deal with phenomena that are both too complex for simple, closed-form analysis and too complex for the imprecision of other approaches. This, perhaps, is the attractiveness of imprecision and journalistic discourse—it gives the impression of understanding without revealing the inherent inadequacies of our ideas. But in the natural sciences, we typically learn things about complex systems through an informal, hands-on, trial-and-error process—a process that is informed not merely by careful empirical analysis but also by the failure of models to perform adequately. In trying to solve real-world problems, the natural scientist and engineer often (always?) confront questions they cannot answer theoretically, and the "solution" is an ad hoc "filling in" of the analysis, sometimes using "good guesses" and other times particularistic experiments that Walt

---

might term “ad hoc” (e.g., wind tunnels), with the hope that over time and through experience, this filling in will become routinized and even generalized. The practicality of the problems treated, combined with accountability (on the part of the researchers themselves) for faulty judgments, imposes a degree of rigor on this process whereby, as if in some giant bookkeeping enterprise, anomalies and empirical regularities are collected and stored, awaiting theoretical refinement and generalization.

Absent this practical accountability and implicit bookkeeping, much of what passes for “theory testing” in political science is, in fact, the search for empirical regularities that warrant theoretical explanation. But more to the point, this practical interplay is largely absent in political science, and we do not consistently confront the real world in a way that involves sanctions for erroneous advice and predictions. Not much is lost to the rest of the world, or even to the authors themselves, if some argument or conceptual scheme proves worthless or wrong, just as not much is lost to anyone if a student of neorealism, Soviet studies, crisis bargaining, or deterrence theory publishes a volume replete with ambiguity, misconceptions, misperceptions, illogical inferences, or just plain dumb advice. In short, political science lacks the feedback among theory, model, and application that characterizes discovery and empirical assessment in other fields, including even economics. But it is here that logical consistency and rigor gain their greatest advantage for advancing political science as a substantively useful, policy-relevant discipline. Imprecision, vagueness, ambiguity, and the like allow researchers to dodge responsibility behind the shelter of a “reformulation” or “reinterpretation” of their ideas: “The analysis was not wrong, only poorly applied or interpreted.” This dodge, however, is far less feasible with rigorous analysis if only because its limitations and failings are, by virtue of its rigor, more readily apparent. And when failure occurs, as it inevitably must, the transparency of formalism allows for the precise cataloging of error, which, of course, is why rigor is valued and why those who eschew it find it uncomfortable.

Finally, we are puzzled most of all by Walt’s assertion that “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)” (p. 46). Even if we were to agree with this statement, we would add that the contributions of Walt’s “other approaches” to this list of security issues escape
us as well. But the list is revealing, for it is the product of someone concerned not with science and empirical regularity as those terms need to be understood for the development of cumulative knowledge, but instead with the commentary and informal discussion we find in newspapers and popular journals that has too long appeared under the label "political science." Such discussion and commentary may be entertaining and even sometimes enlightening, but it remains mere journalism until it can be given the solid scientific grounding that formal theorists pursue.