Rational choice theory has long been criticized, with several recent critiques of applications in political science in general and international relations in particular. In this paper, I seek to clarify discussion about rationality in several ways. I discuss the role of assumptions in theory and the assumption of rationality in rational choice theory. I demonstrate that many critiques of rational choice theory have been misguided, since instrumental rationality—the rationality of rational choice theory—is applicable to a wide variety of social situations, including situations where various psychological, informational, and structural factors claimed to interfere with rational decision-making are present. I also focus on three applications of rational choice theory in international relations that demonstrate ways that rational choice theorists have potentially added to confusion about the rationality assumption and created disillusionment in others about the rational choice enterprise. Finally, I point out that, since rational choice theory is not a single theory, critics’ claims about the non-falsifiability of rational choice are immaterial. An understanding of these issues facilitates movement by the international relations community away from argument over the assumption of rationality and on to more productive and integrative theory building.

KEY WORDS: rational choice theory, assumptions, preferences, instrumental rationality, philosophy of science

Rational choice theory has long been criticized, with several recent critiques directed at both applications in political science in general (Green and Shapiro, 1994; also see Friedman, 1996) and international relations in particular (Walt, 1999a, b). The debate over rational choice theory is by no means limited to political science.

Received for publication 11 April 2001.
A previous version of this paper was presented at the 2000 Annual Meeting of the Midwest Political Science Association, Chicago, Illinois. I would like to thank Frank Zagare, Paul Senese, and Michael Simon for helpful comments and suggestions on various aspects of this paper.
Address correspondence to Stephen L. Quackenbush, University of Missouri, 113 Professional Building, Columbia, MO 65211. E-mail: stephen.quackenbush@us.army.mil

87
Rational choice theory has been used throughout social science, and rationality has been debated recently in other areas of social science as well (e.g., Coleman and Fararo, 1992; Halpern and Stern, 1998).

Green and Shapiro, in *Pathologies of Rational Choice Theory* (1994), assess rational choice applications in American politics. They contend that rational choice theory has been marred by a number of methodological pathologies. Among these are *post hoc* theory development (where theory is modified to account for anomalous instances), poorly formulated empirical tests, and bias in cases selected for testing. They review rational choice explanations of voter turnout, social dilemmas and free-riding, legislative behavior, and electoral competition, and determine that “despite its enormous and growing prestige in the discipline, rational choice theory has yet to deliver on its promise to advance the empirical study of politics” (Green and Shapiro, 1994, p. 7).

Walt (1999a) aims his critique at formal rational choice theory in security studies. He argues that theory should be judged by its logical consistency, originality, and empirical validity. Furthermore, he argues that although “all three are important, the latter two criteria—originality and empirical validity—are especially prized” (Walt, 1999a, p. 13). He then conducts a review of several formal rational choice works in an attempt to demonstrate that they have yielded trivial results, have not been empirically tested, and that empirical tests, when used, have been constructed poorly.

Green and Shapiro’s essay renewed a continuing debate over the nature, successfulness, and usefulness of rational choice theory in political science. Several responses to Green and Shapiro are contained in the volume edited by Friedman, *The Rational Choice Controversy* (1996). Similarly, several responses to Walt’s critique were published in the Fall 1999 issue of *International Security* (Bueno de Mesquita and Morrow, 1999; Martin, 1999; Niou and Ordeshook, 1999; Powell, 1999; Walt, 1999b; Zagare, 1999). Several other contributions to this debate have also appeared in the last decade (Amadae and Bueno de Mesquita, 1999; Cox, 1999; Jones, 1999; Kaufmann, 1994; Opp, 1999; Zagare, 1990).

Despite the lengthy, and at times pointed, debate over rational choice theory, misconceptions of rational choice in the political science literature remain. This paper seeks to advance discussion regarding rationality in several ways. First, it clarifies the role of assumptions in theory by developing a theoretical model highlighting error due to assumptions. It then discusses the nature of rational choice theory and explores the question of whether actors are really rational.

In this paper I argue that much of the criticism of rational choice theory has been based on a basic misunderstanding of the assumption of instrumental rationality—which is, after all, the rationality of rational choice theory. To alleviate this confusion, I explicate the assumption of instrumental rationality and demonstrate that it is consistent with a broad range of motivations and behaviors that critics have claimed contradict it. Next, I focus on three applications of rational choice theory in international relations and demonstrate ways that rational choice theorists themselves have potentially added to confusion about the assumption of rationality and created disillusionment in others about the rational choice enterprise. Finally, I demonstrate that, since rational choice theory is not a single theory, critics’ claims about the universalistic aspirations of rational choice and *post hoc* theory development are unmerited.
Although some of these points have been made by others, many have not. Furthermore, the fact that the conceptual confusion continues to pervade scholarly discourse demonstrates the need to state them again. This is vitally important to the field of international relations, for once the rationality postulate is more properly understood, we can move beyond nonproductive debates about a simple assumption and on to the more meaningful task of explaining international interactions.

THE ROLE OF ASSUMPTIONS IN THEORY

Much of the debate about rational choice theory is fundamentally a debate about an assumption. Before turning to more specific discussion of the rationally assumption, it is useful to consider the role of assumptions in theory. The goal of science is to explain the world in which we live.\(^3\) In international relations, scholars seek to explain why nations go to war, why they align with one another, or any number of other questions dealing with the relations between actors on the international stage. Explanation is accomplished through theory. However, there is much disagreement among political scientists about what theory is. Rapoport (1961) provides a list and explication of various meanings of theory. Although certainly not the only definition, Brodbeck (1959, p. 378) defines theory as

\[
\text{a deductively connected set of laws. Some of these laws, the axioms or postulates of the theory, logically imply others, the theorems. The axioms are such only by virtue of their place in the theory.}
\]

The foundation of a theory, and thus of its theorems, is assumptions. The usefulness of a theory in explaining the world is determined through testing its theorems by observation, whereas the truthfulness of an empirical generalization is only judged logically through theory (Bueno de Mesquita, 1981, pp. 9–10).

There are two types of assumptions (or axioms) in theory that are useful in simplifying the complex world the the theory seeks to explain. The first type of assumption is about questions whose answers cannot be proven. Since the “true” answer cannot be known, an assumption one way or another must be made. The second type of assumption deals with questions whose answers can (at least in principle) be determined empirically. However, these entities are assumed to take particular values in order to simplify theory. Neither type of assumptions is statement of “fact.” As Most and Starr state,

\[
\text{axioms are simply dichotomous variables that analysts either choose or are forced to leave unmeasured. . . . [A]xioms should not be viewed as assertions of universal truth. . . . Once this is seen, one can move quickly beyond debates over the universal truth or falsity of the assumptions themselves (1989, p. 112, emphasis in original).}
\]

The assumptions are not, or at least not known to be, universally true (if something is true, then it does not need to be assumed). How do assumptions aid theory construction? Consider a theory of the form
where $D$ is the dependent variable, $I_i$ represents the independent variables, $i$ varies from 1 to $k$, and $k$ is the number of variables. Presumably, if the model is completely and correctly specified and there is no measurement error, then there would be no random error; the prediction would be deterministic. Since the complete absence of measurement error is highly unlikely, if not impossible, models take the form

$$D = f(I_i) + u(e)$$

where the random error, $u$ is a function of the measurement (and misspecification) error $e$, $j$ varies from 1 to $n$, and $n$ is the number of observations. For the sake of parsimony, assumptions (of the second type) are made. These may be explicitly stated, such as assuming the state is a unitary actor. However, these may also be unstated, and thus (intentionally or not) implicit. Inherent in the selection of some list of variables or explanatory factors is an assumption that nothing else matters (Hempel, 1949). The theoretical model now takes the form

$$D = f(I_i) + u(e, a)$$

where the error term is now a function of measurement error and error due to assumptions, $a$. To the extent that any given assumption is not true for a particular case error is introduced. For example, consider a particular variable that is assumed to equal 1 throughout. In reality, it only equals 1 in 900 of the 1000 cases examined. For each of those 100 cases that the variable does not in fact equal 1, error has been introduced by the assumption.

Possibly recognizing the errors that can result from making assumptions, some scholars argue that assumptions need to be true in order to be useful. For example, Schrodt (1985) provides

not only a treatment of the discrepancies between the assumptions of rational choice models and the empirical evidence, but a rejection of the widely accepted notion that our theoretical assumptions need not be empirically correct as long as they are ‘useful’ in the sense that the models incorporating these assumptions turn out to provide accurate predictions (Singer, 1989, p. 11).

However, arguing that assumptions need to be “empirically correct” in order for the theory to be useful demonstrates a misunderstanding of the role of assumptions in scientific theory. It is also logically incorrect. For example, Van Dyke (1960) argued that if a premise in a deductive argument is false, then the conclusion will be false. However, logic tells us that if “the conclusion of a deductive argument is false, then at least one premise must be false. . . . From knowledge that a premise is false, however, we cannot say anything about the truth or falsity of the conclusion” (Hobbs, 1993, p. 2). For example, consider the argument
My cat is a table

All tables have four legs

Therefore, my cat has four legs.

Notice that each of the premises (which can be thought of as assumptions) is false. However, the conclusion is quite true. Considering the error due to assumptions discussed above, notice that in this example the error introduced by the second premise cancels out the error introduced by the first.

Although this example is clearly simplistic, there is no good reason that false assumptions cannot lead to correct conclusions in more complex deductive arguments (theories). The usefulness of assumptions is determined by the empirical accuracy of the theory that results from them. The accuracy of a theory is determined through empirical testing of its predictions. Standard levels of statistical significance are generally used to determine whether or not a theory is supported by the data. Testing a theory’s assumptions does not provide a test of the theory’s predictions (Friedman, 1953), although an examination of assumptions can be useful in other ways (Morton, 1999).

If the predictions of a theory are not supported, the culprit is the error term in Equation 3. If measurement (and misspecification) errors have been minimized, then the error resulting from assumptions is the cause. There are two solutions for the latter problem. First, the applicability of the theory can be restricted to a subset of the population where the error from assumptions is not too great. For example, this results in what Most and Starr (1989) term “nice laws.” However, it is possible that no such subset exists. Second, one or more assumptions can be changed in order to reduce the error from assumptions. Of course, changing an assumption requires that the theory (if logically consistent) be altered as well.

For example, the vast majority of aerodynamic theory (and fluid dynamics in general) rests on the assumption that air is a continuum. This is, of course, false. Air is actually made up of a number of discrete molecules that bounce off one another. Air is therefore not a continuum. Everyone knows this, but engineers continue to assume that it is. Why? Because (except for at extremely high temperatures and altitudes) it is very useful. The assumption simplifies the real world by eliminating unnecessary complexity and detail. There are a number of other standard assumptions used to study aerodynamics, such as the assumptions that air is frictionless or incompressible, none of which are empirically correct. Virtually all of science and engineering is rooted in assumptions that are untrue. The more that is assumed, the simpler the theory. However, many assumptions also limit the range of applicability of the theory.

To summarize: all theories rest on assumptions, regardless of whether these assumptions are made explicit. Assumptions should be judged not by whether or not they are true, but by their usefulness. The usefulness of assumptions is determined by the empirical validity of the theory that derives from them. If assumptions are too heroic, and the resulting error is too large, the theory cannot be empirically verified. With this understanding of theory and assumptions in mind, it is useful to consider
the nature of rational choice theory.

WHAT IS RATIONAL CHOICE THEORY?

Despite Green and Shapiro’s (1994) characterization of the universalistic claims of rational choice theory, rational choice is not a ‘theory’ in the sense of the term as used above. Rather it is a descriptive phrase used to describe any of a number of individual theories that use the rationality assumption. In addition, rational choice assumes that outcomes are the result of choices made by actors. There are three separate rational choice approaches: non-formal theory, expected-utility theory, and game theory.

It might seem surprising to some that I include nonformal theory as a type of rational choice theory. For example, Walt (1999a) surveys and critiques many recent rational choice works in security studies, but it is immediately clear that he equates “rational choice” with “formal.” However, there is nothing inherent in the assumption of rationality (which is, after all, what rational choice theory is based on) that requires the use of formal models. While game theory and expected-utility theory have proven useful in deducing the logical implications of certain assumptions, rational choice theory can exist without these formal methodologies.

A common criticism of rational choice theory is that real decision-makers are not rational. They are constrained by institutions, cultural influences, or psychological limitations that make the assumption of rationality problematic at best, or foolhardy at worst. This criticism is based on a misunderstanding of both the assumption of rationality in rational choice theory and the use of assumptions in theory in general. Next I illustrate this contention by discussing Abelson’s essay “The Secret Existence of Expressive Behavior.”

The Unsurprising Existence of Expressive Behavior

Abelson (1996) illustrates the misunderstanding of the assumption of rationality. Entering the debate about rational choice theory “with a certain sense of aggravation,” Abelson (1996, p. 26) offers several examples to support his claim that “it is no surprise that rational choice models receive such poor empirical support … {since} the axiomatic requirements of rational choice are too demanding.” Throughout his work, Abelson provides numerous examples to demonstrate that real actors are not rational.

Abelson begins by relating the story of the “doter’s paradox.” A hypothetical rational choice theorist is supposedly at a loss to explain why John Doe dotes on his granddaughter every fall. Because the man does not expect that his doting will make any difference in his granddaughter’s life, Abelson claims that his behavior is inexplicable by rational choice theory. However, “John Doe traveled a long distance for something he claimed was important” (Abelson, 1996, p. 25). By doting, he was simply doing that which he preferred to do; hence, he was acting entirely rationally.

Despite using the term instrumental rationality, Abelson misconstrues the assumption of instrumental rationality. He claims that “political attitudes about many issues, including health care, abortion, welfare, and school busing depend surprisingly more
on symbolic factors than on self-interest” (Abelson, 1996, p. 28). A self-interested actor (in the sense of rational choice theory) is one that acts according to his or her own preferences. Thus, if these symbolic factors determine actor’s preferences, then they determine the actor’s self-interest. Abelson has created a false dichotomy between symbolic factors and self-interest.13

Furthermore, Abelson incorrectly characterizes economic gains as being the only allowable influence on instrumentally rational actors’ preferences. An in-depth discussion of his final example will illuminate these issues. Abelson describes an experimental demonstration where subjects were given

A choice between two proofreading tasks with different rates of pay per unit of text. One task was to proofread an address directory—pretty boring work. The other was proofreading personal ads, a job with an appeal that people might be reluctant to admit. When the unit pay rate was fixed at 24 cents for the directory and 17 cents for the personal ads, subjects overwhelmingly preferred the stodgy but financially superior alternative. However, when the pay rate for the personal ads was made elastic by the instruction “pay rate to be determined later; ranges from 4 cents to 30 cents,” there was a sizeable degree of erosion in the popularity of proofreading the directory in favor of working on the personals (Abelson, 1996, p. 32).

Abelson claims that this finding is particularly damaging to the requirement that actors have transitive preferences. However, a closer inspection demonstrates that this is not so. The three options are: a, proofreading the directory, which pays 24 cents; b, proofreading the personal ads, which pays 17 cents; and c, proofreading the personal ads where the pay is between 4 and 30 cents (and thus, the expected utility is 17 cents).14 Given the distribution of choices when faced with the choice between a and b or a and c, it is clear that many had the preference ordering a cPaPb.15 Since u(a) = 24, u(b) = 17, and EU(c) = 17, but the evidence indicates that c is preferred to a and a is preferred to b, Abelson claims that the above ordering is intransitive.

Abelson’s claim that this preference ordering is intransitive rests on two assumptions: first, that the actors are solely interested in economic gain; and second, that the actors are risk-neutral. This can be easily demonstrated. The reflexive operator, R, indicates that for xRy, x is at least as good as y. Thus, xPy and xLy each indicate xRy.16 Hence, for the above preference ordering, cRa, aRb, and cRb. The requirement for transitivity is that xLy and yRx implies that xRz for all x, y, and z. (Austin-Smith and Banks, 1999, pp. 2–4). Therefore, since aPb and bIc, this should imply aPc (rather than the observed cPa) in order for the preference ordering to be transitive.

However, and this is the key point, this intransitivity results only if the two assumptions above are met. If the subjects’ risk attitudes are accounted for, the mystery disappears. Risk neutral actors are indifferent between a certain outcome and a lottery with the same expected utility. Risk averse actors prefer the certain outcome, and risk acceptant actors prefer the lottery (Morrow, 1994). Risk acceptant actors would prefer c to b (i.e., they would prefer the lottery with expected utility of 17 to the certain outcome of 17), and the ordering cPaPb is transitive.

Furthermore, the only reason that u(a) = 24, u(b) = 17, and EU(c) = 17 is that the players’ utilities were modeled solely as a function of economic gain. This is indicative of a recurring problem throughout Abelson’s essay. In each of his examples, he
demonstrates that people are motivated by “fun” or other noneconomic benefits, and thus he claims that real people are not instrumentally rational. However, throughout Abelson actually gives examples of people choosing according to their preferences—and thus acting rationally.

What Preferences are Rational?

It is quite clear that Abelson misconstrues instrumental rationality. He argues that “expressive” motivations are somehow different from “instrumental” ones. However, nothing about instrumental rationality excludes “expressive” motivations. Instrumental rationality is only concerned when whether people act in accordance with their motivations, regardless of what those motivations may be. Unfortunately, Abelson is not alone in his confusion about this issue. Similarly, Ingelehart (1990) equates rational choice theory to a sole reliance on economic factors. These misconceptions about the nature of rational choice theory have likely resulted because it was first used in economics, and many early rational choice applications in political science modeled players’ utilities in economic terms.

However, there is in no way a requirement that rational actors are only (or at all) motivated by economic benefit. Demonstrating that people do not choose according to economic preferences does not demonstrate that they do not choose rationally; it only means that they are not solely motivated by economic gain. “So if we see a player choosing in a fashion that doesn’t maximize his payoffs as we have modeled them, then we must have incorrectly modelled his payoffs” (Kreps, 1990, p. 26).

Thus, recalling equation 3, it is possible that a given model of players’ payoffs creates an error due to assumption that is too large for the resulting theory to be empirically verified. Hence, the next logical step is to incorporate a better model of the players’ payoffs into the theory. Still misunderstanding, Green and Shapiro state that “[I]t is possible, of course, to preserve a variant of RCT by bending and stretching the conception of rational action to encompass such findings” (1994, p. 85). Changing the assumption about the players’ payoffs is not “bending and stretching the conception of rational action.” However, a more complete discussion of instrumental rationality needs to be presented.

ON INSTRUMENTAL RATIONALITY

Although the assumption of instrumental rationality has been clearly stated time and again (e.g., Luce and Raiffa 1957; Riker and Ordeshook, 1973; Zagare, 1990), it is evident from the previous discussion that it is still misunderstood by many. It appears as if this misunderstanding results from either confusion between procedural and instrumental rationality or an objection to the use of the term “rationality.” Each problem will be taken up in turn.

Procedural rationality (Simon, 1976) is similar to the common, everyday conception of rationality. Under procedural rationality, omniscient actors are said to make a “cool and clearheaded ends-means calculation” (Verba, 1961, p. 95) in the course of considering all available options and choosing the best one. Thus, to the extent that real actors suffer from misperceptions and other psychological and cognitive limita-
tions, they are not procedurally rational. Although procedural rationality may have its uses, it is manifestly not the rationality of rational choice theory.

Rational choice theory is rooted in the assumption of instrumental rationality. A rational actor is one who, when confronted with “two alternatives which give rise to outcomes, . . . . will choose the one which yields the more preferred outcome” (Luce and Raiffa, 1957, p. 50). Additionally, a rational actor must have connected and transitive preferences. For preferences to be connected, for any two alternatives \( x \) and \( y \), either \( xPy \), \( yPx \), or \( xIy \). Transitivity, as discussed above, means that if \( xPy \) and \( yPz \), then \( xPz \).

Instrumental rationality makes no normative judgements about preferences. That is, whether one’s preferences are “good” or “evil,” “instrumental” or “expressive,” or anything else has no impact on one’s instrumental rationality. Thus, a person such as Adolf Hitler who prefers to launch an attempt to conquer entire continents and wipe out entire races, and acts accordingly, is just as rational as someone like Woodrow Wilson who seeks to promote democracy and acts accordingly. Similarly, a job seeker who prefers to minimize his income, and acts accordingly, is just as rational as another who acts according to her preference for maximizing income. Saying that someone is instrumentally rational is not paying them a compliment; it is simply saying that they act according to their preferences, whatever they may be. Even so, many have argued that a variety of psychological, informational, or structural factors interfere with actors’—particularly states’—ability to act rationally. However, these factors only interfere with procedural rationality, not instrumental rationality. Instrumental rationality is compatible with a wide variety of supposedly limiting factors (Zagare, 1990).

Instrumental rationality is an assumption. Testing a rational choice theory (as with any theory) consists of testing the propositions that flow from the theory, not the assumptions. Therefore, Green and Shapiro’s statement that “rational choice theorists seldom set forth a clear statement of what datum or data, if observed, would warrant rejection … their conviction that politics flows from the maximizing behavior of rational actors” (1994, p. 36) is off the mark. Instrumental rationality is consistent with a wide variety of behaviors. Although certain types of behavior may contradict particular theories, that does not mean that they contradict the rationality postulate.

As Green and Shapiro (1994, p. 34) themselves state, “it is not obvious what sorts of behaviors, in principle, could fail to be explained by some variant of rational choice.” It is interesting to note how Green and Shapiro state the point so clearly yet still not understand it. That “it is not obvious what sorts of behaviors, in principle, could fail to be explained by some variant of rational choice theory” is precisely the point, and this constitutes a great strength of rational choice theory. Rational choice theory can, in principle, be applied to virtually any social situation. The suitability of any given rational choice explanation of behavior is an empirical matter that cannot be decided a priori. However, Green and Shapiro (along with many others) seem to want a world divided into behavior that is explainable by rational choice theory and behavior that is not.

Similarly, Kaufmann (1994) tests whether rational or psychological models foreign policy decision-making best explain German actions during the 1905–1906 Moroccan
Crisis. He examines the amount of time that it took each of 19 top-level German decision-makers to reach “an assessment that the expected French concessions could not be obtained” (Kaufmann, 1994, p. 569). He found very limited support for his predictions based on a psychological model. His only real finding is that increased personal interest in the crisis led to a longer time until decision-makers revised their assessment of the French position. Although this may be counter to what one would expect under procedural rationality, it is entirely understandable in terms of instrumental rationality. While he claims to have findings contrary to rational choice theory, he actually found that each of the actors acted according to his own preferences.

In a statement reminiscent of Green and Shapiro’s, Kaufamnn remarks that although

\[\text{a rational process may be consistent with errors, it cannot explain why they should be correlated with the types of errors predicted by psychological models. Indeed, if by some reinterpretation of standards of rationality it could do so, there would cease to be any meaningful distinction between so-called “rational” and “psychological” models (Kaufmann, 1994, p. 578).}\]

It does not require any reinterpretation of the standards of rationality to account for the behavior identified by Kaufmann. While some rational choice theories may assume away psychological factors that others do not, the definition of instrumental rationality remains the same. It is important to note that changing one’s assumption about the relevant utility function (the factors which are important in deciding an actor’s preferences) is not at all the same as changing the standards of rationality. It is simply, as Kreps stated, as attempt to model an actor’s preferences more correctly.

Fundamentally, the issue comes down to the question of whether or not choice matters. From a strictly structural perspective, choice does not matter; for example, Luebbert (1991, p. 306) provides an emphatically structural argument and concludes that “leadership and meaningful choice played no role in the outcomes.” Similarly, other theories such as process theories or dynamic models (e.g. Kadera, 2001) assume that choices have no impact on outcomes. If no role is given to choice within a theory, then clearly the issue of rationality is a moot point.

But what of theories that acknowledge the role of choice? With the assumption of rationality, one has a clear means of connecting actors’ preferences to their actions. But if we do not assume that actors are instrumentally rational, then what guides their decisions? Do they act randomly? Maybe they do, but such an assumption makes explaining their behavior problematic at best, impossible at worst. However, if one accepts that choices matter, but rejects rationality, that is what we are left with.

Although instrumental rationality may seem to be the clear way to explain actors’ choices, a phobia about using the term rationality appears to keep many scholars from recognizing their own implicit assumption of rationality. In studies of electoral behavior, various socioeconomic factors, parental party identification, religious preferences, gender, etc. are examined in order to determine voters’ preferences. Similarly, in psychological theories of foreign policy, various misperceptions and cognitive limitations are explored to determine foreign policy decision makers’ preferences. Why? Because, in each of these cases and many others, it is implicitly as-
sumed that actors will choose according to their preferences; i.e., that they are instrumentally rational. Thus, it is a fallacy for those who offer purposive theories of international politics to counter pose their arguments to rationalist explanations, since their explanations are also rationalist (just less explicit in their treatment of assumptions). Such critiques are, at best, undercutting the foundations of the theories these theorists themselves advocate.

**RATIONAL CHOICE THEORY IN INTERNATIONAL RELATIONS**

The discussion to this point has focused on explicating the assumption of instrumental rationality and demonstrating its compatibility with a wide range of possible motivations and behaviors. That this point is still misunderstood by many could be due to a lack of explicit explanation of the rationality postulate; hopefully the above discussion has served to fill that void. However, rational choice theorists have potentially added to confusion about the assumption of rationality, and create disillusionment in others about the rational choice enterprise in several ways. I now switch gears somewhat to focus on three applications of rational choice theory in international relations that demonstrate the problems associated with unfair tests of alternative theories, insufficient attention given to preferences, and inconsistent use of rationality.

**Conclusions Come from Assumptions**

Rational choice theories assume more than that actors are rational. Depending on the structure and function of the theory, assumptions are made about relevant actors, actors’ preference orderings, the choices available, and so on. It must be remembered that the conclusions of a theory depend on the assumptions of the theory. Formal theory is useful in making these assumptions explicit; however, assumptions are the basis of any theory, whether formal or not. Thus, when considering a theory’s empirical validity it is important to keep in mind that introducing additional assumptions changes the original theory. If the assumptions made are overly heroic, the error due to assumptions will be too large for the theory to be verified. However, if these heroic assumptions are not part of the original theory, it is not the original theory that is falsified but the modified one.

Bueno de Mesquita and Lalman make this mistake in *War and Reason* (1992). In their Realpolotik Proposition 3.1 (1992, p. 60), they show that negotiation and the status quo are the only possible equilibrium outcomes of the realpolitik variant of the international interaction game, under complete information. Further, the Realpolitik Acquiescence Impossibility Proposition 3.2 (1992, p. 67) shows that, with the realpolitik variant, acquiescence by either side is never an equilibrium outcome, regardless of information conditions. Given the numerous wars that have broken out and the 109 cases of acquiescence (out of 707 observations) in their data set, they conclude that the realpolitik variant (and thus realism) does not adequately explain international relations.

This seems reasonable, particularly given the wealth of evidence contradicting realism (Vasquez, 1997, 1999). However, both of these results are (given their structure of the game) rooted in one key assumption: “Assumption 4. All nations prefer to
resolve their differences through negotiation rather than war” (Bueno de Mesquita and Lalman, 1992, p. 40). This assumption logically follows from the implied assumption that the expected benefits from negotiations are the same as the benefits expected from war. Since war involves higher costs than negotiation, negotiation should be favored. However, this assumption seems problematic, given the evidence that Japan preferred to attack Pearl Harbor rather than continuing to negotiate, that Germany preferred to attack Poland rather than negotiate, or other possible historical examples. But more importantly, realism does not assume that negotiation is always preferred to war; thus, Bueno de Mesquita and Lalman have incorporated an additional assumption (not present in realism) into their realpolitik variant.

The disconfirming cases that they cite are contradictory entirely because of this one assumption. If the assumed preference ordering between negotiation and war is removed, then both war and acquiescence can be equilibrium outcomes of the international interaction game (as can easily be seen with backwards induction). Since there is nothing within realism that assumes that negotiation is always preferred to war, it seems that Bueno de Mesquita and Lalman have constructed an unfair test of realists theory and their conclusions are only correct about the realpolitik variant, not about realism.

The Importance of Preferences

The absolute versus relative gains debate provides several examples of misconceptions, and questionable applications of, rational choice theory. Grieco (1988) seeks to defend neorealism against what he dubs neoliberal institutionalism. He contends that neoliberals’ views that states are often in Prisoner’s Dilemma situations is wrong (Axelrod, 1984; Keohane, 1984; Axelrod and Keohane, 1985). Grieco disputes the claim that a state in an iterated Prisoner’s Dilemma responds (as a rational actor must) “solely out of a desire to maximize its individual long-term total payoffs” (1988, p. 496). He claims that such utility maximization is equivalent to maximizing one’s absolute gains, and therefore it is wrong because states are concerned with maximizing their relative gains.

Whether states are more interested in maximizing their absolute or relative gains is irrelevant for the purposes of this paper. Rather, if states are in Prisoner’s Dilemma situations then they must be maximizing their utility. To state it differently, one cannot argue that any particular game is an appropriate model of international relations (and thus use a rational choice perspective) and at the same time argue that the players do not seek to maximize their utilities (and thus violate a fundamental tenant of rational choice). However, utility maximization in no way presupposes that absolute gains are the only ones that matter. To the extent that relative gains matter, that concern needs to be incorporated into the utility functions. Thus, rational actors must maximize their utility regardless of weather that utility is independent of or partially interdependent upon other actors, as Grieco (1988) proposes. Snidal (1991) makes a similar argument and demonstrates that as relative gains become more important, several different types of games become Prisoner’s Dilemmas. Ultimately, if relative gains are the only important factor, then all games are zero-sum.

Powell attempts a “reformulation of the problem of absolute and relative gains
that links] changes in the states’ behavior, the feasibility of cooperation, and especially the states’ concern for relative versus absolute gains explicitly to changes in the constraints facing the states” (1991, p. 1304). He then develops a simple $3 \times 3$ “neoliberal institutional model [that] has been explicitly designed so that (F, F) [cooperation] is a first-period equilibrium outcome” (1991, p. 1308). In his conclusions, Powell (surprisingly enough) finds that cooperation is possible in an anarchic world.

To conclude that a model demonstrates what it has been explicitly designed to demonstrate is to have not said much. Indeed, Powell’s results rest solely on his assumptions about actors’ preferences. He makes rather arbitrary assumptions about states’ preferences, arguing in favor of shifting “the focus of analysis away from preferences to constraints” (Powell, 1991, p. 1304). Although the test of assumptions is the empirical accuracy of the resulting theory, ignoring actors’ preferences by assuming them away is counter to the very idea of rational choice theory. As Frieden (1999, p. 76) observes.

the goal of improved analyses of international politics would be better served if preferences were defined more carefully, if they were derived more methodically, and if their implications for outcomes were stated and evaluated more systematically. . . . [Furthermore,] no analysis of relations among actors … can be undertaken without a notion of the actors’ preferences. . . . Insufficient . . . attention to the role of preferences in international politics has been responsible for many fruitless debates and much poorly designed research.

However, Powell allows no variation in the preferences of states. Leaving preferences open as variables (i.e., allowing for multiple types of players) is far more general and allows for application to a wide variety of situations. Then utility functions for the various actors can be specified and states’ preferences can be measured, according to a particular operationalization. For example, Bueno de Mesquita and Lalman (1992) specify a utility function for the actors in the international interaction game, but they allow states’ preferences to vary. Indeed, they explicitly operationalize and measure states’ preferences.

Rational Choice means Rational Choice

Once rationality is assumed, it must be used consistently. Classical deterrence theorists such as Schelling have argued that deterrence relationships are best modeled by the game of Chicken (Figure 1). The two players, State A and State B, simultaneously choose between cooperating and defecting. For each state, winning is the best outcome (utility = 4) and conflict is the worst (u = 1). The status quo provides the second best outcome, with losing being the second worst. The outcomes A Wins and B wins are Nash equilibria (represented by an asterisk).

Obviously, each state would like to win, but how is this accomplished? Several strategies have been proposed, but only one will be examined here: the irrevocable commitment. Schelling (1966, p. 43) explains the rationale of this strategy:

If you are faced with an enemy who thinks you would turn and run if he kept advancing, and if the bridge is there to run across, he may keep advancing. He may
advance to the point where, if you do not run, a clash is automatic. Calculating what is in your long-run interest, you may turn and cross the bridge. At least, he may expect you to. But if you burn the bridge so that you cannot retreat, and in sheer desperation there is nothing you can do but defend yourself, he has a new calculation to make. He cannot count on what you would prefer to do if he were advancing irresistibly; he must decide what he ought to do if you were incapable of anything but resisting him.

Similarly, Kahn (1962) proposes making an irrevocable commitment by ripping the steering wheel off, thereby forcing the opponent to concede. Given the assumptions that the actors are rational and that the moves are simultaneous, this is impossible. If the moves are simultaneous, an irrevocable commitment can only truly be made (or believed) by an irrational actor. However, Schelling’s description of burning one’s bridges appears to be sequential moves (one player burns his bridges, and then the other responds by not attacking). But sequential moves are not possible in a simultaneous move game (Rapoport, 1964).

As Motyl (1999) points out, concepts must be both clearly defined and used consistently in order to be useful. Thus, if rationality is assumed, irrational choices are unacceptable. Similarly, if the theory uses a simultaneous move game, one player cannot move before the other. The assumption of rationality and formal modeling can be very useful in theory construction; however, they must be used consistently in order for the resulting theory to be logically consistent. In addition, inconsistent usage of rationality creates confusion about the nature and utility of rational choice theory. The presence of logical inconsistencies is a fundamental problem with any theory; therefore, theories that use rationality inconsistently are fundamentally flawed. However, this is not a problem with the assumption of rationality, but with those theories that use rationality inconsistently.

**IS RATIONAL CHOICE ONE THEORY?**

Throughout their work, Green and Shapiro (1994) make much of their claim that rational choice theory is indeed one theory. Although they state what seems to me the
appropriate response, “that we caricature rational choice theory by presenting it as a single theory, that purports to explain everything, when in fact that is not the case” (1994, p. 192), they brush past it, continuing in their claims.

It cannot possibly be the case that rational choice theory is indeed one theory. Few would dispute that balance of power theory and power transition theory are two separate theories. However, there have been rational choice approaches to both balance of power theory (e.g., Niou, Ordeshok, and Rose, 1989; Wagner, 1986, 1994) and power transition theory (e.g., Bueno de Mesquita, 1996; Danilovic, 1996; Zagare, 1996b). Furthermore, other rational choice explanations of conflict (e.g., Bueno de Mesquita, 1981; Bueno de Mesquita and Lalman, 1992; Zagare and Kilgour, 2000) would not be labeled as either balance of power or power transition theory. Both classical deterrence theory (briefly reviewed above) and perfect deterrence theory (Zagare and Kilgour, 2000) are rational choice theories. However, they are unquestionably different theories.

Although other examples of competing rational choice theories are available, the point should be quite clear. Rational choice is not one theory. But what does it matter? Many of Green and Shapiro’s examples of what they call post-hoc theorizing were in fact examples of new theories being developed to improve on previous ones. They characterize rational choice theory as bordering on nonfalsifiability because when one rational choice theory is “falsified” another is created to explain the disconfirming evidence. While there may be some merit to this charge, this would only be a problem if rational choice were one theory. However, since rational choice is an approach to theory and not one theory itself, nonfalsifiability of rational choice is irrelevant. The falsifiability of any given theory (rational choice or otherwise) is the important issue; falsifying the rational choice perspective is not only impossible, it is immaterial.

CONCLUSIONS

In this paper, I have addressed broad issues concerning rational choice theory. First, I have attempted to clarify the role of assumptions in theory. There are two types of assumptions. Assumptions of the first type cannot be proven; they must be assumed. A second type of assumption is used to simplify reality in theory construction. To the extent that the assumptions are not accurate for any given case, they add an error due to assumptions to the theoretical model. The ultimate test of the usefulness of assumptions is the empirical validity of the theory that results from them.

Rational choice theory is a descriptive phrase used to describe any of a number of individual theories that incorporate a rational choice approach. This approach is grounded in the basic assumption that actors make rational choices in an attempt to reach their most preferred outcome. In addition, rational choice assumes that outcomes are the result of choices made by actors. There are three separate rational choice approaches: nonformal theory, expected utility theory, and game theory.

Many critics of rational choice theory have questioned whether actors are really rational. However, I have demonstrated that these questions are based on misunderstandings of the assumption of instrumental rationality. Rational choice theory is consistent with behavior that is constrained by institutions, cultural influences, or
psychological limitations. Furthermore, many so-called nonrational choice theories implicitly assume instrumental rationality. Once this, along with the role of assumptions in theory, is properly recognized, we can move beyond debates about the assumptions of rationality and on to constructing useful theory and explaining international relations.

An exploration of some applications of rational choice in international relations has demonstrated several important points. First, when testing other theories, it must be ensured that new assumptions have not been added; otherwise, no conclusions can be reached about the original theory. Secondly, the extent to which absolute or relative gains matter to states needs to be incorporated into the states’ preferences. If assuming rationality, states must seek to maximize their utility; anything else would be logically inconsistent. Additionally, preferences matter. Assuming particular preferences in order to construct a model to demonstrate a particular conclusion is to have said nothing. Finally, concepts must be defined clearly and used consistently. Therefore, if rationality is assumed, irrational behavior is not allowed. Likewise, if a game is simultaneous, then sequential moves are not allowed.

Finally, rational choice is not one theory. Rational choice is an approach to theory that can result in competing and contradictory individual theories. Thus, Green and Shapiro’s concerns about the nonfalsifiability of the rational choice approach are immaterial. Rather, it is the falsifiability of individual theories that is of concern.

With a proper understanding of the role of assumptions in theory and of the assumption of instrumental rationality, the field of international relations, and political science in general, can move beyond debates about an assumption and on to more productive theory construction and research. The key, of course, is understanding that a variety of motivations and behaviors are consistent with instrumental rationality, which is after all the rationality of rational choice theory.

NOTES

1. I address the question of what constitutes rational choice theory below.
2. O’Neill (1992) provides a summary of game theory models of peace and war, and includes a discussion of the debate about rationality in international relations dating back to the 1960s.
3. Other goals of science include descriptions of past events, prediction of future events, and prescription of solutions for the future. However, the key to these goals is the proper understanding (or explanation) of why things happen. With the foundation of understanding laid, one can give better descriptions of events, make better predictions of future events, and also make better prescriptions for the future.
4. The basic format of this example is inspired by Przeworski and Teune (1982).
5. Notice that it is an assumption that such a deterministic explanation is even possible.
6. The danger of leaving assumptions implicit is that some analysts may not be aware of their assumptions. Although one may claim to assume nothing and simply ‘see what the data say’, this is impossible. In order to examine the data, assumptions must be made concerning the nature of the relationships, the relevant variables, and so on. Indeed, for any given statistical method, a host of assumptions are made.
7. There has, of course, been an extensive discussion of the role of assumptions in economics literature (e.g., Blaug, 1992; Friedman, 1953). Although there has been some discussion of the issue in political science (e.g., Bueno de Mesquita, 1981, chap. 1; 1985), it is clear that the role of assumptions has not received the attention in political science that it has in economics. As well, the discussion of ‘error due to assumptions’ is a novel approach that is useful in highlighting the role of assumptions in theory.
8. Notice, of course, that if statements are “empirically correct” then there is nothing to be assumed; they can simply be stipulated.
9. If theory is “incorrect,” then it will suffer from misspecification error and will not be confirmed regardless of the amount of measurement or assumption error.
10. This is not to say that assumptions are irrelevant. Rather, it is to argue that the ultimate test of a theory is the empirical accuracy of its predictions, and not the accuracy of its assumptions.
11. For more on rational choice and universalism, see text following.
12. Thus, in contrast to Morton’s (1999, p. 19) comment that “distaste for rational approaches … [is] a barrier to the acceptance of much formal modeling in political science,” it is evident that, for some, dislike for formal modeling is a barrier to acceptance of rational choice.
13. For example, O’Neill (1999) considers various ways in which honor and symbols affect the likelihood of war through a rational choice perspective. Thus, he explicitly considers how honor and symbols affect actors’ self-interest. Also, Schuessler (2000) examines the role of expressive choice in the context of electoral participation.
14. The expected utility will only equal 17 if each of the possible outcomes (4, 5, 6, … 29, 30) has an equal probability of occurring.
15. Where \( x \succ y \) indicates that \( x \) is preferred to \( y \), and \( x \equiv y \) indicates that the actor is indifferent between \( x \) and \( y \).
16. Notice also that \( x \equiv y \) implies that \( y \succ x \).
17. Opp (1999) refers to procedural rationality as the narrow version and instrumental rationality as the wide version of the rationality assumption.
18. The requirement for connectedness (or completeness) is from von Neumann and Morgenstern (1944). Aumann (1962) provides an alternative specification of utility theory without the completeness axiom.
19. Cashman (1993, pp. 79–81) provides a good summary of nine key criticisms of rationality along these lines.
20. Although Kaufmann claims to have found limited support for his predictions, it is not clear that he has. His first model (and each of the variables in it) was statistically insignificant. The second model and the two variables in it are significant. However, the results are quite confusing; although he states that the second model uses “only the motivated bias and salient bias variables” (Kaufmann, 1994, p. 576), the variables that appear in the model are for motivated bias and big-stick belief. Nonetheless, it seems from his preceding discussion about the problems with the big-stick belief variable that it is indeed the salience variable that appears in the second model. However, although he predicts that higher salience leads to a reduced time until assessment revision (and thus a negative coefficient), the result is positive and significant (<0.01) and Kaufmann claims that it supports his model!
21. For examination of the role of choice in international relations, see Lake and Powell (1999).
22. Yamamoto (1990) conducted a test of whether rationality or chance best explains war expansion. Unsurprisingly, the rational choice model performed much better than the model based on chance.
23. As previously discussed, these psychological factors are entirely consistent with the assumption of instrumental rationality.
24. Therefore, the following discussion deals with the implicit assumption about the expected benefits of negotiation and war, as well as assumption 4. Assumption 4 is highlighted it is the one that Bueno de Mesquita and Lalman make explicit in their assumptions.
25. Indeed, Hitler later regretted having negotiated over Czechoslovakia in Munich in 1938 rather than attacking (Weinberg, 1995).
26. Although it seems rather obvious, to me at least, that approaching a problem from a rational choice perspective or arguing that states are in a Prisoner’s Dilemma situation in no way implies agreement with, much less subscription to, neoliberalism, this point is not particularly relevant to the current discussion. Nonetheless, I fail to see where Axelrod and Keohane (1985) make a claim to neoliberalism.
27. For example, the fighting cost to each state is 3.5 (later 7) and the payoff to victory is 6. These numbers come from out of the blue. If the costs or payoffs were different, different results would be obtained.
28. Zagare (1996a) and Zagare and Kilgour (2000, ch. 1) provide detailed critiques of classical deterrence theory.
29. Of course, as stated previously, theory does not have to be formal in order to be rational choice. Although nonformal rational choice works certainly exist in these areas, only formal pieces are used
as examples here because they are unquestionably rational choice and clearly demonstrate that rational choice is not one theory.

30. Within American politics, the congressional institutions literature is dominated by formal rational choice works. However, there are distinct disagreements among rational choice theorists over many issues (Shepsle and Weingast, 1995).

**REFERENCES**


Green, Donald P. and Ian Shapiro (1994). *Pathologies of Rational Choice Theory: A Critique*
THE RATIONALITY OF RATIONAL CHOICE THEORY

of Applications in Political Science, New Haven: Yale University Press.


Kahn, Herman (1962). Thinking About the Unthinkable, New York: Horizon Press.


**CONTRIBUTOR**

**Stephen L. Quackenbush** is Assistant Professor of Political Science at the University of Missouri. His research focuses on international conflict processes and deterrence. His work includes an article in the *Journal of Politics* and a forthcoming chapter in *Making Sense of IR Theory*. 