PERFECT DETERRENCE THEORY

INTRODUCTION

In the aftermath of World War II, a large, often contradictory, but nonetheless influential body of literature emerged in the field of security studies that is commonly referred to as *(classical) deterrence theory*.[1] Although its roots can be traced to the contentious debate between realists and liberals during the interwar period about the causes of great power wars in general, and of World War I in particular (Carr, 1939), it is generally understood that Bernard Brodie's (1946, 1959) work is seminal. Shortly after atomic bombs were dropped, first on Hiroshima and then on Nagasaki, Brodie realized that war prevention would become the primary goal of the American security apparatus in what was sure to become a world of nuclear plenty. How and under what conditions that goal might be achieved are the principal questions addressed in this vast literature.

It is quite understandable that the earliest attempts to answer this question relied heavily on extant realist theory. Since the conventional wisdom at the time was that a power asymmetry was an all-but sufficient condition for major power conflicts, it followed that the key to peace was a carefully calibrated strategic balance. Balance-of-power theory, however, did not withstand intense logical and empirical scrutiny. As evidence began to accumulate that both world wars were actually contested under parity conditions (Organski, 1958, ch. 11; Organski & Kugler, 1980; see also Snyder, 2002, pp. 167–168), a new variable was added to explain away the glaring discrepancy between fact and theory: the cost of war. What eventually emerged was a part of the literature that I have elsewhere called *structural deterrence theory* (Zagare, 1996a). Structural deterrence theorists, of whom Kenneth Waltz (1964, 1979, 1981) is the exemplar, focus on the interplay of strategic structure and the cost of conflict to explain the absence of a superpower conflict during the cold war period. In their view, parity, when coupled with sufficiently high war costs, all but guarantees peace (Intriligator & Brito, 1984). By contrast, structural deterrence theorists contend that peace may break down—even under parity—when the costs of conflict are, or are believed to be, low. For Waltz (1993, p. 77), the breakdown of the international system in 1914 occurred precisely because decision makers in Berlin, Vienna, St. Petersburg, Paris, and London thought that a war could be waged on the cheap.[2]

While Waltz and other structural theorists (e.g., Mearsheimer, 1990) explored the systemic conditions that they believed explained the "long peace" (Gaddis, 1986) of the cold war period, a second group of strategic thinkers approached the question from a choice theoretic perspective. Pitched at the micro-level of analysis, this strand of the literature, which I label *decision theoretic deterrence theory*, includes the work of psychological choice, expected utility, and game theorists. The exemplar of this part of the deterrence landscape is Thomas Schelling (1960, 1966), who dabbled in all three genres. Zakaria (2001) credits Schelling as the "inventor" of modern deterrence theory. Others, however, point to Kaufmann (1956).

These two strands in the strategic studies literature, which together constitute the dominant paradigmatic (i.e., realist) articulation of deterrence theory, have much in common. For example, structural deterrence theorists conclude that bilateral nuclear relations are exceedingly stable. In fact, they argue, that as the costs of conflict rise, the probability of a war between nuclear rivals "approaches zero" (Waltz, 1990, p. 740; see also Intriligator & Brito, 1984). Since rational conflicts are

believed to be unlikely, the gravest threat to peace in the nuclear era is seen to be accidental war (Bracken, 1983).

Decision-theoretic deterrence theorists take these conclusions as axiomatic. Indeed, they are built into the very structure of the game that they use as the underlying metaphor of a bilateral nuclear relationship: Chicken (Jervis, 1979, p. 291). In Chicken—see Figure 1—the outcome labeled *Conflict* is a mutually worst outcome. In consequence, rational conflict is precluded. In a strict 2 x 2 ordinal game like Chicken, a mutually worst outcome can never be part of a Nash equilibrium, which is the accepted standard of rational play in a normal-form game. Thus, *Conflict* can occur only when rational players miscalculate—that is, when an accident occurs.

Structural deterrence theory and decision-theoretic deterrence theory share many of the same logical and empirical flaws.[3] For example, structural deterrence theorists are unable to explain the absence of a superpower conflict during the periods of American nuclear superiority without making an ad hoc adjustment to the theory, as Waltz himself points out (1993, p. 47). Even when the motivation exists, states generally do not jump through "windows of opportunity" (Jervis, 1985; Lebow, 1984). The abstract version of the theory clearly implies that the United States should have exploited the obvious strategic advantage that it enjoyed throughout the 1950s and the early 1960s, not only against the Soviet Union, but against lesser powers as well. As Jervis (1988, p. 342) notes, structural deterrence theorists are unable to "explain the fact that the United States did not conquer Canada sometime in the past hundred years." Or as Gaddis (1997, p. 88) more tactfully puts it, "the actions the United States took [during the early Cold War years] failed to fit traditional patterns of great power behavior."

For their part, decision-theoretic deterrence theorists are also hard put to explain the absence of a superpower conflict during the most intense periods of the cold war period without logical contradiction. And it is easy to understand why. *Conflict* is not the only outcome in Chicken that is inconsistent with rational play. The outcome labeled *Status Quo*, which results when both states cooperate and neither attacks the other, is also not a Nash equilibrium. As a result, the long peace can be explained within the confines

		State B: Cooperate (C)	State B: Defect (D)
State A	Cooperate (C)	*Status Quo* (3, 3)	*B Wins* (2, 4)*
State A	Defect (D)	*A Wins* (4, 2)*	*Conflict* (1, 1)

Key: (x,y) = payoff to State A, payoff to State B
4 = best; 3 = next-best; 2 = next-worst; 1 = worst
* = Nash equilibrium

Figure 1. Chicken.

of decision-theoretic deterrence theory only by assuming, simultaneously, that the players are rational (when they are being deterred) and irrational (when they are deterring and threatening mutual destruction).[4] All of which helps to explain why Achen (1987, p. 92) has observed that "far from leaning too heavily on rational choice postulates, 'rational deterrence theory' necessarily assumes that nations are not always self-interestedly rational."[5]

PERFECT DETERRENCE THEORY

Perfect deterrence theory was developed to overcome the empirical and logical problems that plague classical deterrence theory. As developed by Zagare and Kilgour (2000), this theoretical structure is composed of a number of interrelated game models that are analyzed under a common set of preference assumptions. The assumptions are both intuitively obvious and, for the most part, innocuous. In these models, the players are assumed to prefer winning to losing and to do so at the lowest possible cost. Most other preference relationships are taken as strategic variables. For example, some players might prefer *Conflict* to losing. Players with such a preference are assumed to have *credible* threats (i.e., threats that are rational to execute). Other players, with the opposite preference, have threats that lack credibility. The players also may or may not prefer *Status Quo* to *Conflict*. A player whose opponent prefers *Status Quo* to *Conflict* is said to have a capable threat (i.e., a threat that hurts [Schelling, 1966, p. 7]). Threats that do not hurt are considered incapable. Finally, in perfect deterrence theory, the players are not necessarily undifferentiated, as they are in classical deterrence theory (Legro & Moravcsik, 1999, p. 13). Specifically, in perfect deterrence theory, some players, called *Defenders*, prefer *Status Quo* to all other outcomes; by contrast, *Challengers* are those who are motivated to upset it. A bilateral relationship wherein both players are Challengers is said to constitute a *Mutual Deterrence Game*. *Unilateral Deterrence Games* are those in which a satisfied Defender and a dissatisfied Challenger are involved.[6] Perfect deterrence theory examines the strategic impact of various configurations of credible and capable threats under conditions of complete and incomplete information in each of its four constituent games.[7]

The variable nature of these critical preference relationships is, perhaps, the most important way that perfect deterrence theory is distinguished from decision-theoretic deterrence theory (Quackenbush, 2011, p. 747). In either the formal or informal studies that take as their starting point the game of Chicken, all preferences are necessarily fixed. But another important difference between these two theoretically distinct variants of deterrence theory is perfect deterrence theory's strict adherence to Selten's (1975) perfectness criterion, which requires players to make rational choices at every decision point in a game, including those that they might never face.[8] As is well known, some Nash equilibria are supported by threats that are not rational (or credible) to execute. Selten's perfectness criterion eliminates these equilibria as rational strategic possibilities and, not incidentally, the logical inconsistency that is associated with them.

In a dynamic (or extensive-form) game of complete information, Selten's rationality postulate is associated with the concept of a *subgame perfect equilibrium*; in a game of incomplete information, with the concept of a *perfect Bayesian equilibrium* (Harsanyi, 1967–1968). The name *perfect deterrence theory* stems from its reliance on these two accepted measures of rational strategic behavior. Adherence to the perfectness criterion assures logical consistency, which is arguably the most

important characteristic of perfect deterrence theory. Walt (1999) notwithstanding, logical inconsistencies are clearly fatal to the health and well-being of any theory. Logically inconsistent theories are incapable of being falsified. By contrast, because it has clear empirical implications, perfect deterrence theory can, and has been, rigorously tested (Quackenbush, 2010a, 2011). Many of its theoretical propositions are also consistent with the empirical record, as I discuss next.

STRATEGIC VARIABLES

Perfect deterrence theory is a general theory of conflict initiation and resolution. Unlike classical deterrence theory, perfect deterrence theory makes no particular assumption about the cost of conflict. It is, therefore, applicable to a much wider range of strategic relationships. Perfect deterrence theory is simply not a divergent theory of nuclear war avoidance. Rather, it is a universal theory, applicable to both nuclear and nonnuclear interactions. As such, it can be used to help explain why crises occur, why some conflicts escalate and others do not, and when and why limited and all-out wars are waged. Perfect deterrence theory's empirical domain is not even restricted to interstate interactions. As a general theory of strategic interaction, it is potentially applicable to intergroup or interpersonal conflict-of-interest situations, whenever and wherever they may occur.

In perfect deterrence theory, the cost of conflict is, nonetheless, a critical strategic variable. Its value relative to other variables determines both the *capability* and the *credibility* of a deterrent threat. As noted, capable threats are threats that hurt. Threats that hurt are those that leave a player worse off than if the prohibited action was not taken. Operationally, this means that one player's deterrent threat is capable only if the other player (the threatened player) prefers the status quo to the outcome that would result if the threat were executed. Conversely, a deterrent threat will lack capability whenever the threatened player prefers to act even if the threat is carried out (Zagare, 1987, p. 34).

The threat of nuclear retaliation is clearly a capable threat. Classical deterrence theorists contend that capability constitutes a *sufficient* condition for deterrence success, which is why they believe that the mere possession of nuclear weapons all but guarantees a lasting peace (e.g., Bundy, 1983; Levy, 1988, pp. 489–490; Waltz, 1993, pp. 53–54). In perfect deterrence theory, by contrast, deterrence may fail even when all threats are capable. For instance, by definition, both players have a capable threat in the game of Chicken—see Figure 1. Yet the status quo generally does not survive rational play in this game.

On the other hand, within the theoretical confines of perfect deterrence theory, a capable threat constitutes a *necessary* condition for deterrence success—that is, deterrence will always fail whenever a deterrent threat lacks capability (Zagare & Kilgour, 2000, pp. 81–84). To illustrate, consider for now the Unilateral Deterrence Game shown in Figure 2. This game, which is one of perfect deterrence theory's four constituent games, depicts a one-sided (or asymmetric) deterrence relationship in which a Defender seeks to deter a Challenger, but not the other way around.

In the Unilateral Deterrence Game Challenger begins play (at decision node 1) by deciding whether to contest the existing order. If no demand is made, *Status Quo* prevails and deterrence succeeds. But if a demand is made and general deterrence breaks down, Defender must decide (at node 2) whether to resist it or concede the issue. Concession results in the outcome *Defender Concedes*. Resistance brings about a subsequent choice for Challenger at node 3: If Challenger backs down, the outcome labeled *Challenger Defeated* results. If

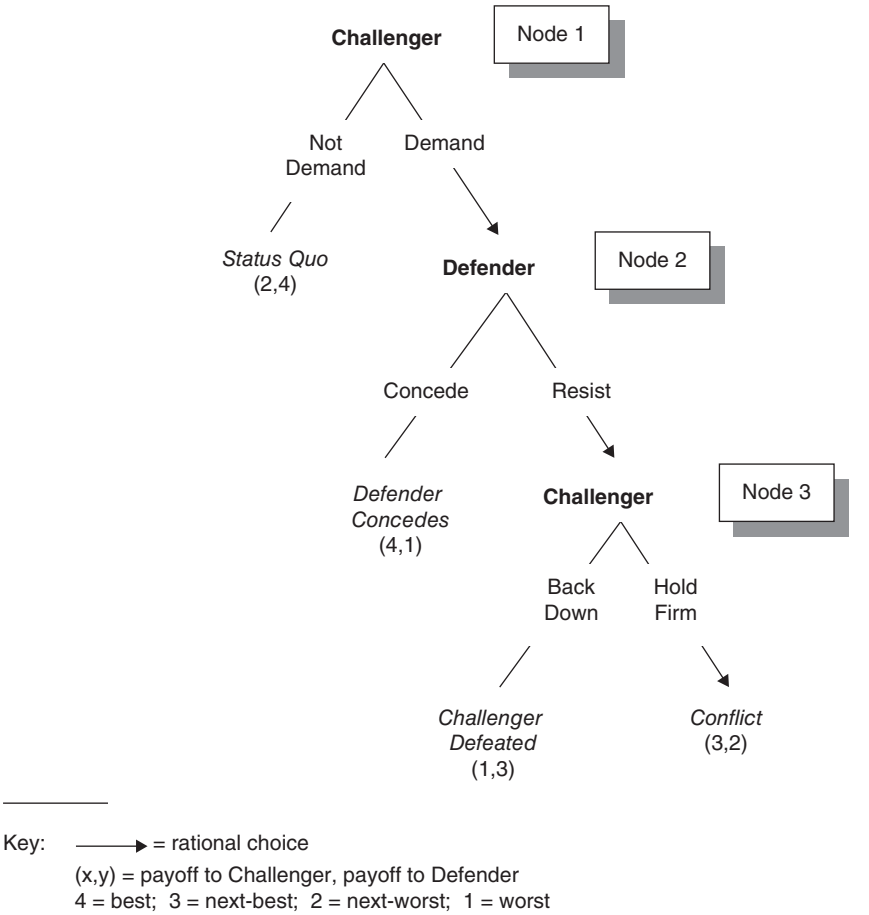


Figure 2. Unilateral deterrence game: Defender's threat is credible but not capable.

Challenger holds firm, there is *Conflict*. The four outcomes in this game correspond, roughly, to the four outcomes in Chicken.

As before, the ordered pair beneath each outcome represents the ordinal payoff to the players, from best (i.e., 4) to worst (i.e., 1). The first entry represents Challenger's evaluation; the second, Defender's. For example, in this variant of the Unilateral Deterrence Game, *Conflict* is Challenger's next best outcome (i.e., 3) and Defender's next worst (i.e., 2). Since the present assumption is that Challenger prefers *Conflict* to *Status Quo*, Defender's threat to resist at node 2 lacks capability, by definition.

The version of the Unilateral Deterrence Game given by Figure 2 is an extensive-form game with complete information. In such a game, rational play is determined by applying backward induction, starting with Challenger's choice at node 3 and working backward up the game tree. The outcome that survives the backward induction process in a game of complete information will be a subgame perfect equilibrium.

As the arrows indicate, a rational Challenger will hold firm at node 3 in order to avoid its worst outcome, *Challenger Defeated*. Since concession results in Defender's worst outcome, it will rationally resist should it face a choice

at node 2. Anticipating both these choices, Challenger will rationally issue a demand at node 1. After all, by assumption, Challenger prefers *Conflict*, the unique subgame perfect equilibrium in this game and the outcome that is implied by the backward induction process, to *Status Quo*. In consequence, deterrence fails—as it always will—whenever a Defender's threat lacks capability.

Notice that in this example that Defender's threat is credible—that is, it prefers *Conflict* to *Defender Concedes*. Yet deterrence does not succeed. Freedman (1989, p. 96) claims that credibility is the "magic ingredient" of deterrence. While there is a modicum of truth to his statement, a credible threat does not constitute a sufficient condition for deterrence success, as the game in Figure 2 demonstrates.

Nor is credibility a necessary condition. (Table 1 summarizes perfect deterrence theory's conclusions about the causal characteristics of deterrent threats.) To see this, consider now the version of the Unilateral Deterrence game depicted in Figure 3. In this example, neither player's threat is credible. Since *Conflict* is a mutually worst outcome, Challenger rationally backs down at node 3. But because it expects a rational Challenger to accept defeat at node 3, Defender intends to resist (because it prefers *Challenger Defeated* to *Defender Concedes*) if and when it is faced with a choice at node 2. Node 2 is never reached in rational play, however, because Challenger, preferring to avoid certain humiliation (i.e., *Challenger Defeated*), will suddenly find *Status Quo* acceptable.

Table 1. Causal Characteristics of Deterrent Threats

	Necessary Condition	**Sufficient Condition**
Capable threats	Yes	No
Credible threats	No	No
Credible and capable threats	No	No

In this case, then, deterrence succeeds even though Defender's threat lacks credibility. All these elements illustrate that it may be the characteristics of Challenger's threat, not Defender's, that ultimately determines rational play in a game. This is not a trivial point. For the most part, classical deterrence theorists have fixated on a defender's threat in order to understand the conditions under which deterrence breaks down.[9] In so doing, they fail to recognize the interactive impact that deterrent threats have. Empirical studies that focus attention on the characteristics of a defender's threat also miss this important dimension of not only direct but of extended deterrence relationships, and in the process, introduce case selection bias into their studies.[10] In a multilevel extended deterrence game where a challenger's strategic level (or end-game) threat lacks credibility, deterrence should prevail regardless of the configuration of a defender's tactical and strategic-level threats (Zagare & Kilgour, 2000, ch. 9).

By now, it should be clear that the relationship between credibility and the operation of deterrence is anything but straightforward. More specifically, deterrence may fail even when all threats are credible; and it may succeed even when all threats lack credibility. But this is not the end of the story. Conditions also exist under which one player's possession of a credible threat will actually undermine the stability of the status quo. The game given in Figure 4 is a case in point.

There is only one difference (and a relatively minor one at that) between the games presented in Figures 3 and 4. In the former case, Challenger's threat lacks credibility, yet deterrence succeeds anyway. In the game in Figure 4, by contrast, Challenger's threat is indeed credible, but as the arrows indicate,

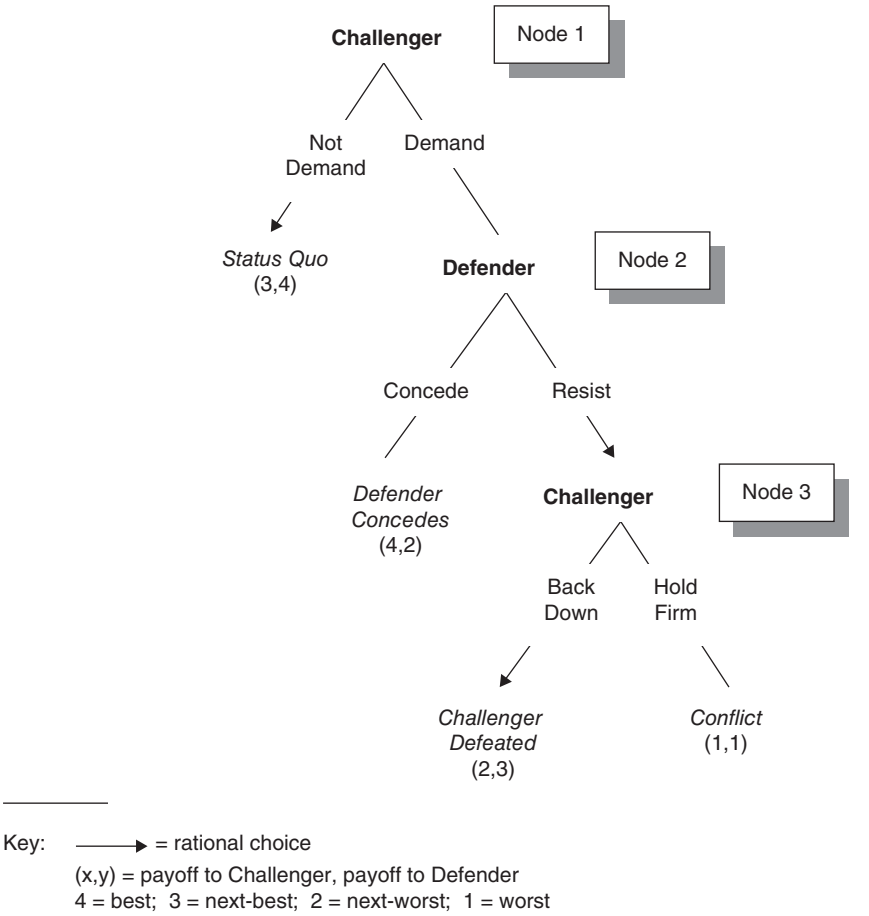


Figure 3. Unilateral deterrence game when neither player has a credible threat.

deterrence rationally fails. So once again, a credible threat (or the lack thereof) is central to the operation of deterrence game, but not in the way that one might suspect. In the game of Figure 4, one player's (i.e., Challenger's) possession of a threat that is rational to execute actually incentivizes a breakdown of deterrence.

As noted previously, classical deterrence theorists tend to focus on the defender's threat characteristics rather than the challenger's. One possible explanation is their mania for explaining deterrence failures and their comparative lack of interest in explaining deterrence success. Their focus on deterrence breakdowns also helps to explain why they are similarly prone to overlook the role played by another important strategic variable—the value of the status quo. This is clearly a significant oversight. The extent to which a player is satisfied (or dissatisfied) with the existing order is not only a critical determinant of a threat's capability, but also the only standard that can be used to establish a player's incentive (or willingness) to overturn it (Most & Starr, 1989). The failure to explore the consequences of this critical benchmark variable is yet another source of case selection bias that plagues almost every empirical examination of classical deterrence theory's central propositions.[11]

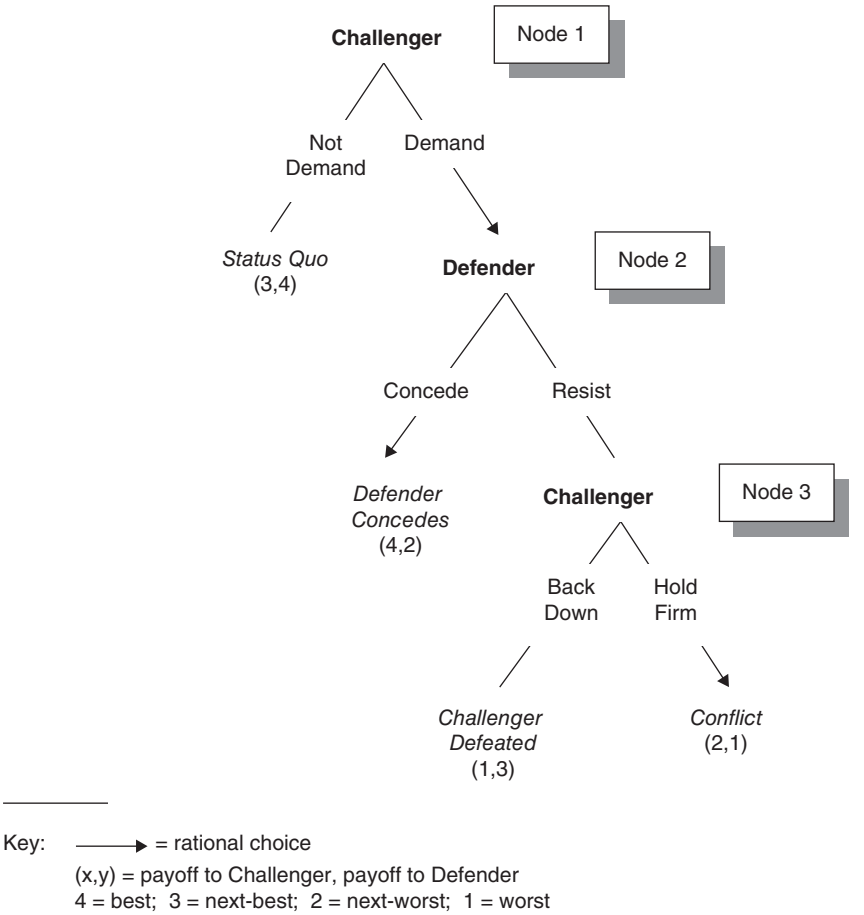


Figure 4. Unilateral deterrence game when only challenger's threat is credible.

RELATIONSHIP PREDICTIONS

In some ways, the theoretical characteristics of classical deterrence theory and perfect deterrence theory are similar. For example, both frameworks assume egotistical or self-interested actors who make rational choices. But there are also some significant differences. As noted, perfect deterrence theory makes consistent use of the rationality postulate, whereas classical deterrence theory does not. In contrast to the standard version of deterrence theory, perfect deterrence theory does not treat all actors as similarly motivated (i.e., as undifferentiated); nor does it take status quo evaluations and threat credibility to be fixed and constant. In perfect deterrence theory, these entities are important strategic variables. As a result, the empirical expectations of these two competing specifications of deterrence theory are significantly at odds.

To be sure, it is not easy to summarize classical deterrence theory's core propositions. Since its sporadic application of the rationality postulate inevitably leads to logical inconsistencies, anything and its opposite can be deduced from its axiomatic base (Martin, 1999, p. 83). Nonetheless, although there are

some exceptions, there also appears to be a general consensus among mainstream strategists. Table 2 summarizes both this confluence of opinion and the competing relationship propositions of perfect deterrence theory.

Several of these differences have already been discussed, but one that has not concerns the overall nature of contentious deterrence relationships. As previously indicated, classical deterrence theorists see strategic (i.e., nuclear) deterrence relationships as more than robust (i.e., exceedingly stable). In their view, rational nuclear conflicts are extremely unlikely. Perfect deterrence theory, however, reaches a different conclusion. In the four incomplete information game models that collectively delineate the theory, a deterrence equilibrium under which no player contests the status quo always exists. But these equilibria are hardly ever unique—that is, more often than not, they coexist with other equilibria, some of which admit the possibility of a complete deterrence breakdown. In other words, the conditions that are most conducive to peace and stability are also consistent with the possibility of war. Depending on leadership preferences and policy choices, either denouement is possible—that is, deterrence failures, limited conflicts, and all-out conflicts can almost never be ruled out as a rational strategic possibility. In perfect deterrence theory, then, deterrence success is highly contingent and anything but robust.

A related difference concerns the strategic characteristics of parity conditions. Classical deterrence theorists see them as generally stable and view asymmetric power relationships as potentially unstable. Once again, perfect deterrence theory reaches a different conclusion. In both theory and principle, it is much easier to stabilize the status quo when only one player is a defender than when both play that role. Since the former category of games is more likely to exist when power is asymmetrically distributed, parity or balance-of-power conditions are judged, *ceteris paribus*, to be less stable than one-sided deterrence relationships, but especially those in which a dominant and highly satisfied great power clearly controls the agenda of the international system.

Finally, classical deterrence theorists see a monotonic and strictly positive relationship

Table 2. Classical Deterrence Theory and Perfect Deterrence Theory: Empirical Propositions

	Classical Deterrence Theory	**Perfect Deterrence Theory**
Propositions:		
Status quo	Unimportant/ignored	Significant
Strategic deterrence	Robust/all but certain	Fragile/contingent
Relationship between conflict costs and deterrence success	Strictly positive and monotonic	Nonmonotonic
Asymmetric power relationships	Unstable	Potentially very stable
Parity relationships	Very stable	Potentially unstable
Capability	Sufficient for deterrence success	Necessary, but not sufficient, for deterrence success
Limited conflicts and escalation spirals	Unexplained	Placed in theoretical context

between the cost of conflict and the probability of deterrence success. Specifically, the higher the war costs, the less likely a deterrence breakdown is, and the converse is also true (e.g., Mearsheimer, 1990, p. 19). In perfect deterrence theory, by contrast, the relationship is neither monotonic nor positive. To be sure, there is a minimum cost threshold below which deterrence success is impossible. This is the point at which a noncapable threat becomes a threat that hurts. But there is also a maximum threshold beyond which additional increases in the cost of conflict are unrelated to the probability of successful deterrence. Moreover, in perfect deterrence theory, extended deterrence relationships are actually more likely to unravel, *ceteris paribus*, as war costs increase.

EMPIRICAL SUPPORT

There is considerable empirical support for both the relationship and, as will be seen, for the equilibrium predictions of perfect deterrence theory (Morton, 1999). In other words, perfect deterrence theory is a more plausible theory of strategic interaction than is classical deterrence theory. In part, this advantage stems from classical deterrence theory's loose formulation and its inconsistent use of the rationality postulate. By contrast, perfect deterrence theory's straightforward empirical expectations are the result of an analytic framework that respects the requirements of strict logic.

Take, for instance, the empirics that support the stabilizing effect of satisfaction with the status quo.[12] For the most part, classical deterrence theorists have ignored this literature precisely because their conceptual lens has blinded them to its importance. This wide and largely straightforward literature, however, is more than consistent with the gestalt of perfect deterrence theory. It should be patently obvious that highly satisfied states are, *ceteris paribus*, less prone to contest the governing structure of either a regional or a systemic structure. Classical deterrence theorists, however, seem not to have noticed.

Much the same can be said about perfect deterrence theory's claim that a capable threat is a necessary (though not sufficient) condition for deterrence success. Harvey's (1998, p. 691) empirical findings "indirectly support [perfect deterrence theory's] claim about the crucial role of capabilities" in deterrence relationships. But so does Bueno de Mesquita's (1981, pp. 154–156) finding that conflict initiators, in general, tend to be more capable then those that they attack, regardless of the belligerents' alignment status or the level of initiated conflict.

There is also impressive empirical support for perfect deterrence theory's claim about the central role that credible threats play in deterrence relationships. For instance, in an important study of extended deterrence, Danilovic (2002) finds that a state's inherent credibility—as reflected in large part by its regional interests—is a far more important predictor of deterrence outcomes than are the coercive bargaining tactics recommended by decision-theoretic deterrence theorists. Huth's (1999) earlier review of the literature reaches a similar conclusion and, at the same time, undermines the empirical foundation of many of Schelling's most well known policy prescriptions. More recently, Johnson, Leeds, and Wu's (2015, p. 309) comprehensive study of strategic alliances clearly shows that "defense pacts with more capability and more credibility reduce the probability that a member state will be the target of a militarized dispute." Their results reinforce both Danilovic's conclusions and perfect deterrence theory's propositions about the pacifying role played by both these variables in extended deterrence relationships. In this context, it should also be mentioned that the vast literature supporting

action-reaction models, reciprocity, and tit-for tat behavior[13] is also consistent with perfect deterrence theory's key finding that conditionally cooperative strategies (or policies) are associated with deterrence success, and that unconditionally hard-line policies are generally prone to failure.

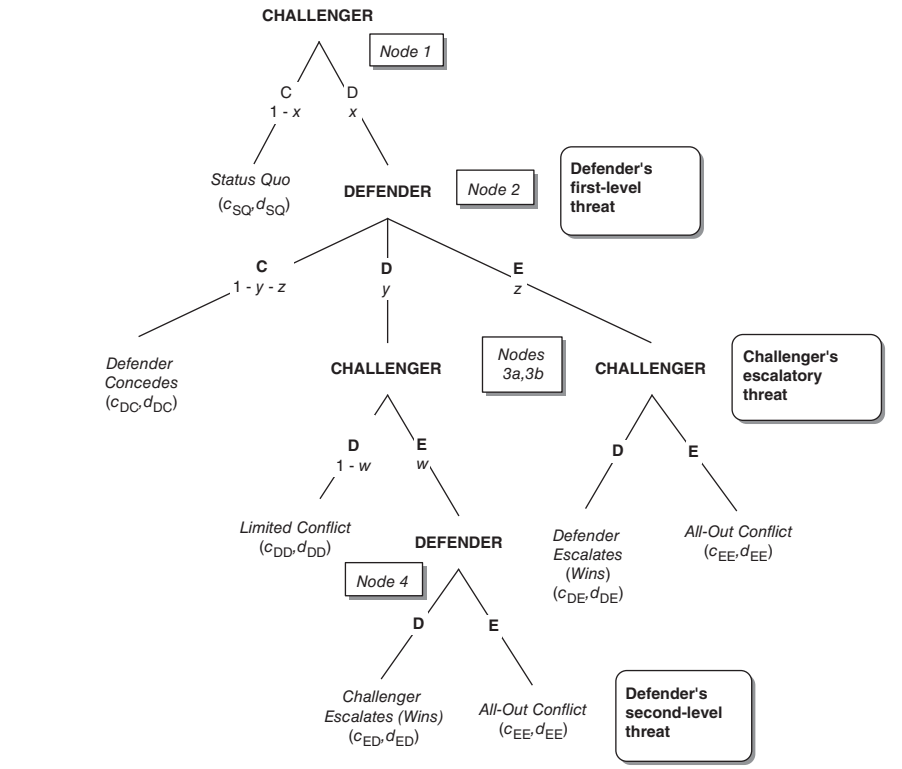
In another significant examination of the dynamics of deterrence, Senese and Quackenbush (2003) studied the long-term consequences of different types of conflict settlements on recurring conflicts. Their hypothesis that imposed settlements, which they contend create one-sided (or unilateral) deterrence relationships, should last longer than negotiated agreements, which they associate with bilateral (or mutual) deterrence situations, was independently deduced from perfect deterrence theory's axiomatic base. In an exhaustive analysis of 2,536 interstate conflict settlements between 1816 and 1992, this now-clear implication of perfect deterrence theory was strongly supported. Senese and Quackenbush's result held even when they controlled for regime type, relative capability, contiguity, decisive outcomes, power shifts, and war, leading them to observe that in their test of one of perfect deterrence theory's relationship predictions, the theory performed "quite well." Follow-up studies by Quackenbush and Venteicher (2008) and Quackenbush (2010b) reach the same conclusion.

Space and other considerations preclude a full discussion of empirical studies that support the relationship predictions of perfect deterrence theory.[14] But mention should be made of one study that attempted to test perfect deterrence theory's equilibrium predictions and also of those analytic narratives that have been constructed to illustrate the "theory in action."

In the literature of deterrence, attempts to test equilibrium predictions of a game theoretic model are virtually nonexistent, in large part because of the numerous methodological hurdles that stand in their way. Quackenbush (2011) is the rare exception in more than one way. Not only is it a direct test of a game model's equilibrium predictions, but it is also one of the few attempts to test a theory of general deterrence. The empirical literature of deterrence is dominated by tests of extended immediate deterrence (Morgan, 1977) in which general deterrence has already failed. Huth (1988) and Danilovic (2002) are the best examples of this literature.

In order to test perfect deterrence theory's equilibrium predictions (and, not coincidentally, minimize case selection bias), Quackenbush had to construct a relevant data set. For this purpose, he first developed a category of relationships that he termed "politically active" (Quackenbush, 2006). He also had to develop measures that reflected the players' utility functions—perhaps the most daunting precondition for a meaningful test of any rational choice theory. Finally, using binary and multinomial logit methods to examine the predictions of various game outcomes, he found that the predictions of perfect deterrence theory were "generally supported by the empirical record" (Quackenbush, 2011, p. 74). He also found that his results provide much stronger support for perfect deterrence theory than the support that Bennett and Stam (2000) found for Bueno de Mesquita and Lalman's (1992) international interaction game.

Because it is a direct test, Quackenbush's study is the most persuasive evidence to date that perfect deterrence theory stands on firm empirical grounds. Less systematic, but significant nonetheless, is the accumulated collection of case studies that have been constructed using the constituent models of perfect deterrence theory as an organizing mechanism. For instance, Zagare (2016) uses the *Asymmetric Escalation Game* (see Figure 5) to explain the strategic decisions of the Soviet



Key:

Challenger: C = Cooperate
D = Demand
E = Escalate

Defender: C = Concede
D = Defy
E = Escalate

x = probability Challenger initiates at node 1
w = probability Challenger escalates at node 3a
y = probability Defender responds in kind at node 2
z = probability Defender escalates at node 2

Figure 5. Asymmetric escalation game.

Union and the United States during the 1962 Cuban missile crisis. He also uses it, along with the *Tripartite Crisis Game*, to examine four critical decisions that he claims account, collectively, for the initiation, escalation, and expansion of World War I.

In the Asymmetric Escalation Game, limited conflicts are most likely to occur under a *Constrained Limited-Response equilibrium*; escalation spirals and all-out conflicts under an *Escalatory Limited-Response equilibrium*.[15] Under either of these equilibria, a Challenger will initiate a crisis, but only when it judges it very likely that Defender will capitulate and that the outcome will be *Defender Concedes*. In either case, if Defender in fact responds, it will come as a complete surprise to Challenger. What distinguishes the two equilibrium forms from one another, however, is the inference that Challenger draws after observing Defender's response. Under a Constrained Limited-Response equilibrium, Challenger believes that Defender will counteresclate if it further intensifies the conflict. Hence, it backs off, and a limited conflict ensues.

Zagare (2016) argues that a political compromise was reached, and the Cuban crisis resolved, precisely because the Soviet leadership (i.e., Premier Nikita Khrushchev) came to believe, even before the blockade was announced, that the United States was prepared to remove the missiles by force if necessary. By contrast, under an Escalatory Limited-Response equilibrium, Challenger believes that escalation will likely force Defender to capitulate. All out conflicts like World War I occur when that belief is mistaken.

In one sense, the explanations that Zagare develops for these and other real-world conflicts can also be considered a test of perfect deterrence theory's equilibrium predictions. Specifically, the perfect Bayesian equilibria of the incomplete information models that he applies make specific predictions about not only the likelihood of certain events, but also about the conditions (i.e., specific beliefs) that are necessary and sufficient to bring them about. Explanations are developed when a plausible connection between these conditions and a real-world outcome is established (Riker, 1990, p. 175). But the explanation can also be considered as additional (albeit limited) evidence of perfect deterrence theory's empirical robustness.

POLICY IMPLICATIONS

Classical deterrence theory and perfect deterrence theory stand in sharp opposition to one another. The classical formulation is plagued by logical inconsistencies and empirical anomalies. By contrast, perfect deterrence theory respects the laws of logic and is not contradicted by the empirical record. In addition, it has performed quite well when tested directly—in fact, better than other theories, formal or not. The distinctions (axiomatic, theoretical and empirical) are not simply arcane academic nitpicks. The policy implications of these two competing formulations of deterrence theory are also widely divergent, as Table 3 shows. In other words, there are significant real-world consequences for which variant is used, not only to understand how interstate relationships operate, but also how best to navigate through them.

One important difference can be traced to the conclusions that the two theoretical frameworks reach about the implications of manipulating the cost of conflict. As already noted, classical deterrence theorists, but especially decision-theoretic deterrence theorists, assumed fixed preferences. Specifically, players are assumed to always prefer an advantage to the status quo, and to always prefer backing

Table 3. Classical Deterrence Theory and Perfect Deterrence Theory: Policy Prescriptions

	Classical Deterrence Theory	**Perfect Deterrence Theory**
Policies:		
Overkill capability	Supports	Opposes
Minimum deterrence	Opposes	Supports
"Significant" arms reductions	Opposes	Supports
Proliferation	Supports	Opposes
Negotiating stances	Coercive, based on increasing war costs and inflexible bargaining tactics	Conditionally cooperative, based on reciprocity

off rather than enduring the costs associated with conflict. As a consequence, any change in the cost of conflict will have the same impact relative to these other outcomes. Since they assume that the probability of successful deterrence increases monotonically as these costs rise, they see no obvious limit to the pacifying impact that can be caused by expanded expenditures for weapons systems that are capable of imposing pain and suffering on an opponent. For this reason, most classical deterrence theorists favor an "overkill capability," opposed "significant" arms control agreements, and stand in opposition to minimum deterrence policies.

In perfect deterrence theory, however, the cost of conflict relative to other variables is not fixed. Rather it is gauged against two other important strategic variables. The first, of course, is the value of the status quo. The second reference point is the value of concession. In perfect deterrence theory, the players may, or may not, prefer to concede rather than execute a deterrent threat. As a consequence, there are two strategically significant cost thresholds—one minimum and the other maximum. The minimum cost threshold is the point below which deterrence cannot succeed. This is the point separating threats that are capable from those that are not. The two theoretical frameworks are not at odds over the strategic impact of increasing conflict costs beyond this minimum cost threshold. Both agree that deterrence success can be achieved only when this threshold is crossed.

In perfect deterrence theory, however, the maximum cost threshold is also strategically significant. Once this point is reached, further increases in the cost of conflict do not contribute to the probability of deterrence success. Rather than an overkill capability, then, the logic of perfect deterrence theory is consistent with a policy of minimum deterrence, which rests on a threat that is costly enough to deter an opponent, but that is not so costly that the threat itself is rendered incredible.

An equally important difference stems from the fact that in perfect deterrence theory, there is no simple monotonic relationship between the cost of conflict and the stability of the status quo, as there is in classical deterrence theory. Since extended deterrence becomes more and more difficult to maintain as conflict costs rise, an increased military capability may actually have destabilizing implications. In other words, in perfect deterrence theory, increased conflict costs can be stabilizing under some circumstances, but under others, they may have the opposite effect. For this and other reasons, perfect deterrence theory stands in opposition to unconstrained armaments programs and an overkill capability, and supports defense postures that are sufficiently capable but not unnecessarily redundant.

Perhaps the most significant policy difference between classical and perfect deterrence theories concerns proliferation policies. Those classical deterrence theorists who fully understand the logical implications of their paradigm favor them. In the past, Mearsheimer (1990, 1993), Posen (1993), Van Evera (1990/1991), and Waltz (1981), *inter alia*, argued that nuclear weapons can help stabilize contentious interstate rivalries and that, therefore, nuclear weapons should be disseminated selectively. More recently, Waltz (2012) made the case that if Iran acquired nuclear weapons, the possibility of an all-out conflict between Iran and Israel would be all but eliminated.[16]

Of course, those deterrence theorists who favor proliferation argue that nuclear technology should not be distributed indiscriminately. An exception, they assert, should be made for so-called rogue states that operate on the fringes of the interstate system and whose behavior is borderline rational. Of course, there is no such category in realism, classical or neo, or in the standard formulations of

deterrence theory, where all states are taken as like units (Waltz, 1979; Mearsheimer, 1990). Nonetheless, this qualifier is frequently called on to help those classical deterrence theorists who favor proliferation avoid a conclusion that they wish to avoid. But the very fact that this anomalous category of states exists is further evidence that realism in general, and classical deterrence theory in particular, are degenerative research programs (Vasquez, 1997).

The proliferation of nuclear and other weapons of mass destruction is not a policy prescription that can be derived from the axioms of perfect deterrence theory. To be sure, the increased costs of conflict that are associated with these weapons imply, *ceteris paribus*, an increased probability of deterrence success. But all things are hardly ever equal in interstate relations. For one thing, the minimum cost threshold that renders a threat capable can also be achieved with more conventional weapons. For another, the increased probability of success that nuclear weapons imply must be balanced against the risks associated with the possibility of a massive breakdown of deterrence. Since these risks are seen as considerable in perfect deterrence theory, it opposes proliferation policies, selective or otherwise.

There is one additional significant policy dispute between classical deterrence theory and perfect deterrence theory that remains to be discussed: how best to manage a crisis and how best to negotiate or bargain with an opponent. This question has received considerable attention from decision-theoretic deterrence theorists from Schelling (1960, 1966) on. In general, decision-theoretic deterrence theorists favor coercive bargaining tactics that involve either increasing the cost of conflict or making an irrevocably commitment to a hard-line policy so that the opponent is forced to make the difficult choice between war and peace. The standard example involves a player forfeiting control in a game of Chicken by pulling the steering wheel off the steering column. Snyder (1972) provides a useful summary of less colorful (but equally counterintuitive) tacit negotiating tactics.

Empirically, the stratagems that decision-theoretic deterrence theorists recommend have no basis in fact (Huth, 1999, p. 74; Danilovic, 2002; Snyder and Diesing, 1977, pp. 489–490; Betts, 1987, p. 30)—and for good reason. As Jervis (1979, p. 292) notes, many of Schelling's policy prescriptions are "contrary to common sense." Rapoport (1992, p. 482) found the tacit bargaining tactics conjured up by decision-theoretic deterrence theorists to be just plain "bizarre," while Morgenstern (1961, p. 105) concluded that they would be "dangerous should they have an influence on policy."

By way of contrast, perfect deterrence theory recommends conditionally cooperative bargaining stances rooted in reciprocity. The large majority of the deterrence equilibria in its constituent models are supported by strategies that offer to cooperate if the other player cooperates, but threatens not to cooperate if the other player does not reciprocate. Hard-line (unconditionally noncooperative) and soft-line (unconditionally cooperative) bargaining stances generally do not lead to either stable relationships or political compromises.

It is noteworthy that the efficacy of conditionally cooperative negotiating stances is well established not only in perfect deterrence theory, but also in standard diplomatic practice. For example, in a large-N statistical analysis of extended deterrence relationships, Huth (1988) found that firm-but-flexible negotiating styles and tit-for-tat deployments are highly correlated with extended deterrence success. Each of these more or less standard diplomatic or military postures involves reciprocity, a behavioral norm that is pervasive in interstate relationships, contentious or not (Cashman, 1993, ch. 6). It is also a norm that

decision-theoretic deterrence theorists have a hard time explaining. It is, in a word, inconsistent with their underlying conceptual framework (Zagare, 2011, p. 55).

CODA

In this article, I have drawn a sharp contrast between classical deterrence theory and perfect deterrence theory. Perfect deterrence theory is a completely general theory of conflict initiation, escalation, and resolution. It is applicable across time and space. Indeed, its empirical domain is not restricted to contentious nuclear relationships. Rather, its analytic framework can be used to understand the full range of situations wherein at least one actor's goal is to preserve the existing distribution of value, which is Arnold Wolfers's (1951) neologism for the status quo. Unlike classical deterrence theory, perfect deterrence theory is logically consistent and in accord with the empirical record. In a direct test, perfect deterrence theory even outperformed what is, according to Bennett and Stam (2000, p. 451), "one of the most important theories of international conflict." Observing all this, Quackenbush (2011, p. 74) recommends that the academic and policy communities take note.

In perfect deterrence theory, a capable threat is a necessary condition for deterrence success; but it is not sufficient, as it is in classical deterrence theory. Threat credibility plays an important role in the operation of both direct and extended deterrence relationships, but it is neither necessary nor sufficient for deterrence to prevail; under certain conditions, the presence of a credible threat may actually precipitate a deterrence failure. In perfect deterrence theory, the cost of conflict and status quo evaluations are also important strategic variables, insofar as their values determine the characteristics of the players' threats.

Significantly different policy recommendations also distinguish perfect deterrence theory from its classical rival. Perfect deterrence theory is consistent with a policy of minimum deterrence, recommends a conditionally cooperative diplomatic approach to resolve disputes, opposes even the selective proliferation of nuclear weapons and other weapons of mass destruction, and supports arms control agreements and other limitations on redundant military expenditures.

REFERENCES

Achen, C. H. (1987). A Darwinian view of deterrence. In J. Kugler & F. C. Zagare (Eds.), *Exploring the stability of deterrence* (pp. 91–105). Denver, CO: Lynne Rienner Publishers.

Bennett, D. S., & Stam, A. C. (2000). A universal test of an expected utility theory of war. *International Studies Quarterly, 43*, 451–480.

Betts, R. K. (1987). *Nuclear blackmail and nuclear* balance. Washington, D.C.: Brookings.

Bracken, P. (1983). *The command and control of nuclear forces*. New Haven, CT: Yale University Press.

Brodie, B. (Ed.). (1946). *The absolute weapon: Atomic power and world order*. New York: Harcourt Brace.

Brodie, B. (1959). The anatomy of deterrence. *World Politics, 11*, 173–179.

Bueno de Mesquita, B. (1981). *The war trap*. New Haven, CT: Yale University Press.

Bueno de Mesquita, B., & Lalman, D. (1992). *War and reason: A confrontation between domestic and international imperatives*. New Haven, CT: Yale University Press.

Bundy, M. (1983). The bishops and the bomb. *New York Review of Books, 30*(10) (June 16), 3–8.

Carr, E. H. (1939). *The twenty years' crisis: 1919–1939*. New York: Harper & Row.

Cashman, G. (1993). *What causes war? An introduction to theories of international conflict*. New York: Lexington Books.

Crescenzi, M. J. C., Best, R. H., & Kwon, B. R. (2010). Reciprocity in international studies. In Robert A. Denemark (Ed.), *The international studies*

encyclopedia (Vol. 9, pp. 6115–6133). Oxford: Wiley-Blackwell.

Danilovic, V. (2002). *When the stakes are high: Deterrence and conflict among major powers*. Ann Arbor: University of Michigan Press.

Förster, S. (1999). Dreams and nightmares: German military leadership and the images of future warfare, 1871–1914. In M. F. Boemeke, R. Chickering, & S. Förster (Eds.), *Anticipating total war: The German and American experiences, 1871–1914* (pp. 343–376). Washington, DC: German Historical Institute.

Freedman, L. (1989). *The evolution of nuclear strategy*. 2d ed. New York: St. Martin's.

Gaddis, J. L. (1986). "The long peace: Elements of stability in the postwar international system." *International Security, 10*, 99–142.

Gaddis, J. L. (1987). *The long peace: Inquiries into the history of the Cold War*. New York: Oxford University Press.

Gaddis, J. L. (1997). *We now know: Rethinking Cold War history*. New York: Oxford University Press.

Geller, D. S., & Singer, J. C. (1998). *Nations at war: A scientific study of international conflict*. Cambridge, U.K.: Cambridge University Press.

Harsanyi, J. C. (1967–1968). Games with incomplete information played by "Bayesian" players (3 parts). *Management Science, 14* (Series A), 159–182, 320–334, 486–502.

Harvey, F. P. (1998). "Rigor mortis, or rigor, more tests: Necessity, sufficiency, and deterrence logic." *International Studies Quarterly, 42*, 675–707.

Huth, P. K. (1988). *Extended deterrence and the prevention of war*. New Haven, CT: Yale University Press.

Huth, P. K. (1999). Deterrence and international conflict: Empirical findings and theoretical debates. *Annual Review of Political Science, 2*, 61–84.

Intriligator, M. D., & Brito, D. L. (1984). Can arms races lead to the outbreak of war? *Journal of Conflict Resolution, 28*, 63–84.

Jervis, R. (1979). Deterrence theory revisited. *World Politics, 31*, 289–324.

Jervis, R. (1985). Introduction. In R. Jervis, R. N. Lebow, & J. Gross Stein (Eds.), *Psychology and deterrence* (pp. 1–12). Baltimore, MD: Johns Hopkins University Press.

Jervis, R. (1988). Realism, game theory, and cooperation. *World Politics, 40*, 317–349.

Johnson, J. C., Leeds, B. A., & Wu, A. (2015). Capability, credibility, and extended deterrence. *International Interactions, 41*, 309–336.

Kaufmann, W. (1956). The requirements of deterrence. In W. Kaufmann (Ed.), *Military policy and national security* (pp. 12–38). Princeton, NJ: Princeton University Press.

Kugler, J. (2012). A world beyond Waltz: Neither Iran nor Israel should have the bomb. http://www.pbs.org/wgbh/pages/frontline/tehranbureau/2012/09/opinion-a-world-beyond-waltz-neither-iran-nor-israel-should-have-the-bomb.html.

Lebow, R. N. (1981). *Between peace and war: The nature of international crisis*. Baltimore, MD: Johns Hopkins University Press.

Lebow, R. N. (1984). Windows of opportunity: Do states jump through them? *International Security, 9*, 147–186.

Legro, J. W., & Moravcsik, A. (1999). Is anybody still a realist? *International Security, 24*, 5–55.

Levy, J. S. (1988). "When do deterrent threats work?" *British Journal of Political Science, 18*, 485–512.

Martin, L. (1999). The contributions of rational choice: A defense of pluralism. *International Security, 24*, 74–83.

Mearsheimer, J. J. (1983). *Conventional deterrence*. Ithaca, NY: Cornell University Press.

Mearsheimer, J. J. (1990). Back to the future: Instability in Europe after the Cold War. *International Security, 15*, 5–56.

Mearsheimer, J. J. (1993). The case for a Ukrainian nuclear deterrent. *Foreign Affairs, 72*, 50–66.

Morgan, P. M. (1977). *Deterrence: A conceptual analysis*. Beverly Hills, CA: SAGE.

Morgenstern, O. (1961). Review of *The strategy of conflict*. *Southern Economic Journal, 28*, 105.

Morton, R. B. (1999). *Methods and models: A guide to the empirical analysis of formal models in political science*. Cambridge, U.K.: Cambridge University Press.

Most, B. A., & Starr, H. (1989). *Inquiry, logic, and international politics*. Columbia, SC: University of South Carolina Press.

Organski, A. F. K. (1958). *World politics*. New York: Knopf.

Organski, A. F. K., & Kugler, J. (1980). *The war ledger*. Chicago: University of Chicago Press.

Posen, B. R. (1993). The security dilemma and ethnic conflict. *Survival, 35*, 27–47.

Powell, R. (1990). *Nuclear deterrence theory: The search for credibility*. New York: Cambridge University Press.

Quackenbush, S. L. (2006). "Identifying Opportunity for Conflict: Politically Active Dyads." *Conflict Management and Peace Science, 23*(1), 37–51.

Quackenbush, S. L. (2010a). General deterrence and international conflict: Testing perfect deterrence theory, *International Interactions, 36*, 60–85.

Quackenbush, S. L. (2010b). Territorial issues and recurrent conflict. *Conflict Management and Peace Science, 27*, 239–252.

Quackenbush, S. L. (2011). *Understanding general deterrence: Theory and application*. New York: Palgrave Macmillan.

Quackenbush, S. L., & Venteicher, J. F., II (2008). Settlements, outcomes, and the recurrence of conflict, *Journal of Peace Research, 45*, 723–742.

Rapoport, A. (1992). Comments on "Rationality and misperceptions in deterrence theory." *Journal of Theoretical Politics, 4*, 479–484.

Riker, W. H. (1990). Political science and rational choice. In J. E. Alt & K. A. Shepsle (Eds.), *Perspectives on positive political economy* (pp. 163–181). Cambridge, U.K.: Cambridge University Press.

Schelling, T. C. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Schelling, T. C. (1966). *Arms and influence*. New Haven, CT: Yale University Press.

Selten, R. (1975). A re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory, 4*, 25–55.

Senese, P. D., & Quackenbush, S. L. (2003). Sowing the seeds of conflict: The effect of dispute settlements on durations of peace. *Journal of Politics, 65*, 696–717.

Snyder, G. H. (1972). Crisis bargaining. In C. F. Hermann (Ed.), *International crises: Insights from behavioral research* (pp. 217–256). New York: Free Press.

Snyder, G. H. (2002). Mearsheimer's world: Offensive realism and the struggle for security. *International Security, 27*, 149–173.

Snyder, G. H., & Diesing, P. (1977). *Conflict among nations: Bargaining, decision making and system structure in international crises*. Princeton, NJ: Princeton University Press.

Sullivan, M. P. (2002). *Theories of international relations: Transition vs. persistence*. New York: Palgrave Macmillan.

Van Evera S. (1990/1991). Primed for peace: Europe after the Cold War. *International Security, 15*, 7–57.

Vasquez, J. A. (1997). The realist paradigm and degenerative versus progressive research programs: An appraisal of neotraditional research on Waltz's balancing proposition. *American Political Science Review, 91* (December), 899–912.

Walt, S. M. (1999). Rigor or rigor mortis? Rational choice and security studies. *International Security, 23*, 5–48.

Waltz, K. N. (1964). The stability of the bipolar world. *Daedalus, 93*, 881–909.

Waltz, K. N. (1979). *Theory of international politics*. Reading, MA: Addison-Wesley.

Waltz, K. N. (1981). The spread of nuclear weapons: More may be better. *Adelphi Paper No. 171*. London: International Institute for Strategic Studies.

Waltz, K. N. (1990). Nuclear myths and political realities. *American Political Science Review, 84*, 731–745.

Waltz, K. N. (1993). The emerging world structure of international politics. *International Security, 18*, 44–79.

Waltz, K. N. 2012. Why Iran should get the bomb: Nuclear balancing would mean stability. *Foreign Affairs, 91*(4), 2–5.

Wolfers, A. (1951). The pole of power and the pole of indifference. *World Politics, 4*, 39–63.

Zagare, F. C. (1987). *The dynamics of deterrence*. Chicago: University of Chicago Press.

Zagare, F. C. (1996a). Classical deterrence theory: A critical assessment, *International Interactions*, 21, 365–387.

Zagare, F. C. (1996b). The rites of passage: Parity, nuclear deterrence, and power transitions. In J. Kugler & D. Lemke (Eds.), *Parity and war: Evaluations and extensions of the war ledger* (pp. 249–268). Ann Arbor: University of Michigan Press.

Zagare, F. C. (2007). Toward a unified theory of interstate conflict. *International Interactions, 33*, 305–327.

Zagare, F. C. (2011). *The games of July: Explaining the Great War*. Ann Arbor: University of Michigan Press.

Zagare, F. C. (2016). A general explanation of the Cuban missile crisis. *International Journal of Peace Economics and Peace Science*, *1*, 91–118.

Zagare, F. C., & Kilgour, D. M. (2000). *Perfect deterrence*. Cambridge, U.K.: Cambridge University Press.

Zagare, F. C., & Kilgour, D. M. (2003). Alignment patterns, crisis bargaining, and extended deterrence: A game-theoretic analysis. *International Studies Quarterly*, *47*, 587–615.

Zakaria, F. (2001). Don't oversell missile defense: The old theory of nuclear deterrence still makes sense. Just ask the man who invented it. *Newsweek*, May 14, p. 34.

NOTES

1. Classical deterrence theory is also sometimes referred to as *rational deterrence theory.* However, there are other theories of deterrence based on rational choice and therefore seem to fall under the umbrella of rational deterrence theory.
2. The empirical basis for this claim is weak; see, for example, Förster (1999).
3. For a full discussion, see Zagare (1996a).
4. For a discussion of several unsuccessful attempts to eliminate this logical inconsistency, see Zagare and Kilgour (2000, ch. 2).
5. For specific examples of this theoretical pirouette, see Schelling (1966, p. 37) or Brodie (1959, p. 293).
6. The assumption of differentiated actors is not ad hoc in perfect deterrence theory, as it is in most manifestations of classical deterrence theory. Perfect deterrence theory is connected, theoretically, with power transition theory (Organski and Kugler, 1980), which sees the international system as hierarchical rather than anarchistic. In a hierarchical system, the dominant state and its allies are generally content with the status quo. For a discussion of the linkage between power transition theory and perfect deterrence theory, see Zagare (1996b, 2007).
7. The Asymmetric Escalation Game and the Tripartite Crisis Game (Zagare and Kilgour, 2003) are also component parts of the theory.
8. Powell (1990), whose work is in the tradition of Schelling and other decision-theoretic deterrence theorists, is an exception to the rule. His analysis of nuclear deterrence is consistent with Selten's criterion. But there are other problems with his explanation of the long peace. For a discussion of this point, see Zagare and Kilgour (2000, pp. 54–57).
9. Lebow's (1981, p. 85) conclusion is typical. For Lebow "four conditions emerge as crucial to successful deterrence. Nations must (1) define their commitment clearly, (2) communicate its existence to possible adversaries, (3) develop the means to defend it, or to punish adversaries who challenge it, and (4) demonstrate their resolve to carry out the actions this entails."
10. Not surprisingly, Lebow (1981) is a case in point.
11. See, for example, Mearsheimer (1983).
12. For a detailed overview, see Geller and Singer (1998).
13. For overviews, see Cashman (1993, ch. 6) and Sullivan (2002, ch. 9). For details about reciprocity, see Crescenzi, Best, and Kwon (2010).
14. For a detailed discussion, see Zagare (2011, pp. 184–186) or Zagare and Kilgour (2000, ch. 10).
15. Both games are examined under incomplete information. Hence, these are perfect Bayesian equilibria.
16. For a trenchant rebuttal, see Kugler (2012).

Frank C. Zagare

THE POLIHEURISTIC THEORY OF POLITICAL DECISION-MAKING

INTRODUCTION

The poliheuristic theory of political decision-making provides the most thorough framework for examining process and choice in government and politics. Poliheuristic theory combines the two emphases of process and choice to provide testable hypotheses concerning how personal and situational variables