

## Oxford Handbooks Online

### Psychological Approaches

James Goldgeier and Philip E. Tetlock

The Oxford Handbook of International Relations

*Edited by Christian Reus-Smit and Duncan Snidal*

Print Publication Date: Aug 2008

Subject: Political Science, International Relations, Political Methodology

Online Publication Date: Sep 2009 DOI: 10.1093/oxfordhb/9780199219322.003.0027

### Abstract and Keywords

This article addresses the following question: Under what conditions, in what ways, and to what degree should various strands of theorizing at the international level — structural realism, institutionalism, constructivism — incorporate various categories of psychological theory? The article samples the diverse ways in which such interweaving of levels of analysis has either already begun or could be readily initiated given recent empirical and theoretical developments. The most promising candidates for conceptual integration are organized into four broad categories which identify: the appropriate boundary conditions for the applicability of clashing hypotheses within the structural-realist tradition; the factors that facilitate and impede the creation of international institutions and norm enforcement mechanisms within the institutionalist tradition; the factors that determine whether policy-makers and epistemic communities frame issues in terms of the logic of consequential action or the logic of obligatory action within the constructivist tradition; and the price that international relations theorists pay for placing a hedgehog-style premium on theoretical parsimony and the value of adopting a more flexible, foxlike, contextualist style in future theory-building exercises.

Keywords: international relations, psychological theory, structural realism, institutionalism, constructivism

---

DEBATES over whether international relations theories are reducible—in some mysterious metaphysical sense—to psychology should have faded long ago into intellectual history (Waltz 1959). As we have noted in earlier essays (Tetlock and Goldgeier 2000; Goldgeier and Tetlock 2001) and as others have patiently pointed out over the last few decades (Kelman and Bloom 1973; Jervis 1976; George 1980), the right question is more nuanced and less polarizing: under what conditions, in what ways, and to what degree should various strands of theorizing at the international level—structural realism, institutionalism, constructivism—incorporate various categories of psychological theory?

With this as our starting point, our goal in this chapter is to sample the diverse ways in which such interweaving of levels of analysis has either already begun or could be readily initiated given recent empirical and theoretical developments. We organize the most promising candidates for conceptual integration into four broad categories: (1) identifying the appropriate boundary conditions for the applicability of clashing hypotheses within the structural—realist tradition (for example, when and why do states pursue more or less risk-seeking security policies?); (2) identifying the factors that facilitate and impede the creation of international institutions and norm enforcement mechanisms within the institutionalist tradition (for example, when and why do states make commitments that apparently reduce their sovereignty and freedom of action?); (3) identifying the factors (p. 463) that determine whether policy-makers and epistemic communities frame issues in terms of the logic of consequential action or the logic of obligatory action within the constructivist tradition (for example, when and why do powerful norms—such as the taboos against using nuclear and chemical weapons—take hold?); and (4) highlighting the price that international relations theorists pay for placing a hedgehog-style premium on theoretical parsimony and the value of adopting a more flexible, foxlike, contextualist style in future theory-building exercises (for example, when and why are expert observers most vulnerable to contingency blindness?).

Before launching this exercise, we should, however, reassure readers who worry that psychology is too fragmented—insufficiently paradigmatic—to be a reliable guide for revising international relations theories. To be sure, the most relevant fields of psychology—cognition, social, personality—have their controversies. Disagreements persist on basic issues such as the degree to which human behavior is under the control of stable personality dispositions or shifting situational cues (Mischel 1999) or whether alleged cognitive biases represent departures from rationality or adaptive responses to the environment (Gigerenzer and Goldstein 1996; Kahneman and Tversky 2000; Gigerenzer and Selten 2001). But we take no extreme theoretical stands here and see little disagreement on the circumscribed and well-replicated findings that we import into our analysis. And, in view of past borrowing practices in international relations theory (Waltz 1979), we suspect that future historians of ideas will find it curious if international relations theorists should now prove choosier about feeder disciplines than microeconomists who have already made extensive use of many of the findings we highlight (Rabin 1998; Camerer 2003; Fehr, Fischbacher, and Kosfeld 2005).

## 1 Realism

Thanks partly to Jervis (1976) and partly to the high costs associated with miscalculations over war and peace issues, psychological approaches have made greater inroads into realist thinking than in the other major international relations traditions. Often these works are implicit in their use of psychology, relying on an unspecified notion of “perceptions” to fill in gaps in realist thinking—for example, Walt's transformation (1987)

of balance-of-power theory into balance-of-threat theory or Christensen and Snyder's efforts (1990) using perceptions of the offense—defense balance to explain when states tie themselves too closely to the fate of other states as opposed to when they fail to do enough to assist others in the international system.

(p. 464)

A recent attempt to marry explicit psychological theorizing with realist theory is Taliaferro (2004), who draws heavily on prospect theory to explain why great powers often initiate risky ventures in peripheral areas and persist even as prospects for victory dim and costs escalate.

Prospect theory, developed by Kahneman and Tversky (1979; 2000) and their collaborators, shows that individuals make decisions under risk differently depending on whether they confront prospective gains or losses, and that individuals are much more risk acceptant when faced with losses and risk averse in the domain of gains. Problems that are logically equivalent but framed differently (for example, number of lives saved versus numbers lost) are treated sharply differently by respondents. Experimental evidence shows that prospective losses hurt more than gains gratify by roughly two to one, and much more than that when the endowments are charged with moral and emotional symbolism (McGraw, Tetlock, and Kristel 2003).

Behavioral economists (e.g. Rabin 1998; Thaler 2000; Camerer 2003) have embraced prospect theory's findings to reshape their field. Not so political scientists, even though, as McDermott (1998; 2004) notes, the theory is based not on personality predispositions but on situationally activated response tendencies, which should appeal to structural theorists in international relations. Indeed, recent work has shown that people think about costs and benefits with different parts of their brain, rather than making a single calculation as utility theorists would assume (Cowen 2006), thus providing a neurological substrate for prospect theory and underscoring how far certain lines of reductionist inquiry can be taken.

Applications of prospect theory to international relations have grown (e.g. Levy 1992; Farnham 1994; McDermott 1998; 2004; Taliaferro 2004; Mercer 2005), although skeptics have stressed the difficulty of specifying in a nontautological fashion decision-makers' reference points and their calculations of gains and losses (Tetlock 1999b). Taliaferro (2004) focuses on aversion to losses in states' relative power, international status, or prestige as driving the risky behavior, thereby combining prospect theory's emphasis on framing with defensive realism's emphasis on relative power, and ducking the complexities of gauging the role of domestic politics in shaping how leaders frame problems as prospective gains or losses.

Prospect theory applies directly to a central debate within realism—that between the defensive realists (Waltz 1979) and the offensive realists (Mearsheimer 2001). Waltz and his followers have argued that states seek to maximize their security in an anarchic state system that requires them to practice balance-of-power politics. Their main goal is to

assure their survival, which requires accruing power by either building military might or finding allies, but states can become satisfied with the status quo if they feel sufficiently powerful vis-à-vis other states. For Mearsheimer, no power advantage can ever be big enough, and thus states can be modeled as power maximizers pursuing hegemonic aspirations. In either case, states have to form accurate representations of the world and respond to the actions of other (p. 465) powers in a timely manner, which makes them deeply suspicious and receptive to worse-case assumptions about the motives of other states in the system.

Waltzian states can become satisfied with their position in the system. Such states will be loss averse and disinclined to pursue expansionist policies that might trigger counterbalancing by other states. In Mearsheimer's world, states are never satisfied short of hegemony. Prospect theory would suggest that Waltz is more correct when states are faced with prospective gains; Mearsheimer when states are facing losses. The United States was reluctant to suffer casualties in interventions abroad in the 1990s (Haiti, Somalia, Bosnia, Kosovo) as it sat confidently atop the international system, and acted in a highly risk-averse manner, as Waltz would predict. But, in the wake of the attacks of 11 September, it became more willing to take risks. To invoke much more extreme examples, in which regime type and leader personalities also play powerful roles, Nazi Germany in 1939 and Imperial Japan in 1941 were willing to roll the dice in environments in which they had either incurred steep losses or faced the prospect of such losses and pursued high-risk efforts to reverse or avoid those fates, as Mearsheimer would predict. Rather than seeing the feud within realism as one that can be definitively resolved in favor of one approach or another, psychology helps us to reconnoiter boundary conditions: to appreciate that defensive realism is more likely to apply when decision-makers frame problems as potential gains, and offensive realism is more likely to apply when states confront potential losses.

Prospect theory and associated work on the endowment effect (the increase in value one places on something once one possesses it) also explains a major difference between deterrence and compellence—namely, that it is far more difficult to induce a state to give up something that it already possesses than to prevent it from taking something that it does not possess (Schelling 1967). The Clinton administration should have paid heed in the 1990s. Bosnia was not part of Serbia, and Slobodan Milošević was willing to negotiate a settlement over something he did not possess in 1995. Believing the lesson from that case was that he could be shaken by a brief demonstration of force, the Clinton administration thought in 1999 that it would not take much to get him to cease and desist in Kosovo. But Kosovo was part of Serbia, and Milosevic was willing to take much higher risks to keep something in his possession than he had been to acquire new territory (Chollet and Goldgeier 2002). Similarly, the United States was willing to use force when the North Koreans attacked the South in 1950 in order to ensure the viability of containment but was not willing to roll-back Communism to liberate Hungary in 1956. Psychologists will predict that deterrence is roughly twice as easy as compellence in analogous situations—with two key exceptions: (1) when the target state feels a deep emotional attachment to the territory in question, in which case compellence becomes

many orders of magnitude more difficult than deterrence; and (2) when the target state has not internalized the territory as part of its self-definition (its “endowment”), in which case compellence may not be appreciably more difficult than deterrence.

(p. 466)

Given the widespread influence of prospect theory in economics, it is curious that the theory has not had comparable impact in political science, particularly in international relations. Mercer (2005) has noted that of the 2,000 most recent citations to Kahneman and Tversky's article (1979) on prospect theory, half come from economics and related fields, one-third from psychology, and only one-twentieth from political science; the article received a mere eight citations in the *American Political Science Review* between 1985 and 2003. Mercer (2005, 18) reasons that choices made by decision-makers in the world of international relations are not the same as choices between gambles in the lab, but he also concludes that, given how well the theory has been utilized in economics, “it seems likely that the resistance is to the discipline of psychology rather than to problems unique to prospect theory.”

## 2 Institutionalism

The institutionalist challenge arose in response to the realists' pessimistic appraisal of the prospects for cooperation in an anarchic international system. The micro-economic variant of institutionalism articulated by Keohane (1984) accepted that the international system is anarchic and populated by unitary, egoistic actors, but argued that those actors can create institutional frameworks that, by lowering their fear of being exploited by free riders, permit them to secure the gains obtainable through cooperation.

In the neoliberal variant of the institutionalist model, when actors recognize that they confront a repeated play situation and that the transaction costs for continuing to improvise ad hoc solutions are high, they quite rationally enter into binding compacts that can persist despite shifting balances of power among member states. Psychology has something to say about when and why states seek to create institutions in the first place, when and why states abide by institutional norms and principles, and whether and when institutions can over time transform how decision-makers view their own state's interests.

A key psychological element in the formation of institutions is transparency, which includes both the ability to discern basic facts and the ability to draw sound causal inferences. Epistemic communities (Haas 1997) can both advance and reflect causal transparency. At one end of the transparency continuum are policies shaped by epistemic communities possessing great authority and prestige by virtue of their scientific achievements in teasing apart cause—effect relationships. This is especially likely in domains where investigators can test theoretical hypotheses in controlled and replicable experiments (and there is minimal reliance on what (p. 467) Tetlock and Belkin (1996)

call speculative counterfactual control groups). Examples include the near consensus among biomedical specialists on the right strategies (or methods of developing strategies) for coping with transnational epidemics, among climatologists on the effect of chlorofluorocarbons (CFCs) on the ozone layer of the atmosphere, or among physicists and engineers on the prerequisites for nuclear proliferation. Political elites in the advanced countries are usually disinclined to question advice emanating from these communities on their issues of expertise.

But this deference does not hold in those policy domains in which the epistemic communities are either deeply divided or relatively undeveloped, as was the case with respect to the sources, scope, pace, and effects of global warming until recently. But, as evidence mounts, and fears that a tipping point may be closer than previously recognized, just as has been the case in the policy domains mentioned above, even political elites who dislike regulation are becoming less willing to dismiss scientists who present evidence that global warming requires coordinated international action (Hansen 2006).

Realists often argue that institutions typically come about because the most powerful state in the system decides to create them as a means of pursuing its interests with greater legitimacy and/or efficiency. But we should also expect that, for psychological reasons, institutions are more likely to be created in those domains where there is a consensus among experts with epistemic authority, and that in the absence of such consensus, political obstacles to the formation of international institutions will often be insurmountable.

Once institutions are created, questions arise about why states comply with the demands of the institutions and whether over time institutions can transform how decision-makers view their state's interests (Simmons 1998). Here we think it useful to posit a social influence continuum that runs from compliance to internalization (Kelman 1958). At the compliance end, decision-makers abide by the rules of the institution based purely on utilitarian calculations of rewards and punishment. Institutions in which compliance is the likely source of rule-following include military alliances formed to counter direct threats and the World Trade Organization, which lowers trade barriers for members and provides sanctioning mechanisms to punish norm and rule violators.

Toward the middle of the continuum are decision-makers who do what is expected of them because they seek to establish particular social identities in the eyes of domestic or international audiences whose opinion they value, and because they are unwilling to bear the reputation costs of defecting. This form of social influence—sometimes called identification by social psychologists—should be less context specific than mere compliance. Identification should persist even when it is materially inconvenient, but it still falls short of full-fledged internalization because it will not persist when the policy ceases to serve its function of enhancing a desired social identity. Examples include states with dubious civil liberties practices or suspect environmental records signing human rights or environmental agreements in order to be part of the “civilized”

international community—decisions that can produce normative entanglements with real political consequences (for example, the 1975 Helsinki Accords basket three provisions for protecting human rights that the Soviet Union and its Warsaw Pact allies signed).

As we continue toward the internalization end of the continuum, we find states doing the right thing for what they perceive to be the right reasons. The calculation is no longer utilitarian but guided by moral, religious, or ideological ideals—in other words, decision-makers are following a logic of appropriate action rather than the logic of consequential action (March and Olsen 1998). The humanitarian assistance efforts of Scandinavian countries come to mind here.

Psychology can help delineate the conditions under which decision-makers are likely to go from compliance to internalization. Cognitive dissonance and self-perception research using forced-compliance paradigms suggest that decision-makers are especially likely to internalize attitudes consistent with their behavior when they believe they had some elements of free choice (see Bem 1967; Larson 1985). Most psychologists would agree that political actors (psychopaths excluded) are likely to internalize the norms of fair play implicit in international institutions, and these norms can become autonomous from the original interest-based actions. This internalization process should be especially reliable in democracies, in which leaders must justify departures from widely held norms of fair play to their constituents.

Rational choice approaches to norm enforcement stress the problem of free-riding; in the absence of a sovereign, no single party wants to incur the costs of punishing norm violators, so defectors escape sanction (Coleman 1990). A growing body of work in experimental social psychology and microeconomics has revealed, however, that people are often willing to make substantial sacrifices to punish cheaters—and those who fail to punish cheaters (Fehr and Schmidt 1999)—and that this self-sacrificial behavior is driven, in part, by moralistic anger (desire for retribution) aimed at those who violate shared foundational values (Tetlock et al. 2007). Indeed, brain-imaging studies have shown that people derive pleasure from inflicting pain on norm violators (Fehr, Fischbacher, and Kosfeld 2005). From this standpoint, international institutions built on shared values of distributive and procedural justice should be able to perpetuate themselves in the absence of strong central authority as long as members perceive the specific norms and norm-enforcement agencies to be anchored in those shared values. Psychology does not tell us this will be easy; there will be many occasions when the human tendency to engage in motivated reasoning causes participants to “see” evidence of wrong-doing that the alleged wrong-doers do not see (Kunda 1999) and delicate mediational and perspective-taking exercises will be required (Rouhana and Kelman 1994). Nonetheless, putting the complexities of establishing and maintaining legitimacy to the side, the more egregious the normative transgression, the more willing people (p. 469) will be to punish both transgressors and those who fail to punish transgressors (Axelrod 1984; Coleman 1990; Crawford 2000).

### 3 Constructivism

For constructivists, actors often make decisions not using the logic of consequential action, the cost—benefit analyses of realists and institutionalists, but rather using the logic of obligatory action, the normative guidelines for legitimating conduct that are specified by a given social order that is itself the emergent byproduct of interactions among actors in the system (Wendt 1999). Examples of norm-following logics include the demise of dueling and slave trading (Mueller 1989); the rise of human rights norms, including the sensitivity to racism (Risse, Ropp, and Sikkink 1999); the nonuse of weapons of mass destruction and other constraints on the conduct of war (Legro 1995; Price 1997; Tannenwald 1999); and the changing purpose of military intervention over time (Finnemore 2003).

From a constructivist perspective, every policy debate carries a social-identity subtext: What kind of people are we, the advocates of this policy, claiming to be—and what kind of people do they, the opponents of this policy, portray us to be? In the case of the ferocious debate over the wisdom of invading Iraq in 2003, many advocates laid claim to the following identities: principled enforcers of international law, prudent defenders of national interests, and even heroic liberators of the Middle East. By contrast, opponents depicted proponents as neocolonialists interested only in oil and geopolitical leverage or, somewhat more charitably, as irrational and overly emotional Americans whose judgment has been clouded by the trauma of 11 September. These identity claims were, of course, just the initial moves: Each side would do its best—in response to events and moves of the other side—to bolster its own credibility and dismantle the other side's credibility.

From a psychological perspective, debates over the “true motives” underlying policies are generally rather fruitless: People have, at best, a shaky introspective grasp of the factors driving their own behavior—and there is not much reason for supposing that people are better at divining the true drivers of the behavior of others (Bargh and Williams 2006).

In short, psychology is likely to be unhelpful in determining which masters of political spin are closer to the “truth.” The competition for the social-identity high ground in debates will be with us into the indefinite future. Psychology can, however, be helpful in determining which masters of political spin are more likely to prevail in particular battles for the identity high ground. For instance, psychological research on the implicit theories of leadership suggests that, *ceteris paribus*, people (p. 470) associate leadership with firmly upholding principles and staying the course more than they do with flexibility and midcourse corrections (Staw 1980; Tetlock 1999a), which gave George W. Bush an advantage in his 2004 re-election campaign against John Kerry.

Another realm in which psychology can help to determine who has the rhetorical advantage is that of trade-off reasoning. There is a profound disjuncture between constructivism and the more rationalist strands of international relations theory (realism and institutionalism) in how they approach trade-off reasoning (Katzenstein, Keohane,



and Krasner 1998; Ruggie 1998). Realists and institutionalists expect states rationally to measure costs and benefits. By contrast, there are large classes of issues for which constructivists should expect trade-off reasoning to be extremely difficult and politically toxic. Indeed, the very willingness to consider certain categories of trade-offs is taken as a sign in many political cultures that one is not adequately committed to core cultural values and identities.

A key issue in many identity contests is how people “frame” the various types of trade-offs underlying policy options: routine, taboo, and tragic. A routine trade-off is the type of reasoning one uses when shopping and comparing the relative importance of secular values. A trade-off is taboo if it pits a secular value such as money against a sacred value such as protection of human rights or ecosystems—a value that members in good standing of moral communities are supposed to treat as beyond monetization. Finally, a trade-off is tragic if it pits two sacred values against each other—for example, protection of endangered species against protection of indigenous cultures (Tetlock 2003).

Decision-makers caught making taboo trade-offs are often ostracized by their moral communities, and thus they go to great lengths to portray their decision process as free of any taint of taboo trade-offs; their adversaries meanwhile attempt to convince key constituencies that the boundaries of the thinkable have been breached. We can observe this dynamic process at work in debates as disparate as those over Israeli withdrawal from parts of the West Bank and the sacred city of Jerusalem, Pakistani and Indian willingness to compromise over Kashmir, or, in the United States, the solvency of Social Security or Medicare or the wisdom of drilling in the Arctic National Wildlife Refuge in Alaska.

As a general psychological proposition, adversaries should be able to block compromise whenever they can plausibly portray the policy they oppose as implicating taboo trade-offs that degrade sacred values by mixing them with secular values of money, comfort, and convenience. Those in the position of defending policies that opponents characterize as involving taboo trade-offs must work hard to claim the social identities of flexible and reasonable while avoiding the identities of immoral and unprincipled. Those in the position of opposing these policies must work hard to claim the social identities of resolute and principled while avoiding the identities of dogmatic and fanatical. All other things being equal, the advantage falls to the opponents of taboo trade-offs. The exceptions arise when political leaders with (p. 471) unusually solid credentials as defenders of the sacred value decisively change sides—as in the resolute anticommunist President Richard Nixon establishing relations with China (and abandoning a trusted ally, Taiwan) or the Israeli hardline Prime Minister Ariel Sharon withdrawing from Gaza.

From a realist or institutionalist standpoint, resistance to making taboo trade-offs (among the mass public and sometimes among policy elites as well) is an unwelcome complication. Instead of adjusting in a timely fashion to the geopolitical or economic realities of the moment—for example, compromising on security policy or agricultural

price supports—states may appear to outsiders to be displaying an ostrich-like obliviousness to trade-offs. But one theorist's bothersome anomaly can be another theorist's productive obsession.

## 4 The Need for a Fourth Perspective: Reflexivism (Theorizing About the Theorists)

Seizing on an enigmatic epigram from the classical Greek poet Archilocus, Isaiah Berlin (1953) divided thinkers into two admittedly crude categories: the hedgehogs who know one big thing and the foxes who know many things. Crossing that classification with our threefold discussion of international relations theorists yields a six-category system: hedgehog and fox realists, institutionalists, and constructivists.

We suspect that few international relations scholars qualify as true hedgehogs—so wedded to their favorite framework that they eschew anything that smacks of foxlike opportunism or borrowing from other frameworks. From a psychological perspective, cognitive styles are a matter of degree, not kind (Kruglanski 1996; Sued-feld and Tetlock 2001). Some of us are closer to the ideal-type hedgehog, faithful to our vision, and some of us closer to the ideal-type fox, prepared to improvise whatever patchwork quilt of ideas best fits the peculiarities of the current situation. And there is good reason to suppose that, just as cognitive stylistic orientations shape how people think in other domains, so they also influence professional observers of world politics.

Tetlock (2005) underscores just how applicable psychological research on cognitive styles and biases is to the judgments that professional observers make of possible pasts and futures in world politics. Drawing on a sample of almost 300, mostly Ph.D.-level professionals, he assessed the following variables: (a) theoretical allegiances to various schools of thought in social science (including realism and institutionalism); (b) cognitive styles (including the hedgehog—fox index that (p. 472) gauged the value that respondents placed on parsimony, comprehensiveness, and explanatory closure); (c) plausibility judgments of historical counterfactuals that undercut, reinforced, or were irrelevant to theoretical allegiances; (d) probability judgments of possible futures conditional on their theoretical understanding of the underlying forces at work being correct or conditional on alternative views being correct; (e) willingness to change one's mind in a roughly Bayesian fashion in response to learning that relatively expected or unexpected events either did or did not transpire.

Turning first to judgments of historical counterfactuals, we find evidence for some basic cognitive processes in theoretical reasoning about international relations. Cognitive theory leads us to expect that, confronted by the quirkily complex, path-dependent meanderings of history (Pierson 2004), even experts need to resort to simplifying strategies. One obvious strategy is to rely on one's causal schemata to impose order by portraying what we all know did happen as the inevitable, or at least highly probable, result of the operation of one's favorite covering laws on well-defined antecedent conditions (such as a bipolar or multipolar world)—in the process ruling out large classes of alternative outcomes (historical counterfactuals) as impossible, or at least highly improbable. Many experts do exactly that. For instance, experts sympathetic to structural

realism and to nuclear-deterrence theory are more likely to dismiss dissonant counterfactual scenarios that imply that the cold war could easily have become a hot war. In their view, structural realism and nuclear-deterrence theories are well positioned to explain the general pattern of observed outcomes: the bipolar stability of the cold war and the reluctance of the superpowers to provoke each other in regional contests.

This fit provides support for the theory only insofar as we are prepared to grant that, if we could rerun the history of the cold war hundreds of times, each time manipulating theoretically trivial background conditions such as the location of American and Soviet ships during the Cuban missile crisis, we would get pretty much the same pattern of outcomes. Perhaps the theoretically committed observers are right and this proposition is true, but no one has devised the time machinery necessary for checking on the veridicality of such claims, either in the affirmative (yes, history would have come back on to roughly the same trajectory) or in the negative (no, history would have diverged dramatically from what we know as international reality).

Experts vary in the degree to which they find parsimony attractive—and ambiguity aversive (hedgehogs more so than foxes in both cases). It follows that hedgehogs should be especially inclined to dismiss historical counterfactuals that mess up their causal models of the past, whatever those models might be. Again, the prediction holds. For instance, structural realists with hedgehog epistemological preferences are more likely than their fox colleagues to dismiss counterfactuals that imply that the First World War (some kind of major war among European powers in the early twentieth century) could easily have been averted by manipulating theoretically (p. 473) trivial antecedent conditions, such as whether Archduke Franz Ferdinand's carriage driver had not made a fateful wrong turn after the initial assassination attempt had failed.

We suspect much the same would be true for institutionalist hedgehogs and foxes confronted by counterfactuals that imply the European Union could easily have never emerged or could have come undone. We are less confident about constructivists, because the symbolic interactionist fluidity of much of their theorizing would seem more compatible with accepting indeterminacy in historical outcomes. But it would be worth checking whether hedgehog versus fox constructivists differ markedly in their openness to the possibility that basic features of the normative landscape of world politics today could easily have been undone by tweaking coincidences from the past. For instance, would the nuclear taboo be as taboo as it is if the United States had used such weapons in Korea, as Dwight Eisenhower seriously considered?

The term “theory” carries deterministic connotations: a framework that allows us to deduce a relatively immutable chain of causality from first premises. It is, therefore, discomfiting to contemplate how much international relations theory depends on contingency, on speculative counterfactual judgments of what was or was not once historically possible—and on what, looking forward, might now be possible. What should happen to our confidence in international relations theories if it turns out that many aspects of the international landscape that they are well positioned to explain could easily

have worked out otherwise, but for minor accidents of personalities and timing? Or, more modestly, what should happen to our confidence in these theories if it turns out that we just do not know whether these supposedly stable aspects of the landscape really are stable?

Shifting from experts' judgments of possible pasts to possible futures, we can derive a strikingly similar series of predictions from cognitive theory. For instance, we should expect the greater hedgehog aversion to ambiguity to manifest itself in imposing greater theoretical order on the long-range future than foxes are inclined to impose (or than, for that matter, may be justified given the inherent unpredictability of politics). Looking at a broad range of forecasts, Tetlock found evidence that hedgehog forecasters were indeed systematically more overconfident than were foxes in their longer-range predictions. Perhaps most serious, Tetlock also found that hedgehog forecasters were less inclined to acknowledge mistakes and more inclined to resist changing their minds as much as they should according to Bayesian benchmarks of belief updating. Tetlock documents in detail the dissonance-reduction strategies that experts—especially hedgehogs making longer-range forecasts—tended to invoke when the relatively unexpected happened. For instance, experts who were most surprised by the demise of the Soviet Union were also most inclined to embrace the close-call counterfactual defense that the disintegration of the Soviet state was the result of the odd or reckless (from an orthodox Communist point of view) liberalization policies pursued by Mikhail (p. 474) Gorbachev—and that the disintegration of the state could have been prevented even as late as the coup attempt of August 1991 (if the coup plotters had been more resolute and less intoxicated, they just might have prevailed in reinstalling the old regime).

Some international theorists may believe it is inappropriate to use psychological theory to explain the behavior of theorists—bordering on the *ad hominem*. But from a cognitive perspective, theorists are human beings attempting to make sense of a bizarrely complicated world that forces each and every one of us to rely on speculative counterfactual thought experiments to infer causality. Under such conditions, we have no alternative but to resort to simplifying strategies. Psychology can shed light on what types of strategies we use—and on whether we may be at risk of overusing them, as arguably many of the hedgehogs in the Tetlock sample seem to do.

## 5 The Influence of the Real World—and Efforts to Influence It

Like other international relations theorists, those drawing heavily on psychological approaches have often been influenced by the central issues of contemporary international politics. During the cold war, for example, policy-makers and scholars alike most feared the miscalculation that might result in nuclear conflagration, particularly after the Cuban Missile Crisis, when President Kennedy opined that the chances of a

nuclear exchange had been between one in three and even. The Nixon—Kissinger détente policy was predicated on the notion that the US—Soviet competition had to be bounded in some way so that conflicts did not escalate even while each side vied for advantage vis-à-vis the other.

Given the dangers, those using psychology to understand international relations during the cold war focused their attention on the US—Soviet relationship and particularly on crisis bargaining (e.g. Snyder and Diesing 1978; Lebow 1981; George 1983; Goldgeier 1994). Crises lend themselves to psychological approaches quite well. Decision-makers are pressed for time, they are under tremendous stress, and they have to sort through voluminous and ambiguous information. They are prone to using analogies that might or might not be appropriate (Khong 1992), to falling prey to wishful thinking about how things might turn out, and to discounting information that does not fit with their preconceptions of the adversary (Larson 1985). During the cold war, leaders often found themselves having to make decisions that carried great risks, both to their political futures and to their country's (p. 475) survival. And so many who studied cold war decision-making used psychological approaches, particularly in explaining why crises over Korea and Taiwan, Cuba, and the Middle East turned out the way they did.

Among the most influential scholars of the cold war period, Alexander George sought to use theory to inform and improve policy on a range of topics such as coercive diplomacy (1971), deterrence (with Smoke 1974), and the organization of presidential advisory systems (1980). His work on the most important topics of the cold war was infused with an effort to understand how leaders' belief systems shaped their approach to politics, how value trade-offs got resolved psychologically, and the reasons why distortions were introduced into the policy process. His pioneering work on qualitative methods that continued into the post-cold war era (George and Bennett 2005) was never an end in itself but rather a means by which one could use the careful study of historical examples to provide guidance in current and future policy problems of a similar nature.

As the cold war ended, and Yugoslavia erupted in violence and the Soviet Union broke apart along ethnic lines, there was an increasing emphasis on identity issues, and thus a resurgence of work utilizing social identity theory and work on in-groups and out-groups (Mercer 1995; Goldgeier 1997). How do individuals learn to distinguish themselves as part of a group that is dissimilar to another, even in cases where to outsiders there is hardly any basis for understanding why deep divisions should occur?

During the cold war, we discussed how best to convey credible commitments in order to ensure effective deterrence, and the psychological problems included ensuring that signals were sent and received clearly and that decision-makers had sufficient reputations to be believed by their audience. After 11 September, the question is whether suicide bombers can be deterred, with some arguing they are the ultimate irrational actors, whereas other scholars are suggesting that deterrence is possible even with

respect to terrorists, and thus many of the old theories of deterrence developed during the cold war remain relevant (Lebovic 2006).

Finally, let us consider the decision to go to war in Iraq in the context of psychological theories of error and bias. Over the past three decades, we have learned a lot in experimental social psychology and behavioral economics about the social and market forces that can either amplify or attenuate deviations from rationality. Many experimental researchers assume that the root cause of error and bias is widespread over-reliance on simple, easy-to-execute heuristics that give people unjustifiable confidence in their judgments and decisions. If correct, then social, political, and economic systems that encourage actors to engage in more self-critical and reflective forms of information processing should have the net effect of attenuating bias. Conversely, social, political, and economic systems that encourage mindless conformity, defensiveness, and the perpetuation of shared misconceptions should have the net effect of amplifying bias.

(p. 476)

Competitive market pressures can also reduce error and bias. Certain classes of error and biases—especially breakdowns in consistency and transitivity produced by reliance on simple (lexicographic) decision rules—can be corrected when the choice process moves into open and transparent market settings that provide for repeated play, interaction with attentive competitors, and rapid, unequivocal feedback on the consequences of one's choices (Kagel and Roth 1995; Camerer and Hogarth 1999).

It is now possible to identify an ideal set of conditions under which the error-and-bias portrait of the decision-maker should be of minimal predictive usefulness. The inner advisory group deliberates under the norms of multiple advocacy, is accountable to the institutions of a self-correcting democratic polity, and is making decisions in a policy domain in which critical information is readily available. Mistakes are quickly and publicly punished.

Although the full story of the 2003 Bush administration's decision to go to war in Iraq remains to be told, we know enough to suspect that the conditions for eliminating error and bias were not present. Complaints have abounded, particularly from senior officials at the State Department (Wilkerson 2006), that a “cabal” led by Vice-President Dick Cheney and Secretary of Defense Donald Rumsfeld overrode any objections to the rosy scenarios they painted regarding the war and its subsequent occupation. The US Congress barely debated the authorization to use force, and failed to ask the tough questions that mark ideal domestic accountability networks. And, in terms of available information, we now know that the core of Secretary of State Colin Powell's presentation to the United Nations regarding the intelligence information on weapons of mass destruction programs was hardly robust in its definitiveness. It was nowhere near the “slam dunk” described to the president by CIA director George Tenet (Woodward 2004), and as the story is now being told, was based on evidence provided by a mentally

unstable Iraqi engineer, codenamed Curveball. In the war's aftermath, mistakes were neither quickly nor publicly punished, even as casualties tragically mounted, as the president effectively painted his critics as soft on terror.

## 6 Closing Remarks: Beyond Reductionism

Psychological explanations work best when seamlessly integrated into more macro-organizational, societal, economic, and systemic approaches to international relations. Properly integrated, psychological explanations fill logical holes in existing theoretical coverage and offer salutary reminders of the dangers of (p. 477) passionate hedgehog-style theoretical commitments and the benefits of detached foxlike eclecticism. Parsimony is not a trump value: Theories should be as simple as possible, but no simpler.

## References

AXELROD, R. M. 1984. *The Evolution of Cooperation*. New York: Basic Books.

BARGH, J. A., and WILLIAMS, E. L. 2006. The automacity of social life. *Current Directions in Psychological Science*, 15: 94–8.

BEM, D. 1967. Self-perception: an alternative interpretation of cognitive dissonance phenomenon. *Psychological Review*, 74: 183–200.

BERLIN, I. 1953. *The Hedgehog and the Fox: An Essay on Tolstoy's View of History*. New York: Simon and Schuster.

CAMERER, C. F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

— and HOGARTH, R. M. 1999. The effects of financial incentives in experiments: a review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19: 1–3, 7–42.

CHOLLET, D., and GOLDGEIER, J. 2002. The scholarship of decision-making: do we know how we decide? Pp. 153–80 in *Foreign Policy Decision-Making (Revisited)*, Richard C. Snyder et al. New York: Palgrave.

CHRISTENSEN, T. J., and SNYDER, J. 1990. Chain gangs and passed bucks: predicting alliance patterns in multipolarity. *International Organization*, 44: 137–68.

COLEMAN, J. S. 1990. *Foundations of Social Theory*. Cambridge, Mass.: Belknap Press.

COWEN, T. 2006. Enter the neuro-economists: why do investors do what they do. *New York Times*, 20 Apr.: C3.



## Psychological Approaches

---

**CRAWFORD**, N. C. 2000. The passion of world politics: propositions on emotion and emotional relationships. *International Security*, 24: 116–56.

**FARNHAM**, B. R. 1994. *Avoiding Losses/Taking Risks: Prospect Theory and International Conflict*. Ann Arbor: University of Michigan Press.

**FEHR**, E., **FISCHBACHER**, U., and **KOSFELD**, M. 2005. Neuroeconomic foundations of trust and social preferences. *American Economic Review*, 95: 346–51.

— and **SCHMIDT**, K. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114: 817–68.

**FINNEMORE**, M. 2003. *The Purpose of Intervention: Changing Beliefs about the Use of Force*. Ithaca, NY: Cornell University Press.

**GEORGE**, A. L. 1971. *Limits of Coercive Diplomacy*. New York: Little, Brown.

— 1980. *Presidential Decisionmaking in Foreign Policy: The Effective Use of Information and Advice*. Boulder, Colo.: Westview.

— 1983. *Managing US–Soviet Rivalry: Problems of Crisis Prevention*. Boulder, Colo.: Westview.

— and **BENNETT**, A. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, Mass.: MIT Press.

— and **SMOKE**, R. 1974. *Deterrence in American Foreign Policy*. New York: Columbia University Press.

(p. 478) **GIGERENZER**, G., and **GOLDSTEIN**, D. G. 1996. Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, 103: 650–69.

— and **SELTEN**, R. (eds.) 2001. *Bounded Rationality: The Adaptive Toolbox*. Cambridge, Mass.: MIT Press.

**GOLDGEIER**, J. M. 1994. *Leadership Style and Soviet Foreign Policy: Stalin, Khrushchev, Brezhnev, Gorbachev*. Baltimore: Johns Hopkins University Press.

— 1997. Psychology and security. *Security Studies*, 6: 137–66.

— and **TETLOCK**, P. E. 2001. Psychology and international relations theory. *Annual Review of Political Science*, 4: 67–92.

**HANSEN**, J. 2006. The threat to the planet. *New York Review of Books*, 53: 12–16.

**HAAS**, P. M. (ed.) 1997. *Knowledge, Power and International Policy Coordination*. Columbia: University of South Carolina Press.

## Psychological Approaches

---

JERVIS, R. 1976. *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.

KAGEL, J. H., and ROTH, A. E. 1995. *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.

KAHNEMAN, D., and TVERSKY, A. 1979. Prospect theory: an analysis of decision under risk. *Econometrica*, 47: 263–91.

—— — (eds.) 2000. *Choices, Values, and Frames*. New York: Cambridge University Press.

KATZENSTEIN, P. J., KEOHANE, R. O., and KRASNER, S. D. 1998. International organization and the study of world politics. *International Organization*, 52: 645–85.

KELMAN, H. C. 1958. Compliance, identification, and internalization: three processes of attitude change. *Journal of Conflict Resolution*, 2: 51–60.

—— and BLOOM, A. 1973. Assumptive frameworks in international politics. Pp. 261–95 in *Handbook of Political Psychology*, ed. J. Knutson. San Francisco: Jossey-Bass.

KEOHANE, R. O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton, NJ: Princeton University Press.

KHONG, Y. F. 1992. *Analogies at War: Korea, Munich, Dien Bien Phu and the Vietnam Decisions of 1965*. Princeton, NJ: Princeton University Press.

KRUGLANSKI, A. 1996. Motivated freezing of the mind: seizing and freezing. *Psychological Review*, 103: 263–83.

KUNDA, Z. 1999. *Social Cognition: Making Sense of People*. Cambridge, Mass.: MIT Press.

LARSON, D. W. 1985. *Origins of Containment*. Princeton, NJ: Princeton University Press.

LEBOVIC, J. 2006. *Deterring International Terrorism: US National Security Policy after 9/11*. New York: Routledge.

LEBOW, R. N. 1981. *Between Peace and War: The Nature of International Crisis*. Baltimore: Johns Hopkins University Press.

LEGRO, J. W. 1995. *Cooperation under Fire: Anglo-German Restraint during World War II*. Ithaca, NY: Cornell University Press.

LEVY, J. 1992. Prospect theory and international relations: theoretical applications and analytical problems. *Political Psychology*, 13: 283–310.

MARCH, J. G., and OLSEN, J. P. 1998. The institutional dynamics of international political orders. *International Organization*, 52: 943–70.

## Psychological Approaches

---

MC DERMOTT, R. 1998. *Risk-taking in International Politics: Prospect Theory in American Foreign Policy*. Ann Arbor: University of Michigan Press.

—— 2004. Prospect theory in political science: gains and losses from the first decade. *Political Psychology*, 25: 289–312.

(p. 479) MC GRAW, P., TETLOCK, P. E., and KRISTEL, O. 2003. The limits of fungibility: relational schemata and the value of things. *Journal of Consumer Research*, 30: 219–29.

MEARSHEIMER, J. J. 2001. *The Tragedy of Great Power Politics*. New York: Norton.

MERCER, J. 1995. Anarchy and identity. *International Organization*, 49: 229–52.

—— 2005. Prospect theory and political science. *Annual Review of Political Science*, 8: 1–21.

MISCHEL, W. 1999. Implications of person-situation interaction: getting over the field's borderline personality. *European Journal of Personality (special issue)*, 13: 455–61.

MUELLER, J. 1989. *Retreat from Doomsday: The Obsolescence of Major War*. New York: Basic Books.

PIERSON, P. 2004. *Politics in Time: History, Institution, and Social Analysis*. Princeton, NJ: Princeton University Press.

PRICE, R. 1997. *The Chemical Weapons Taboo*. Ithaca, NY: Cornell University Press.

RABIN, M. 1998. Psychology and economics. *Journal of Economic Literature*, 36: 11–46.

RISSE, T., ROPP, S. C., and SIKKINK, K. (eds.) 1999. *The Power of Human Rights: International Norms and Domestic Change*. New York: Cambridge University Press.

ROUHANA, N. N., and KELMAN, H. C. 1994. Non-official interaction processes in the resolution of international conflicts: promoting joint Israeli–Palestinian thinking through a continuing workshop. *Journal of Social Issues*, 50: 157–78.

RUGGIE, J. G. 1998. *Constructing the World Polity: Essays on International Institutionalization*. London: Routledge.

SCHELLING, T. C. 1967. *Arms and Influence*. New Haven, Conn.: Yale University Press.

SIMMONS, B. A. 1998. Compliance with international agreements. *Annual Review of Political Science*, 1: 75–93.

SNYDER, G. H., and DIESING, P. D. 1978. *Conflict among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton, NJ: Princeton University Press.

STAW, B. M. 1980. Rationality and justification in organizational life. Pp. 45–80 in *Research in Organizational Behavior*, vol. ii, ed. B. M. Staw and L. Cummings. Greenwich, Conn.: JAI Press.

SUEDFELD, P., and TETLOCK, P. E. 2001. Individual differences in information processing. Pp. 284–304 in *Blackwell International Handbook of Social Psychology: Intraindividual Processes*, vol. i, ed. A. Tesser and N. Schwartz. London: Blackwell.

TALIAFERRO, J. 2004. *Balancing Risks: Great Power Intervention in the Periphery*. Ithaca, NY: Cornell University Press.

TANNENWALD, N. 1999. The nuclear taboo: the United States and the normative basis of nuclear non-use. *International Organization*, 53: 433–68.

TETLOCK, P. E. 1999a. Theory-driven reasoning about possible pasts and probable futures: are we prisoners of our preconceptions? *American Journal of Political Science*, 43: 335–66.

—1999b. Prospecting for evidence for prospect theory. Review of R. McDermott, *Risk-taking in International Politics: Prospect Theory in American Foreign Policy*. *Contemporary Psychology*, 44: 399–401.

—2003. Thinking about the unthinkable: coping with secular encroachments on sacred values. *Trends in Cognitive Science*, 7: 320–4.

—2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton University Press.

—and BELKIN, A. (eds.) 1996. *Counterfactual Thought Experiments in World Politics*. Princeton, NJ: Princeton University Press.

(p. 480) TETLOCK, P. E., and GOLDGEIER, J. 2000. Human nature and world politics: cognition, influence, and identity. *International Journal of Psychology*, 35: 87–96.

—VISSER, P., SINGH, R., POLIFRONI, M., SCOTT, A., ELSON, B., MAZZOCCO, P., and RESCOBER, P. 2007. People as intuitive prosecutors: the impact of social-control motives on attributions of responsibility. *Journal of Experimental Social Psychology*, 43: 195–209.

THALER, R. H. 2000. From homo economicus to homo sapiens. *Journal of Economic Perspectives*, 14: 133–41.

WALT, S. M. 1987. *The Origins of Alliances*. Ithaca, NY: Cornell University Press.

WALTZ, K. N. 1959. *Man, the State and War*. New York: Columbia University Press.

—1979. *Theory of International Politics*. Reading, Mass.: Addison-Wesley.

WENDT, A. 1999. *Social Theory of International Politics*. New York: Cambridge University Press.

WILKERSON, L. 2006. Weighing the uniqueness of the Bush administration's national security decision-making process: boon or danger to American democracy? The New America Foundation, at **[www.thewashingtonnote.com](http://www.thewashingtonnote.com)**.

WOODWARD, B. 2004. *Plan of Attack*. New York: Simon and Schuster.

### **James Goldgeier**

James Goldgeier is Professor of Political Science and International Affairs at George Washington University and a senior fellow at the Council on Foreign Relations.

### **Philip E. Tetlock**

Philip Tetlock holds the Mitchell Endowed Chair at the University of California, Berkeley.

