
The Micro-Foundations of International Relations Theory: Psychology and Behavioral Economics

Janice Gross Stein

Abstract Almost forty years ago, a small group of scholars drew on cognitive psychology to explain anomalous patterns of behavior by leaders on issues of international security. Although it made significant contributions to theory and research, that scholarship did not diffuse broadly into the field. Drawing on concepts in psychology and behavioral economics, research that uses new methods is now producing a wave of scholarship in international relations exemplified by the work in this special issue. Analysis of the use of prospect theory over the last three decades identifies the scope conditions that enable the predictions of rational choice and psychological theories. These scope conditions motivate the focus on the heterogeneity of decision makers that is at the core of current contributions. Future research will move beyond the now-sterile debate between rational choice and psychology.

Psychology and International Relations: Early Obstacles to Diffusion

Almost four decades ago, a small group of scholars in international relations, intrigued by the cognitive turn in psychology, explored how these new models could be applied to big questions in international security.¹ Cognitive psychologists had identified multiple patterns of thinking that systematically deviated even from relaxed concepts of rational choice. Political scientists drew on these results to examine how loss aversion informed anomalies in the behavior of the parties in a deterrence relationship, how cognitive biases explained counterintuitive outcomes in bargaining, and how poor information processing and probability estimation led to unplanned wars and crisis escalation. They also explored the impact of group dynamics on collective choice that led to crisis escalation and an inability to anticipate unlikely but damaging consequences.² Scholars in international relations continued to develop and deepen psychological explanations of international behavior but remained a minority within a corner of the larger field. Why did this early scholarship

1. Herrmann 1985; Jervis 1976, 1998; Jervis, Lebow, and Stein 1985; Lebow 1981.

2. Janis 1972; Jervis 1998.

not diffuse more widely in research and journals? Both analytic and methodological factors seem to have been at play.³

Rational choice, the predominant analytical model in the field that enabled formalization and the clarity of a deductive structure, provided a valuable normative baseline for evaluating choice. These are considerable advantages in a discipline that does not have a unified theory. Although cognitive psychology had established empirical support for a large number of systematic deviations from rational inference and choice, it was difficult to establish which bias would matter when and to specify its impact on choice. The failure to identify scope conditions was common to both rational choice and psychological theories, but it did not trouble theorists who developed deductive rather than empirical theories of choice.

A new wave of experimental research is establishing scope conditions for different biases and, in some cases, identifying thresholds when biases can be reversed. Research is also identifying heterogeneity across situations and individuals. The articles in this special issue address these questions to move beyond the sterile debate between rational choice and psychology. Instead they ask: under what conditions are decision makers likely to be rational and when are they likely to behave in ways that behavioral theories expect?

Several of the papers in this issue identify the conditions that enable the predictions of behavioral theories of choice. Renshon, Lee, and Tingley use citizens as a convenience sample for leaders and show that emotional arousal reduces the rate at which individuals select strategically optimal policies. Other contributors analyze the impact of identity on information processing and choice. Hermann finds that strong attachment to national identities leads people to interpret information through an emotional lens and individuals with distinct identities are therefore unlikely to converge as they are exposed to congruent information. Bayram demonstrates that actors with cosmopolitan social identities are more likely to accept the legitimacy of international law and their belief in legal obligation then attenuates their sensitivity to compliance costs. Strategic rationality in compliance, she argues, applies only to a particular set of actors. Closely related, Rathbun, Kertzer, and Paradis find that individuals with a pro-social orientation are less likely to behave in rational self-interested ways than are those with a pro-self orientation.⁴

A number of papers explore the conditions that predispose individuals to behave as rationalists expect. Experience is one of these conditions. Saunders shows that inexperience biases information processing and inference.⁵ Even when sufficient information exists in their strategic environment, less-experienced decision makers have more difficulty than those with greater experience in acquiring and in updating information. Rathbun, Kertzer, and Paradis also demonstrate in their experiments that experience attenuates the relationship between low epistemic motivation and

3. Elms 2008.

4. Renshon, Lee, and Tingley 2017; Herrmann 2017; Bayram 2017; Rathbun, Kertzer, and Paradis 2017.

5. Saunders 2017.

biased choice. Kertzer shows that variation in individuals' time and risk preferences affects how individuals update their beliefs about the desirability of conflict when they are confronted with new information about its associated costs.⁶

I draw on one of these enabling conditions—the variation in risk preferences—and show how experimental research using a behavioral theory of risk has diffused outside the laboratory and deepened over the last three decades by identifying scope conditions. I use “prospect theory,” a theory developed through experimental research by psychologists and economists, as a case study for several reasons. First, prospect theory was central to the first wave of behavioral research and therefore provides a useful case study of how behavioral theories diffused in the analysis of international politics. Second, it deals explicitly with risk preferences and choice, a central concern in international relations theory, and speaks directly to rational choice. Third, it makes predictions based on situational and individual variables, engaging both structure and agent.

Finally, a close examination of the applications of prospect theory to puzzles in international security demonstrates the hard challenge of applying theory developed through experimentation outside the laboratory. Much of the current experimental work that is fueling the development of behavioral theories will face this challenge of external validity. Drawing on the example of prospect theory, I argue that moving experimental theories outside the laboratory is slow and demanding work. The story is nevertheless encouraging; prospect theory has diffused over the last three decades and now informs the way many important puzzles in international security are analyzed. Ongoing experimental research is also developing an increasingly robust set of scope conditions of the relevance of prospect theory outside the laboratory, a critical requirement for the more general advance of behavioral theories in the study of international relations.

Prospect Theory and International Security

“Prospect theory” is considered to be the best description of how people evaluate risk in experimental settings.⁷ Its four central elements are well known: reference dependence, loss aversion, diminishing sensitivity, and probability weighting. Kahneman and Tversky demonstrated that people's choices among prospects are shaped by “framing effects,” the method for or sequence of how options are presented.⁸ People frame their choices around a reference point and consider relative gains and losses from that reference point rather than estimate the net expected value of their assets. They are also much more sensitive to losses than they are to equivalent gains; the value function is concave in the region of gains but convex in the region

6. Kertzer 2017.

7. Barberis 2012.

8. Kahneman and Tversky 1979, 1984; Tversky and Kahneman 1981, 1992.

of losses. The impact of loss aversion is amplified by systematic distortions in the weighting of probabilities. Assessment of risks becomes skewed when outcome probabilities are very high or very low. Depending on their reference points, individuals tend either to overweight or underweight these kinds of probabilities.⁹

Moving experimental results to explain puzzles outside the laboratory is difficult. As scholars of international relations drew on psychological concepts that were generated through laboratory experiments, they faced four issues. First, economists and psychologists provide the frame in prospect theory in their experiments and then study the impact; outside the laboratory, the frame is not given and prospect theory provides only a partial account of framing through theories of “mental accounting.”¹⁰ The absence of a strong theory of framing has slowed the diffusion of prospect theory in the field.¹¹ Second, reference points must be defined *ex ante*, independently of outcomes, to properly test theories.¹² Third, risk has emotional as well as probability dimensions. Emotion influences the choice of reference points and processes of probability estimation.¹³ Risks that are particularly vivid or salient are systematically overestimated. Emotion also influences loss aversion directly; people *feel* the pain of loss more intensely than they *feel* the pleasure of equivalent gain and are more likely to take risks to recover a loss and less likely to see concessions as compensating for the injury when emotion inflates the loss and feelings of injury.¹⁴ Finally, prospect theory, developed in the lab, paid little attention to the scope conditions that would shape its impact outside.

Scholars in international relations first used prospect theory to explain decision-making behavior during crises.¹⁵ Using archival material, memoirs, and interviews, early research demonstrated the counterintuitive impact of loss aversion on deterrence

9. Tversky and Kahneman 1992; See also Gonzalez and Wu 1999; Abdellaoui 2000; Bruhin, Fehr-Duda, and Epper 2010.

10. Mental accounting “is a set of rules people use to choose reference points and categories for comparing various gains and losses.” Camerer and Kunreuther 1989, 573; Thaler 1985. These rules of mental accounting often violate the rules of economic decision making. Other things being equal, people prefer the status quo because of the endowment effect. Knetsch and Sinden 1984; Thaler 1985. Mental accounting is only one among several possible theoretical explanations for the shifts in reference points.

11. In economics, the absence of a strong theory of framing made it difficult for economists to know *how* to test its propositions outside the lab, but economists are now deriving the predictions of prospect theory under a variety of plausible definitions of gains and losses and testing these predictions both in the lab and in the field across domains (Barberis 2012). Koszegi and Rabin consider that the reference point people use to compute gains and losses is their expectations or “beliefs ... held in the past about outcomes” (2006, 2007, 2009). They propose that people derive utility from the difference between consumption and expected consumption.

12. Taliafero argues that a careful test of prospect theory requires three observations: “(a) decision makers evaluated outcomes in terms of the reference point adopted at *t*; (b) decision makers perceive themselves as facing gains or losses relative to that reference point at *t + n*; and (c) the group’s risk-taking behavior is in the predicted direction” (1998, 109). See also Davis 2000; McDermott 2004a.

13. Carnevale 2008; Hall and Ross 2015; Kahneman 2011, 326–37; Loewenstein et al. 2001, 267; Mercer 2005a, 2010; Slovic et al. 2004.

14. McDermott 2009.

15. Levy 1992a, 1992b.

strategy and examined its implication for deterrence theory.¹⁶ For example, Farnham explained the changes in President Roosevelt's reaction to the Munich crisis and McDermott compared the decision making of Prime Minister Eden in Suez in 1956 and President Carter during the Iran hostage crisis.¹⁷

That early research met strong resistance from rational choice theorists who discounted the impact of single case studies. But new work developed more robust tests of loss aversion's impact. In the analysis of deterrence outcomes, scholars in international relations used careful process tracing, first to identify the reference point of decision makers *ex ante*, then to establish the gains and losses that decision makers perceived themselves facing relative to that reference point, and then to assess whether decision makers' behavior was as predicted.¹⁸ A second promising approach has used the hedonic tone of the problem to suggest an obvious or natural frame.¹⁹ Analysts then assess the material consequences of policy choices as gains when challengers are likely to judge that they are doing well and as losses when they are likely to do poorly. Both formulations find support for the proposition that loss aversion promotes challenges to deterrence that are not expected by rational choice theorists.²⁰ In the analysis of deterrence, scholars now routinely use concepts of loss aversion to develop theory and loss aversion has now been built into policy.²¹

A second important development was the effort to specify the scope conditions of the theory. Experimental research has demonstrated that the strength of framing effects varies across conditions.²² Domain has an impact on the strength of loss aversion: the tendency toward risk seeking is greater in human life problems than in money problems.²³ Prospect theory is more likely to be relevant, therefore, to decisions about war than it is to decisions about finance.²⁴ When their survival or the survival of those they represent is not assured, decision makers are especially likely to be risk acceptant; their predisposition to risk is in part a function of the domain.²⁵

16. Stein 1985.

17. Farnham 1992, 1997; McDermott 1998, 2004a.

18. Stein 1985.

19. Kuhberger 1998; Mandel 2001.

20. Berejikian 2004; Mercer 2005b.

21. The United States initiated an important shift to "tailored" deterrence strategies that built in the psychology of the would-be challenger to "get into the heads" of adversaries. US Department of Defense 2006, 21. Official documents make clear that US strategy has expanded rational choice models of deterrence to incorporate prospect theory and loss aversion into the design of strategy.

22. Framing effects vary by domain—for example, medical diagnoses, betting, escalation of commitment; by the information contained in the prospects presented to decision makers; by the descriptors associated with particular outcomes; and by the outcome probabilities. Boettcher 2004; Levin, Schneider, and Gaeth 1998; Mandel 2001.

23. Mandel 2001, 71. Kuhberger 1998 identified a wide range of framing effects across nine different domains.

24. Carnevale 2008; Kanner 2004; McDermott 2009.

25. In modeling outcomes of crisis bargaining, when decision makers are risk averse or risk neutral, the predictions of rational choice and prospect theory are likely to converge. They are likely to diverge, however, when leaders on both sides are risk acceptant because they feel that their survival is at stake. Both domain and individual heterogeneity matter. McDermott, Fowler, and Smirnov 2008, 345; Schaub 2004.

Missing information, a common attribute of problems in international security, also enhances framing effects.²⁶ Outcome probabilities also affect the strength of framing effects: certain/risky bimodal choices produce greater effects than choices between two risky outcomes.²⁷ The condition of a certain/risky choice, combined with the endowment effect, helps to explain the conditions that favor deterrence over compellence success. Behavioral economists have identified the endowment effect that leads people to overvalue what they currently possess and therefore exaggerate the cost and losses of concessions.²⁸ The endowment effect amplifies loss aversion in the context of compellence. Other things being equal, the endowment effect makes it more difficult for leaders to give up something that they already possess than to prevent them from taking something they do not currently have. The threat of loss of status has been used as well to explain leaders' tendency to commit additional resources to a conflict once costs are sunk.²⁹ Experiments demonstrate that when leaders are in the domain of loss, they tend to make the risk-acceptant choice of committing additional resources.³⁰

The use of prospect theory to explain puzzles in international security demonstrates that the diffusion of psychological concepts developed in the laboratory has been slow and bumpy. The challenges in moving experimental results from highly structured problems in the lab to ill-structured problems in international security have been real. There is still no consensus on how to theorize reference points. Complex multi-dimensional choices, the need to construct options, and imperfect information make the transfer out of the lab difficult. Yet progress in testing alternative theories of framing has been significant, in micro-economics as well as in international

26. Kuhberger 1995; Mandel 2001, 60.

27. Boettcher 2004, 338; Kuhberger 1998, 36. Prospect theory deals only with the framing effects of utility-equivalent choices. People are more risk acceptant when they frame options so that a risky gamble entails the possibility, however unlikely, of avoiding a losing outcome. They are more risk averse if they frame options so that one promises potential gains and the other has a probability, however low, of a losing outcome.

28. Goldgeier and Tetlock 2001; Jervis 1992; Thaler 1985.

29. Renshon 2015. There is individual heterogeneity here as well; the risk of loss of status is more acute in low-power than in high-power decision makers.

30. In the last decade, scholars have also worked widely with framing and prospect theory to examine a broad range of problems in international political economy. A shift in framing explained a change in debates over elements of Trade-Related Aspects of Intellectual Property (TRIPS) in the WTO. Sell and Prakash 2004. Prospect theory was also used to test the proposition that policy-makers are less likely to back down in trade disputes where they seek to minimize losses than they are in disputes where they seek to maximize expected gains. Using data from 100 cases of trade disputes initiated by the United States in a large-*N* study, Berejikian and Early find strong support for an explanation of loss aversion (2013). Drawing on decision makers' tendency to underweight low-probability events, Poulsen and Aisbett find that leaders in developing countries tended to ignore the risks of bilateral investment treaties until they themselves became subject to an investment treaty claim (2013). Risk acceptance has also provided a compelling explanation of the pace and timing of domestic economic reforms in Latin American states. Weyland 1996, 2002. Leaders in Argentina, Brazil, and Peru who confronted economic crises chose risky and costly reform strategies while Chile's leaders, who faced better economic conditions, chose more cautious strategies.

security.³¹ Scholars have developed alternative formulations of these issues and, working with a range of methodologies from large-*N* studies to archival case studies to an increasing emphasis on experiments both in the laboratory and the field, are using prospect theory to explain a large number of puzzles. Especially important has been the deepening knowledge about the scope conditions, both situational and individual, for the selection of reference points and the thresholds of framing effects.³² Experiments have demonstrated that domain, missing information, outcome descriptors, outcome probabilities, and individual differences in emotion, identities, experience, and risk preferences can shift the selection of reference points and the impact of framing effects in systematic ways.³³ Identification of these scope conditions, encompassing both situational and individual differences, and of threshold effects helps to establish when individuals are likely to behave as psychological theories predict and when they are more likely to approximate rational choosers.

Aggregation and Disaggregation

Moving experimental results outside the laboratory is not the only significant challenge behavioral theories face. As Powell's contribution to this issue demonstrates, processes of aggregation and disaggregation are generally considered an obstacle to the diffusion of psychological explanations of international behavior.³⁴ In the analysis of international behavior, the units of analysis tend to be the state, international institutions, or the system as a whole, although increasingly individuals are the unit of analysis in the micro-analysis of the behavior of states with strong executive systems. Psychologists—and behavioral economists—are methodological individualists; in cognitive psychology and micro-economics, explanations generally remain at the level of the individual and the problem of inference goes away. But theoretical propositions drawn from individual-level analysis do not move easily to “higher-level” units such as states.

Psychological theories are not alone, however, in facing the challenge of aggregation. Rational choice theorists continue to assume, despite the now overwhelmingly

31. Barberis 2012.

32. Economists have advanced their knowledge of the threshold at which nonlinear probability weighting shifts. Barberis 2012. Individuals tend to be nonlinear in the way they weight probabilities; they shift from systematically overweighting to systematically underweighting the probability of outcomes somewhere between .30 and .40. Camerer and Ho 1994; Tversky and Kahneman 1992; Wu and Gonzalez 1996. In set of experiments, Tversky and Kahneman find a distinctive fourfold pattern of attitudes to risk: risk aversion for gains and risk acceptance for losses of high probability and risk seeking for gains and risk aversion for losses of low probability. Tversky and Kahneman 1992. See also Elms 2008, 247; Kahneman and Tversky 1984, 345; Schaub 2004, 399.

33. Individual differences matter in explaining cooperative behavior as well. Pro-social people are more cooperative in a loss frame whereas pro-self people are more cooperative in a gain rather than a loss frame. De Dreu and McCusker 1997. Tversky and Kahneman 1992 also highlight how individuals vary in their likelihood to take risk beyond situational framing effects.

34. Powell 2017.

contradictory empirical evidence, that the state is a unitary rational actor. Or, as the introduction suggests, they use a “boxes-in-boxes” approach where preferences—and politics—at one level connect to and constrain preferences at the next level.³⁵ The advantages of this approach are clear: rationalists have to deal with only one source of bias—politics—while psychological approaches have to deal with two—the impact of politics at one level on another and the impact of psychological processes that may vary across individuals and situations. This adds orders of magnitude of complexity to the specification of psychological models of international behavior.³⁶

However, the difficulty remains that rational choice models are also inherently individual when they specify how actors estimate and choose. States cannot think, process information, estimate probabilities, or calculate; only their leaders can. Yet rational choice theorists routinely assume a unitary decision maker and go up the ladder of analysis to attribute these properties to the state. In so doing, rational choice theorists assume that the impact of psychological processes is small and can be easily managed through standard “as if” assumptions. Yet the evidence increasingly suggests their impact is large, that it varies across individuals and conditions, and that under certain conditions, it is strong enough not only to affect but even to reverse preferences. There is then a considerable cost to assuming away the impact of psychological processes in theorizing choice.³⁷

Second, scholars are increasingly paying explicit attention to when and how psychological processes are amplified or diminished in collective settings.³⁸ It is important to ask how individual heterogeneity is aggregated into collective choice. In a group context, choice shifts are specific cases of the more general phenomenon of group polarization where group discussion leads to adoption of a more extreme position than the group average in the initially preferred direction.³⁹ Drawing on social identity theory, theories of social contagion, strong pro-social orientations, social comparison, and persuasive arguments, researchers have specified how individual biases are modified or exaggerated by participation in different kinds of collective settings.⁴⁰ Social comparison explanations of aggregation assume that individuals are motivated to present themselves as socially desirable and these processes of amplification become mutually reinforcing within the group; the average at the end of the group process has shifted toward a more extreme collective choice.⁴¹ Theories of persuasive arguments

35. Lake and Powell 1999, 14–16.

36. Powell 2017.

37. Until rational choice and psychologists specify what their theories predict and assemble the empirical evidence on the explanatory power of each, it is impossible to evaluate the merits of the two approaches. That systematic comparison has not yet been done.

38. Janis 1972; t’Hart Stern, and Sundelius 1997. A meta-analysis of the influence of framing effects on risky decisions finds that the unit of analysis—the individual or the group—had little impact. Kuhberger 1998, 23.

39. Boettcher 2004; Lamm 1988, 807.

40. Hafner-Burton, Hughes, and Victor 2013; Shiller 2003; Tajfel 1982; Tajfel and Turner 1986; t’Hart, Stern, and Sundelius 1997.

41. Boettcher 2004, 335; Isenberg 1986, 1142–45.

emphasize the validity and novelty of the content of argument as two factors that shift individual choices in a group; novel arguments that are valid are especially persuasive.⁴² Theories of social contagion expect that information, evaluation, and affect circulate among members of a group, building upon the fertile ground of the “group within the individual” to forge loyalty and consensus.⁴³ Saunders explicitly links individual-level variation in bias with principal-agent models to explain why some biases are expressed or mitigated in certain cases rather than others.⁴⁴ Rather than ignore these processes in the name of parsimony and assign individual-level attributes to states, as Powell suggests, social and political psychologists are modeling alternative formulations of the impact of aggregation processes on shifts in both the direction and magnitude of choice.

Conclusion: How Different Is this Time?

The diffusion and deepening of knowledge that I have shown in the use of prospect theory to explain important puzzles in international security suggests that, at the intersection of psychology, behavioral economics, and international politics, empirical theories can continue to develop a fruitful research agenda of puzzles using new methods and designs. Scholars of international relations now have the opportunity to design and execute experimental studies, both in the laboratory and in the field. Social psychologists have long worked with colleagues in international politics on experimental studies, but they were restricted to the lab, relied almost exclusively on students, and were very few in number. As a next best approximation, psychologists, political scientists, and historians worked together to analyze counterfactuals, but satisfactory controls were difficult to achieve.⁴⁵ It is far easier today to do web-based experiments with a broad range of groups: in this issue, Herrmann embeds a series of experiments in a national survey in the United States to demonstrate that beliefs are not independent of preferences; Rho and Tomz use web-based experiments to analyze the impact of information on voter preferences about trade policy; and Renshon, Lee, and Tingley use experimental vignettes to examine voters’ responses to shifts in bargaining power globally.⁴⁶ These experimental studies allow direct assessments of “interventions” as well as tighter control of confounding variables than other kinds of methodologies that scholars of international relations use.

The internal validity of the results from well-designed experiments tends to be high, but their external validity, as I have argued, is much more challenging. There is good reason to be careful about the inferences that are drawn from tightly structured experimental studies and applied in very different contexts, across situations and

42. Boettcher 2004, 335; Isenberg 1986, 1145.

43. Hall and Ross 2015, 856–60 trace bottom-up, top-down, and horizontal processes of aggregation that work in large part through processes of social contagion.

44. Saunders 2017.

45. Tetlock and Lebow 2001; Tetlock, Lebow, and Parker 2006.

46. Herrmann; Renshon, Lee, and Tingley; and Rho and Tomz all in this issue.

individuals. When scholars are assessing the impact of beliefs and information processes on voter preferences toward international issues, there is every reason to be confident in the external validity of the results of carefully designed experiments where participants represent the broader population. Better understanding of how publics think matters to the analysis of the constraints that decision makers face and to their assessment of audience costs. It is also central to the specification of a process of aggregation in a “box-to-box” approach.⁴⁷

The challenge comes in drawing inferences from these kinds of experiments to the behavior of leaders. Just as psychologists have had to be careful not to claim too much from experiments with undergraduate students, scholars of international politics must be careful about the claims they make about leaders’ behavior based on experiments with students or undifferentiated publics. Saunders and Tetlock and Gardner have shown that experience can profoundly influence judgment.⁴⁸ Field studies and interviews with experienced decision makers, as well as archival work, will continue to be important to validate generalizations from laboratory experiments. Scholars are now using large-*N* studies to test the impact of psychological concepts on leaders’ behavior. Multiple streams of evidence will help to refine theories about the impact of psychological concepts on international behavior. Finally, as I have argued, it is critically important to specify the scope conditions of theories that are designed and tested in the laboratory.

The opportunities for future research at the intersection of psychology and rational choice are considerable. In the last wave of research, rational choice was the default option against which deviance was measured. There was value to that kind of research to establish empirically the robustness of different patterns of thinking and choosing. That work is now largely accomplished and the next generation of research can examine how and when rational and psychological processes explain choice and behavior.⁴⁹ Behavioral game theory is now drawing on psychology and behavioral economics to model the processes that deepen or dampen escalatory behavior. The analysis of a broad range of collective action problems can benefit from this kind of integrated approach.

Psychological approaches will also be useful in uncovering the micro-foundations of instrumentally rational strategic behavior. In an example of exactly this kind of research, Rathbun, Kertzer, and Paradis bridge the rational and psychological approaches to explore individual differences.⁵⁰ Using both an experimental bargaining game and archival research, they find that the psychological micro-foundations of

47. Lake and Powell 1999, 14–16.

48. Saunders 2017; Tetlock and Gardner 2015.

49. Neuroscientists as well as social scientists argue that a strict divide between rationality and emotion is not supported by the evidence. Feelings are important carriers of information and act as short circuits to processes of judgment. More sophisticated models of inference and judgment now consider that rationality presupposes and works through emotion. The binary between even soft rationality and psychological processes is empirically not supported. Holmes 2013; Kahneman 2011; McDermott 2004b; Mercer 2005a, 2010.

50. Rathbun, Kertzer, and Paradis 2017.

strategic rationality are a pro-self, social-value orientation and a high level of epistemic motivation. In a related study, Kertzer and Rathbun find that social preferences regarding reciprocity and fairness explain why some actors are better able than others to avoid bargaining failure.⁵¹ Bayram finds that cosmopolitans comply with international law because of their sense of legal obligation, unlike those who are instrumentally rational and sensitive to the payoff structure of compliance.⁵² All these studies, drawing on cognitive and social psychology, explore individual differences and engage the puzzle of who is instrumentally rational and when.

A third useful direction for research across the two approaches will be to establish thresholds for when psychological processes change the instrumentally rational outcome. Psychologists and behavioral economists have found, through experimental methods, the threshold when people shift from systematically overweighting to systematically underweighting the probability of outcomes. Using puzzles in international politics, experimental work can be done on other important cognitive biases to establish the thresholds when decisions shift.⁵³ These kinds of experiments can begin to answer the question of how sensitive rational choice models are to empirically established processes of judgment and choice.

The research agenda for those using psychological approaches to explain anomalies in international behavior extends beyond the fruitful dialogue with rational choice. By moving beyond the assumption of universal self-interested maximizing behavior, behavioral theories suggest sources of cooperation other than those in the standard bargaining literature, including the impact of a broader set of social preferences, the emotional attachments to these preferences, and collective identities and moods on behavior. Scholars can examine when and how these preferences and identities mitigate the tragedy of the commons, a ubiquitous problem in a loosely governed international system where governance varies across sectors. Psychological approaches can also be helpful in explaining high levels of compliance even when reporting mechanisms are weak and enforcement mechanisms are negligible.

Generally, building more complex models of judgment and choice that are tested through combinations of experiments, large-*N* studies, and archival research can only improve the accuracy of the baselines we use to identify the important anomalies and puzzles in international politics.⁵⁴ What is anomalous or puzzling can only be determined against a baseline, and the more empirically grounded the theories that specify the baselines, the better the choice of puzzles and the more productive the research agenda.

51. Kertzer and Rathbun 2015.

52. Bayram 2017.

53. Colleagues are currently running experiments to examine when the priming of various anomalies changes expected outcomes. By conducting these experiments across subjects with different degrees of experience in decision making, we can also examine the extent to which elite decision makers vary from the general population. Empirical evidence of the strength of these biases has been generated largely through studies of the general population. Stein, Loewen, and Sheffer, *in progress*.

54. Jervis 1998, 989–90; Odell 2002; Rabin 2002.

References

- Abdellaoui, Mohammed. 2000. Parameter-free Elicitation of Utility and Probability Weighting Functions. *Management Science* 46 (11):1497–1513.
- Barberis, Nicholas C. 2012. Thirty Years of Prospect Theory in Economics: A Review and Assessment. *NBER Working Paper* 18621.
- Bayram, A. Burcu. 2017. Due Deference: Cosmopolitan Social Identity and the Psychology of Legal Obligation in International Politics. In *The Behavioral Revolution and International Relations*, supplement, *International Organization* 71 (S1):S137–S163.
- Berejikian, Jeffrey. 2004. *International Relations Under Risk: Framing State Choice*. Albany: State University of New York Press.
- Berejikian, Jeffrey, and Bryan R. Early. 2013. Loss Aversion and Foreign Policy Resolve. *Political Psychology* 34 (5):649–71.
- Boettcher, William. 2004. The Prospects for Prospect Theory: An Empirical Evaluation of International Relations' Applications of Framing and Loss Aversion. *Political Psychology* 25 (3):331–62.
- Bruhin, Adrian, Helga Fehr-Duda, and Thomas Epper. 2010. Risk and Rationality: Uncovering Heterogeneity in Probability Distortion. *Econometrica* 78 (4):1375–1412.
- Camerer, Colin, and T.G. Ho. 1994. Violations of the Betweenness Axiom and Nonlinearity in Probability. *Journal of Risk and Uncertainty* 8 (2):167–96.
- Camerer, Colin, and Howard Kunreuther. 1989. Decision Processes for Low Probability Events: Policy Implications. *Journal of Policy Analysis and Management* 8 (4):565–92.
- Carnevale, Peter J. 2008. Positive Affect and Decision Frame in Negotiation. *Group Decision and Negotiation* 17 (1):51–63.
- Davis, James W. 2000. *Threats and Promises: The Pursuit of International Influence*. Baltimore, MD: Johns Hopkins University Press.
- De Dreu, Carsten K.W., and Christopher McCusker. 1997. Gain-loss Frames on Cooperation in Two-Person Social Dilemmas: A Transformational Analysis. *Journal of Personality and Social Psychology* 72 (5):1093–106.
- Elms, Deborah Kay. 2008. New Directions for IPE: Drawing from Behavioral Economics. *International Studies Review* 10 (2):239–65.
- Farnham, Barbara. 1992. Roosevelt and the Munich Crisis: Insights from Prospect Theory. *Political Psychology* 13 (2):205–35.
- . 1997. *Roosevelt and the Munich Crisis*. Princeton, NJ: Princeton University Press.
- Goldgeier, James, and Philip E. Tetlock, 2001. Psychology and International Relations Theory. *Annual Review of Political Science* 4 (1):67–92.
- Gonzalez, Richard, and George Wu. 1999. On the Shape of the Probability Weighting Function. *Cognitive Psychology* 38 (1):129–66.
- Hafner-Burton, Emilie M., D. Alex Hughes, and David G. Victor. 2013. The Cognitive Revolution and the Political Psychology of Elite Decision Making. *Perspectives on Politics* 11 (2):368–86.
- Hall, Todd H., and Andrew A.G. Ross. 2015. Affective Politics after 9/11. *International Organization* 69 (4):847–79.
- Herrmann, Richard K. 1985. American Perceptions of Soviet Foreign Policy: Reconsidering Three Competing Perspectives. *Political Psychology* 6 (3):375–411.
- . 2017. How Attachments to the Nation Shape Belief About the World: A Theory of Motivated Reasoning. In *The Behavioral Revolution and International Relations*, supplement, *International Organization* 71 (S1):S61–S84.
- Holmes, Marcus. 2013. The Force of Face-to-Face Diplomacy: Mirror Neurons and the Problem of Intentions. *International Organization* 67 (4):829–61.
- Isenberg, Daniel J. 1986. Group Polarization: A Critical Review and Meta-analysis. *Journal of Personality and Social Psychology* 50 (6):1141–51.

- Janis, Irving L. 1972. *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes*. Boston, MA: Houghton Mifflin.
- Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.
- . 1992. Political Implications of Loss Aversion. *Political Psychology* 13 (2):187–204.
- . 1998. *System Effects: Complexity in Social and Political Life*. Princeton, NJ: Princeton University Press.
- Jervis, Robert, Richard Ned Lebow, and Janice Gross Stein. 1985. *Psychology and Deterrence*. Baltimore, MD: Johns Hopkins University Press.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus, and Giroux.
- Kahneman, Daniel, and Amos Tversky. 1979. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica* 47 (2):263–92.
- . 1984. Choices, Values, and Frames. *American Psychologist* 39 (4):341–50.
- Kanner, Michael. 2004. Framing and the Role of the Second Actor: An Application of Prospect Theory to Bargaining. *Political Psychology* 25 (2):213–39.
- Kertzer, Joshua D. 2017. Resolve, Time, and Risk. In *The Behavioral Revolution and International Relations*, supplement, *International Organization* 71 (S1):S109–S136.
- Kertzer, Joshua D., and Brian C. Rathbun. 2015. Fair Is Fair: Social Preferences and Reciprocity in International Politics. *World Politics* 67 (4):613–55.
- Knetsch, Jack L., and John A. Sinden. 1984. Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value. *Quarterly Journal of Economics* 99 (3):507–21.
- Koszegi, Botond, and Matthew Rabin. 2006. A Model of Reference-Dependent Preferences. *Quarterly Journal of Economics* 121 (4):1133–65.
- . 2007. Reference-Dependent Risk Attitudes. *American Economic Review* 97 (4):1047–73.
- . 2009. Reference-Dependent Consumption Plans. *American Economic Review* 99 (3):909–36.
- Kuhberger, Anton. 1995. The Framing of Decisions: A New Look at Old Problems. *Organizational Behavior and Human Decision Processes* 62 (2):230–245.
- . 1998. The Influence of Framing on Risky Decisions: A Meta-Analysis. *Organizational Behavior and Human Decision Processes* 75 (1):23–55.
- Lake, David A., and Robert Powell. 1999. International Relations: A Strategic Choice Approach. In *Strategic Choice and International Relations*, edited by David A. Lake and Robert Powell, 3–38. Princeton, NJ: Princeton University Press.
- Lamm, Helmut. 1988. A Review of Our Research on Group Polarization: Eleven Experiments on the Effects of Group Discussion on Risk Acceptance, Probability Estimation, and Negotiation Positions. *Psychological Reports* 62 (3):807–13.
- Lebow, Richard Ned. 1981. *Between Peace and War: The Nature of International Crisis*. Baltimore, MD: Johns Hopkins University Press.
- Levin, Irwin P., Sandra L. Schneider, and Gary J. Gaeth. 1998. All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects. *Organizational Behavior and Human Decision Processes* 76 (2):149–88.
- Levy, Jack. 1992a. An Introduction to Prospect Theory. *Political Psychology* 13 (2):171–86.
- . 1992b. Prospect Theory and International Relations: Theoretical Applications and Analytical Problems. *Political Psychology* 13 (2):283–310.
- Loewenstein, George F., Elke U. Weber, Christopher K. Hsee, and Ned Welch. 2001. Risk As Feelings. *Psychological Bulletin* 127 (2):267–86.
- Mandel, David R. 2001. Gain-Loss Framing and Choice: Separating Outcome Formulations From Descriptor Formulations. *Organizational Behavior and Human Decision Processes* 85 (1):56–76.
- McDermott, Rose. 1998. *Risk-Taking in International Politics: Prospect Theory in American Foreign Policy*. Ann Arbor: University of Michigan Press.
- . 2004a. Prospect Theory in Political Science: Gains and Losses from the First Decade. *Political Psychology* 25 (2):289–312.

- . 2004b. The Feeling of Rationality: The Meaning of Neuroscientific Advances for Political Science. *Perspectives on Politics* 2 (4):691–706.
- . 2009. Prospect Theory and Negotiation. In *Negotiated Risks*, edited by Rudolf Avenhaus and Gunnar Sjöstedt, 87–109. Berlin: Springer-Verlag.
- McDermott, Rose, James H. Fowler, and Oleg Smirnov. 2008. On the Evolutionary Origin of Prospect Theory Preferences. *The Journal of Politics* 70 (2):335–50.
- Mercer, Jonathan. 2005a. Rationality and Psychology in International Relations. *International Organization* 59 (1):77–106.
- . 2005b. Prospect Theory and Political Science. *Annual Review of Political Science* 8:1–21.
- . 2010. Emotional Beliefs. *International Organization* 64 (1):1–31.
- Odell, John. 2002. Bounded Rationality and the World Political Economy. In *Governing the World's Money*, edited by David Andrews, C. Randall Henning, and Louis W. Pauly, 163–93. Ithaca, NY: Cornell University Press.
- Poulsen, Lauge N. Skovgard, and Emma Aisbett. 2013. When the Claim Hits: Bilateral Investment Treaties and Bounded Rational Learning. *World Politics* 65 (2):273–313.
- Powell, Robert. 2017. Research Bets and Behavioral IR. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S265–S277.
- Rabin, Matthew. 2002. A Perspective on Psychology and Economics. *European Economic Review* 46 (4): 657–85.
- Rathbun, Brian C., Joshua D. Kertzer, and Mark Paradis. 2017. *Homo Diplomaticus*: Mixed-Method Evidence of Variation in Strategic Rationality. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S33–S59.
- Renshon, Jonathan. 2015. Losing Face and Sinking Costs: Experimental Evidence on the Judgment of Political and Military Leaders. *International Organization* 69 (3):1–37.
- Renshon, Jonathan, Julia J. Lee and Dustin Tingley. 2017. Emotions and the Micro-Foundations of Commitment Problems. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S189–S217.
- Rho, Sungmin, and Michael Tomz. 2017. Why Don't Trade Preferences Reflect Economic Self-Interest? In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1): S85–S108.
- Saunders, Elizabeth N. 2017. No Substitute for Experience: Presidents, Advisers, and Information in Group Decision Making. In The Behavioral Revolution and International Relations, supplement, *International Organization* 71 (S1):S219–S247.
- Schaub, Gary Jr. 2004. Deterrence, Compellence, and Prospect Theory. *Political Psychology* 25 (3):389–411.
- Sell, Susan, and Aseem Prakash. 2004. Using Ideas Strategically: The Contest Between Business and NGO Networks in Intellectual Property Rights. *International Studies Quarterly* 48 (1):143–75.
- Shiller, Robert. 2003. From Efficient Markets Theory to Behavioral Finance. *Journal of Economic Perspectives* 17 (1):28–38.
- Slovic, Paul, Melissa L. Finucane, Ellen Peters, and Donald G. MacGregor. 2004. Risk as Analysis and Risk as Feelings: Some Thoughts About Affect, Reason, Risk, and Rationality. *Risk Analysis* 24 (2): 311–22.
- Stein, Janice Gross. 1985. Calculation, Miscalculation, and Conventional Deterrence I: The View from Cairo. In *Psychology and Deterrence*, edited by Robert Jervis, Richard Ned Lebow, and Janice Gross Stein, 34–59. Baltimore, MD: Johns Hopkins University Press.
- Stein, Janice Gross, Peter Loewen, and Lior Sheffer. In progress. Experimental Evidence on the Sensitivity of Rational Choice Models to Behavioral Anomalies across Populations.
- Tajfel, Henri. 1982. *Social Identity and Intergroup Relations*. New York: Cambridge University Press.
- Tajfel, Henri, and John C. Turner. 1986. The Social Identity Theory of Intergroup Behavior. In *Psychology of Intergroup Relations*, edited by Stephen Worchel and William G. Austin, 2nd ed., 7–24. Chicago: Nelson-Hall.

- Taliafero, Jeffrey W. 1998. Quagmires in the Periphery: Foreign Wars and Escalating Commitment in International Conflict. *Security Studies* 7 (3):94–144.
- Tetlock, Philip E., and Dan Gardner. 2015. *Superforecasting: The Art and Science of Prediction*. London: Random House.
- Tetlock, Philip E., and Richard Ned Lebow. 2001. Poking Counterfactual Holes in Covering Laws: Cognitive Styles and Historical Reasoning. *American Political Science Review* 95 (4):829–43.
- Tetlock, Philip E., Richard Ned Lebow, and Neil Geoffrey Parker, eds. 2006. *Unmaking the West: "What-If" Scenarios that Rewrite World History*. Ann Arbor: University of Michigan Press.
- Thaler, Richard. 1985. Mental Accounting and Consumer Choice. *Marketing Science* 4 (3):199–214.
- t'Hart, Paul, Stern, Erik K., and Sundelius, Bengt. 1997. Foreign Policy Making at the Top: Political Group Dynamics. In *Beyond Group Think: Political Group Dynamics and Foreign Policy Making*, edited by Paul t'Hart, Erik K. Stern, and Bengt Sundelius, 3–34. Michigan: University of Michigan Press.
- Tversky, Amos, and Daniel Kahneman. 1981. The Framing of Decisions and the Psychology of Choice. *Science* 211 (4481):453–58.
- . 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty* 5 (4):297–323.
- US Department of Defense. 2006. Deterrence Operations, Joint Operating Concept. Version 2.0. Available at <www.dtic.mil/doctrine/concepts/joint_concepts/joc_deterrence.pdf>. Accessed 27 January 2017.
- Weyland, Kurt. 1996. Risk Taking in Latin American Economic Restructuring: Lessons from Prospect Theory. *International Studies Quarterly* 40 (2):185–208.
- . 2002. *Politics of Market Reform in Fragile Democracies: Argentina, Brazil, Peru and Venezuela*. Princeton, NJ: Princeton University Press.
- Wu, George, and Richard Gonzalez. 1996. Curvature of the Probability Weighting Function. *Management Science* 42 (12):1676–90.