Photo by Ingrid Truemper

Real-World Nuclear Decision Making:
Using Behavioral Economics Insights to Adjust Nonproliferation and Deterrence Policies to Predictable Deviations from Rationality

A Report on a Workshop Organized by the James Martin Center for Nonproliferation Studies

By Jeffrey W. Knopf, Anne I. Harrington, and Miles Pomper

January 2016

CNS

James Martin Center for Nonproliferation Studies

Real-World Nuclear Decision Making: Using Behavioral Economics Insights to Adjust Nonproliferation and Deterrence Policies to Predictable Deviations from Rationality

Jeffrey W. Knopf
Anne I. Harrington
Miles Pomper

January 2016

Acknowledgments

This publication was developed under work supported by the Naval Postgraduate School (NPS) Project on Advanced Systems and Concepts for Countering WMD (PASCC) awarded by the NAVSUP Fleet Logistics Center San Diego (NAVSUP FLC San Diego), with funding from the Defense Threat Reduction Agency (DTRA). It has not been formally reviewed by NPS.

The three authors of this report were the co-principal investigators on the grant proposal to PASCC. The co-authors submitted the grant proposal under the auspices of their affiliations with the James Martin Center for Nonproliferation Studies (CNS). The views expressed herein are those of the authors and do not necessarily reflect the official policy or position of the Naval Postgraduate School, the Defense Threat Reduction Agency, the Department of Defense, or the United States Government.

© The President and Trustees of Middlebury College, January 2016

About CNS

The James Martin Center for Nonproliferation Studies (CNS) strives to combat the spread of weapons of mass destruction by training the next generation of nonproliferation specialists and disseminating timely information and analysis. CNS is the largest nongovernmental organization in the United States devoted exclusively to research and training on nonproliferation issues. CNS is based at the Middlebury Institute of International Studies at Monterey (www.miis.edu) and also has offices in Washington, DC, and Vienna, Austria. Individual authors who may be CNS staff or affiliated experts write CNS reports in their personal capacity. Thus, these reports do not represent a CNS institutional position on policy issues. All statements of fact and expressions of opinion contained in this report are the sole responsibility of the authors.

For further information, please contact:

James Martin Center for Nonproliferation Studies
Middlebury Institute of International Studies at Monterey
460 Pierce Street
Monterey, CA 93940, USA
Tel: +1 (831) 647-4154
Fax: +1 (831) 647-3519
nonproliferation.org

About PASCC

The Naval Postgraduate School Center on Contemporary Conflict (CCC) is the research wing of the Department of National Security Affairs (NSA) and specializes in the study of international relations, security policy, and regional studies. One of the CCC's programs is the Project on Advanced Systems and Concepts for Countering WMD (PASCC). PASCC operates as a program planning and implementation office, research center, and intellectual clearinghouse for the execution of analysis and future-oriented studies and dialogues for the Defense Threat Reduction Agency.

For further information, please contact:

The Center on Contemporary Conflict
Naval Postgraduate School
1411 Cunningham Road
Monterey, CA 93943, USA
pascc@nps.edu

TABLE OF CONTENTS

Executive Summary i
Introduction 1
Behavioral Economics: A Brief Review 3
Strand I: Psychology and Decision Making 6
Strand II: Neuroscience and Emotions 8
Workshop Summary 10
Basic Research 11
Empirical Evidence 13
Country-Specific Case Studies 16
Applications to Nonproliferation 17
Philosophical Reflections 20
Workshop Discussion Areas 23
Conclusions and Recommendations 24
Appendix: Workshop Agenda 27
Appendix: Workshop Participants 30

EXECUTIVE SUMMARY

Research in fields such as psychology and neuroscience has shown that human thinking and decision making often fail to match the assumption of rationality. Indeed, people can depart in quite systematic ways from the predictions of a rational actor model. These findings form the basis of behavioral economics, an approach to understanding human behavior that has attracted enormous attention in recent years. In this project, we explored the implications of behavioral economics, and the research that informs it, for policies and strategies designed to deal with the challenges posed by nuclear weapons. Papers commissioned for this project examined topics ranging from deterrence to economic sanctions, the nuclear nonproliferation regime, and U.S. domestic debates about ballistic missile defense.

We find that behavioral research has generated a number of specific insights that it would be useful to be aware of and keep in mind in strategic planning and decision making. At the same time, it proved difficult to integrate these into a single, comprehensive framework. We therefore conclude that behavioral economics does not yet provide a coherent and predictive model that can by itself serve as a basis for designing strategies. While it suggests important insights, behavioral economics does not on its own give us an analytical framework sufficient for identifying policies that will minimize nuclear dangers. For this reason, policymakers and defense planners are likely to still find value in taking into consideration the rational self-interest and strategic cultures of other actors. At the same time, behavioral research reveals important limitations to relying too heavily on assumptions of either rational or cultural determinants of behavior. Predictions that others will behave rationally can often be frustrated. And while human behavior is influenced by culture, it is a product of many other factors as well. Behavioral approaches reinforce long-standing intuitions about the importance of understanding other actors, but in ways that go beyond an appreciation for strategic culture.

Behavioral research suggests ways to add new depth to the ancient insight of Sun Tzu about the value of knowing both ourselves and others. In particular, it is vital to learn as much as possible about how others understand their situations. The framing for choice can have a huge impact on the decisions people make. Cultural legacies are often one ingredient in such framings, but the key question is how individuals frame the situation or context in which they have to make a decision, and this can be affected by their present circumstances and recent diplomatic interactions as well as by historical legacies. Just as it is important to learn as much as possible about how others frame their circumstances and choice options, it is also important to examine ourselves and try to figure out our own biases. The better we know ourselves and know others, the better we can begin to anticipate how some common psychological tendencies might influence the course of our interactions.

Even if behavioral economics does not provide a reliable basis for prediction, the research that informs it suggests several insights that can be valuable to keep in mind. First, emotions can exert a powerful impact on human decisions and behavior. We should not assume that decisions about nuclear weapons will be shaped solely by cool calculations of cost and benefit. Second, people are often motivated more strongly by the desire to avoid or minimize loss than by the pursuit of gain. In situations where we perceive an aggressive actor out to expand its territory or influence, it is worth considering whether in its own thinking that actor views itself as seeking to avoid a future loss or reverse a past loss. Third, people also tend to care about considerations of justice and fairness, but often in self-serving ways. This makes it important to explore the possibilities for developing common standards of fairness that might pave the way for a successful diplomatic negotiation or a stable deterrent relationship. Fourth, time horizons can change how people think. In particular, longer time horizons may allow space for more deliberative reasoning, which in turn can reduce the risks of rash or hasty decision making. Fifth, and finally, it is not easy to find optimal or permanent solutions to problems. We should not be unduly dismissive of the merits of "muddling through" when such efforts manage to keep problems from becoming much worse. When it comes to dealing with nuclear weapons, there is much to be said for avoiding worst-case outcomes even when the solutions are less than perfect.

INTRODUCTION

The invention of nuclear weapons created unprecedented challenges for the world. Even today, seventy years after the first atomic weapons test, the effort to find effective policies and strategies for dealing with nuclear weapons remains a daunting challenge. From early in the nuclear age, attention focused on deterrence as a strategy to prevent nuclear war. By the 1960s, key states were also seeking to limit the growth of nuclear arsenals and spread of nuclear arms through tools such as arms control and nonproliferation. More recently, global efforts have also encompassed nuclear security, or measures to keep bomb-making materials out of the hands of non-state actors such as terrorist groups. While international agreements are central to advancing nonproliferation and nuclear security goals, states sometimes find it advantageous to use other policy tools such as economic sanctions or diplomatic engagement to reinforce their efforts in these areas. And in the United States and a few other countries, efforts to develop ballistic missile defenses also remain an ongoing policy goal.

Given the stakes, it is important to understand when these policy tools do or do not work to reduce nuclear dangers and how to make them as effective as possible. Historically, research has focused primarily on deterrence, and other policy tools have not been the subject of as much systematic investigation. In addition, thinking about deterrence has often been based on a rational actor model. In the first decades of the nuclear age, models based on assuming a generic, rational actor proved remarkably productive in generating some crucial insights into the likely workings of nuclear deterrence.[1]

Eventually, however, dissatisfaction began to develop with rational theories of deterrence. One line of criticism argued that the Soviet Union did not think about nuclear weapons in the same way as the United States.[2] This critique targeted the notion of a generic, universal form of rationality.[3] It suggested that different actors can think quite differently, putting a premium on learning as much as possible about what the other side values most. This line of work led to the concept of strategic culture and an assumption that different countries develop different strategic cultures.[4] This thinking, in turn,

---

[1] The most influential work remains that of Thomas C. Schelling. See his *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1960) and *Arms and Influence* (New Haven: Yale University Press, 1966). For an attempt to formalize earlier work that takes advantage of advances in game theory, see Robert Powell, *Nuclear Deterrence Theory: The Search for Credibility* (Cambridge: Cambridge University Press, 1990).

[2] Richard Pipes, "Why the Soviet Union Thinks It Could Fight and Win a Nuclear War," *Commentary* (July 1977).

[3] Keith Payne, "The Fallacies of Cold War Deterrence and a New Direction," *Comparative Strategy* 22, no. 5 (2003).

[4] Jack L. Snyder, *The Soviet Strategic Culture: Implications for Limited Nuclear Operations*, RAND report R-2154-AF (Santa Monica: RAND, 1977).

provided much of the impetus for the emphasis on "tailored deterrence" in recent U.S. strategic doctrine.[5]

Strategic culture approaches are consistent with an assumption of rationality. They object to the assumption that all actors share the same values and hence would make the same rational calculations. But once the different value systems and hence utility functions of different actors are understood, it should be possible to predict how each will calculate the relative costs and benefits of acquiring or using nuclear weapons.

A different line of criticism targets the underlying assumption of rationality more directly. By the late 1960s, scholars began to question whether states (and the leaders making decisions for those states) could actually live up to the demanding requirements of the rational actor model.[6] Drawing on findings from psychology and organization theory, these critics argued that deterrence might fail more frequently than we would otherwise expect because of the various limitations on human rationality. Research identified a range of cognitive (or unmotivated) and motivated biases that could lead to misperceptions or miscalculations that might undermine deterrence.[7]

This "psychology and deterrence" research program peaked in the mid-1980s. Although some research along these lines continued after that date, increasingly attention turned elsewhere. The winding down of the Cold War and collapse of the Soviet Union made nuclear deterrence seem like a less urgent problem. Where academic research continued, much of it involved a renewed emphasis on rational actor models. Game theoretic work on situations involving incomplete or asymmetric information attracted particular interest and led to a new wave of deterrence research based on formal models.[8] The critics of rational actor approaches, in turn, found inspiration in the rise of social constructivism in the field of International Relations and began exploring how processes of social construction affect deterrence relationships.[9] Finally, for analysts concerned with contemporary policy problems, attention shifted to the challenges posed by rogue states and, after 9/11, to the very difficult problem of whether terrorism by non-state actors can be deterred. Collectively, these trends have been

---

[5] See M. Elaine Bunn, "Can Deterrence Be Tailored?" *Strategic Forum* (Jan. 2007) for a good short summary of this approach. On the relation between strategic culture and tailored deterrence, see Jeffrey S. Lantis, "Strategic Culture and Tailored Deterrence," *Contemporary Security Policy* (Dec. 2009).

[6] Stephen Maxwell, "Rationality in Deterrence," *Adelphi Papers* no. 50 (1968); Alexander L. George and Richard Smoke, *Deterrence in American Foreign Policy: Theory and Practice* (New York: Columbia University Press, 1974).

[7] Richard Ned Lebow, *Between Peace and War: The Nature of International Crisis* (Baltimore: Johns Hopkins University Press, 1981).

[8] For a good review, see Vesna Danilovic and Joe Clare, "Deterrence and Crisis Bargaining," in *International Studies Compendium*, ed. Robert E. Denemark (New York: Wiley-Blackwell, 2010).

[9] Amir Lupovici, "The Emerging Fourth Wave of Deterrence Theory—Toward a New Research Agenda," *International Studies Quarterly* (September 2010).

referred to as a "fourth wave" in deterrence research.[10] Notably, however, they all involved a shift of focus away from psychological and organizational constraints on rationality.

At the same time, policy tools other than deterrence never received the same level of systematic attention and remained theoretically underdeveloped. Some scholars did note the use of tools like positive incentives, economic sanctions, diplomatic engagement, or reassurance. Academic research focused largely on basic questions such as whether or not such policy tools work. For the most part, they did not have much to say about how limitations on human rationality might affect the operation of such policy tools.[11] Many of the relevant studies also examined the effectiveness of sanctions or engagement in general, without considering how these might operate specifically in the nuclear policy domain, for example as tools for preventing nuclear proliferation.[12]

Given this background, the research for this project had three goals. First, it sought to update the earlier "psychology and deterrence" literature. Since the mid-1980s, there have been significant developments in psychology, neuroscience, and related fields, many of which have informed the field of behavioral economics. This study aimed to take note of some of the key developments and explore how they might affect our understanding of deterrence and coercion. Second, this study sought to extend the reach of behavioral insights into the analysis of other policy tools, particularly as those tools are used to promote nuclear nonproliferation. Do the factors that affect deterrence outcomes also affect the operation of economic sanctions or the process of diplomatic negotiations? Third, at the same time that it sought to highlight the potential benefits of applying behavioral economics to strategic questions, this project also sought to assess the limits and potential pitfalls of this approach. Behavioral economics holds out the promise of being able to predict deviations from rational behavior. This makes it important to consider the possible limits on how much predictability it offers and the potential pitfalls of applying a science that grew out of the study of individual human behavior to a bureaucratic actor like the state.

BEHAVIORAL ECONOMICS: A BRIEF REVIEW

Research in neuroscience and related branches of psychology has revolutionized our understanding of human decision-making in situations involving risk or uncertainty. This research has

---

[10] Jeffrey W. Knopf, "The Fourth Wave in Deterrence Research," *Contemporary Security Policy* (Apr. 2010).

[11] An important exception is the work of Janice Gross Stein on reassurance. See her "Deterrence and Reassurance," in Philip Tetlock et al., eds., *Behavior, Society and Nuclear War*, vol. 2 (New York: Oxford University Press, 1991).

[12] There are exceptions. For a study on the impact of positive incentives on nonproliferation, Thomas Bernauer and Dieter Ruloff, eds., *The Politics of Positive Incentives in Arms Control* (Columbia, SC: University of South Carolina Press 1999). For studies of the impact of sanctions on nonproliferation, see Etel Solingen, ed., *Sanctions, Statecraft and Nuclear Proliferation* (Cambridge: Cambridge University Press, 2012); Nicholas L. Miller, "The Secret Success of Nonproliferation Sanctions," *International Organization* (fall 2014).

begun to influence other social science disciplines, most notably through the rise of the "behavioral economics" perspective in economics. Many of the key findings in this body of research have been nicely summarized for a general audience in Daniel Kahneman's landmark book, *Thinking, Fast and Slow*. Kahneman's book and other similar works have created a golden opportunity to leverage advances made in the field of behavioral economics—and in particular its critique of the rational actor assumption—to improve our understanding of nuclear deterrence and nonproliferation policies.[13]

This body of literature starts with the presumption that rational theories of human behavior are wrong to dismiss how individuals make decisions. Rather than assuming that actors are rational, behavioral economists use surveys and experiments to observe how individuals actually behave when confronted with a choice. In contrast, rational theories of human behavior, such as neoclassical economic theory or traditional deterrence theory, are based on the assumption that actors are "procedurally" rational, by which they mean that actors respond to incentive structures in predictable ways because their preferences are internally consistent (if they prefer A to B and B to C then they also prefer A to C). Not to be confused with reasoning or deliberation, procedural rationality is not meant to be an accurate description of what decision makers do. It describes an ideal type that allows for simple and elegant explanations of complex policy problems. Although no individual behaves rationally all the time, the argument goes, in the aggregate enough people behave rationally enough of the time to warrant the assumption of rationality. Therefore individual acts of irrationality can be dismissed at the level of general theory.

Behavioral economists argue that the ideal type of procedural rationality not only fails to accurately predict outcomes, it is also blind to the ways in which these failures themselves are foreseeable. In laboratory experiments, the judgments subjects make violate the assumption of procedural rationality in predictable, law-like ways. Humans are capable of rational thought, but it requires calculations and abstract thinking. These activities absorb time and energy that most people do not have to spare. Instead, humans rely on short-cuts. Rather than calculating probabilities, they reason in terms of averages, norms, and heuristics. These short-cuts consistently lead individuals to express preferences that are not logically consistent (i.e., they claim to prefer A to B and B to C, but then when asked to make a choice choose C over A).

Given that both deterrence theory and nonproliferation policy draw heavily from the field of economics for their conceptual foundations, it is surprising that so little work has been done thus far to draw out the implications of behavioral economics for the nuclear field. The assumption of rationality has long been identified as a weak link in the logic of nuclear deterrence theory. Unlike in

[13] Daniel Kahneman, *Thinking Fast and Slow* (New York: Farrar, Straus and Giroux, 2011); Richard Thaler and Cass Sunstein, *Nudge* (New Haven: Yale University Press, 2008); Daniel Gilbert, *Stumbling on Happiness* (New York: Vintage Books, 2007); Dan Ariely, *Predictably Irrational* (New York: Harper Collins, 2008); Nassim Taleb, *The Black Swan* (New York: Penguin, 2008).

economic theory, individual acts of irrationality cannot be dismissed as irrelevant to deterrence theory. The consequences of even a single deterrence failure are too costly.

Starting in the 1960s deterrence theorists were already looking to psychology for alternatives to rational choice models. A body of academic research utilizing a decision-making approach based heavily on research in psychology developed. This research emphasized biases in human decision making, described as cognitive (or unmotivated) and motivated. Cognitive biases reflect the impact of images and beliefs that people already hold. They lead people to filter out information inconsistent with those beliefs so that they see what they already expect to see, even when that image is inaccurate. Motivated biases reflect underlying needs and desires, including needs that political leaders might have that reflect their desire to stay in power domestically. Motivated biases produce wishful thinking, leading people to see what they want to see. Both kinds of bias can lead to deterrence failures as well as missed opportunities for negotiation. The high point of this research was the 1985 publication of *Psychology and Deterrence* by Robert Jervis, Ned Lebow, and Janice Stein.[14] Unfortunately, work on deterrence in this research tradition has flagged since then and has not fully kept up with developments in behavioral economics.

In the policy domain, the most influential critique of the rational actor model held that, due to differences in history and form of government, the Soviet Union did not think about nuclear war in the same way as the United States. Because of these differences in strategic culture, threats that would deter the United States might not deter Soviet leaders. Strategic culture models capture these cultural differences at the level of theory. Like other rational actor models, they also entail cost-benefit calculations, but assume these are not the same for all actors; instead, these calculations are based on culturally specific value systems that provide a yardstick for assessing costs and benefits specific to each actor. Policies influenced by these models seek to "tailor" deterrence to the strategic culture of the target actor.

Behavioral economics offers an alternative perspective to strategic culture models. Work in the behavioral tradition shows common patterns of deviation from the predictions of rational actor models, but many of these patterns are not culturally bounded and hence do not vary with strategic culture. For this reason, behavioral research offers a promising avenue for the analysis of decision making relevant to deterrent planning and to a range of other policy tools commonly used to combat WMD proliferation.

There are two major strands of behavioral economics. The first strand incorporates insights from social psychologists, economists, and decision theorists. The second strand draws heavily from

[14] Robert Jervis, Ned Lebow, and Janice Stein, *Psychology and Deterrence* (Baltimore: Johns Hopkins University Press, 1985).

psychology as well, but also includes findings from the field of neuroscience and incorporates emotions and what some have labeled "sacred values."

Strand I: Psychology and Decision Making

Humans can and do engage in the calculation and analysis required for rational thought—processes that Kahneman refers to as "slow thinking." However, humans are much more likely to fall back on short cuts in order to "think fast," especially when under pressure.[15] Results from laboratory experiments have yielded a list of identifiable ways in which these short cuts, or "heuristics," cause human decision making to deviate systematically from the ideal-type rational actor model. Some of these heuristics are consistent with and extend the research on cognitive biases that informed an earlier generation of research on psychology and deterrence. Many of the most provocative findings, however, reflect the emergence of prospect theory, an approach developed by Kahneman and his colleague Amos Tversky. Some potentially relevant findings include:

- Loss aversion: Subjects are more sensitive to losses than they are to gains. Therefore, even when the probabilities of different outcomes can be known, subjects make decisions based on perceived losses and gains rather than the value of the final outcome. In particular, actors will engage in highly risky gambles in an attempt to avoid losses, but they tend to behave cautiously when they are in the domain of gains. While this finding may appear to be common sense, it flies in the face of rational utility models that treat costs and benefits of equal magnitude as being equal in value.[16]

- Narrow framing: Subjects who consider cases individually are more prone to be loss averse. Instead, it is preferable to frame the problem broadly and generate a risk policy for all similar decisions (i.e., buying an extended warranty is tempting because you want to protect your new investment, but having a household policy of accepting the risk that items will fail is the economically rational choice). Elites are generally better at avoiding narrow framing and looking at individual losses in the context of a broader pattern of costs and benefits. However, even experienced policymakers can succumb to narrow framing (consciously or unconsciously) when taking the broader view conflicts with their individual preferences and interests.[17]

- Anchoring effects: When asked to offer a numerical estimate of an unknown value or quantity, the subject's estimate will be influenced by the prior suggestion of a number even if that number is randomly generated and obviously irrelevant to the question at hand. In a famous example, students were asked to write the last two digits of their social security numbers. They were then asked to

[15] Kahneman, "Part 1: Two Systems", *Thinking Fast and Slow*, pp. 19-105.
[16] *Ibid.*, 282-86.
[17] *Ibid.*, 234-41.

estimate the value of a bottle of wine. The results showed a correlation between the estimates and the social security numbers.[18] As a potential hypothesis suggested by anchoring effects, one could ask whether existing arsenal levels affect perceptions of how many nuclear weapons are necessary to maintain a credible deterrent.

- Excessive coherence: This pattern refers to the common tendency to only consider evidence that fits within an existing narrative or explanation, dismissing evidence for counterarguments as irrelevant or less trustworthy. Excessive coherence leads to overconfidence, a frequent Achilles heel of otherwise rational elite decision makers.[19]

- Base-rate neglect: When asked to identify how likely an outcome is, a rational actor would reason by making inferences from statistical probabilities. However, in practice, this is not what subjects do. Instead, they use heuristics, or simple rules of thumb, to make decisions when faced with complex or uncertain environments. For instance, if asked how likely it is that an individual is a terrorist, they might reason by representative characteristics, or the stereotype of a terrorist (ethnicity, nationality, religion, gender, skin color) and overestimate the likelihood rather than thinking in terms of probabilities (i.e., the total population of people with those characteristics is likely very high and the number of terrorists relatively low).[20]

Of the forms of systematic deviation on this list, loss aversion has received the most attention from academics working in the field of foreign policy because it forms the foundation of prospect theory. Prospect theory formalizes the empirical findings on loss aversion into a descriptive model of decision-making behavior that produces an "S" shaped curve on a four-quadrant graph (rather than the straight line at a forty-five degree angle produced by rational utility theory). That curve means that behavior deviates more sharply from rational expectations at the extremes. People skew toward risk-aversion when facing the possibility of a gain, and risk-seeking when facing the possibility of a loss. In other words, people would rather hold on to what they have than accept even a small risk of losing it in order to gain more. However, the opposite is true when people are faced with the possibility of a loss. They are much more likely to take very high risks to avoid suffering a loss, even if it is likely that they would be better off just accepting the loss and moving on.[21]

---

[18] *Ibid.*, 119-28.

[19] *Ibid.*, 75-76.

[20] *Ibid.*, 166-73.

[21] Daniel Kahneman and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, XLVII (1979), 263-291.

There have been several books and articles written on the application of prospect theory to international relations, most of which were published in the 1990's.[22] The empirical case studies primarily focus on applications of framing principles to international bargaining, although Jeffrey Berejikian (a contributor to this project) wrote an article exploring the potential application of prospect theory to deterrence.[23] Some more recent works on foreign policy and prospect theory show the potential for applying behavioral insights to the analysis of decision making surrounding WMD acquisition or use. For example, Emilie Hafner-Burton, Alex Hughes, and David Victor have written an article on elite decision making in which they use insights from behavioral economics to analyze negotiations over the North Korean nuclear program.[24] The potential exists to extend this line of inquiry into the study of specific policy instruments. In this project, Etel Solingen has considered how prospect theory could be tapped for extending existing research on economic sanctions.[25]

Strand 2: Neuroscience and Emotions

A second strand of research within behavioral economics draws heavily on findings from neuroscience that supplement lines of research in psychology. Some of the most interesting and important findings in those bodies of research in the last two decades concern the impact of emotions. Three interrelated findings have especially attracted notice in behavioral economics: rationality and emotions are not necessarily opposed; different emotions have different effects; and fairness and justice matter. Recently, some IR scholars have begun to explore the implications of these findings for issues in international relations that are relevant to this project.

First, most specialists now reject the idea that rationality and emotions should be treated as entirely separate and opposed forces. It is true that sometimes strong emotions can override rational calculations. But, importantly, emotion and reason can also interact in the process of decision making. Considerable research suggests that the ability to make rational decisions depends to some extent on emotions. The emotional feelings people have about alternative outcomes feed into the process of

---

[22] Jan Faber, "On Bounded Rationality and the Framing of Decision in International Relations: Towards a Dynamic Network Model of World Politics," *Journal of Peace Research* 27, no. 3 (1990); Janice Gross Stein and Louis W. Pauly, eds., *Choosing to Co-operate: How States Avoid Loss* (Baltimore: Johns Hopkins University Press, 1993); Barbara Farnham, ed., *Avoiding Losses/Taking Risks: Prospect Theory in International Politics* (Ann Arbor: University of Michigan Press, 1994); Jack Levy, "Prospect Theory, Rational Choice, and International Relations," *International Studies Quarterly* 41, no. 1 (1997); Robert Jervis, "The Implications of Prospect Theory for Human Nature," *Political Psychology* 25, no. 2 (2002); Jeffrey Berejikian, "The Gains Debate: Framing State Choice." *American Political Science Review* 91, no. 4 (1997); Rose McDermott, *Risk Taking in International Politics: Prospect Theory in American Foreign Policy* (Ann Arbor: University of Michigan Press, 1998); Rose McDermott, "Prospect Theory in Political Science: Gains and Losses from the First Decade," *Political Psychology* 25, no. 2 (2004).

[23] Jeffrey Berejikian, "A Cognitive Theory of Deterrence." *Journal of Peace Research* 39, no. 2 (2002).

[24] Emilie M. Hafner-Burton, D. Alex Hughes, David G. Victor, "The Cognitive Revolution and the Political Psychology of Elite Decision Making" *Perspectives on Politics* 11, no. 2 (June 2013).

[25] Her paper builds on her earlier study, Solingen, *Sanctions, Statecraft, and Nuclear Proliferation*.

assigning value – or utility – to those outcomes. When people have distinct likes or dislikes, these become part of the yardstick by which they evaluate which choices are better or worse. In contrast, when people have brain damage that reduces their ability to feel emotions, they often cycle endlessly through alternative options and find it difficult to choose one.[26]

In the field of proliferation studies, this provides a possible underpinning for the work of Jacques Hymans. Hymans argues that some state leaders pursue nuclear weapons because they simply fall in love with the idea of having the bomb.[27] Hymans follows the traditional dichotomy of describing this as an emotional reaction beyond the realm of rational calculation, but behavioral economics suggests a reinterpretation in which emotional reactions to the bomb become one factor in the cost-benefit calculations leaders make about whether to seek nuclear weapons. Beyond its relation to Hymans' theory of proliferation, one can imagine a number of possible implications of the intertwining of rationality and emotions. The way a target state responds to an offer of positive incentives, for example, might be affected by the recipient's feelings about the sender, about the type of good being offered, or even about the acceptability of letting oneself be bribed.

Second, scholars now talk about the impact of "emotions" in the plural, rather than "emotion" in the singular.[28] This reflects the fact that people can experience different emotions. Several distinct emotional states are possible, such as happiness, sadness, or anger. Research shows that different emotional states have different effects. One of the most important distinctions involves the differences between anger and fear. Fear tends to induce caution, whereas anger makes people more likely to take risks and act without much concern for the consequences. This has potential implications for a strategy like deterrence. A deterrent threat that creates a measure of fear in the target, or that takes advantage of a preexisting fear, has a decent chance of working effectively. In contrast, deterrent threats that anger the target are more likely to backfire and provoke escalation.

Third, several lines of research suggest that people give significant weight to considerations involving fairness or justice. When people believe they are treated unfairly, this creates particularly strong reactions. Experiments involving the well-known ultimatum game show that people consistently reject highly unequal distributions of benefit even when, in strictly economic terms of expected utility, these choices actually leave them worse off.[29] This research finding is rich with potential implications

---

[26] Jonathan Mercer, "Rationality and Psychology in International Politics," *International Organization* 59, no. 1 (winter 2005).

[27] Jacques E.C. Hymans, *The Psychology of Nuclear Proliferation* (Cambridge: Cambridge University Press, 2006).

[28] Janice Gross Stein, "The Psychology of Assurance: An Emotional Tale," in Jeffrey W. Knopf, ed., *Security Assurances and Nuclear Nonproliferation* (Stanford, CA: Stanford University Press, 2012).

[29] David A. Welch, *Justice and the Genesis of War* (Cambridge and New York: Cambridge University Press, 1993); Jonathan Mercer, "Emotional Beliefs," *International Organization*, Vol. 64, No. 1 (Winter 2010), pp. 10-11; Stein, "The Psychology of Assurance."

for WMD nonproliferation. Knopf, for example, has suggested that this might account for the importance some non-nuclear weapon states attach to the disarmament issue in NPT Review Conferences. The NPT involves an inherent inequality between nuclear haves and have-nots. Given this background, perceptions that nuclear weapon states are not complying with nuclear disarmament commitments associated with Article VI and hence not fulfilling their end of the NPT bargain are especially likely to rankle non-nuclear nations.[30]

Finally, a somewhat different line of work on "sacred values" suggests some similar conclusions.[31] This concept refers to core values or beliefs that people hold that are so deeply internalized that they become part of a person's identity. These sacred values are usually beyond the reach of cost-benefit calculations. Indeed, the attempt to treat them as questions of material cost and benefit can be perceived as offensive and insulting. The classic example involves the idea of offering somebody money to change their religion. Such an offer would be likely to provoke moral outrage and anger at the person making the offer. Rational actor models will typically fail to predict behavior in connection with sacred values, because this is an area dominated by beliefs about right and wrong, not by calculations of cost and benefit. It is possible that there are other areas, which might not fit the concept of sacred values, where actors do not engage in traditional cost-benefit calculations. This is an issue worth considering in explorations of deterrence and nonproliferation.

In the next section, we summarize individual papers commissioned for this project that explore some of the issues raised above.

WORKSHOP SUMMARY

For this project, we did not seek to impose a standard social science methodology. We believed it was more important to invite leading experts in the field and let them build on their most recent research. Hence, unlike projects that provide paper writers with a well-defined list of questions or hypotheses in advance, or that assign a specific case study to each author, we chose to leave our commissioned paper writers with a certain amount of discretion. We provided the invited participants with a summary of the project and the narrative from the grant proposal, then let them propose paper topics. We sought to ensure coverage of a broad range of topics related to nuclear weapons policy, and we ended up with papers that address issues ranging from deterrence to differences in national

---

[30] Jeffrey W. Knopf, "Disarmament and Nonproliferation: Examining the Linkage Hypothesis," *International Security* 37, no. 3 (winter 2012/13).

[31] For a nice summary, see Frank Rose, "Don't Mess with My 'Sacred Values,'" *New York Times*, Sunday Review, Nov. 16, 2013. Relevant studies include J. Baron and M. Spranca, "Protected Values," *Organizational Behavior and Human Decision Processes* 70 (1997); Philipp Tetlock, "Thinking the Unthinkable: Sacred Values and Taboo Cognitions," *Trends in Cognitive Sciences* 7 (2003); Scott Atran and Robert Axelrod, "Reframing Sacred Values," *Negotiation Journal,* 24 (July 2008).

decision-making styles to the nonproliferation regime as a whole. More important, we sought to bring together an interdisciplinary group of specialists. Although the majority of workshop participants came from a political science background, the participants also included a neuroscientist, a well-known behavioral economist, and scholars from fields such as philosophy, organization theory, and the study of science, technology, and society (STS).

For purposes of summarizing the workshop papers, we have grouped them into the following topics that were addressed: basic research, empirical tests of deterrence, country-specific studies, the nonproliferation regime, and observations from philosophy and STS.

Basic Research

The workshop began with a keynote address by George Loewenstein, a professor at Carnegie Mellon University who is one of the leading figures in the field of behavioral economics. Professor Loewenstein summarized the findings of a large range of experiments that illustrate some major findings in behavioral economics. He organized his presentation around three main topics: emotions, the role of self-serving biases in perceptions of fairness, and belief-based conflict.

Starting with emotions, Loewenstein noted that emotions can reprogram how people perceive and calculate their options. For example, artificially inducing emotions of either sadness or disgust can lead people to greatly discount the price at which they are willing to sell something. Some of the most interesting research in this area involves "hot-cold empathy gaps." The idea here is that people try to predict the future states of mind of both themselves and others by empathically imagining themselves in the circumstances that they or others might find themselves in later. But people tend to project from their current emotional state to assume that this will be the prevailing emotional state in the future circumstances. Hence, people who are hungry when they go shopping buy more food because they anticipate that they will still be very hungry later, even though this will not be true after they have eaten. Empathy gaps with respect to others are even more pronounced. People in a calm emotional state can find it very hard to understand how others will feel in heightened emotional state – that is, they find it hard to predict from cold to hot emotions. Loewenstein proposed that emotions are a double-edged sword for nuclear decision making. They increase the chances of impulsive, short-sighted actions that could prove disastrous. But the fear of possible nuclear destruction in the future could also help make MAD effective as a deterrent.

Next, Loewenstein discussed research that shows people care a great deal about being treated fairly. Unfortunately, we also tend to perceive fairness through self-serving biases that lead us to evaluate outcomes that are better for us as also being fairer. Again, this has mixed implications. On the negative side, it makes it harder for negotiations to succeed, as the two parties might not agree on what constitutes a fair settlement. The self-serving bias also makes nuclear threats more dangerous. A threat

that one side perceives as a legitimate defense of its interests could be perceived by the other side as unjust, provoking an angry response. On the positive side, fairness is not an inherently aggressive or hostile motivation, and a belief in fairness opens the door to mechanisms like third-party mediation that might enable states involved in disputes to reach agreements.

Finally, Loewenstein drew attention to the possibility of belief-based conflict. This refers to conflicts that arise not over material conflicts of interest, but instead from the fact two actors promote different values and beliefs, such as different religions. Although this is more speculative, Loewenstein hypothesized that actors might be especially inclined to escalate conflicts that involve perceived threats to their core beliefs. Tying together the different threads, Loewenstein reiterated the key implication that game theoretic and other rational approaches to studying nuclear interactions might prove misleading in practice.

The remaining papers for the workshop were presented in four panels. Now that we have a clearer understanding of the themes that emerged, we have slightly reorganized the order of presentation to better reflect those themes. Like Loewenstein, two other papers, by McDermott and Writght, emphasized basic research in psychology and neuroscience. These papers are discussed next, while the following sections discuss papers dealing with empirical evidence, country case studies, applications to nonproliferation, and broader philosophical reflections, respectively.

Rose McDermott of Brown University drew on the field of evolutionary psychology to derive some general implications for the operation of coercion (encompassing both deterrence and compellence). Evolutionary psychology argues that some human psychological tendencies reflect our evolutionary history. Psychological traits that were adaptive for the human needs to survive and reproduce are likely to be part of our basic psychological architecture. This is not a claim that genetics determines behavior. Rather, certain types of circumstances can trigger certain hard-wired behavioral tendencies, but the specific form that the behavior takes might come out of a repertoire of options.

McDermott observed that coercion was likely important in human evolution, but in two different ways. People needed to master not only the ability to coerce others but also the ability to resist coercion. McDermott also noted that coercion need not involve physical force, but could also involve social sanctions such as shaming or ostracism. According to McDermott, the psychology associated with coercion is motivational rather than purely calculating. McDermott suggested that some behaviors emphasized in conventional wisdom about deterrence, such as the value of forming a reputation for toughness or engaging in costly signaling, could have an evolutionary basis. But she noted that other behaviors less emphasized in deterrence theory, such as appeasing threats or entering into larger coalitions for self-protection, could also be consistent with evolutionary needs. McDermott concluded that efforts to craft deterrence should draw more on what we know about the psychology of coercion and focus less on manipulating cost-benefit calculations in an economic sense.

Nicholas Wright, a trained neuroscientist who has an affiliation with the Carnegie Endowment for International Peace, also drew heavily on basic research for his paper. Wright drew on findings from neurobiology as a way to examine doctrinal publications in the United States and China concerned with deterrence. Wright noted that both China and the United States recognize a psychological component to deterrence, seeing it as a state of mind brought about in one's adversary. But Wright suggested that neither country may have the psychology quite right, in that their predictions of how the other side will behave in response to deterrent threats or actions might not be accurate.

Wright highlighted three potential problems. The first is prediction error, or the impact of an incorrect prediction on later behavior. Both China and the United States emphasize the value of achieving surprise. But for the other side, a surprise would represent a significant prediction error (i.e., they would not have predicted it), and this could lead them to an extreme action in response. Second, as with several other authors, Wright emphasized the importance of fairness motivations and the costs actors are prepared to pay to reject unfair outcomes. But he pointed out that U.S. doctrine, as reflected in the *Deterrence Operations Joint Operating Concept* (DO-JOC) for example, assumes the other side makes straightforward cost-benefit calculations and does not take into account the motivation to be treated fairly.[32] Finally, Wright pointed out that both countries have multiple different decision systems that could be involved in a decision about using nuclear weapons. This makes unitary actor models, even if they are informed by psychology and neuroscience, potentially problematic. Wright ended with a note of caution about embracing findings from neurobiology. He suggested looking for findings that have been well-established through multiple research studies, that have sizable enough implications to be worth taking into account, and that have high potential to generalize across different cultures.

Empirical Evidence

Some papers for the project sought to evaluate behavioral research findings against empirical evidence. Jeffrey Berejikian and Florian Justwan from the University of Georgia carried out a statistical analysis to test some of the implications of prospect theory for deterrence. Prospect theory predicts that people's risk orientation will vary depending on whether they perceive themselves to be in a domain of gains versus losses. When people expect their alternative options will generally leave them better off (for example with respect to the status quo) they tend to be risk averse; they will choose a safe bet that returns modest gains over a gamble that could produce greater gains but at the risk of getting nothing. In the domain of losses, however, people flip to become risk acceptant; they often prefer an option that has the potential to eliminate their losses even if it comes with some risk of greater losses.

[32] U.S. Strategic Command, *Deterrence Operations Joint Operating Concept*, version 2.0 (Dec. 2006).

Berejikian and Justwan noted that this prediction has implications for the U.S. strategy of tailored deterrence. The 2006 DO-JOC mentions the need to consider which adversaries might be risk acceptant, but it does not offer a framework for determining which potential targets of deterrence are likely to have this risk orientation. Prospect theory provides one possible approach to predicting when adversaries will be risk acceptant. This is important because one can predict that deterrence is more likely to fail against risk-acceptant actors. When a state is risk averse, even a modestly credible deterrent threat should be sufficient to deter it, but a much higher level of credibility is likely to be required to check an actor that is risk acceptant. If we can identify when an actor perceives itself to be in a losses frame, this could serve as a leading indicator for predicting deterrence breakdowns.

To test this, Berejikian and Justwan analyzed a dataset of strategic rivalries from 1816 to 1999. For each dyad (i.e., pair of states), the authors considered the possible range of alliance partners each state might have had, from best case to worst case. States whose actual alliance portfolio was above the 50 percent point in this range were coded as being in the domain of gains; states below the hypothetical 50 percent point were coded as being in the domain of losses. In a series of statistical tests, some of which controlled for other factors that can influence deterrence outcomes, the authors found that deterrence fails more frequently when one or both states are in a losses frame and that states whose external security environment is getting worse become more likely to initiate military disputes. Berejikian and Justwan pointed out that this is contrary to conventional wisdom, which assumes that rising powers are the most likely to challenge deterrence, because it suggests that states are more likely to challenge when their relative power is going down rather than when they are gaining in power. The authors also considered a subset of cases where the deterring state has nuclear weapons and found this did not change the results. In fact, states in a losses frame were even more likely to challenge deterrence when the defender possessed a nuclear deterrent. Berejikian and Justwan drew an important policy implication from these findings: in general, states should try to avoid doing things that could make the other side feel more insecure, as that could put it in a losses frame that makes it more willing to engage in risky behavior that challenges deterrence.

A paper by Janice Gross Stein from the University of Toronto and Morielle Lotan from IDC-Herzliya in Israel sought to interpret a surprising empirical finding. In previous research, Lotan found that military personnel at lower levels of the nuclear chain of command were often reluctant to take actions that could lead to a launch of nuclear weapons.[33] Personnel in command centers that monitor possible warnings of nuclear attack did not automatically pass along apparent warnings of nuclear attack out of concern those warnings would turn out to be false alarms. Personnel who had their fingers on the button similarly seem to have been hesitant to carry out launch orders for the same reason, and there is evidence that superiors also sometimes ordered lower-level personnel not to respond immediately to warnings of attack.

[33] Morielle I. Lotan, "Strategic Dilemmas of WMD Operators," *Comparative Strategy* 34, no. 4 (2015).

Stein and Lotan identified several possible explanations for this behavior. It is possible that nuclear operators internalized the emerging nuclear taboo and this made them reluctant to launch in response to anything less than incontrovertible proof of a nuclear attack. It is also possible that some early warnings turned out to be due to a computer error or other glitch, leading nuclear operators to develop a heuristic that led them to assume that future warnings during peacetime conditions were similarly likely to be false alarms arising from some human or machine malfunction. They also noted that research on fear shows that experiencing this emotion tends to make people cautious and risk-averse. Fear of potentially starting a nuclear war in response to a false alarm could certainly induce a great deal of hesitation. Stein and Lotan discounted prospect theory as an explanation, arguing that warning of nuclear attack would put people in the domain of losses and make them willing to take the risky gamble of launching nuclear retaliation. But this is not necessarily the case. Motivation to avoid the enormous destruction that would accompany a nuclear exchange might lead operators to take the potentially risky gamble of delaying a response in the hope that the warning turns out to be false.

Interestingly, Stein and Lotan disagreed about the policy implications of their research. Stein argued that hesitancy to act on initial warning reduced the danger of accidental nuclear war but in a way that undermined the automaticity of nuclear retaliation that might be necessary for effective deterrence. Lotan, in contrast, suggested that understanding how nuclear operators think might open the door to tailored deterrent approaches that reinforce the reluctance of operators to initiate the use of weapons of mass destruction.

Zachary Zwald, who is now at the University of Houston, also presented a paper with an empirical focus. Zwald examined domestic U.S. debates about ballistic missile defense. He did this in the context of a larger question about how states choose which new military technologies to develop. Zwald wanted to know if there is a particular approach to making such decisions that would improve the odds of making a good choice about when to develop a candidate military technology.

Zwald noted that conventional wisdom on this question involves a "science vs. politics" framing. This approach assumes that those who have parochial interests at stake, such as military services or defense contractors that would benefit from missile defense projects, will give biased judgments that exaggerate the technical feasibility and strategic utility of missile defense projects. In contrast, technical experts from the scientific community are expected to give unbiased judgments. In this view, scientific experts have frequently identified flaws in proposed missile defense programs, but special interests have had enough political influence to keep missile defense programs funded despite the technical problems identified by experts.

Zwald challenged this conventional wisdom on two grounds. First, some key participants in missile defense debates do not have obvious parochial interests at stake. For many policymakers and outside analysts, their views reflect other general beliefs they hold rather than narrow self-interest. For

example, support for missile defense can be affected by beliefs about the scenarios most likely to lead to nuclear war (e.g., deliberate attack or inadvertent escalation) or about the degree to which decisive technical breakthroughs are possible. Second, research in psychology has shown that all people are prone to certain biases and tend to rely on heuristics. Once people come to hold certain beliefs, their subsequent judgments will be biased by their preexisting beliefs. This does not mean that those with special interests are unbiased, but rather that scientists and political figures on both sides of the missile defense debate are also likely to have biases. Zwald found support for this in the history of the U.S. debate about candidate missile defense technologies. For example, when a particular test of a system under development had ambiguous results, missile defense supporters tended to interpret the test favorably and critics unfavorably.

Zwald concluded by recommending a more "pragmatist" approach. This approach takes note of the tremendous uncertainty and complexity involved in predicting whether missile defenses will ultimately help U.S. national security. And it suggests that the democratic political process can be helpful. Rather than try to insulate military technology decisions from politics and turn them over to supposedly neutral experts, it might be better to encourage open political debate. This would require parties to the debate to better explain their positions and might allow them to look for common ground.

Country-Specific Case Studies

Two of the papers did deep dives on individual countries. Joshua Pollack, then with SAIC, examined North Korea's leadership. He noted the common observation that North Korea is a particularly opaque country, which makes it hard to ascertain what the DPRK's leaders are thinking. Pollack observed that people sometimes respond to this constraint by drawing conclusions based on comparing North Korea to other countries, but he argued this could be problematic given idiosyncratic features of the North Korean regime. Pollack set out to show that it is possible to glean insight into North Korea's thinking.

Pollack identified several sources of information into North Korean decision making. First, there are several memoirs written by former officials who defected. Second, there are archival materials from the former Soviet Union concerning Soviet relations with North Korea, but these end in 1991. There are also occasional leaked documents, but their authenticity is hard to verify. Speeches by leaders that are made public represent another source. Finally, Pollack suggested that a process akin to old-style Kremlinology remains relevant. This process includes both the close reading of ideological texts to detect subtle shifts in wording and the analysis of photos, including who is photographed standing close to the top leader. Based on these sources of information, Pollack concluded that North Korea under Kim Jung Un remains a highly unitary regime, but with a greater level of uncertainty than

before about how strongly the young Kim has consolidated power and who in the inner circle has his ear.

Nikolai Sokov from the Center for Nonproliferation Studies at the Middlebury Institute of International Studies wrote a case study on Russia's national security policies and negotiating style. Sokov put an emphasis on culture and history as sources of policy leanings and negotiating style. Sokov claimed that many studies of cultural influences are too deductive in nature, and he argued for an inductive approach based on examining history over a long period of time. Sokov said that culture can bring about deviations from the predictions of rational actor models, but in directions that are consistent and predictable.

Turning to the Russian case, Sokov observed that there are some commonalities across czarist times, the Soviet era, and post-Soviet Russia, but also that more recent historical lessons have also had an impact. Sokov said that Russia tends to respond on a more elongated time frame than countries in the West. Rather than an immediate tit-for-tat or action-reaction pattern, Russia tends to wait passively for a time before responding actively and all at once to a whole series of moves by outside powers. For example, while Russia's intervention in Ukraine appeared to Western countries to be a sudden and unprovoked escalation in Russian behavior, to Russia's leaders it was a long overdue reaction to a series of Western moves viewed as unfavorable to Russia's interests.

In addition, Sokov described Russia as tending to follow *realpolitik*, with an emphasis on power and national interest and a strong drive for equal status with other great powers. These interest calculations can sometimes be overridden, however, by a tendency to identify with fellow Slavs who are perceived as under threat in other countries. Russians, Sokov argued, also tend to become emotionally attached to policies, which can make them slow to alter course when policies are not working well. In addition, according to Sokov, Russians place a great deal of emphasis on developing trust through personal relationships. This can create complications in negotiations with the United States, which has a practice of frequently rotating personnel. Collectively, all of these cultural tendencies make Russia not an easy country to deal with in negotiations. But, Sokov concluded, it is still better to take the time to learn about other country's particular negotiating styles than to simply assume a generic, rational actor.

Applications to Nonproliferation

Three papers dealt specifically with aspects of nuclear nonproliferation. Etel Solingen from the University of California-Irvine updated her previous work on the impact of economic sanctions. Solingen said most research on sanctions has assumed the target state is a unitary state actor. In contrast, Solingen's work has emphasized domestic political coalitions. Solingen has found that ruling coalitions that are outward looking and interested in having their state participate actively in the global

economy are more likely to be sensitive to sanctions. In contrast, inward-looking leaderships tend to emphasize nationalism and economic self-reliance and so are less susceptible to being influenced by economic sanctions.

In her paper for this project, Solingen added insights from psychology, and especially prospect theory, to her previous analysis of domestic coalitions. She suggested that prospect theory reinforces the coalitional dynamics she has identified. States looking to integrate into the global economy will see this as a domain of gains and so be risk averse. This should make them more willing to adjust behavior to avoid the potential negative impact of sanctions. States that resist economic integration are probably viewing increased ties to the world economy as a potential loss, based on the idea that ruling elites are tied to a domestic base of support that involves groups that either could not compete successfully in the international economy or that fear the impacts of exposure to the modern world. Leaders in such states may see the risky gamble of forging ahead with a nuclear program as a last-gasp hope of saving a ruling regime from crumbling in the face of international pressures, which makes it harder to use sanctions to dissuade them from a nuclear path.

Solingen also hypothesized that the timing of sanctions can be important. Prospect theory has identified an endowment effect, in which actors value something more once they possess it, meaning they will demand a higher payment to give it up than they would be willing to offer to buy the asset before they possessed it. Nuclear programs, as they progress, could create endowment effects. States will become more reluctant to give up capabilities they have achieved than capabilities they have not yet mastered. Solingen suggested that enrichment capabilities could be such a threshold. Sanctions will have to be more severe to get states to give up a uranium enrichment capacity they have achieved than to get them for forswear such a capacity before it is realized.

Solingen also hypothesized that common psychological biases could affect how outside analysts assess the outcomes of individual cases. In many cases where sanctions are imposed, the outcome is not a clear-cut success or failure. In some countries, nuclear programs are delayed or scaled back or moved onto a different approach to technology development but without being fully terminated. Solingen noted that analysts who believe that "sanctions don't work" will typically interpret such cases as yet another failure, whereas analysts more favorably disposed toward sanctions will see such cases as at least a partial success. Solingen called for more self-awareness among analysts in how our own views can affect our assessment of the impact of sanctions in specific cases.

The other two nonproliferation papers both focused on research concerning the importance that people attach to considerations of justice and fairness. George Perkovich from the Carnegie Endowment for International Peace analyzed the nuclear negotiations with Iran as a case study of how claims about fairness can enter into nonproliferation discussions. Because the workshop was held before the conclusion of the Iran nuclear deal in summer 2015, Perkovich's paper dealt with the talks

from 2003 through fall 2014. Perkovich noted that the United States has tended to describe the talks as being about compliance, whereas Iranian negotiators have been quite explicit about invoking fairness considerations.

According to Perkovich, compliance concerns can be related to questions of justice. They involve a claim that it is just to abide by international legal commitments. Perkovich added that such justice claims ultimately depend on the perceived legitimacy of the legal commitments. The more legitimate the NPT is perceived to be, the more weight will be carried by claims about the importance of compliance with that treaty. The potential problem here is that the NPT is widely perceived to be an unequal treaty, and this can work against claims about the treaty's legitimacy.

In Iran, Perkovich claimed, the discourse of justice is pervasive. It is reflected in the names of political parties and fits into a narrative in which Shias perceive themselves as victims of a historic injustice. Iran has long been vocal about pointing out what it regards as unfair elements of the NPT. Iran regularly claims that the nuclear weapon states have not fulfilled their commitments under NPT Article IV (regarding peaceful use) and Article VI (regarding disarmament). Iran also objects to the memberships of the IAEA Board of Governors and the UN Security Council as giving disproportionate weight to the nuclear weapon states and their allies. Finally, Iran points to double standards with respect to enforcing compliance, as illustrated by the U.S.-India nuclear deal for example. Perkovich suggested that Iran's rhetoric was effective in transforming the situation from one strictly about enforcing compliance to one involving bargaining between equals.

Perkovich next examined what it might take for the Iran nuclear negotiations to succeed in light of Iran's fairness-related claims. Perkovich looked at the memoir of Hossein Mousavian, a former Iran nuclear negotiator, which identified twelve key issues in the negotiations. Perkovich noted that many of these issues involved a fairness element. Analyzing these fairness concerns, Perkovich predicted (fairly accurately as it turned out) that Iran might agree to a nuclear deal if that deal enabled Iran to keep some enrichment capability and specified some limits on international inspections.

Perkovich concluded by returning to the question of compliance. He argued that compliance is necessary for a rule-based international order. But for that order to be perceived as fair and legitimate, the obligations imposed by the rules have to be balanced and compliance has to be universal. If key powers selectively exempt themselves from some of the rules, the legitimacy of subsequent compliance demands will be compromised, making it harder to maintain international order.

The final paper dealing with nonproliferation was by Harald Müller, executive director of the Peace Research Institute Frankfurt. Müller focused on how justice concerns might affect the nonproliferation regime. Müller summarized several lines of research that lead to a conclusion that human beings seem hardwired to care about issues of justice and fairness. Drawing on work by Nancy

Fraser, Müller noted that justice concerns can be about distributional, procedural, or recognition (i.e., are people treated as having equal standing) questions.

If people have innate concerns about fairness, this can be problematic for the nonproliferation regime. The foundation for global nonproliferation efforts is the NPT, which is in important ways an unequal treaty: it recognizes five states as nuclear weapon states while requiring all other states to join as non-nuclear weapon states. Müller identified several ways in which justice concerns have contributed to frictions in the NPT context. One example is complaints by non-nuclear weapon states about a lack of progress toward nuclear disarmament by the nuclear weapon states. These complaints involve both distributional concerns (some states have nuclear weapons but others don't) and recognition concerns (some states have an identity that allows them to be nuclear-armed but others do not get equal status). Other issues that have justice dimensions include access to peaceful uses of nuclear technology and questions related to the lack of universal membership in the NPT, in particular the perception that India and Israel have received special treatment despite not having signed the NPT. Müller also considered procedural justice issues, including questions about which decisions can be made by bodies in which not all states are members, such as the IAEA Board of Governors.

Müller's analysis was not wholly pessimistic, however. He noted that justice and fairness concerns can also at times provide a basis for finding common ground. For example, certain norms can become widely accepted and provide a basis for agreement. Müller suggested that concerns about nuclear security became largely accepted once they were taken out of the "haves vs. have-nots" context of many NPT debates. In addition, progress on one dimension of justice can help alleviate concerns on other dimensions. As an example, Müller cited procedural innovations in NPT Review Conferences that helped give non-nuclear weapon states a better sense that their views were represented. In some years, these have helped pave the way for agreements on certain action steps that would help address the concerns on non-nuclear weapon states in substantive areas such as disarmament and peaceful uses. Müller also noted that states can engage in a kind of issue-linkage around justice issues, trading concessions on a justice claim that others care about more in return for concessions on justice issues that they care about. Müller concluded that justice concerns have both negative and positive implications for nonproliferation. They can add to the cleavages among states on nonproliferation issues, but they can also create new mechanisms for finding common ground on how to move forward.

Philosophical Reflections

A final set of papers stepped back from the realm of behavioral economics to put some of the issues into a larger context. Specialists from fields such as philosophy, organization theory, and the study of science, technology, and society brought the perspectives of their disciplines to questions concerning rationality and behavior when it comes to nuclear deterrence. Jean-Pierre Dupuy, a

philosopher who has affiliations with both Stanford University and the Ecole Polytechnique in Paris, reexamined some classic deterrence paradoxes that previous analysts had sought to address during the Cold War era. He noted that many analysts found it difficult to square deterrence through mutual assured destruction (MAD) with assumptions of rational choice: if one side actually was destroyed in a nuclear attack, it would no longer be rational for that actor to retaliate, and the other side's recognition of this should undermine the credibility of deterrence.

Dupuy discussed a series of possible solutions. It can help if the two actors no longer think in terms of playing the game against each other but instead come to view themselves as playing jointly against a fictitious third entity. This third party would represent the common fate the two actors would share if deterrence failed and they both suffered catastrophic destruction. This potential shared fate creates an interest in the two parties in both making sure deterrence succeeds. According to Dupuy, however, this solution still suffers from the problem that it does not appear credible for either side to believe the other. To make it possible for mutual deterrence to work, Dupuy suggested two additional moves.

Part of the solution, said Dupuy, lies in how we think about time. Dupuy argued we need a notion of "projected time." Normally, we think in terms of "occurring time," in which we can make choices that would take us down different branches to alternative futures. In projected time, in contrast, we treat the future as if it is fixed – in this case fixed on a future in which the two sides avoid nuclear war. This feeds back into the present in a way that leads actors to react to this known future in ways that help produce it. Even this does not fully solve the problem, however. By the logic of backward induction familiar from game theory, if one side knows the other side will be perfectly deterred from using nuclear weapons it will feel free to behave recklessly.

For this reason, Dupuy argued we also need an element of indeterminacy. Dupuy made an analogy to Schrodinger's cat, which while it remains unobserved inside a box is famously both alive and dead. We can think of this as living in a world in which nuclear Armageddon both does and does not happen in the future. But we want to also think in terms of a fixed future in which the outcome is no nuclear war. To get there, according to Dupuy, requires nearly experiencing the disastrous alternative. It was precisely the experience of going through several near misses during the Cold War that led the two sides to exercise the caution that made it possible for mutual deterrence to work. This leads to the uncomfortable conclusion that perhaps close calls were not failures of deterrence but rather necessary conditions for deterrence to succeed.

Alfred Nordmann, a professor of both philosophy and history of science at Darmstadt Technical University, wrote a paper with the intriguing title "Four Horsemen and a Rotten Apple." The four horsemen reference involves George Shultz, Henry Kissinger, Sam Nunn, and William Perry, whose series of op-eds in the *Wall Street Journal* brought renewed attention to the idea of nuclear

disarmament.[34] The "rotten apple" is a reference to the Austro-Hungarian empire before World War I. Nordmann's point was that efforts to maintain a stable nuclear order can be compared to the efforts to keep the Austro-Hungarian empire from falling apart. In the latter case, doing nothing was not an option, but efforts at far-reaching reform ran the risk of bringing about exactly the result they were intended to prevent.

Nordmann suggested that in nuclear affairs and pre-1914 Europe the prevailing approach had been a kind of technically informed "muddling through." He said there are different ways of thinking about the "rightness" of particular policies or actions. We can think of these as being based on claims about what is true or false or on claims about what is morally right or wrong. But there is also a third option: claims about what is technically right or wrong to keep a given system functioning. In cases of a "rotten apple," or a mechanism that is decaying over time, this requires a good feeling for the mechanism and an ability to improvise or jerry-rig solutions. Nordmann concluded that this kind of fine tuning of the mechanisms of deterrence and nuclear security might be the most important factor in avoiding nuclear war. Rather than big emotions like fear of nuclear war or grand principles such as the desirability of disarmament, efforts to avoid nuclear war might rest primarily on finding adequate technical and managerial approaches. Nordmann expressed concern, however, that the nuclear order might be so complex that no one fully understands the mechanisms necessary to keep it functioning.

The final paper presented at the workshop was co-authored by John Downer from the University of Bristol and Anne Harrington, one of the project PI's who was working for CNS at the time the grant was awarded but now has a faculty position at the University of Cardiff. Downer and Harrington set themselves the task of "red teaming" behavioral approaches, or offering a critique of behavioral economics as a foundation for understanding nuclear decision making. Their paper raised two related criticisms. The first concerns the degree to which behavioral economics can be used as a basis for prediction. Downer and Harrington noted that behavioral economics has been attractive to analysts who agree with its critique of the rational actor model at the heart of traditional economics. But, they argued, behavioral economics shares a problem with mainstream economics. They both make predictions about the expected average behavior of a large number of actors. The models developed by behavioral economists are based mainly on the findings of experiments conducted on groups of people. Within any experiment, some people do not make the same choices as the majority. The predictions of behavioral economics are probabilistic. They suggest what the average behavior will be in a large group of people; they do not predict the choices or behavior of any given individual. In this sense, although behavioral economics has identified predictable deviations from rationality, they are deviations that can be observed only in the aggregate. We still cannot predict if any given actor in a given situation will deviate in that way from making a rational choice. Yet, when it comes to

[34] George P. Shultz, William J. Perry, Henry A. Kissinger and Sam Nunn, "A World Free of Nuclear Weapons," *Wall Street Journal*, Jan. 4, 2007.

deterrence, what we really want to know is whether a particular actor in a particular set of circumstances will be deterred. Behavioral economics might be able to assist us in assessing the probabilities, but it cannot generate a reliable prediction.

Second, applying behavioral economics to nuclear decision making involves a form of category error, such as mistaking a set of buildings for being equivalent to "the university." Behavioral economics applies to decisions by individuals. These, however, will not necessarily be the same thing as eventual behaviors by states. As suggested by the well-known "level of analysis" issue in international relations, there is a difference between an individual and a state. The latter is a collective actor made up of multiple organizational units. The fact that bureaucratic organizations are involved in state decision making means that predictions about individual decisions will not necessarily translate directly into state decisions. Downer and Harrington suggested that there might be a parallel to Kanheman's distinction between thinking fast and slow. An individual leader might generally follow the pattern of thinking fast, but bureaucratic processes impart the state with elements of thinking slow as well.

Workshop Discussion Areas

Some themes addressed in workshop presentations generated relatively little controversy, while others provoked a fair amount of discussion and debate. Because the issue of innate concerns with fairness came up in several papers, there was considerable discussion of this topic at the workshop. One participant strongly challenged the notion that considerations of justice or fairness would so strongly stand out from other types of value-laden beliefs in human decision making. This participant suggested that people could be equally or more motivated by other values, such as peace or freedom or meritocracy. Most workshop participants disagreed with this claim, arguing that the cumulative research is too solid to dismiss the idea of an innate concern with fairness. Another criticism of focusing on justice concerns was that a belief in justice could lead to a desire for vengeance, which is hardly helpful for reducing nuclear dangers. One other point, raised by several participants, involved how to distinguish between sincere and strategic uses of fairness arguments. When the Iranians argue they have been treated unfairly, for example, do they really believe it or are they using this rhetoric to win sympathy from others? If the latter were true and it worked, however, it would still suggest there is something about fairness concerns that resonates with people.

In discussion following the Downer and Harrington "red teaming" paper as well as others, some of the other workshop participants agreed with the cautionary note about behavioral economics. One agreed that behavioral economics is trying to stick too closely to the axiomatic approach of mainstream economics. He argued for making use of individual findings that are well supported and helpful for policy purposes without necessarily embracing the entire approach.

There was also some disagreement, however, about which findings are well supported. Prospect theory is a case in point. Some papers relied heavily on its hypotheses concerning loss aversion. Others expressed some skepticism about prospect theory. One noted that if fear makes people more cautious, this might be inconsistent with the expectation that being in the domain of losses makes people more risk acceptant. Another pointed out that prospect theory depends heavily on framing, and if we don't understand how people frame their reference point we cannot tell whether they are in the domain of gains or losses. None of the workshop participants, however, wanted to jettison behavioral economics altogether. Most seemed to conclude that it contains valuable insights that might help us to better understand nuclear decision making.

CONCLUSIONS AND RECOMMENDATIONS

Behavioral economics, and the bodies of research that inform it, give us good reasons to doubt that human beings can fully live up to the assumptions of a rational actor model. The participants in this project generally agreed with this skepticism, and they believe it extends to the realm of decision making about nuclear weapons. For this reason, our first conclusion is neither new nor controversial, but is still worth remembering: it would not be prudent to rely on theories or models of deterrence or nonproliferation that assume states will make rational choices.

Beyond this starting point, this project did not generate a strong consensus around specific conclusions, so the remaining conclusions identified here are modest. While we should not assume perfect rationality in real-world situations, it would be a mistake to jettison expectations of rationality so fully that we end up assuming that other actors are simply irrational. Fortunately, truly irrational actors are fairly rare. The difficulties lie elsewhere. First, rational choice theories often require making assumptions about what actors value (as in the assumption that firms want to maximize profits).[35] But these assumptions can also be inaccurate. This was the essence of the critique of deterrence theory from the perspective of strategic culture. It argued that we have to understand what other actors value most and tailor deterrence to hold those values at risk.

But behavioral economics suggests a second difficulty. Actors strive to be rational, but they are subject to biases and rely on heuristics that lead to systematic deviations from the predictions of rational actor models. The allure of behavioral economics rests in part on this hope for predictability. If we understand the ways in which people systematically deviate from rationality, we can still predict their behavior. Unfortunately, at the end of this project, we conclude that this promise cannot be fully realized. There are disagreements among the proponents of a behavioral approach about which

[35] Herbert Simon referred to these as "auxiliary assumptions" and argued that they often do much of the work in rational choice theories (see Herbert A. Simon, "Human Nature in Politics: The Dialogue of Psychology with Political Science," *American Political Science Review* 79, no. 2 [June 1985]).

predictions are most valid, and because the research findings hold true for the majority of people but not all there are likely to be exceptions to almost every prediction. Although behavioral research has generated a number of specific insights that it would be useful to be aware of and keep in mind in strategic planning and decision making, it is difficult to integrate these into a single, comprehensive framework.

We conclude, therefore, that behavioral economics does not yet provide a coherent and predictive model that can by itself serve as a basis for designing strategies. While it suggests important insights, behavioral economics does not on its own give us an analytical framework sufficient for identifying policies that will minimize nuclear dangers. For this reason, policymakers and defense planners are likely to still find value in taking into consideration the rational self-interest and strategic cultures of other actors. At the same time, behavioral research reveals important limitations to relying too heavily on assumptions of either rational or cultural determinants of behavior. Predictions based on the assumption that others will behave rationally can often be frustrated. And while human behavior is influenced by culture, it is a product of many other factors as well. Behavioral approaches reinforce long-standing intuitions about the importance of understanding other actors, but in ways that go beyond an appreciation for strategic culture.

Behavioral research suggests ways to add new depth to the ancient insight of Sun Tzu about the value of knowing both ourselves and others. In particular, it is vital to learn as much as possible about how others understand their situations. If any theme runs across multiple papers for this project, it is the effect of how people frame their situations. The framing for choice can have a huge impact on the decisions people make. Frames can affect whether people perceive a question as involving only a simple interest calculation or instead involving a matter of justice or fairness. Frames also affect whether actors perceive themselves to be in a domain of gains or losses. Cultural legacies are often one ingredient in such framings, but the key question is how individuals frame the situation or context in which they have to make a decision, and this can be affected by their present circumstances and recent diplomatic interactions as well as by historical legacies. Just as it is important to learn as much as possible about how others frame their circumstances and choice options, it is also important to examine ourselves and try to figure out our own biases. The better we know ourselves and know others, the better we can begin to anticipate how some common psychological tendencies might influence the course of our interactions.

Even if behavioral economics does not provide a reliable basis for prediction, the research that informs it suggests several insights that can be valuable to keep in mind. First, emotions can exert a powerful impact on human decisions and behavior. We should not assume that decisions about nuclear weapons will be shaped solely by cool calculations of cost and benefit. Emotions can also be differentiated, and in most situations it will be advisable to avoid taking actions that provoke anger in the other side.

Second, people are often motivated more strongly by the desire to avoid or minimize loss than by the pursuit of gain. In situations where we perceive an aggressive actor out to expand its territory or influence, it is worth considering whether in its own thinking that actor views itself as seeking to avoid a future loss or reverse a past loss. As with anger, it is also generally a good idea not to take actions that will put other actors in a losses frame.

Third, people also tend to care about considerations of justice and fairness, but often in self-serving ways. This makes it important to explore the possibilities for developing common standards of fairness. Such agreements, which might be about fair procedures and not just substantive outcomes, could pave the way for a successful diplomatic negotiation or a stable deterrent relationship.

Fourth, time horizons can change how people think. In particular, longer time horizons may allow space for more deliberative reasoning, which in turn can reduce the risks of rash or hasty decision making. This may be particularly the case with respect to keeping in mind the catastrophic destruction that could result from use of nuclear weapons and imparting a sense of shared danger or shared fate with respect to the urgency of avoiding any use of nuclear weapons in the future.

Fifth, and finally, it is not easy to find optimal or permanent solutions to problems. We should not be unduly dismissive of the merits of "muddling through" when such efforts manage to keep problems from becoming much worse. When it comes to dealing with nuclear weapons, there is much to be said for avoiding worst-case outcomes even when the solutions are less than perfect.

Even if behavioral economics does not provide an obvious solution to minimizing nuclear risks or a ready formula for making strategic choices, it does contain a number of separate insights that are useful to know and keep in mind. It is our hope that this study will help to create greater awareness of those insights.

REAL-WORLD NUCLEAR DECISION MAKING WORKSHOP

Final Agenda

DATES:
Monday, October 27, 5:00-8:00pm, and Tuesday, October 28, 8:30am-8:00pm

LOCATION:
Center for Nonproliferation Studies, Middlebury Washington DC Office
1400 K Street, 12th Floor

DRESS: Business casual

PROGRAM: Monday,October27

5:00-8:00pm
Reception
Welcome remarks by Sandy Spector, CNS Deputy Director

6:00-7:00pm
Keynote address by George Loewenstein, Carnegie Mellon University

Tuesday,October28

8:00am
Coffee and light breakfast

8:45-10:15am
Panel 1: Shifting the Deterrence Paradigm: From Rational Choice to Behavioral Sciences

Chair: Anne Harrington

Rose McDermott, Brown University
Title: The Evolved Psychology of Coercion

Janice Stein, University of Toronto, and Morielle Lotan, IDC-Herzliya
Title: Disabling Deterrence and Preventing War: Decision Making at the End of the Nuclear Chain

Jeffrey Berejikian, University of Georgia
Title: Deterrence and Risk

10:30am – 12:00pm
Panel 2: Contemporary Case Studies in Real World Nuclear Decision Making

Chair: Jon Wolfsthal, CNS

Nicholas Wright, Neuroscientist & Stanton fellow, Carnegie Endowment for International Peace
Title: The Neurobiology of Deterrence: Lessons for Current United States' and Chinese Doctrine

Joshua Pollack, SAIC
Title: North Korean Leadership and Decision-Making: Do We Know Anything About It?

Zachary Zwald, Air War College
Title: Constructing Power: A Pragmatist Approach to Military Technology Innovation

12:00-1:00pm
Lunch

1:15-3:15pm
Panel 3: Sanctions, Negotiations, and the Nonproliferation Regime

Chair: Jeff Knopf

Harald Muller, Peace Research Institute-Frankfurt
Title: Nuclear Non-proliferation and the J-word: Justice as the Key for Fixing or Destroying the Regime

Miles Pomper, CNS (presenting a paper authored by Nikolai Sokov)
Title: The Impact of Culture on National Security Policy and Negotiations: The Case of Russia

George Perkovich, Carnegie Endowment for International Peace
Title: Framing the Iran Negotiations: Compliance vs. Justice

Etel Solingen, UC-Irvine
Title: Sanctions, Sequences, and Statecraft: Insights from Behavioral Economics

3:30-5:00pm
Panel 4: Adding Insights from Philosophy and Science and Technology Studies

Chair: Miles Pomper

Jean-Pierre Dupuy, Stanford
Title: Apocalypse Now: Rational Choice Before the Unthinkable

Alfred Nordmann, Technical University Darmstadt
Title: Four Horsemen and a Rotten Apple

John Downer, University of Bristol, and Anne Harrington, CNS and ETH-Zurich
Title: Red Teaming Behavioral Approaches to Nuclear Decision Making

5:00-5:30pm
Discussion of next steps/future plans?

6:00-8:00pm
Dinner

REAL-WORLD NUCLEAR DECISION MAKING WORKSHOP

Participant Bios

Jeffrey D. Berejikian is a Josiah Meigs Distinguished Teaching Professor at the University of Georgia, Associate Professor in the Department of International Affairs, and a Faculty Fellow at the Center for Trade and Security. He teaches classes in international politics and foreign policy. In his research Dr. Berejikian uses findings from cognitive psychology and neuroscience to better understand contemporary foreign policy, including international conflict and military deterrence. His work appears in the top journals in the discipline, including the *American Political Science Review*, *Political Psychology*, and the *Journal of Conflict Resolution*. His book, *International Relations Under Risk: Framing State Choice* was published by SUNY Press, and has been translated into Korean. Dr. Berejikian received his Ph.D. and M.S. in Political Science from the University of Oregon and B.S. in Political Science from California State University at San Luis Obispo.

John Downer After undergraduate training in sociology, and the philosophy of science from Edinburgh and Cambridge universities respectively, John Downer received his PhD in 2007 from Cornell University's Department of Science and Technology Studies. His thesis, entitled **"**The Burden of Proof: Regulating Ultra-High Reliability in Civil Aviation**"** looked at the epistemological complexities of critical-technology assessment, and was awarded Cornell's Guilford Prize for writing. On graduating, he worked for three years as a researcher at The London School of Economics' ESRC Centre for Analysis of Risk and Regulation (CARR), and then Stanford University where he lectured for the Science, Technology and Society program and undertook research at the Center for International Security And Cooperation (CISAC) as a Zuckerman, and then Stanton, Fellow. On returning to the UK in late 2012, he joined the faculty at the University of Bristol's department of Sociology, Politics and International Studies (SPAIS), where he is affiliated with the Global Insecurities Centre. His publications look at a range of questions relating to technology regulation, risk management, and the structural causes of disaster in complex, high-reliability systems. Primarily using case studies from the aviation and nuclear spheres, and drawing heavily on the STS literature, they look at issues pertaining to the limits of knowledge and expertise: the inherent ambiguities of formal assessments and the policy implications that arise from understanding those ambiguities. His most recent publication is entitled "Disowning Fukushima: Managing the Credibility of Nuclear Reliability Assessment In The Wake Of Disaster." It can be found in the July 2014 edition of *Regulation and Governance*.

Jean-Pierre Dupuy is Professor Emeritus of Social and Political Philosophy, Ecole Polytechnique, Paris and Professor of Political Science, Stanford University. He is a member of the French Academy of Technology, a spinoff of the Academy of Sciences, and of the Conseil Général des Mines, the French High Magistracy that oversees and regulates industry, energy and the environment. He chairs the Ethics Committee of the French High Authority on Nuclear Safety and Security. He is the Director of the Research Program of Imitatio, a new foundation devoted to the dissemination and discussion of René Girard's mimetic theory.

His most recent work has dealt with the topic of catastrophe. Among his recent publications in English: *The Mechanization of the Mind* (Princeton University Press, 2000); *On the Origins of Cognitive Science* (The MIT Press, 2009); *The Mark of the Sacred* (Stanford University Press, 2013); *Economy and the Future. A Crisis of Faith* (Michigan State University Press, 2014); *A Short Treatise on the Metaphysics of Tsunamis* (Michigan State University Press, in press.)

Anne I. Harrington is on the faculty at the University of Cardiff in Wales, an appointment that began in January 2016. Before moving to Cardiff, she was a Researcher in the Nuclear Policy Group at ETH Zurich, and prior to joining ETH she was a Cybersecurity Fellow at the Congressional Research Service. Since earning her PhD from the University of Chicago in 2010, Dr. Harrington has held fellowships at Stanford University and the Center for Nonproliferation Studies (in Monterey, CA and Washington, DC). In 2013 she was awarded an American Political Science Association Fellowship to work on Capitol Hill where she staffed Senator Kirsten Gillibrand (D-NY). In Senator Gillibrand's office she worked on military personnel issues including sexual assault prevention and response, the integration of women in combat, and the implementation of the repeal of Don't Ask Don't Tell. She also handled the Senator's cyber portfolio, focusing on DOD cyberworkforce issues and critical infrastructure protection. She is married to Col Brenda Cartier, USAF Special Operations, Vice Commander, Air Warfare Center, Hurlburt Field, FL.

Jeffrey W. Knopf received a Ph.D. in Political Science from Stanford University, where he worked most closely with Alexander L. George and Scott Sagan. Dr. Knopf has taught at the University of Southern California, the University of California-Santa Cruz, and the Naval Postgraduate School. Since August 2012, he has been Professor and Program Chair of Nonproliferation and Terrorism Studies at the Middlebury Institute of International Studies at Monterey, and a Senior Research Associate at the Institute's Center for Nonproliferation Studies (CNS). Dr. Knopf is a former Editor of *The Nonproliferation Review* and has published extensively on nuclear weapons-related issues. He has been a PI or team member for several research projects sponsored by DTRA. He was also part of the team for a project on Influencing Violent Extremist Organizations sponsored by the Strategic Multilayer Assessment program in OSD, and he has been an invited speaker at the US Strategic Command Deterrence Symposium. He received the Bernard Brodie Prize for the best article in 2010 in the journal *Contemporary Security Policy* for his essay on "The Fourth Wave in Deterrence Research." His most recent publication is as editor of a book volume on *International Cooperation on WMD Nonproliferation* published by University of Georgia Press.

George Loewenstein is the Herbert A. Simon University Professor of Economics and Psychology at Carnegie Mellon University. He received his PhD from Yale University in 1985 and since then has held academic positions at The University of Chicago and Carnegie Mellon University, and fellowships at Center for Advanced Study in the Behavioral Sciences, The Institute for Advanced Study in Princeton, The Russell Sage Foundation, The Institute for Advanced Study (Wissenschaftskolleg) in Berlin, and the London School of Economics. His research focuses on applications of psychology to economics, and his specific interests include

decision making over time, bargaining and negotiations, psychology and health, law and economics, the psychology of adaptation, the role of emotion in decision making, the psychology of curiosity, conflict of interest, and "out of control" behaviors such as impulsive violent crime and drug addiction. He is one of the founders of the fields of behavioral economics and neuroeconomics.

Morielle I. Lotan is the assistant academic director of the Honors program in Strategy and Decision-Making at the Interdisciplinary Center (IDC) in Israel. Ms. Lotan is also a teaching and research assistant at the school. While completing her Masters degree at the IDC and in the years since then, Ms. Lotan has been a member of the Comparative National Security Project (IDC) which aims to explore theoretical, historical and empirical questions pertaining to international security studies. Her individual research explores nuclear operators and their role as decision makers in nuclear strategy. Her article entitled “Fingers on the trigger: Dilemmas of WMD operators and Nuclear Strategy” will be published in *Comparative Strategy* in 2015. Ms. Lotan holds an M.A with honors in Conflict resolution and Diplomacy and a B.A with Honors in Political Science and psychology.

Rose McDermott is the David and Mariana Fisher University Professor of International Relations at Brown University and a Fellow in the American Academy of Arts and Sciences. She received her Ph.D.(Political Science) and M.A. (Experimental Social Psychology) from Stanford University and has taught at Cornell, UCSB and Harvard. She has held numerous fellowships, including the Radcliffe Institute for Advanced Study, the Olin Institute for Strategic Studies and the Women and Public Policy Program, all at Harvard University. She was also a fellow at the Stanford Center for Advanced Studies in the Behavioral Sciences. She is the author of three books, a co-editor of two additional volumes, and author of over ninety academic articles across a wide variety of disciplines encompassing topics such as experimentation, emotion and decision making, and the biological and genetic bases of political behavior.

Harald Müller is Executive Director of the Peace Research Institute Frankfurt (PRIF), which is Member of the Leibniz Association. He also is Professor of International Relations at the Goethe University Frankfurt. As a visiting professor he regularly teaches at the Johns Hopkins University Center for International Relations, Bologna, Italy. From 1999 to 2005 he was member of the Advisory Board on Disarmament Matters of the UN Secretary General, chairing the Board in 2004. Between 2004 and 2005 he was appointed member of the Expert Group on Multilateral Fuel Arrangements of the International Atomic Energy Agency.
From 1999 on, he has been co-chairing the Working Group on Peace and Conflict at the German Foreign Office’s Planning Staff.

Since 2007 he is member of the Board of Directors of the Frankfurt University’s Cluster of Excellence called “The Formation of Normative Orders” and since 2010 Vice-President of the EU Consortium for Non-proliferation and Disarmament. His research focuses on disarmament issues, theories of democratic peace, present great power relations and transatlantic relations. His most recent book is *Norm Dynamics in Multilateral Arms Control: Interests, Conflicts, and*

Justice (ed. with Carmen Wunderlich).

Alfred Nordmann After receiving his Ph.D. in Hamburg (1986) and serving on the faculty of the Philosophy Department at the University of South Carolina (1988-2002), Alfred Nordmann became Professor of Philosophy and History of Science and of Technoscience at Darmstadt Technical University. He remains affiliated with the University of South Carolina.

Regarding the history and philosophy of science, his historical interests concern conceptions of scientific objectivity in the formation of fields of knowledge such as theories of electricity and chemistry in the 18th century, mechanics, evolutionary biology, and sociology in the 19th century, nursing science and nanoscale research in the 20th century. His epistemological interests concern the philosophies of Kant, Peirce and Wittgenstein and their influence on 20th century philosophy of science.

Regarding the history and philosophy of technoscience, Nordmann has been studying the changing relationship of science, technology, nature and society. He considered philosophical and societal dimensions of nanoscience and converging technologies, also of synthetic biology. Where the philosophy of science investigates the relation of theory and reality, the philosophy of technoscience seeks to explicate the relation of making and knowing – incorporating insights from the philosophy of technology into an understanding of the research process.

With Davis Baird he initiated the first NSF-sponsored research team on philosophical and societal dimensions of nanotechnology. He served as rapporteur for the EU expert group Converging Technologies – Shaping the Future of European Societies (2004). With Martin Carrier he coordinated In 2006/2007 the research group /Science in the Context of Application/ at the ZiF of Bielefeld University. In cooperation with Astrid Schwarz and Bernadette Bensaude-Vincent he conducted from 2010 to 2013 the German-French research project /Genesis and Ontology of Technoscientific Objects/. Since 2013 is editor of the book series /History and Philosophy of Technoscience/.

George Perkovich is vice president for studies at the Carnegie Endowment for International Peace. His research focuses on nuclear strategy and nonproliferation, with a concentration on South Asia, Iran, and the problem of justice in the international political economy.

Perkovich is author of the award-winning book *India's Nuclear Bomb* (University of California Press, 2001) and co-author of the Adelphi Paper "Abolishing Nuclear Weapons," published in September 2008 by the International Institute for Strategic Studies. This paper is the basis of the book *Abolishing Nuclear Weapons: A Debate*, which includes seventeen critiques by thirteen eminent international commentators. He also co-wrote a major Carnegie report entitled "Universal Compliance: A Strategy for Nuclear Security," a blueprint for rethinking the international nuclear-nonproliferation regime. The report offers a fresh approach to dealing with states, terrorists, nuclear weapons, and fissile materials to ensure global safety and security.

He served as a speechwriter and foreign policy adviser to Senator Joe Biden from 1989 to 1990.

Perkovich is an adviser to the International Commission on Nuclear Nonproliferation and Disarmament and a member of the Council on Foreign Relations' task force on U.S. nuclear policy.

Joshua Pollack is a consultant to the US government. At the time of the workshop, he was a senior analyst at SAIC. He has conducted studies in several areas, including arms control, verification technologies, proliferation, deterrence, intelligence, homeland security, counterterrorism, and regional security affairs. He is a regular contributor to the *Bulletin of the Atomic Scientists* and the prominent blog Arms Control Wonk, focusing primarily on current challenges to the nuclear nonproliferation regime. He also has written recently about issues surrounding emerging non-nuclear strategic forces, including conventional prompt global strike weapons and strategic missile defenses. He is a graduate of Vassar College and the University of Maryland, where he attended the Maryland School of Public Policy.

Miles Pomper is a Senior Research Associate in the Washington DC office of CNS. His work focuses on nuclear energy, nuclear nonproliferation, nuclear security, and nuclear arms control. He holds a master's degree in international affairs from Columbia University and a master's degree in journalism from Northwestern University. Before joining CNS he served as Editor-in-Chief of *Arms Control Today* from 2003-2009. Previously, he was the lead foreign policy reporter for CQ Weekly and Legi-Slate News Service, where he covered the full range of national security issues before Congress, and a Foreign Service Officer with the US Information Agency.

Nikolai N. Sokov is a senior fellow at the Center for Nonproliferation Studies. He has a Ph.D. from the University of Michigan (1996) and a Candidate of Historical Sciences degree (the Soviet equivalent of a Ph.D) from the Institute of World Economy and International Relations (1986). He graduated from Moscow State University in 1981.

Beginning in 1981, Nikolai worked at the Institute of US and Canadian Studies and the Institute of World Economy and International Relations in Moscow. From 1987-92 he worked at the Ministry for Foreign Affairs of the Soviet Union and later Russia, and participated in START I and START II negotiations as well as in a number of summit and ministerial meetings. He has been invited to speak at numerous conferences and provide briefings to many organizations and governmental agencies.

Nikolai is the author or co-author of several monographs, including:

- *Delegitimizing Nuclear Weapons: Examining the Validity of Nuclear Deterrence*(2010)
- *Reducing and Regulating Tactical (Nonstrategic) Nuclear Weapons in Europe*(2009)
- *Engaging China and Russia on Nuclear Disarmament* (2009)
- *Evolution of Nuclear Strategy in U.S. and Russia and its Implications for Arms Control* (2003)
- *Tactical Nuclear Weapons: Options for Control* (2000)
- *Russian Strategic Modernization: Past and Future* (2000)

- *Russian Policy Toward the Baltics: What the West Can Expect and What It Could Do* (1999)

He is also co-author and co-editor of the first Russian-language college-level textbook on nuclear nonproliferation (*Yadernoe Nerasprostranenie*, Vol.I-II, PIR Center, 1st edition 2000, 2nd edition 2002).

Etel Solingen is the Thomas T. and Elizabeth C. Tierney Chair in Peace and Conflict Studies and recent Chancellor's Professor at the University of California Irvine. She was president of the International Studies Association, Chair of the Steering Committee of the University of California's Institute on Global Conflict and Cooperation, president of ISA's International Political Economy Section and of the American Political Science Association's International History and Politics Section, and member of the APSA's presidential Taskforce on U.S. Standing in World Affairs. Her book *Nuclear Logics: Contrasting Paths in East Asia and the Middle East* (Princeton 2007) received the APSA's Woodrow Wilson award for the best book published in 2008 in political science and the APSA's Robert Jervis and Paul Schroeder Award for best book on International History and Politics. She received a MacArthur Foundation Research and Writing Award on Peace and International Cooperation, Social Science Research Council-Mac Arthur Foundation Fellowship on Peace and Security in a Changing World, Social Science Research Council/Japan Foundation Abe Fellowship, and Center for Global Partnership/Japan Foundation fellowship, APSA Excellence in Mentorship Award, and Distinguished Teaching Award from UCI's Academic Senate, among others. She also authored *Regional Orders at Century's Dawn: Global and Domestic Influences on Grand Strategy* (Princeton 1998), *Industrial Policy, Technology, and International Bargaining* (Stanford U. Press 1996), and *Comparative Regionalism* (Routledge 2014); and edited *Sanctions, Statecraft, and Nuclear Proliferation* (Cambridge 2012), *Scientists and the State* (U. Michigan 1994) and co-edited the special presidential issue of *International Studies Review* (2014). Her articles appeared in *International Security, American Political Science Review, International Organization, International Studies Quarterly, Comparative Politics, Journal of Peace Research, Journal of Theoretical Politics, Global Governance, Review of International Studies, Journal of Democracy, International Relations of the Asia-Pacific, and New Political Economy*, among others. She serves in various editorial boards and was Review Essay Editor for International Organization.

Janice Gross Stein is the Belzberg Professor of Conflict Management in the Department of Political Science and the Director of the Munk School for Global Affairs at the University of Toronto. She is a Fellow of the Royal Society of Canada and an Honorary Foreign Member of the American Academy of Arts and Sciences.. She is the co-author, with Eugene Lang, of the prize-winning *The Unexpected War: Canada in Kandahar*. Her most recent book is *Diplomacy in the Digital Age*. She was the Massey Lecturer in 2001 and a Trudeau Fellow. She was awarded the Molson Prize by the Canada Council for an outstanding contribution by a social scientist to public debate. She has received an Honorary Doctorate of Laws from the University of Alberta, the University of Cape Breton, McMaster University, and the Hebrew University of Jerusalem. She is a member of the Order of Canada and the Order of Ontario.

Nicholas Wright is an Associate in the Nuclear Policy Program, Carnegie Endowment for International Peace, Washington DC. He applies insights from neuroscience and psychology to decision-making in international confrontations. Prior to joining Carnegie, he examined decision-making using functional brain imaging at University College London and in the Department of Government at the London School of Economics. He has published academically (e.g.*Proceedings of the Royal Society*), in general publications such as *the Atlantic*, and with the Joint Staff at the Pentagon (see www.nicholasdwright.com/publications). He has briefed multiple times at the Pentagon, at the UK MoD and elsewhere, and appeared on the BBC and CNN.

Zachary Zwald was, at the time of the workshop, an Assistant Professor jointly appointed in the US Air War College and the USAF Center for Unconventional Weapons Studies. In fall 2015, he moved to a faculty position at the University of Houston. He received his PhD from UC Berkeley in Political Science, with an emphasis in International Relations. Zwald was a Post-Doctoral Fellow at the Mershon Center for International Security Studies at the Ohio State University and previously taught at UC Santa Cruz. During the 2013-2014 academic year, he was the Stanton Foundation Nuclear Security Junior Faculty Fellow at MIT's Security Studies Program.

Zwald's research examines judgment- and decision-making processes on issues at the intersection of technology and international security. More specifically, he is currently engaged in a number of projects that are motivated by three questions: (1) During the military technology innovation process, what determines the efficiency with which learning occurs? (2) In a nuclear weapon state, how does a policymaker determine whether a potential doctrine or weapon program will help or harm the credibility of his/her state's deterrent? (3) How can members of the policymaking community—expert and layperson—increase the accuracy of judgments and decisions on national security policies with science & technology at their core? To date, this research has focused on theory generation (synthesizing findings in psychology, behavioral economics, organizational theory, and science and technology studies) along with plausibility tests of the models' causal logic via qualitative analysis in the US context—i.e., in-depth process tracing through archival research and interviews.