H-Diplo | Robert Jervis International Security Studies Forum

Policy Roundtable II-5

The Psychology of Nuclear Brinkmanship

4 August 2023 |PDF: https://issforum.org/to/jprII-5 | Website: rjissf.org

Editor: Diane Labrosse | Commissioning Editor: Diane Labrosse | Production Editor: Christopher Ball

Contents

Introduction by Jacques E. C. Hymans, University of Southern California

At the tail end of the Cold War, the journal *International Security* published a brilliant article by historian Marc Trachtenberg demolishing the widely held "idea that the First World War came about because statesmen were overwhelmed by military imperatives and thus 'lost control' of the situation."[1] Quite to the contrary, he wrote, "The most remarkable thing about all these claims that support the conclusion about events moving 'out of control' in 1914 is how little basis in fact they actually have."[2] Trachtenberg's argument was not that Germany or any other great power was hell-bent for war in 1914, but rather that after weeks of highly stressful negotiations during the July Crisis, German Chancellor Theobald von Bethmann Hollweg and other top leaders finally just gave up the quest for peace: "There had been no 'loss of control,' only an abdication of control. Bethmann had chosen not to act. He had decided to let events take their course—and thus to take his 'leap into the dark.'"[3]

In his article, Trachtenberg addressed contemporary international relations scholars and practitioners as well as diplomatic historians. He pointed out that the IR "theory of 'inadvertent war' is in turn rooted, to a quite extraordinary degree, in a specific interpretation of a single historical episode": the 1914 July Crisis.[4] And the theory of inadvertent war holds enormous political sway: "The idea that a great war need not be the product of deliberate decision—that it can come because statesmen 'lose control' of events—is one of the most basic and common notions in contemporary American strategic thought….Many important conclusions, about the risk of nuclear war, and thus about the political meaning of nuclear forces, rest on this fundamental idea."[5] For example, the assumption of a strong possibility of inadvertent war is the basis for the highly influential nuclear brinkmanship strategy of "the threat that leaves something to chance" that was originally developed by Thomas Schelling.[6] Trachtenberg emphasized how dangerous it is for scholars and strategists to deal in lofty abstractions about "great impersonal forces" spiraling out of control, which provide comfort to self-interested politicians seeking to evade historical responsibility for the negative consequences of their actions—whereas the anticipation of bearing historical responsibility is a major prophylactic against careless or selfish decisionmaking.[7]

The major collective rethink of American Cold War nuclear strategy that Trachtenberg yearned for did not materialize. Various individuals made vigorous efforts, but after 1990 the general tide of academic and policy discussion simply turned away from nuclear strategy to other concerns. Therefore, mid-twentieth century strategic ideas such as Schelling's "threat that leaves something to chance" lived on in syllabi, policies, and mentalities, not because they were really believed in, but simply because they were already there.

[1] Marc Trachtenberg, "The Meaning of Mobilization in 1914," *International Security* Vol. 15, No. 3 (Winter 1990/91): 120-150, here 148. This article later served as the basis for Chapter 2 of Marc Trachtenberg, *History and Strategy* (Princeton, NJ: Princeton University Press, 1991).

[2] Trachtenberg, "The Meaning of Mobilization," 150.

[3] Trachtenberg, "The Meaning of Mobilization," 143.

[4] Trachtenberg, "The Meaning of Mobilization," 120. Note that Trachtenberg's interpretation of the July Crisis and its implications for IR theory is far from uncontested. See esp. Jack S. Levy, Thomas J. Christensen, and Marc Trachtenberg, "Mobilization and Inadvertence in the July Crisis," *International Security* Vol. 16, No. 1 (Summer 1991): 189-203. For a discussion of the latest historical scholarship on the question, see Richard Ned Lebow, "What Can International Relations Theory Learn from the Origins of World War I," *International Relations* Vol. 28, No. 4 (2014): 387-410.

[5] Trachtenberg, "The Meaning of Mobilization," 120.

[6] Thomas C. Schelling, *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1960), esp. ch. 8.

[7] Trachtenberg, "The Meaning of Mobilization," 148-149.

Nearly a quarter-century after Trachtenberg's pathbreaking *International Security* article, Reid B. C. Pauly and Rose McDermott have written a new *International Security* article that picks up his torch.[8] Like Trachtenberg, Pauly and McDermott argue that the strategists' standard understanding of the political dynamics of nuclear crises is flawed, or at least incomplete. The mere fact that nuclear crises are subject to the workings of "chance" does not therefore mean that states lose their freedom of "choice."[9] Also like Trachtenberg, Pauly and McDermott emphasize that the mantle of choice in nuclear crises falls heavily on the shoulders of individual political leaders. The effectiveness of strategies of nuclear brinksmanship will depend on the character and state of mind of the individuals who are deploying them, not just on the strategies themselves. Therefore, we should closely scrutinize every tic and tendency of the specific leaders who are involved.

Pauly and McDermott come to these insights through a theoretical analysis of Schelling's approach to nuclear strategy complemented with findings from experimental psychology. In particular, they contend that Schelling's proposed mechanisms of how choice gives way to chance—accidents, loss of self-control, and loss of control over others—actually still leave room for individual human historical agency. Regarding nuclear accidents or mechanical failures, Pauly and McDermott argue, "The decider has agency in the aftermath of the risk-generating event."[10] Regarding the decisionmaker's loss of self-control, they argue that this may represent a departure from Schelling's assumption of *rational* choice, but not from choice altogether. And regarding the decisionmaker's loss of control over others, they argue that in such a situation the opportunity and burden of historical agency has simply passed on to someone else. The key point is that "decisionmakers are humans, and all humans have emotions, and those emotions can be idiosyncratic and unpredictable. Therefore, these psychological issues become part of the calculus and introduce inherent uncertainty that Schelling overlooked."[11]

The fact that Trachtenberg on the one hand and Pauly and McDermott on the other—scholars from different disciplines writing in different decades—have converged, apparently coincidentally,[12] on similar general conclusions about what is missing from our understanding of nuclear crisis politics should strengthen our confidence that their arguments have merit.[13] And today, with nuclear war clouds gathering over Ukraine, the international security field can no longer postpone this important discussion.

In this H-Diplo|RJISSF Policy Roundtable, highly accomplished IR scholars come together to discuss Pauly and McDermott's theoretical perspective, both in general and in the specific case of the ongoing war in Ukraine. Pauly and McDermott kick off the proceedings with a summary of their article. Then we are treated to essays by Marika Landau-Wells of UC Berkeley, Joshua Rovner of American University, and Janice Gross Stein of the University of Toronto. Finally, Pauly and McDermott offer a reply.

The essays provide a great buffet of ideas. It is surely unfair to summarize the entire forum in one sentence, but here goes: while all the contributors lavishly praise Pauly and McDermott for making an important

[8] Reid B. C. Pauly and Rose McDermott, "The Psychology of Nuclear Brinkmanship," *International Security* Vol. 47, No. 3 (Winter 2022/23): 9-51.

[9] Pauly and McDermott, "The Psychology of Nuclear Brinkmanship," 9.

[10] Pauly and McDermott, "The Psychology of Nuclear Brinkmanship," 24.

[11] Pauly and McDermott, "The Psychology of Nuclear Brinkmanship," 27.

[12] Pauly and McDermott's article omits to cite Trachtenberg's article. They do cite a different article by Trachtenberg on the Cuban missile crisis, which leads them to call him a "nuclear coercionist" (Pauly and McDermott, "The Psychology of Nuclear Brinkmanship," 17). That label is an oversimplification.

[13] Note that Trachtenberg is certainly not the only other previous scholar who criticized Schelling for overlooking the importance of individual-level psychological factors for the course and outcome of nuclear crises. For instance, here is Ole Holsti, writing in 1972: "Thomas Schelling and others have demonstrated that the game of 'chicken' played by teenaged hot-rodders has its counterparts in many other areas of life, including international politics. But 'chicken' is a dangerous game under the best of circumstances; for reasons that have been emphasized throughout this book, it is doubly perilous when played in high-stress situations." Ole R. Holsti, *Crisis, Escalation, War* (Montreal and London: McGill-Queen's University Press, 1972), 227.

theoretical advance beyond Schelling, they are less than fully satisfied by Pauly and McDermott's continued reliance on Schelling's basic model of nuclear crisis politics. However much one may admire Schelling's genius, is the best way forward for the security studies field really to construct one more layer on top of a theoretical structure from the 1950s and 60s? An analogy: how many people think that the best way forward for the automobile industry today is to make some incremental improvements to the '57 Chevy? This is the roundtable's most profound critique of the Pauly-McDermott research agenda, so I focus attention on it—but I reiterate that it is only one of many threads in the very rich and wide-ranging discussion.

The contributors make their common critique of Pauly and McDermott's continuing attachment to Schelling in a variety of different contexts. Landau-Wells argues that "Pauly and McDermott engage with Schelling on his preferred terms, modifying his 'tethered mountain-climbers' scenario topographically to capture the implications of their arguments for his model. The mental morphing required does not actually pay off, however, because Schelling's initial psychological model is so flimsy. Rather, the authors miss the opportunity to lay out the decision-related psychology that gives rise to, and maintains, brinkmanship situations in the real world."

Rovner similarly wonders why Pauly and McDermott seem so determined to hang on to Schelling's core idea of using brinkmanship to leverage nuclear weapons for political gain. Instead, he writes, "The last possibility, then, is that brinkmanship threats are inconsequential. Threats leaving something to chance do not generate leverage nor do they cause escalation. They are not useful coercive instruments, but they are not terribly dangerous either."

Stein also relativizes Pauly and McDermott's focus on Schellingesque nuclear threats as the means to gain leverage in a tense crisis situation. She stresses that the US has neither responded in kind to Russian President Vladimir Putin's nuclear saber-rattling, nor has it caved in to it. Instead, it has warned "that the use of a tactical nuclear weapon would be met with a severe *conventional* response" (italics added). Stein observes, "By making explicit what the United States would do if Russia were to use an unconventional weapon, policy makers in Washington hoped both to influence Russia's decisionmaking and put a break in an escalatory spiral if Russia nevertheless were to go ahead. This is an example of the finely differentiated and sequenced strategies of escalation management that will become much more important as great power competition among nuclear powers returns with a vengeance."

Pauly and McDermott respond to these constructive criticisms with equally constructive thinking of their own. Indeed, their reply is a model of scholarly modesty. Regarding Landau-Wells' abovementioned critique that they are trying too hard to work with Schelling's grab bag of cute but questionable analogies, Pauly and McDermott write, "In critiquing Schelling but building on his analogies, we may have indeed tied our hands by debating on his terms. Perhaps that approach is inherently limited when it comes to psychology. Still, Schelling's abstraction may also be a source of his staying power in the field and his influence on policy." Regarding Rovner's accounting of the many reasons why "threats that leave something to chance" are likely to be ineffective, they write, "We agree in principle but think that the net effect is as yet unclear." Finally, regarding Stein's critique that states can often gain more leverage over crisis situations with demonstrations of nuclear restraint than nuclear threats, they write, "Schelling did not advise signaling your disciplined pragmatism in nuclear crisis bargaining….And yet it seems to have led to some stability in Ukraine about the expectations of either side—no direct NATO intervention, no Russian attacks on NATO targets, a hurting equilibrium of Western military assistance and Russian willingness to bear high costs. Perhaps Russia is not doing a good job of leaving things to chance (certainly a possible answer), or maybe pragmatists are better at playing (or avoiding) the game of chicken than we think. Or perhaps luck has played a bigger role than anyone is comfortable acknowledging."

In conclusion, this forum processes a highly significant contemporary issue with a vigorous scholarly discussion that succeeds in generating much more light than heat.

Contributors:

Reid Pauly is an Assistant Professor of Political Science at Brown University and the Dean's Assistant Professor of Nuclear Security and Policy at the Watson Institute for International and Public Affairs.

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.

Jacques E.C. Hymans is an Associate Professor of Political Science and International Relations at the University of Southern California.

Marika Landau-Wells is an Assistant Professor in the Travers Department of Political Science at the University of California, Berkeley. She received her PhD in Political Science from the Massachusetts Institute of Technology where she was also a Postdoctoral Research Fellow in the Department of Brain and Cognitive Sciences.

Joshua Rovner is an Associate Professor in the School of International Service at American University.

Janice Gross Stein is Belzberg Professor of Conflict Management and Founding Director of the Munk School of Global Affairs and Public Policy at the University of Toronto.

"The Psychology of Nuclear Brinkmanship"
by Reid B. C. Pauly and Rose McDermott, Brown University

As the shadow of nuclear war grows darker once more, scholars are busy updating and questioning inherited theory from the Cold War bipolar nuclear rivalry, wondering what to keep or jettison as policymakers grapple with how to prevent nuclear escalation on the Korean Peninsula, in South Asia, over Taiwan, or in Ukraine.[1] This piece is one such attempt to modernize an idea that is core to the nuclear security studies canon: the "threat that leaves something to chance" in nuclear brinkmanship.

The theory of nuclear brinkmanship explains how states can manipulate the risk of disaster to compete under the shadow of a nuclear war that threatens their mutual survival. A leader may have to be irrational to start a strategic nuclear war, but Thomas Schelling recognized that they do not have to be irrational to *risk* one. Schelling offered threats that leave something to chance as a solution to this problem of agency in coercion. Nuclear threats can produce bargaining leverage if made credibly. But since a rational decisionmaker should never choose mutually assured destruction (MAD), canonical theories of brinkmanship remove leaders from the process of escalation.[2] Scholars subsequently studied the causes of nuclear risk, such as bureaucratic and organizational pathologies,[3] accidents,[4] or inadvertent escalation,[5] while minimizing the role of the individual decisionmaker.

In the Winter 2022/2023 issue of *International Security,* we brought the study of psychology and emotion to bear on this puzzle of how "chance" can generate coercive leverage in nuclear crises while leaders still retain agency over the "choice" to escalate.[6] We argue that chance can coexist with choice. In a MAD world, it is indeed irrational to accept more risk of nuclear escalation than the stakes are worth to you; but emotional or psychological biases explain how brinkmanship can operate even when leaders retain control over their nuclear forces. We identify three mechanisms of risk in nuclear brinkmanship—accidents, self-control, and control of others. Each can be influenced by universal aspects of human decisionmaking architecture, including emotional responses. The ability to recognize and control these responses differs profoundly across individuals, especially during times of stress or in the face of loss.

Bringing Humans Back In

Schelling's best brinkmanship analogy involved two mountaineers chained together at the edge of a cliff.[7] One climber cannot credibly threaten to push the other off the cliff, because that would doom them both; just as if one state threatens nuclear war, unacceptable retaliation might ensue. But each climber could still generate leverage over the other by taking risks. One climber could take a step closer to the edge, stand on one foot, even dance around on the smooth slope or loose gravel below their feet. In this way, the climbers

---

[1] Research and arguments included in this essay are drawn from a larger article, Reid B. C. Pauly and Rose McDermott, "The Psychology of Nuclear Brinkmanship," in *International Security*, Volume 47, Issue 3 (2023): 9-51. https://doi.org/10.1162/isec_a_00451. We thank the editors of IS for granting us permission to publish a condensed version of the essay on H-Diplo|RJISSF.

[2] Thomas C. Schelling, *The Strategy of Conflict* (Cambridge: Harvard University Press, 1960); Schelling, *Arms and Influence* (New Haven: Yale University Press, 1966).

[3] Graham T. Allison, "Conceptual Models and the Cuban Missile Crisis," *American Political Science Review* 63, 3 (1969): 689–718; https://doi.org/10.2307/1954423.

[4] Scott D. Sagan, *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons* (Princeton: Princeton University Press, 1993).

[5] Barry R. Posen, *Inadvertent Escalation: Conventional War and Nuclear Risks* (Ithaca: Cornell University Press, 1991).

[6] Pauly and McDermott, "The Psychology of Nuclear Brinkmanship."

[7] Schelling, *Arms and Influence*, 99.

are engaged in brinkmanship—a competition in risk-taking or manipulating the chance of mutual disaster. As both climbers move further down the path toward oblivion, a rationalist theory predicts that whichever actor is less resolved will concede when the more resolved actor generates sufficient risk.

Schelling's theory was simple and powerful. Yet it made an assumption that, for brinkmanship to work, humans must lose agency over the choice to engage in nuclear war. If strategic nuclear war is irrational but decisionmakers are assumed to be rational, then scholars must discount the decisionmaking actor and assign to events a momentum of their own. Escalation in the aftermath of an accident or a false warning is unnecessarily presumed to have a quality of automaticity. And subsequent theorizing about the coercive utility of nuclear weapons mostly baked in this assumption.[8]

We argue for bringing the humans back in. Even in the aftermath of accidents, leaders still retain a choice over how to respond. Advances in the study of human psychology since Schelling's writing have made this missing perspective even more noticeable. We locate the real sources of risk when agency remains in the architecture of human psychological decisionmaking.

In table 1 below, we summarize three distinct mechanisms by which chance can be converted into leverage in nuclear crises: (1) accidents; (2) a lack of self-control; and (3) a loss of control over others. The categories may bleed together in the real world, and they are not mutually exclusive, but they are analytically distinct. Each may apply to any actor on any side of a crisis.[9] Our original contribution lies especially in the right-most column of table 1, and in how the mechanisms interact. Mechanism 1—the chance of an accident—is generally understood in the literature to be a source of risk in threats that leave something to chance. But accidents do not remove the need for either side to decide whether to use nuclear weapons. Rather, they generate risk because psychological and emotional factors, which are discussed in mechanisms 2 and 3, make the outcomes of those decisions more uncertain than a cost-benefit calculus would suggest.

To understand brinkmanship and the choices that surround threats that leave something to chance, we argue that it is critical to delineate the relationship between choice and chance. Schelling wrote of brinkmanship: "Where does the uncertain element in the decision come from? It must come from somewhere outside of the threatener's control. Whether we call it 'chance,' accident, third-party influence, imperfection in the machinery of decision, or just processes that we do not entirely understand, it is an ingredient in the situation that neither we nor the party we threaten can entirely control."[10] The choice over whether and when a nuclear war will break out is understood to be a hindrance to effective brinkmanship, since it cannot be made credible. But chance and choice are not as contradictory as they appear, or as Schelling thought. A leader does not have to lose control to threaten to make an irrational choice in crisis. Psychological and emotional factors make the outcomes of their decisions more uncertain than a cost-benefit calculus would suggest.

---

8. Robert Powell, "The Theoretical Foundations of Strategic Nuclear Deterrence," *Political Science Quarterly* 100, 1 (Spring 1985): 75–96; https://www.jstor.org/stable/2150861; Todd S. Sechser and Matthew Fuhrmann, *Nuclear Weapons and Coercive Diplomacy* (Cambridge: Cambridge University Press, 2017); Matthew Kroenig, *The Logic of American Nuclear Strategy: Why Strategic Superiority Matters* (New York: Oxford University Press, 2018). On organizations and accidents, see Bruce G. Blair, *The Logic of Accidental Nuclear War* (Washington, DC: Brookings Institution, 1993); Paul J. Bracken, *The Command and Control of Nuclear Forces* (New Haven: Yale University Press, 1983); Peter D. Feaver, *Guarding the Guardians: Civilian Control of Nuclear Weapons in the United States* (Ithaca: Cornell University Press, 1992); Sagan, *The Limits of Safety*. Robert Jervis's work stands as an important exception, particularly Jervis, *Perception and Misperception in International Politics* (Princeton: Princeton University Press, 1976).

9. For simplicity, each mechanism is conceived of—in this article and in Schelling's presentation—as the behavior of just one belligerent in a crisis, even though, in reality, both sides may engage in brinkmanship simultaneously.

10. Schelling, *The Strategy of Conflict,* 188. See also Daniel Kahneman, Olivier Sibony, and Cass R. Sunstein, *Noise: A Flaw in Human Judgment* (New York: Little, Brown Spark, 2021).

Decisionmakers are humans, and all humans have emotions, and those emotions can be both idiosyncratic and unpredictable.

Table 1. Disentangling Chance and Choice in Nuclear Brinkmanship

Mechanisms	Concept mentioned in Schelling's *Strategy of Conflict*	Relationship to choice
accidents	accident mechanical failure	The decider has agency in the aftermath of the risk-generating event.
self-control	panic madness	The decider has agency but cannot control themselves.
	false alarm misapprehension of enemy intentions	The decider has agency but acts on misperception.
control of others	pre-delegation mischief	The decider does not have agency, a third party or unauthorized decider does.
	limited war as a generator of risk	The decider retains a choice, but the adversary also has agency over the decision for nuclear war.

Mechanism 1: Accidents

The first mechanism—accidents—is the category most familiar to the brinkmanship literature. Sometimes, escalation occurs despite the best intentions of the belligerents after an accidental launch, a mechanical failure that causes two planes to collide in the air, or misaimed warning shot. The longer a crisis goes on, the more likely an unexpected event could happen. Schelling conceived of this possibility as a kind of iterated Russian roulette. However unlikely they may seem, as Sagan writes, "things that have never happened before happen all the time."[11] Thus, all crises have some baseline level of risk.

Leaders can manipulate the magnitude of these accidental risks in crises by mobilizing more military forces, placing them on ever higher alert, forward deploying them, or ordering them to operate near the enemy. A leader who approves "buzzing" tactics by pilots, for instance, is delegating to agents the risk of accidents occurring in the field. A fighter jet buzzing another aircraft, as when Russian pilots over the Black Sea downed an American drone in March 2023,[12] was for Schelling "the purest real-life example I can think of in international affairs" of a threat that leaves something to chance.[13]

[11] Sagan, *The Limits of Safety*.

[12] Eric Schmitt, "Russian Warplane Hits American Drone over Black Sea, U.S. Says," *New York Times*, March 14, 2023; https://www.nytimes.com/2023/03/14/us/politics/russia-us-drone-black-sea.html.

[13] Schelling, *Arms and Influence,* 104.

Leaders can also manipulate the magnitude of these risks through decisions about how much risk to bake into their nuclear postures.[14] In general, leaders who delegate more authority to use nuclear weapons, have more mobile and numerous nuclear forces, and do not institute negative controls, such as permissive action links, raise the risk of accidents in a crisis. For instance, Russia's intention to forward deploy nuclear weapons to the territory of Belarus increases however slightly the attendant risk of accidents.[15]

Bureaucratic and organizational structures also affect the risk of accidents. Organization theory on "normal accidents" suggests that accidents and mechanical failures are sure to happen eventually in complex, tightly coupled systems like the military organizations that manage nuclear weapons.[16] The Cuban Missile Crisis is also replete with examples of how military standard operating procedures raised the level of risk beyond that which leaders intended.[17]

All these factors can augment the risk of accidents or close calls. But they do not remove choice entirely. And, most importantly, choice does not have to be removed for brinkmanship to function. Accidents do not dispatch the problem of choice. For a nuclear war to begin after an accident or mechanical failure causes a crisis to escalate, one leader must decide to be the first to launch deliberately. Accidents themselves do not eliminate leaders, their credibility, or their deficiencies from the equation.

In his original climbing analogy, accidents are what made Schelling describe the brink as a curved slope. But it is more accurate to think of these chance events as ledges on the cliff. Falling all the way down the abyss at once is unlikely. If the climber falls, she might find herself on a lower ledge, perhaps injured but still with the acumen to decide whether to plummet another ledge down or try to climb back up to the top. The abyss never ceases to exist, so it remains irrational at any ledge to simply leap into the bottomless chasm.

In the aftermath of accidents, choices remain. Accidents are insufficient as a sole mechanism of brinkmanship. They beg another solution to the puzzle of agency or choice and must work in tandem with the other mechanisms.

Mechanism 2: Self-Control

The second mechanism is that of losing self-control. The risk of escalation comes from within the system, not outside it. Here we see two subcategories. First, a decider may choose to escalate out of panic or madness. We explore the possibility of these irrational choices below. Second, Schelling cited false alarms and the misapprehension of enemy intentions (or a correct apprehension of the enemy's misapprehension of its adversary) as sources of biased decisionmaking in crisis. In such circumstances, the decider still chooses war, but this decision is based on incorrect information or misperception about the enemy's intentions and actions.

Most leaders control their emotions most of the time, and some people certainly have greater emotional awareness and emotional regulation.[18] But many people are not able to completely or consistently control

[14] Peter D. Feaver and David Arceneaux, "The Fulcrum of Fragility: Command and Control in Regional Nuclear Powers," in Scott D. Sagan and Vipin Narang, eds., *The Fragile Balance of Terror: Deterrence in the New Nuclear Age* (Ithaca: Cornell University Press, 2023); Narang, *Nuclear Strategy in the Modern Era* (Princeton: Princeton University Press, 2014).

[15] Andrew Osbourn, "Belarus Says it Will Host Russian Nuclear Weapons to Counter NATO," *Reuters,* March 28, 2023.

[16] Charles Perrow, *Normal Accidents: Living with High-Risk Technologies* (Princeton: Princeton University Press, 1999); Sagan, *The Limits of Safety.*

[17] Allison, "Conceptual Models and the Cuban Missile Crisis."

[18] James J. Gross, "Emotion Regulation: Current Status and Future Prospects," *Psychological Inquiry* 26, 1 (2015): 1–26, https://doi.org/10.1080/1047840X.2014.940781.

their feelings. In the language of Daniel Kahneman's dual process theory of cognition, because of time pressure in times of crisis leaders are likely to disproportionately face decisions that require fast rather than slow thinking processes. Fast processes rely disproportionately on emotion, although both types are prone to specific kinds of errors and biases.[19] In times of stress and crisis, one of the most destabilizing of these pressures is the desire for vengeance in the face of an attack.[20] The desire for revenge, as universal as it may be, is not the only emotion that might decisively affect a leader's decisionmaking calculus. Pride, shame, envy, status-seeking, or a desire to defend one's own or one's family's honor might ignite aggression as well.

Anyone who revels in other's misfortune or who feels the hormonal rush that follows winning a physical fight or intense competition knows that inflicting suffering on others can generate endogenous pleasure, particularly if those others have caused them harm. Almost all violence is perceived by the perpetrators to be virtuous, and individuals thus feel quite justified in attacking others whom they consider bad and blameworthy.[21] The German word for this feeling of pleasure at another person's harm, *schadenfreude*, is well intuited by most. What is more, the physical rush that accompanies a victorious fight—even if verbal and not physical—is automatic, effortless, immediate, and endogenous.[22] The feeling cannot be bought, nor can it be tamped down easily. There are precious few other ways to produce it. Victory feels great; defeat feels awful. Those who win like the feeling and want to fight again; those who lose are less inclined to try again.[23]

A rationalist approach to brinkmanship also fails to account for the perceived emotional and psychological, as well as material, values of retaliation, and the strong and instinctual drive for revenge against those who cause us harm. The psychology of vengeance underlies the stability of nuclear deterrence far more than a rational theory of the nuclear revolution appreciates.[24] In the nuclear era, these psychological patterns bear on the question of whether limited nuclear war can be terminated or whether, as we suggest, leaders are unlikely to succeed in escalating to deescalate and instead continue the use of nuclear weapons once instigated.

A second way in which the theory of brinkmanship must better account for human psychology is in its observation of the value of irrationality. Schelling wrote, "Another paradox of deterrence is that it does not always help to be, or to be believed to be, fully rational, cool-headed, and in control of oneself or of one's country."[25] Additional work in psychology validates the notion that people weigh unpredictable threats more heavily in making decisions. For example, Paul Slovic finds that people are most afraid of so-called "dread"

[19] Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus, and Giroux, 2011).

[20] Rose McDermott, Anthony C. Lopez, and Peter K. Hatemi, "'Blunt Not the Heart, Enrage It': The Psychology of Revenge and Deterrence," *Texas National Security Review,* Vol. 1, No. 1 (2017).

[21] Slovic et al. "Virtuous Violence from the War Room to Death Row."

[22] Pranjal H. Mehta, Amanda C. Jones, and Robert A. Josephs, "The Social Endocrinology of Dominance: Basal Testosterone Predicts Cortisol Changes and Behavior Following Victory and Defeat," *Journal of Personality and Social Psychology* 94, 6 (2008): 1078; https://doi.org/10.1037/0022-3514.94.6.1078.

[23] Brian A. Gladue, Michael Boechler, and Kevin D. McCaul, "Hormonal Response to Competition in Human Males," *Aggressive Behavior* 15, 6 (1989): 409–422; https://psycnet.apa.org/doi/10.1002/1098-2337(1989)15:6%3C409::AID-AB2480150602%3E3.0.CO;2-P; Michael Elias, "Serum Cortisol, Testosterone, and Testosterone-Binding Globulin Responses to Competitive Fighting in Human Males," *Aggressive Behavior* 7, 3 (1981): 215–224; https://doi.org/10.1002/1098-2337(1981)7:3%3C215::AID-AB2480070305%3E3.0.CO;2-M; Pranjal H. Mehta and Robert A. Josephs, "Testosterone and Cortisol Jointly Regulate Dominance: Evidence for a Dual-Hormone Hypothesis," *Hormones and Behavior* 58, 5 (2010): 898–906; https://doi.org/10.1016/j.yhbeh.2010.08.020; Pranjal H. Mehta and Robert A. Josephs, "Testosterone Change After Losing Predicts the Decision to Compete Again," *Hormones and Behavior* 50, 5 (2006): 684–692; https://doi.org/10.1016/j.yhbeh.2006.07.001.

[24] McDermott, Lopez, and Hatemi, "'Blunt Not the Heart, Enrage It'."

[25] McDermott, Lopez, and Hatemi, "'Blunt Not the Heart, Enrage It'." 37

risks, those that feel uncontrollable or that appear to be harbingers of unpredictably worse things yet to come, such as with climate change or nuclear war.[26]

Yet the effect of emotion on decisionmaking is nuanced and sometimes idiosyncratic. It does not operate in the way that a rational theory of brinkmanship would favor. Emotion is necessary to generate motivation;[27] without it, people tend not to possess the will to expend energy and act. In this way, emotions that look irrational from a modern standpoint, such as the drive for status over wealth, may serve humans well in terms of survival and reproduction but may not be rational from an economic standpoint. For example, the best way to protect oneself from an enemy may in fact be to annihilate the entire group; that assures that future trouble from that adversary will not be forthcoming. This strategy indeed appears to have been common even in the early modern era. In the age of nuclear war, such a strategy is more dangerous, and threats that leave something to chance risk widespread annihilation.

Finally, a satisfying theory of brinkmanship must account for the fact that different emotions trigger different perceptions of risk. For example, anger makes people more risk seeking, and thus more likely to downplay the nature of the risks that they confront. It also makes them much more likely to believe that they will be victorious in a conflict, even if such a belief rests on pure overconfidence. Conversely, fearful people have more pessimistic risk assessments and thus prove more risk averse in their choices and behaviors.[28] Gender differences emerge in these tendencies as well. For example, men are more prone to anger, which predicts more support for punitive political policies. In contrast, some studies find that women tend to be more fearful and are thus much more likely to support rehabilitative policies.[29]

The processes by which leaders are selected and self-selected make them more, rather than less, immune from the kinds of biases that might affect their decisionmaking in negative ways. Traits such as narcissism, for example, are more likely to emerge in leaders than in other groups.[30] These individuals are much more likely to exaggerate their own skill and sense of control. For instance, Russian President Vladimir Putin's 2022 invasion of Ukraine, especially his doubling down after the unprecedented sanctions imposed in its wake and the underperformance of the Russian military, would hardly seem rational from a cost-benefit perspective. But his actions appear more comprehensible when explained in terms of his view of his own role in Russian history.[31]

Mechanism 3: Control of Others

The third mechanism of brinkmanship is that of a decider losing control because others make the choice to escalate without that leader's input or knowledge. Here again there are two subcategories. First, Schelling identified mischief as a means of war onset. Someone else (who is perhaps irrational but he left that unsaid) makes the choice for war instead of the leader. When the authority or ability to escalate is shared among

[26] Paul Slovic, "Perception of Risk," *Science* 236, 4799 (1987): 280–285, https://doi.org/10.1126/science.3563507.

[27] Leda Cosmides and John Tooby, "Evolutionary Psychology and the Emotions," in Michael Lewis and Jeannette M. Haviland-Jones, eds., *Handbook of Emotions*, 2nd ed. (New York: Guilford, 2000): 91–115.

[28] Jennifer S. Lerner and Dacher Keltner, "Fear, Anger, and Risk," *Journal of Personality and Social Psychology* 81, 1 (2001): 146, https://psycnet.apa.org/doi/10.1037/0022-3514.81.1.146.

[29] Barbara A. Gault and John Sabini, "The Roles of Empathy, Anger, and Gender in Predicting Attitudes toward Punitive, Reparative, and Preventative Public Policies," *Cognition and Emotion* 14, 4 (2000): 495–520, https://doi.org/10.1080/026999300402772.

[30] Zoltán Fazekas and Peter K. Hatemi, "Narcissism in Political Participation," *Personality and Social Psychology Bulletin* 47, 3 (2021): 347–361, https://psycnet.apa.org/doi/10.1177/0146167220919212.

[31] Anton Troianovski, "Putin the Great? Russia's President Likens Himself to Famous Czar," *New York Times,* June 9, 2022; Andrew Roth, "Putin compares himself to Peter the Great in quest to take back Russian lands," *The Guardian*, June 10, 2022.

multiple actors, these risks increase. A leader may or may not be aware of having created these risks. For instance, nuclear command-and-control arrangements that pre-delegate to military commanders the authority to use nuclear weapons exacerbate these risks by introducing more vectors for others' individual choices. And those individuals are just as susceptible to the irrational decisionmaking described in the prior section.

Some might conceive of this risk as a principal-agent problem, whereby the principal loses control of an agent. In the military domain it has been more specifically theorized as a lack of civilian control over military forces.[32] An overzealous local commander with the ability to use nuclear weapons fits the bill—a *Dr. Strangelove* scenario. But short of nuclear war, military commanders have routinely stretched the purview of their offensive missions beyond the authorization of civilian leaders.[33] Even within military organizations, significant commands have been disobeyed or countermanded by individual leaders. Air Marshal Sir Charles Portal gave clear instructions to Sir Arthur "Bomber" Harris in September 1944 to concentrate Royal Air Force bombers on oil and transportation targets in Nazi Germany. Harris refused and continued with his effort to turn German cities into rubble.[34]

The second subcategory is limited war as a generator of risk. Although it stretches our scope conditions,[35] we place this mechanism into the category of controlling others because deciders may be fully in control of their own forces, while the element of chance lies in the possibility that they could unknowingly cross one of the adversary's red lines; and deciders may not even give sufficient weight to the probability that their actions may do so. The choice to escalate further in response now rests with the enemy.

Schelling considered limited nuclear war as brinkmanship in this way.[36] Deliberately crossing the nuclear threshold demonstrates resolve, even if not enough resolve to jump off the cliff completely. The initial use of limited nuclear force is conceived of by the escalator as rational. In terms of the mountaineering analogy, one climber intentionally jumps down to a lower ledge (as opposed to falling by accident), dragging the other climber with them. But just because one climber jumped to a lower ledge and dragged the other down does not necessarily make inevitable the second jump into the abyss to doom them both. The first use of a nuclear weapon does not automatically lead to strategic nuclear exchange; it is merely an action that is deliberately risky.

Going Forward

By distinguishing between "chance" and "choice," our work aims to elevate the long-hidden psychological and emotional elements that remain central to a more comprehensive understanding of nuclear brinkmanship. Three mechanisms for how chance becomes leverage in crisis—accidents, self-control, and control of others—do not have to wrest choice away from leaders in crises to generate leverage. Psychological and emotional variables are sources of risk and uncertainty, even when leaders retain the option to use nuclear weapons. The point of decision does not have to be eliminated for the risk of catastrophic

[32] Military organizations prefer offensive operations, sometimes without the approval of civilians. See Posen, *Inadvertent Escalation*; Stephen Van Evera, "The Cult of the Offensive and the Origins of the First World War," *International Security* 9, 1 (Summer 1984): 58–107; https://doi.org/10.2307/2538636; Jack Snyder, "Civil-Military Relations and the Cult of the Offensive, 1914 and 1984," *International Security* 9, 1 (Summer 1984): 108–46; https://doi.org/10.2307/2538637.

[33] Posen, *Inadvertent Escalation*, 18–19.

[34] Charles Webster and Noble Frankland, *The Strategic Air Offensive Against Germany 1939–1945,* vol. 3, *Victory, Part 5* (London: Her Majesty's Stationary Office, 1961), 58, 67. We thank Williamson Murray for this example.

[35] We find it within our scope to explain the decisions of either side to engage in limited war, rather than the consequences of their choices in combination.

[36] Schelling, *Arms and Influence*, 108–109; Schelling, *Strategy of Conflict*, 193. See also Robert Powell, "Nuclear Brinkmanship, Limited War, and Military Power," *International Organization* 69, 3 (Summer 2015): 589-626; https://www.jstor.org/stable/24758314.

destruction to remain. The structure of the situation imposes uncertainty and instability and inevitable human emotional responses to threats enhance and strengthen these effects. In short, crises are even more unstable than the traditional theory of brinkmanship posits.

Our approach is useful for policy as well. A focus on psychological models of decisionmaking suggests interventions that constrain decisionmakers from having unilateral authority over nuclear weapon release. It underscores the importance of eliminating launch-on-warning procedures that prioritize speed over sounds decisionmaking, expanding two-person rules to every point in a decision chain, implementing lengthier protocols to check and recheck orders, and maintaining hotlines for crisis communications between adversaries. Reducing haste is important but the augmented decisionmaking time must also be used to disrupt irrational processes.

Finally, as policymakers manage the escalation potential of the war in Ukraine, they should keep in mind the fragility of leadership psychology. Putin seems to be a narcissist bent on irredentist conquest for imperial restoration.[37] While plenty of analysis has focused on whether Putin's red lines are bluffs,[38] our work suggests that Putin may not in fact know ahead of time where his red lines are or what he will do if they are crossed.[39] Emotions like anger and fear, desires for revenge, and biases like psychic numbing and virtuous violence could shape his decisionmaking if he feels he is backed into a corner.

This work is certainly not the last on the topic of brinkmanship. Indeed, it opens up even more interesting avenues to explore. For instance, some studies of brinkmanship conclude that threats that leave something to chance are empirically rare because in times of crisis leaders seek to maintain control. Matthew Fuhrmann and Todd Sechser go so far as to call it "the brinkmanship myth."[40] Our approach suggests looking for evidence elsewhere. Chance can be all around us, even when we try to exercise choice over its consequences.

[37] Nathan Hodge, "Restoration of Empire is the Endgame for Russia's Vladimir Putin," CNN, June 11, 2022; Ishaan Tharoor, "Putin makes his Imperial Pretensions Clear," *Washington Post,* June 13, 2022; Joseph Brugo, "Vladimir Putin, Narcissist?" *Atlantic,* April 15, 2015; Isabel van Brugen, "Putin Branded a Narcissist in Leaked FSB papers." *Newsweek,* November 26, 2022.

[38] Julian E. Barnes and David E. Sanger, "Fears of Russian Nuclear Weapons Use Have Diminished, but Could Re-emerge," *New York Times*, February 3, 2022; Joseph Cirincione, "Why Hasn't Putin Used Nuclear Weapons?" *Daily Beast*, February 9, 2023; Andrew F. Krepinevich, Jr., "Is Putin a Rational Actor?" *Foreign Affairs,* November 22, 2022.

[39] See also Janice Gross Stein, "Escalation Management in Ukraine: 'Learning by Doing' in Response to the 'Threat that Leaves Something to Chance,'" *Texas National Security Review*, Vol. 6, Iss. 3 (Summer 2023), https://tnsr.org/2023/06/escalation-management-in-ukraine-learning-by-doing-in-response-to-the-threat-that-leaves-something-to-chance/.

[40] Sechser and Fuhrmann, "The Madman Myth."

"Expanding the Realm of Plausibility"
by Marika Landau-Wells, University of California, Berkeley

In "The Psychology of Nuclear Brinkmanship," Reid B.C. Pauly and Rose McDermott conduct a vital investigation into the logic of Thomas Schelling's "threats that leave something to chance," which are, for Schelling, the way to reconcile a world of rational actors with coercive behaviors that risk mutual catastrophe.[1] When such behavior—"brinkmanship"[2]—extends to the nuclear realm, we should all be concerned. As the authors note, there are a number of places in the world where these brinkmanship dynamics are relevant today (6).

For those monitoring such situations, Pauly and McDermott provide both good and bad news. The good news is that, while under-theorized by Schelling, "pushing the button" is still a *choice* and it is always possible for a person to choose not to do something. The bad news is that human beings, also under-theorized by Schelling, may have more reasons to make that choice in the affirmative than previously considered. The more plausible pathways there are to "pushing the button," the more credible the threat to do so will be. In their estimation, "[C]rises are even more unstable than the traditional theory of brinkmanship posits" (13).

In this review, I first consider Schelling's formulation of brinkmanship and "threats that leave something to chance" to highlight where Pauly and McDermott break with conventional wisdom. I then discuss their article's primary contribution: a significant expansion of the pathways by which "threats that leave something to chance" can achieve coercive credibility. I then consider their arguments in light of other work on the psychological underpinnings of coercion. Finally, I consider the opportunities the authors miss by giving Schelling's analogical reasoning too much credit.

In Schelling's original formulation, brinkmanship was part of a class of problems where the *threat* of violence (including war) could be used by one party against another to gain something of value.[3] The effectiveness of such threats rested on their credibility, from the perspective of the party being coerced.[4] That party *must* believe that whatever violence is threated might actually come to pass. But, according to Schelling, if Actor A attempts to coerce Actor B by threatening a mutually catastrophic outcome, then the threat lacks credibility as long as B believes A to be rational.[5] In the symmetric, strategic nuclear case, "rational" need only mean that Actor B believes there are no stakes for Actor A that are worth certain annihilation and thus no situation in which A would *intentionally* make the choice to launch a strategic first-strike.

How then can Actor A ever hope to coerce Actor B in a high-stakes contest using the threat of nuclear first-use? Schelling's answer was the "threat that leaves something to chance."[6] That is, Actor A threatens the possibility of first-use (or catastrophic escalation) in such a way that Actor B *believes* there are circumstances that *might* arise in which a first-strike would occur. Actor B's *belief* in the *chance* of catastrophe then opens the door to their being coerced because many outcomes will appear more attractive to Actor B than (probabilistic) annihilation, assuming the same minimal definition of "rationality" applies.[7]

---

[1] Reid B. C. Pauly and Rose McDermott, "The Psychology of Nuclear Brinkmanship," *International Security* 47, no. 3 (January 1, 2023): 9–51, https://doi.org/10.1162/isec_a_00451. This review is based on the shortened version of this article published on H-Diplo. All parenthetical page numbers refer to the H-Diplo version.

[2] Thomas C Schelling, *Arms and Influence*, 2nd ed. (New Haven, C.T.: Yale University Press, 2008), 99.

[3] Thomas C. Schelling, *The Strategy of Conflict.* (Cambridge: Harvard University Press, 1960), 5–6, https://catalog.hathitrust.org/Record/000003282.

[4] Schelling, *The Strategy of Conflict.*, 187.

[5] Schelling, *The Strategy of Conflict.*, 35.

[6] Schelling, *The Strategy of Conflict.*, 187–93.

[7] Pauly and McDermott use the term "leverage" to capture Actor A's power to alter the behavior of Actor B. I avoid the term here because a number of factors can enhance Actor A's leverage, independent of threats (e.g., whether A

Schelling elaborates on what "chance" means across several discussions in both *Strategy of Conflict* and *Arms and Influence*. In *Arms and Influence*, he uses the analogy of introducing a die roll into a game of chess, transforming a game where actions have certain outcomes to one where outcomes are probabilistic.[8] Pauly and McDermott lay out many of Schelling's manifestations of chance in Table 1 (column two). Not all are equivalent—some risks are inherent in the technologies involved (e.g., mechanical failure) and some are self-imposed (e.g., pre-delegation). As Pauly and McDermott note, concerned parties and scholars have paid a great deal of attention to these sources of "chance" (6). If the goal is to reduce the appeal of nuclear brinkmanship for Actor A, then the fewer ways in which "chance" can induce Actor B's *belief* in a plausible first-strike the better.

Pauly and McDermott's first major contribution—and break with convention—is to point out that a focus on the various sources and types of "chance" has left the rest of Schelling's logical chain under-theorized. A psychological perspective forces us to question the cascade of events Schelling sketches. In their words: "We locate the real sources of risk when agency remains in the architecture of human psychological decisionmaking" (7).

To illustrate this perspective in action, consider the near-miss "buzzing" of enemy aircraft, which Schelling himself identified as his model case of the "threat that leaves something to chance."[9] That is, this is a case where a line between the threat and the probability of catastrophe should be obvious. But, as Pauly and McDermott show, it is not. *Why* should it be the case that an accident from an activity such as "buzzing" leads to escalation when a proportional response or non-response are both options? As they put it, there is "a quality of automaticity" assumed in Schelling's escalatory story that is never given explicit justification (8). Pauly and McDermott do not deny that "chance" can open up a world of escalatory choices. On the contrary, their second major contribution is to detail the many ways in which that can occur.

Before turning to the substance of Pauly and McDermott's argument, however, it is important to remember that brinkmanship as a theory of coercion relies on the *belief of the target of coercion* (Actor B), not on the coercer's (Actor A's) actual actions. In the nuclear first-use case, what matters is whether Actor B *believes* that Actor A's choice to "push the button" is plausible. This belief is what lends credibility to A's threat. The set of plausible reasons that Actor A might think of for "pushing the button" is irrelevant for coercion if Actor B cannot also conjure up a similar set when imagining A's next move. Nor is it safe to assume that A's imagined set and B's imagined set are symmetric, shared, or updated simultaneously. The true measure of how viable nuclear brinkmanship is as a coercive tool rests in the minds of those who might be coerced. The fewer plausible pathways Actor B sees to A's affirmative choice to push the button, the less effective brinkmanship will be.

I highlight this aspect of brinkmanship because—while Pauly and McDermott focus their attention on specifying the reasons Actor A might escalate—their appeal to psychology also extends their theory to the domain of Actor B's beliefs. When Actor B conjures up the set of reasons why Actor A might "push the button," Actor B is also relying on their intuitive understanding of psychology.[10] Actor B need not have any special information about Actor A for a psychologically grounded model of brinkmanship to work; they only need to apply their own psychological model of another human to identify plausible pathways to first-use.

Pauly and McDermott's second major contribution is to subtract from, and add to, this set of plausible pathways to first-use with a detailed analysis of the psychology of the "choice" to escalate. They orient their

---

has anything B wants). Instead, I focus specifically on B's beliefs about A's threatened use of a first-strike, which is the mechanism by which nuclear brinkmanship must operate.

[8] Schelling, *Arms and Influence*, 102.

[9] Schelling, *Arms and Influence*, 104, fn. 3.

[10] Chris Frith and Uta Frith, "Theory of Mind," *Current Biology* 15, no. 17 (September 6, 2005): R644–45, https://doi.org/10.1016/j.cub.2005.08.041.

discussion of "choice" around three mechanisms of "chance" that introduce risk into a crisis situation: accidents, self-control, and control of others. These mechanisms are not mutually exclusive, but rather encompass different types of risk that inhere in adversarial confrontations.

Pauly and McDermott conclude that the first mechanism—accidents—is "insufficient as a sole mechanism of brinkmanship" (9). The type of accident does not affect their conclusion. Whether the result of independent probabilities (e.g., mechanical failure) or the product of deliberate baiting (e.g., "buzzing" enemy aircraft, deploying forces to train near the enemy), there is still a long chain of decisionmaking between an accident and a first-strike. In general terms, this mechanism is insufficient as long as Actor B does not believe that the full cascade from accident to launch is *inevitable.* Might mechanical failure affect systems which are so critical that Actor A generates an erroneous launch? Perhaps. But that risk exists with or without A's threat and so is irrelevant for coercion. In this class of "chance" then, the reminder that "choice" prevails is reassuring. The plausible paths to first-use that might trouble Actor B actually *decrease* when we follow the authors' advice to "bring humans back in."

The news is not all good though. Human decisionmaking psychology also *increases* the number of plausible pathways to first-use via Pauly and McDermott's second mechanism: losing self-control. In particular, they focus on the relationship between emotions and behavior. Emotions are part of what makes us human and play a crucial role in decisionmaking processes.[11] In my minimalist formulation of Schelling's rational actor, I note that the only real requirement was that the actors be "survival-maximizers." But, as Pauly and McDermott show through a discussion of the emotions literature, the pay-offs to acting on certain emotional impulses can undermine that priority (9-11). Everyday emotions and motivations—anger, revenge, and pride—can all cause a short-sighted, hedonic approach to decisionmaking with adverse long-term consequences.[12] Here, the fact that Actor A has a choice is not entirely reassuring for Actor B. Indeed, if Actor B believes that Actor A might prioritize emotional satisfaction over long-term goals, then Actor B must consider A's threats credible. This discussion of emotions, their variability, and their myriad influences on decisionmaking is the crux of Pauly and McDermott's argument for "bringing humans back in," and rightly so.

The authors' final mechanism—loss of control of others—reminds us that there are many humans involved in the conduct of diplomacy and the waging of war. This has two consequences when we consider the role of choice and the pathways to plausible first-use. First, the number of humans whose choices are relevant can expand significantly. As they note, "nuclear command-and-control arrangements that pre-delegate… exacerbate these risks by introducing more vectors for others' individual choices" (13) Second, the kinds of "manipulations of risk" that might seem implausible for a leader might seem more likely when considered under the heading of third-party "mischief" (12).

Pauly and McDermott also include the case of limited (nuclear) war in this mechanism. In their reading, the decision to launch a strategic first-strike is *still* a choice, even if one has been the victim of a limited (tactical) nuclear strike. This was also Schelling's position, though he disagreed that such a conflict could be labeled "tactical."[13] Yet, this situation strains the definition of brinkmanship, in which the *threat* of mutual catastrophe provides coercive leverage. In this setting, Actor B has *already used nuclear weapons* against Actor A. Actor A's *threat*, to fit within the brinkmanship model, must be to respond with a strategic nuclear strike in order to force Actor B to concede something. But Actor B has already used a nuclear weapon—in fact, they were the first to do so. To have done so implies that Actor B already believed that Actor A's willingness to "go strategic" was near zero. Otherwise, why invite nuclear out-bidding? Put differently, what "threat that leaves

[11] Antonio Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York, NY: Penguin Books, 2005).

[12] Jennifer S. Lerner, Ye Li, Piercarlo Valdesolo, and Karim S. Kassam, "Emotion and Decision Making," *Annual Review of Psychology* 66, no. 1 (2015): 799–823, https://doi.org/10.1146/annurev-psych-010213-115043.

[13] Schelling, *Arms and Influence*, 109–16.

something to chance" can Actor A issue that will change Actor B's beliefs? While Pauly and McDermott claim this situation is within the scope of their argument, it looks more like a coercive failure where other dynamics take over. Though they attempt to include it by extending Schelling's climber analogy, I argue below that this is not a persuasive rhetorical strategy.

On balance, Pauly and McDermott's three mechanisms illuminate more plausible pathways to nuclear first-use than were granted in Schelling's original conception of "threats that leave something to chance." This unfortunate finding is the result of the fact that the authors take human decisionmaking psychology seriously and insert it into the space between "chance" and escalatory "choice." Moreover, as long as the humans standing in for Actor B *believe* that these many pathways exist, nuclear brinkmanship will be a viable coercive tool.

Does this mean that human decisionmaking psychology is inherently destabilizing? McDermott's own work suggests that is not always the case. In a deterrence framework—where Actor A and B can share the preferred outcome of non-action—the human expectation of revenge fueled by hatred keeps state actors from nuclear first-use.[14] In particular, McDermott and her colleagues argue it is the *mutual understanding shared by all humans* that revenge will follow strategic first-use that makes deterrence stable.[15] But, in a compellence framework—where Actor A's preferred outcome is for Actor B to do something they would rather not—there are destabilizing psychological tendencies, even when "chance" is not at work. For example, across seven experiments, a recent study demonstrated that if the subjects were allowed to respond reflexively, they chose "negative escalation" in a fully transparent strategic game when they felt victimized by the other player.[16]

As Pauly and McDermott convincingly demonstrate, integrating human psychology into the study of coercion provides much needed grounding for theories that can otherwise remain too abstract.[17] Yet, their article does not fully capitalize on this intellectual move. Instead, the authors try to reconcile their new insights with Schelling's analogy-ridden discussion of brinkmanship and "threats that leave something to chance." In doing so, they miss an opportunity to directly explicate the psychological mechanisms that generate and sustain the adversarial relationship at the heart of brinkmanship. Their engagements with Schelling's mountain-climber analogy are the weak links of the piece, both in terms of exposition and persuasion. Oddly, this is because they grant Schelling too much credit when it comes to his models of human psychology.

Schelling deployed analogies because real-world examples struck him as "valid in certain circumstances; but seeing through their strangeness and comprehending the logic behind them is often a good deal easier if one has formalized the problem, studied it in the abstract, and identified analogies in other contexts where the strangeness is less of an obstacle to comprehension."[18] I strongly disagree with Schelling's statement, particularly when trying to link theory to real-world cases as we wish to do with nuclear brinkmanship. Pauly and McDermott implicitly endorse this sentiment, however, since they introduce the psychological situation of "brinkmanship" using only Schelling's tethered climber analogy for reference (6). But very little about Schelling's climber analogy makes sense if one wants to understand how two states (or leaders) find themselves in a coercive, adversarial situation, or how that situation might unfold, or ultimately be resolved.

[14] Rose McDermott, Anthony C. Lopez, and Peter K. Hatemi, "'Blunt Not the Heart, Enrage It': The Psychology of Revenge and Deterrence," *Texas National Security Review* 1, no. 1 (2017): 69, http://hdl.handle.net/2152/63934.

[15] McDermott, Lopez, and Hatemi, "'Blunt Not the Heart, Enrage It': The Psychology of Revenge and Deterrence," 71–72.

[16] James Vandermeer, Christine Hosey, Nicholas Epley, and Boaz Keysar, "Escalation of Negative Social Exchange: Reflexive Punishment or Deliberative Deterrence?," *Journal of Experimental Social Psychology* 84 (September 1, 2019): 103823, https://doi.org/10.1016/j.jesp.2019.103823.

[17] Abstraction has its place. See Actors A and B.

[18] Schelling, *The Strategy of Conflict*, 19.

Do two mutually dependent mountain climbers generally start out with zero-sum goals and antagonistic attitudes? Probably not. Does how brinkmanship starts matter for deriving a theory of its dynamics? Probably.

Even for Schelling, the climber analogy did not reveal enough about brinkmanship or coercive bargaining. In the section of *Arms and Influence* that deals with brinkmanship, Schelling deploys at least eight separate analogies to describe coercive scenarios: two mountain climbers tied together; a group of arthritic seniors at a cocktail party; a modified game of chess; delinquent teenagers playing "chicken" in cars; charioteers playing a different, ancient form of "chicken"; a woman pushing a baby in a stroller through an intersection; drivers jockeying through highway traffic; and drivers of wide American cars on narrow European streets.[19]

This list illuminates the inadequacy of analogical reasoning. Yet, Pauly and McDermott engage with Schelling on his preferred terms, modifying his tethered mountain-climber scenario topographically to capture the implications of their arguments for his model (7, 12). The mental morphing required does not actually pay off, however, because Schelling's initial psychological model is so flimsy. Rather, the authors miss the opportunity to lay out the decision-related psychology that gives rise to, and maintains, brinkmanship situations in the real world.

Grappling with Schelling is no easy task. In their article, Pauly and McDermott provide a much needed deep-dive into the gap between "chance" and "choice" and the logic that sits behind "threats that leave something to chance." Though their arguments are not great news for crisis stability, the corrective is still very welcome.

[19] Schelling, *Arms and Influence*, 99–125.; the eight scenarios occur on 99-103 and 116 to 118.

"The Special Fear of Nuclear Weapons"
by Joshua Rovner, American University

The soul of brinkmanship is a threat that leaves something to chance. Such threats do not guarantee certain punishment for noncompliance. Indeed, they don't guarantee anything. The point is to threaten in a way that increases uncertainty for everyone involved, and to arrange matters such that no one knows exactly how a crisis will play out. It might work out fine, or it might end in calamity. The deliberate use of chance has been particularly important in the nuclear era, because deliberate threats of nuclear attack are hard to believe, especially against rival powers that can retaliate in kind. No political goal is worth enough to risk annihilation. But states can take a series of lesser actions—deploying forces, raising alert levels, recalling diplomats, and so on—to create a sense of unease. These tacit threats suggest that events might spin out of control. They are unsettling.

Importantly, however, such threats do not leave *everything* to chance. Leaders do not throw their hands in the air during crises and allow fate to decide the outcome. Instead, they make a series of choices about how to generate uncertainty, and how to respond when their rivals do likewise. Their decisions, write Reid Pauly and Rose McDermott, reflect particular psychological traits as much as pure rational judgment. Leaders try to make sense of contingencies, interpret their adversaries, and act accordingly. Some do it better than others.

Because not all humans are the same, not all leaders are similarly equipped for crises. Those who are unfazed by pressure may be able to triumph in a test of resolve under conditions of ambiguous information and high stress. Such leaders use chance to their advantage by injecting friction into their opponents' calculations while floating blissfully above the fray. They are also better at interpreting signals, which gives them the opportunity to know just how far to push. Others are less capable. Incompetent leaders fail to issue practical policy guidance when it is most needed. Paranoid leaders assume that their opponents are out to get them even when they are not. Narcissists assume that their opponents will understand their signals even when they do not. In all of these cases, disaster looms.

In this article Pauly and McDermott explore the weird interplay between chance and choice among nuclear powers.[1] In so doing they provide a brilliant deductive characterization of brinkmanship. Leaders have many options for manipulating uncertainty during a crisis. They can pretend to tie their hands by pre-delegating decisions, or they can pretend to be insane. They can deploy forces in ways that make accidents more likely, or they can use limited violence to create the sense that events are spinning out of control. Leverage comes from influencing perceptions and emotions, playing on human psychology. The goal is not outwitting rational strategic adversaries but manipulating real people. In the ideal, their rivals will start to believe in unbelievable threats and obsess about the consequences. They should back down under these conditions, accepting coercive defeat rather than suffering the horrible consequences of nuclear war.

The danger is that they will not. The same psychological pressures that create leverage also increase the risk of misperceptions, bad predictions, and rash decisions. Leaders under pressure might back down or they might lash out. Those who engage in brinkmanship may be none the wiser, especially if they become captivated by their own cleverness. Overconfident leaders inadvertently conspire with desperate enemies and put everyone in jeopardy. "Deficits in emotional self-awareness and control result in the inability to make choices

[1] See the original article, Reid B.C. Pauly and Rose McDermott, "The Psychology of Nuclear Brinkmanship," *International Security* 47:3 (Winter 2022/23): 9–51.

consistent with best interests," Pauly and McDermott warn. "The consequence is not only numerous lost opportunities but increased risk of devastating outcomes" (45).

Coercion

How do leaders transform chance into leverage during nuclear confrontations? Pauly and McDermott provide several answers. Leaders can gain leverage, for example, by increasing the risk of accidents. Putting nuclear forces on alert or deploying large numbers of conventional forces in a crisis creates similar dangers. Raising the alert level increases the risk of an accidental launch among crews who are suffering from stress and fatigue. Deploying conventional forces in a crisis increases the risk of a conventional strike that enemies will interpret as a prelude to a counterforce strike. Such is the path towards inadvertent escalation.[2]

Yet Pauly and McDermott point out that psychological biases often work against accidental brinkmanship as a source of leverage. Especially relevant is Robert Jervis's observation that leaders "tend to overestimate the adversary's unity and control" (31). If statesmen assume that their adversaries are cohesive and that rival leaders exercise strict control over the armed forces, then it becomes hard to imagine that the adversary will suffer accidents with its most important strategic assets. The default assumption is that adversary action, nuclear or conventional, is deliberate and organized. Paradoxically, a reputation for competence undermines a leader's ability to coerce. Convincing rivals that accidents are possible is difficult if the same rivals do not believe that a leader is capable of such errors.

Leaders can also put on displays of recklessness. Such efforts are designed to convince observers that they are overcome by panic or madness, and as a result are more likely to take extraordinary risks. Leaders are in control of nuclear weapons but not their own senses. Judicious decisionmaking is in short supply when individuals have lost the semblance of sobriety and restraint. De-escalatory diplomacy among well-meaning diplomats is well-nigh impossible when their bosses have lost the plot. But coercion may be possible for precisely the same reason. Why take on a leader who is crazy enough to use nuclear weapons?

Interest in the madman theory of nuclear coercion has a long history. Chairman Mao Zedong spoke in outrageously cavalier terms about nuclear war, causing some in the United States to wonder if he could ever be trusted with the bomb.[3] President Richard Nixon toyed with madman tactics in hopes of gaining nuclear leverage at time in which US conventional forces were struggling in Vietnam.[4] Most recently, some have taken President Vladimir Putin's nuclear saber-rattling seriously, given his track-record of strategic blunders and his reliance on advisors steeped in conspiracy theories.[5]

Madness, real or feigned, generates more leverage than increasing the risk of accidents. The coercive value of accidents is limited, as Pauly and McDermott rightly note, because *somebody* must ultimately decide whether to respond by authorizing nuclear use. Accidents do not make decisions, people do, and it stands to reason that leaders with a loose grip on reality are more likely to decide on nuclear attacks. That said, the madman theory doesn't have a lot of empirical support. Statesmen worry about apparently unstable counterparts, to be sure, but they tend not to overreact. The United States learned to live with Soviet leader Joseph Stalin and Mao, and Communist adversaries learned to live with Nixon.

---

[2] Barry R. Posen, *Inadvertent Escalation: Conventional War and Nuclear Risks* (Cornell University Press, 1992).

[3] William Burr and Jeffrey T. Richelson, "Whether to 'Strangle the Baby in the Cradle': The United States and the Chinese Nuclear Program, 1960-1964," *International Security* 25/3 (Winter 2000/01), 54–99.

[4] Scott D. Sagan and Jeremi Suri, "The Madman Nuclear Alert: Secrecy, Signaling, and Safety in October 1969," *International Security*, 27/4 (Spring 2003), 150-183.

[5] Gideon Rachman, "Putin, Ukraine, and the Madman Theory of Diplomacy," *Financial Times*, February 7, 2022; https://www.ft.com/content/3d8b94e9-0db7-4aa5-ac6a-9fef2ce43ab6

The latter case is particularly revealing. One reason that coercion failed is that Nixon's rivals assumed he was a cagey manipulator and not genuinely insane. Had they thought otherwise, there would have been no reason to back down at all, because Nixon would not have been able to convince them that he would later restrain US forces (39-40). Effective coercion requires credible threats and reassurances. That is, adversaries must believe that their adversaries will not take advantage of them in the aftermath of a crisis. If they are not reassured, they have no reason to comply. Mad men cannot provide reassurance, however, because this requires getting people to believe that they will return to sanity.

If leaders struggle to convince adversaries that they lack control over their own minds, perhaps they can gain leverage by yielding control over their forces. Pre-delegating launch authority, or relaxing launch protocols, reduces leaders' agency. Looser command-and-control arrangements create opportunities for lower officials to make decisions about nuclear use. Rivals have to consider the possibility that one or more of them might decide to use them, even if the state's leader is willing to back away from the brink. Similarly, leaders might engage in limited violence to generate the same sense of dangerous momentum. The best way to avoid a needless tragedy is to accept the adversary's coercive demands in the early stage of a crisis or war, when leaders are most likely to maintain control and have the political authority to order their subordinates to stand down.

It is not clear that yielding to the "control of others" (45) helps to transform chance into leverage. Pauly and McDermott note several historical cases of military officers taking matters into their own hands to illustrate the problem for civilian control, but they do not show that this had any coercive value. If anything, the discussion suggests that such actions provoke retaliation among the enraged victims (45-49). To the extent that emotion mediates coercive diplomacy in cases of delegated authority, it probably makes coercion less likely. The victims have no reason to trust their attackers or make concessions. They would rather take their chances on the battlefield, where they at least stand to gain a little revenge.

Emotion and psychology clearly play a role in brinkmanship. This is not surprising, as nuclear crises are highly stressful moments. Yet is not clear that the psychological mechanisms that are outlined in the article show how chance becomes leverage. Emotional and psychological factors seem to limit the potential for coercion in crises, not increase it. Leaders may hope to gain leverage through a variety of threats that leave something to chance, yet it is hard to find successes.[6] The psychology of the targets suggests why: individual leaders exaggerate their adversaries' competence, cohesion, and cunning. These beliefs cut against efforts to gain leverage by increasing the danger of accidents, by leaders who pretend to be insane, or who delegate responsibility.

Catastrophe

Pauly and McDermott imply that the targets of coercive threats offer a kind of grudging respect to their rivals. They do not make excuses for them by pointing to disunity or madness. Instead, they assume that adversaries are competent strategic actors. They may be ruthless, but they are not out of control. These assumptions, however, do not preclude other emotions from affecting crisis decisionmaking. Individuals can give their rivals the benefit of the doubt in some respects, for example, and still hate them.

[6] Todd S. Sechser and Matthew Fuhrman, *Nuclear Weapons and Coercive Diplomacy* (Cambridge University Press, 2017). Matthew Kroenig comes to a different conclusion, finding that the possession of superior nuclear forces creates bargaining leverage in crises. Critics, however, have taken issue with his historical case studies, finding the evidence less convincing. See Matthew Kroenig, *The Logic of American Nuclear Strategy: Why Strategic Superiority Matters* (Oxford University Press, 2018); and James Goldgeier, Charles Glaser, T. Negeen Pegahi, Rachel Elizabeth Whitlark, and Matthew Kroenig, "Roundtable on *The Logic of American Nuclear Strategy*," *H-Diplo International Security Studies Forum,* Roundtable 10/25 (2019); https://networks.h-net.org/node/28443/discussions/3917131/issf-roundtable-10-25-logic-american-nuclear-strategy-why

Brinkmanship arouses fear and loathing, and crises among longtime rivals with a history of violence are likely to trigger bad memories and reignite a desire for revenge. Particularly intense emotions make diplomacy difficult. Rational solutions that serve everyone's material interests are hard to come by when individual leaders believe that their enemies are out for blood. Psychological factors compound the problem by making it hard to assess the possible results of one's policies, and to interpret the decisions of others. Pauly and McDermott's analysis of crisis psychology leads to a very unsettling conclusion: "…crises are even more unstable than the traditional theory of brinkmanship posits" (50).

This is a powerful argument and it deserves close attention from scholars and policymakers alike. It is especially pertinent given the ongoing war in Ukraine, which has featured the kind of ambiguous nuclear signaling that leaves a lot to the imagination. As of this writing, the conflict has bogged down into a grinding war of attrition, but sudden reversals on the battlefield might lead Putin to escalate his nuclear threats. Responsible policymakers have no choice but to take them seriously, even if they suspect that Putin will not follow through. The danger that a crisis will produce nuclear escalation feels as relevant today as it did in the late Cold War, and Pauly and McDermott's grim warning is well-timed.

That said, this is not the first time that nuclear-armed great powers have found themselves approaching such a crisis. We are now eight decades into the nuclear age, and we have not suffered a nuclear war. Scholars have identified a number of logical pathways to disaster, including the psychological factors described by Pauly and McDermott, but states have not gone down any of them. Why not?

Reluctance to escalate is noteworthy. The same leaders who are quite aggressive with conventional arms become surprisingly cautious when given the opportunity to go nuclear. Perhaps they are terrified of reprisals in kind. Perhaps they are morally repulsed by the idea. Perhaps they cannot bear the psychological weight of breaking the nuclear taboo. It may be that this mix of emotions is a powerful barrier to catastrophe. Despite having many opportunities to use nuclear weapons in anger, they have consistently rejected the idea.

Leaders have repeatedly blanched when presented with war plans requiring preemptive nuclear use. President Dwight Eisenhower, for example, backed away from massive retaliation soon after floating the idea. In private conversations during the first Taiwan Straits crisis, he and his secretary of state, John Foster Dulles, worried about the consequences of following through on nuclear threats.[7] Leaders in Moscow went through the same gut-wrenching experience. When presented with estimates about the effects of a nuclear war in Europe, Soviet General Secretary Leonid Brezhnev reportedly went pale.[8] The advent of nuclear weapons had mixed effects on international politics during the Cold War, and it did not change as much as some early observers had expected. Yet in one sense they were revolutionary: leaders were unwilling to use them, despite universal recognition of their killing power. This may be a first in the history of technology and war. No other weapons have inspired such caution.

Threats that Fizzle

The psychology of brinkmanship creates pathways to disaster, but the peculiar emotions attached to nuclear weapons serve as firebreaks. How do we reconcile these two perspectives? And what does it mean for our understanding of violence and coercion in a domain of contingency and chance?

[7] Eisenhower took nuclear weapons seriously, but as the crisis neared its culmination, he became increasingly anxious for other options, including abandoning the offshore islands that were the source of the dispute. Gordon H. Chang, "To the Nuclear Brink: Eisenhower, Dulles, and the Quemoy-Matsu Crisis," *International Security* 12/4 (1988), 96-122.

[8] Nicholas Thompson, "Nuclear War and Nuclear Fear in the 1970s and 1980s," *Journal of Contemporary History* 6/1, 136–149, at 138.

Pauly and McDermott develop Thomas Schelling's insight about threats that leave something to chance, mapping out the possibilities for translating uncertainty into leverage.[9] The prospects for coercion may depend not only on the nature of the threat but on the psychology of the target. Individuals who are relatively tolerant of uncertainty, or resigned to the limits of their own control over events, may not be sensitive to such threats. Because they are comfortable with chance, they will not feel pressure to make concessions and end the crisis. By contrast, those who are obsessed with control will be especially uncomfortable when confronted with threats that leave something to chance. These leaders cannot abide Schelling-esque brinkmanship for very long. Instead, they will feel a powerful compulsion to end the crisis.

Yet there are reasons to expect that these are not particularly powerful coercive tools. Pauly and McDermott note that states tend to exaggerate their adversaries' competence, making it difficult to convince their adversaries that events are truly beyond anyone's control. There is also a deeper problem. Threats that leave something to chance create uncertainty, but in so doing they also convey a lack of resolve. When individuals are unwilling to make clear threats, they conjure nightmare visions of imaginary accidents, insane leaders, or out-of-control subordinates. The historical record is incomplete, to be sure, but it is hard to find cases in which targets interpret such threats correctly or take them seriously. Clever attempts to manipulate uncertainty are a reminder that nuclear threats are inherently incredible.

If coercion is unlikely, what about escalation? Pauly and McDermott describe reasons for concern, noting the particular dangers that come from a lack of emotional self-awareness and an inability to accurately interpret adversary threats. That said, the long history of nuclear non-use suggests that a different emotion is at work: the special fear of nuclear weapons. These are, after all, the only genuine weapons of mass destruction. Their destructive power is orders of magnitude higher than anything else humans have devised. Perhaps the emotional story is simple. Leaders are unwilling to unleash nuclear weapons because they are afraid of them.

The last possibility, then, is that brinkmanship threats are inconsequential. Threats leaving something to chance do not generate leverage nor do they cause escalation. They are not useful coercive instruments, but they are not terribly dangerous either. This might be the case because subtle and nuanced threats fail to register in a crisis. Leaders might not sense their meaning; indeed, they might not sense a threat at all. Leaders who are under pressure have a lot to think about, and they might not have the time or inclination to parse the meaning of these oblique signals. In other cases, the mere presence of a nuclear option might cause leaders on both sides to seek peace, or to pursue their competition at lower levels. In peacetime, leaders might be intrigued by the idea of threats that leave something to chance, seeing brinkmanship as an appealing approach that favors cunning and resolve over brute force and bluster. In a crisis, they may simply get cold feet.

[9] Thomas C. Schelling, *The Strategy of Conflict* (Harvard University Press, 1960); Schelling, *Arms and Influence* (Yale University Press, 1966).

"Manipulating and Reducing Uncertainty Over Time: Sequencing Escalation Management Between the Nuclear Powers"

by Janice Gross Stein, Munk School of Global Affairs and Public Policy, University of Toronto

Reid Pauly and Rose McDermott make a singularly important contribution to the analysis of nuclear brinksmanship by sequencing choice after chance. Nuclear strategists have long struggled with a fundamental dilemma: making nuclear threats credible is very difficult when the use of nuclear weapons that could escalate to nuclear war is inherently irrational. Sixty years ago, Thomas Schelling worked through the logic of manipulating risk through threats that "leave something to chance."[1] If you give up some control and communicate clearly to your adversary that you are no longer fully in control, he claimed, your adversary is more likely to capitulate. Even if nuclear weapons were unusable, the threat to use them could be useful.

Pauly and McDermott bring the individual decisionmaker back in the immediate aftermath of one party manipulating the risk of nuclear war. The other party, they correctly argue, has a choice in how it responds to that threat. It makes choice consistent with chance by sequencing the play.

Pauly and McDermott unpack the dynamics of chance to expand the space for choice. Raising a question that Schelling skipped over lightly, they ask *how* a threat leaves something to chance. What are the mechanisms through which chance enters the story, either for the party that issues the threat or for the party that responds to the threat? They identify three: (1) accidents; (2) a lack of self-control; and (3) a loss of control over others. And each of these three can affect either party.

Should an accident happen, both parties have agency in how they respond. All sorts of communication and coordination problems can arise, and therein lies the risk, but Pauly and McDermott are correct that the parties have choices and that these choices have not been well theorized. The second mechanism is in some sense more worrying. One—or both—of the parties lack self-control. Emotions flood the decisionmaking system in the wake of a threat and one party responds in anger with an aggressive response. Or, as Robert Jervis argued, they misperceive the other's intentions and a spiral model develops that culminates in escalation.[2] Finally, decisionmakers may panic badly or be so disconnected from reality that they cannot form a coherent response. Under any of these conditions, Schelling's strategy of manipulating risk would be unlikely to do very well.

The third condition that Pauly and McDermott identify—control of others—comes closest to capturing the element of chance that Schelling intuited. The authority to fire nuclear weapons is predelegated so that the decision maker has given up choice, or someone lower down in the chain of command assumes authority even when they are not authorized to do so. Or, in a familiar sequence, the adversary makes the next move and that move is determining of whether the conflict escalates to nuclear war. Threats that leave something to chance build in some danger by design. It is the danger inherent in giving up some control that conveys the threatener's commitment to an adversary.

There are, however, dangers beyond the danger that is deliberately designed into a threat that leaves something to chance to convey commitment. It is not only when the other side is "not watching"—is not paying attention or misperceives the signal—that both parties die. If both parties try the strategy at the same

[1] Thomas C. Schelling, *The Strategy of Conflict* (Cambridge: Harvard University Press, 1960), Ch. 8; *Arms and Influence* (New Haven: Yale University Press, 1966); and the recently released Schelling, *The Threat that Leaves Something to Chance* (Santa Monica: Rand Historical Document HD-A1631-1, 2021).

[2] Robert Jervis, *Perception and Misperception in International Politics* (Princeton: Princeton University Press, 2nd edition, 2017).

time, both will also die. And if a threat requires the threatener to give up *all* control, the threatener risks unilaterally conceding.[3] In manipulating uncertainty through threats that leave something, but not everything, to chance, success depends on sequenced rather than simultaneous moves and on retaining at least enough control to avoid unilateral disaster. Uncertainty and control sit in uneasy tension with one another.

By opening up the black box of decisionmaking once a threat is made, Pauly and McDermott make an important contribution to theorizing the set of conditions under which a threat that leaves something to chance is unlikely to lead an adversary to capitulate. Schelling was well aware that the strategy was risky and could fail, in disastrous ways, but he saw no alternative that would make the threat to use nuclear weapons credible, especially when the conventional balance of capabilities was unfavorable. But he paid almost no attention to the psychological and biological attributes of human decisionmaking that could defeat his strategy. Pauly and McDermott have shone a long overdue spotlight on precisely these factors.

Since Russia invaded Ukraine, we have watched a real-time experiment of managing nuclear escalation unfold over time. When we consider the sequencing of strategic choices over time, the story becomes even more complicated.[4] What does the evidence from this one case that has unfolded over time suggest? Russia has used a weak version of "the threat that leaves something to chance" several times in an effort to deter NATO from committing forces to the battlefield on behalf of Ukraine. Moscow has manipulated the risk of nuclear war but only through veiled verbal threats; it has not put any of its strategic nuclear forces on alert or removed any tactical weapons from storage. Moscow has not yet given up any control so, in this sense, it is not a fair test of Schelling's strategy. Even though it is not a fair test, the evidence suggests that no explicit threat, much less one that left something to chance, was necessary. As the Biden administration became more confident that Russia, the largest nuclear power, was going to attack Ukraine, it made the decision that no US forces, either in the air on the ground, would be committed in Ukraine to avoid the possibility of any direct combat between Russian and American forces. Direct combat, the president insisted, could escalate to World War III.

The Biden administration and NATO leaders responded with a pragmatic, experimental, inductive strategy that I call "learning by doing."[5] The administration first imposed constraints to structure an ill-structured problem and then tested the limits of these boundaries by repeatedly probing, evaluating the response, and then probing again.[6] That pragmatic strategy enabled the United States to increase its tolerance of uncertainty

[3] In the game of chicken, if one drives throws their steering wheel out the window to signal resolve and force the other driver to capitulate, the driver's car without the steering wheel can crash and that driver could die.

[4] See my "Escalation Management in Ukraine: 'Learning by Doing' Responds to the 'Threat that Leaves Something to Chance'," forthcoming *Texas National Security Review,* for an analysis of the interactive sequencing of strategies by Russia and the United States since the Russian invasion of Ukraine. See also my "Escalation Management in Ukraine: Responding to the Manipulation of the Risk of Nuclear Escalation," Kissinger Center Papers series, 1:1, forthcoming, https://sais.jhu.edu/kissinger,

[5] Janice Gross Stein. "Political Learning by Doing: Gorbachev as Uncommitted Thinker and Motivated Learner." *International Organization* 48, 2 (1994): 155-83**,** https://www.cambridge.org/core/journals/international-organization/article/abs/political-learning-by-doing-gorbachev-as-uncommitted-thinker-and-motivated-learner/D2C748D82B83EC3DC044D34C64611788. . Learning by doing draws upon the central tenets of pragmatism. A rich tradition that encompasses a wide range of approaches, pragmatism is defined broadly as a set of dispositions for making sense of the world and what works in the world. It treats the world as relational and contingent where people adjust their habits to the contingent. For applications of pragmatist theories to international politics, see Simon Frank Pratt, Sebastian Schmidt, Deborah Avant, Molly Corcoran, Patrick Thaddeus Jackson, Henry Farrell, Jack Knight and Gunther Hellman, "Pragmatism in IR: The Prospects for Substantive Theorizing," *International Studies Review* 23 (2021): 1933-1958, https://academic.oup.com/isr/article-abstract/23/4/1933/6275420; and Molly Cochran, "A Story of Closure and Opening," *European Journal of Pragmatism and American Philosophy* 2 (2012), https://doi.org/10.4000/ejpap.777

[6] A problem is well-structured when it has a clear goal, known constraints, and identified possible outcomes. None of these conditions existed in the wake of Russia's invasion of Ukraine that shattered the norms as well as the rules and created a condition of deep uncertainty. See Herbert A. Simon, "The Structure of ill-Structured Problems,"

over time and gradually expand the range of the possible, all within the limits of the constraints that Washington set even before the invasion began.

The administration's response to a threat that in a weaker version left "something" but not everything to chance was a pragmatic probe of the limits of risk-taking in the context of the stability-instability paradox. The paradoxical effect of mutually assured destruction between nuclear powers is to make proxy conventional wars between them safer.[7] Russia's invasion of Ukraine is not a proxy war—Russia is a direct combatant—but the general finding holds, at least through the first year. In the context of conflict with a nuclear power, the United States repeatedly tried to reduce uncertainty about the consequences of its incremental options and actions. These two strategies—Moscow's threat that left something to chance and Washington's disciplined pragmatism—constituted a dialectic that enabled "iterative negotiation and definition of the outer limits of indirect military conflict among two nuclear peers."[8]

Stepping back from the evidence of the first year of escalation management in Ukraine, several important theoretical questions arise. These questions are framed by the radical uncertainty that Russia's invasion of Ukraine created. It is challenging to theorize interactive and dynamic strategies of managing escalation between nuclear powers, but it is even more difficult to do so under conditions of radical uncertainty.[9]

Schelling assumes at least some rough knowledge of an adversary's preferences. Without that knowledge, it is difficult for one nuclear power to judge how far to go to compel another nuclear power. The evidence shows that US government officials, and many outside experts, did not have that kind of knowledge. Changing estimates of Russian President Vladimir Putin's intention to launch a full-scale invasion of Ukraine were driven upward largely by intelligence analysis of Russia's growing capabilities that were deployed up against Ukraine's borders. Right up until the attack began, a few policymakers in Washington and many in Berlin and Paris argued that it would be irrational for Putin would launch an all-out attack and that he was bluffing. In the aftermath of the attack, many openly expressed their shock. Absent the necessary knowledge of Putin's preferences, a pragmatic, disciplined strategy of learning by doing was a sensible choice.

The challenge lies in the dialectic between these dynamic and interactive strategies. The learning that comes from doing may be limited by changes in underlying conditions that then lead to changes in preferences. But decisionmakers may not be aware that their adversary's preferences have evolved. Preferences are not stable over time and strategies can change preferences in several different ways.

First, when one set of decisionmakers acts, they may change the situation in the course of acting. Through their actions, leaders not only experiment and test to see if they are bumping up against limits, but they also

---

Artificial Intelligence 4 (October 1973): 181-201, https://www.sciencedirect.com/science/article/abs/pii/0004370273900118; James F. Voss and Timothy A. Post, "On the Solving of Ill-structured Problems," in Micheline H. Chi, Robert Glaser and Marshal J. Farr, eds. *The Nature of Expertise* (Hillsdale, N.J.: Lawrence Erlbaum, 1988): 261-285.

[7] Glenn Snyder, *The Balance of Power and the Balance of Terror* (New York: Chandler, 1965); Robert Rauchaus, "Evaluating the Nuclear Peace Hypothesis—A Quantitative Approach," *Journal of Conflict Resolution. 53, 2(2009): 258–277.* doi:10.1177/0022002708330387.

[8] Author interview, senior U.S. official, March 12, 2023.

[9] Uncertainty and risk are often conflated but are distinct. Decisionmaking under risk occurs when probability distributions are available to estimate the likelihoods of consequences of known outcomes. In international politics, in many cases the requirements for probability distributions—reliable data and large numbers of trials—are not present. Decisionmaking under uncertainty occurs either when the consequences of options are known, but there are no relevant probability distributions to estimate their likelihood—unknown knowns—or neither the consequences of options nor their likelihoods are known—unknown unknowns, a condition or radical uncertainty. People find uncertainty very uncomfortable psychologically and therefore transform uncertainty into risk by making subjective estimates of likelihood even when there are no underlying probability distributions to draw on.

change the situation and possibly those limits. When the United States decided to supply Ukraine with High Mobility Artillery Rocket System (HIMARS) multiple rocket launchers, they probed to see whether rockets that could not reach Russian territory would avoid provoking Russian escalation. Russia did not escalate in response, either horizontally beyond Ukraine or vertically to unconventional weapons. What should decision makers "learn" from this sequence?

Was it the limited range of the weapons that led Russian decision makers not to escalate? Or is Russia so anxious to avoid confrontation with superior NATO forces that it will not escalate no matter what equipment NATO supplies? There is no unequivocal answer to this question and the answer can be known only after the fact, should NATO cross Russia's threshold and Putin then escalate. There is a risk of escalation even in a pragmatic, calibrated strategy of the kind the Biden administration has used. Leaders can "overlearn" from past successes that will tell them little about a threshold they might cross in the future, especially if conditions have changed. The risks of overlearning and overconfidence are significant because leaders do not know their adversary's preferences.

A second set of obstacles to theorizing strategies of nuclear escalation is even more daunting. There is good evidence that leaders not only do not know the preferences of their adversary as conditions change, but that at times they may not know their *own* preferences. Economists and political psychologists have demonstrated that, at times, people do not know their preferences. They only discover them by making a choice and then inferring their preferences from their behavior.[10] Individuals come to "know" their own attitudes, emotions, and preferences by inferring them from observations of their own behavior or the circumstances in which their behavior occurs.[11] Here preferences do not dictate strategy; rather, strategy shapes preferences.

There is no better description of the unpredictability of decisionmakers in dynamic strategic environments than the one written by Thomas Schelling for American policy makers more than sixty years ago. His analysis of how decision makers cannot anticipate what they will do applies equally well to Putin and his small circle of advisors today:

> We are not making the enemy unpredictable only to the extent that we can deceive him; our response is unpredictable to him because it is unpredictable, in some significant degree, *even to us*. We are not threatening that we may surprise him because we can calculate his expectations better than he can calculate ours; we may surprise him for the same reason that we may surprise ourselves. He cannot expect to foretell our behavior *in contingencies so complex that we cannot ourselves exactly foretell our response to them.*[12]

These challenges of knowing one's own as well as an adversary's preferences are made more difficult in dynamic environments where non-linear actions have unpredictable results. The condition most likely to approximate that environment would occur should Russia suffer rapid, unanticipated, and severe reverses on the battlefield that its leaders interpret as posing the risk of strategic defeat. Here the argument rejoins directly and compliments the arguments of both agency and emotion made by Pauly and McDermott.

A rapid and adverse change in battlefield conditions that surprises Russia's leaders is more likely to engender panic than the slow, grinding war of attrition that has been fought so far. Were that to happen, Russia's leaders would have little time to normalize for the losses. That panic could be reinforced were Russia's leaders to believe that they were facing strategic defeat. Panic can lead to a "flight" response, which might translate

[10] See Paul A. Samuelson, 1948. "Consumption Theory in Terms of Revealed Preference," *Economic*a, 15: 243-253, https://www.jstor.org/stable/2549561.

[11] Daryl J. Bem, "Self-Perception Theory," *Advances in Experimental Social Psychology* 6 (1972):1-62, https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=daryl+bem+self+perception+theory&oq=Daryl+Bem.

[12] Schelling, *The Threat that Leaves Something to Chance*, 26. Emphasis added.

into a willingness to enter into a process of political accommodation. It can also lead to a "fight" response that would translate into an escalation to an unconventional weapon that Russia's leaders hope would freeze the conventional fighting long enough for their armed forces to regroup. It is worth noting that rational models focused on cost-benefit calculations are more likely to identify the first rather than the second response to a dramatic and adverse shift in battlefield conditions. Arguments focused on emotion would pay attention to both.[13]

Pauly and McDermott identify emotional responses other than fear that a rapid and dramatic increase in the perceived likelihood of a severe defeat might provoke. Principal among them is anger at "the West," an anger that is fueled by a perception that the West seeks to encircle, humiliate, and even dismember Russia. Citing Russian sources, Pauly and McDermott also suggest that Putin may be a narcissist. Narcissists, when thwarted by others, tend to lash out in rage against the injury they are experiencing. These two arguments converge to predict escalation, as do arguments that draw on prospect theory that predicts risk-acceptance when leaders confront a high probability of a large loss.[14]

Four of the five emotional responses that Pauly and McDermott predict as a response to a threat that leaves something to chance would also predict that Putin—who has made these kinds of threats—is likely to escalate should he confront a quick and dramatic prospect of strategic defeat. Arguably, under these conditions, Putin could manipulate the risk of war and leave more to chance than he has done in the past.[15]

This analysis suggests that it is important to identify not only the menu of emotional responses to a threat that leaves something to chance, as Pauly and McDermott do, but also to identify the conditions that are likely to enable each of these and then connect each of them to a behavioral response. Admittedly, this is challenging work, given that analysts are operating in a world of uncertainty rather than risk, but even specifying sequences of responses over several iterations would be helpful to those who are responsible for designing strategies of escalation management over time in dynamic and interactive environments.

Policymakers in Washington have already done this for one sequence. In a major turn that reversed long-standing American strategic policy, US officials signalled informally that should Russia use a tactical nuclear weapon, either in response to NATO's supply of increasingly sophisticated weapons to Ukraine or in response to what Russian leaders might consider an impending strategic defeat, the administration would not retaliate with a nuclear weapon.[16] The US response would be conventional, with grave consequences for Russia, to "signal immediate de-escalation" that would then be followed by international condemnation.

The widespread signalling that the use of a tactical nuclear weapon would be met with a severe conventional response is part of a US strategy of escalation management that was designed with several steps in sequence.

---

[13] A third possible response is freeze. Freeze is more likely to occur at a moment of intense fear engendered by a sense of acute danger. It is less likely to persist over time, especially in a collective environment.

[14] This is the prediction that prospect theory makes. See Amos Tversky and Daniel Kahneman, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5 (1992):297-323, https://link.springer.com/article/10.1007/BF00122574.

[15] Kevin Ryan, "Why Putin Will Use Nuclear Weapons in Ukraine," argues that during the past twelve months, Putin has laid the groundwork for using a tactical nuclear weapon in Ukraine and estimates that Putin will use a tactical weapon, https//www.russiamatters.org/analysis/why-putin-will-use-nuclear-weapons-ukraine. Avril Haines, US Director of National Intelligence, told a Senate hearing in April that it was "very unlikely" that Russia would use nuclear weapons. See Committee on Armed Services United States Senate Hearing to Receive Testimony on Worldwide Threats, Thursday, May 4, 2023, Washington, D.C, https://armed-services.senate.gov/imo/media/doc/23-44_05-04-2023.pdf. There is no underlying probability distribution for the use of nuclear weapons. In both cases, analysts have converted uncertainty into subjective estimates of risk.

[16] Pavel Podvig emphasized the importance of making public this commitment to respond to the use of a tactical nuclear weapon with conventional weapons. Author Interview of Pavel Podvig, January 18, 2023.

By making explicit what the United States would do if Russia were to use an unconventional weapon, policy-makers in Washington hoped both to influence Russia's decisionmaking and put a break in an escalatory spiral if Russia nevertheless were to go ahead.[17] This is an example of the finely differentiated and sequenced strategies of escalation management that will become much more important as great power competition among nuclear powers returns with a vengeance.

[17] Senior US officials ran a top-secret table top nuclear exercise in 2016 where an adversary used a tactical nuclear weapon. Senior officials in the Obama administration urged a nuclear response, but the deputies, including Avril Haines, currently the director of national intelligence, and Colin Kahl, currently the under secretary of defense for policy, advocated responding with conventional weapons. Their response was met at the time with resistance by those who worried about the consequences for extended deterrence. See Sanger and Broad, "Putin's Threats Highlight the Dangers of a New, Riskier Nuclear Era*,"* and Scott Sagan, "The World's Most Dangerous Man: Putin's Unconstrained Power over Russia's Nuclear Arsenal," *Foreign Affairs*, March 16, 2022, https://www.foreignaffairs.com/articles/russian-federation/2022-03-16/worlds-most-dangerous-man. Sagan concludes that "The deputies had the better strategy, one that was firm but less likely to provoke thermonuclear catastrophe."

"Let's Leave Less to Chance"
Response by Reid B. C. Pauly and Rose McDermott, Brown University

We are extremely grateful for such a thoughtful and constructive conversation about nuclear brinkmanship with three distinguished commentators. We are also grateful to Jacques Hymans for writing the introduction to this forum. We respond to their thoughts on our article across three areas: lingering theoretical questions and puzzles, applications to the current war in Ukraine, and prospects for future empirical work on brinkmanship.

Theoretical Questions and Puzzles

Each of the commentators points to various ways in which Thomas Schelling's original work on threats that leave something to chance, while hugely important, remained under-theorized. This was part of the challenge we sought to address in our original piece. Much more theoretical work remains to be done. As Janice Gross Stein observes, choice in general has not been well theorized in this literature, although of course it has been in other disciplines, most notably behavioral economics. We highlight some of the lingering puzzles that this forum illuminates.

First, Schelling's "threat that leaves something to chance" is a theory of coercion, and we have unpacked in our article some of the psychological pathways by which it might work. Yet, for many psychological reasons as well, brinkmanship may have limited coercive value. Joshua Rovner points, for instance, to the psychology of targets and how they are apt to "exaggerate their adversaries' competence, cohesion, and cunning." He goes on to suggest that "emotional and psychological factors seem to limit the potential for coercion in crises, not increase it." We agree in principle but think that the net effect is as yet unclear. Rovner is exactly right that threats do not guarantee anything. There may be situations were threats move the needle on an adversary's behavior, but there are other situations where they do not, and it may not be possible to distinguish ahead of time, raising the risk for inadvertent escalation if threats do not lead to capitulation. Marika Landau-Wells similarly calls readers' attention to the fact that credibility itself is a perception. Even insincere threats can be credible if they are perceived to be so by their targets.[1] As she writes, "the true measure of how viable nuclear brinkmanship is as a coercive tool rests in the minds of those who might be coerced." This insight highlights the reality that even when both sides have beliefs that they understand to be true, they might not have the *same* beliefs. In other words, each side can think they are behaving in a manner consistent with what they believe to be true, but that truth may not be shared across adversaries. This can quickly lead to miscommunication, especially if each side believes that the other shares their understanding of the situation they confront.

Rovner also emphasizes the importance of the role of memory in such interactions. It is rarely the case that adversaries approach a conflict with a blank slate; each side holds beliefs and grudges, if not actual trauma, from past interactions. As he writes, "Brinkmanship arouses fear and loathing, and crises among longtime rivals with a history of violence are likely to trigger bad memories and reignite a desire for revenge." This dynamic not only makes accommodation less likely, and reconciliation more challenging, but may also raise the stakes in any attempts to negotiate a settlement during conflict. Endemic conflicts from the Middle East to South Asia to Northern Ireland demonstrate how long the shadow of history may cast. Rovner thus correctly suggests that the standard literature on brinkmanship lacks a critical temporal dimension, and assumes both a more immediate and a more static interaction than in fact may be the case.

Second, Stein advises complicating matters further by considering interaction effects and sequencing in brinkmanship. In particular, Stein explores, here and in her own published work, the efficacy of counter-

[1] See also Kelly M. Greenhill and Peter Krause, *Coercion: The Power to Hurt in International Politics* (Oxford University Press, 2018).

brinkmanship strategies.[2] She calls attention to the Western "learning by doing" approach to escalation management during the War in Ukraine, which we will return to in the next section. As a theoretical matter, if success in brinkmanship requires "sequenced rather than simultaneous moves" (Stein), it suggests that another Schelling concept of the "last clear chance" is even more critical to brinkmanship than we identified. Forcing sequenced choices, with all of their attendant psychological foibles and renewed opportunities for miscalculation, is the name of the game in brinkmanship.

Third, Schelling was clear to articulate the value of surrendering choice to others, in particular in his discussion of limited war as a generator of risk. Yet even within this illustrious group of scholars, there is disagreement about our inclusion of limited war as a mechanism of brinkmanship. We still think it merits inclusion, but it does raise important questions about whether and how conventional forces affect risk perceptions in nuclear brinkmanship. Rovner offers a theory that is rooted in conventional strength: that the fear of conventional counterforce drives risk perceptions in brinkmanship crises. Yet Stein offers a theory that is rooted in conventional weakness: substituting nuclear for conventional threats and demonstrating resolve through vulnerability. We think this is a fruitful area for more theorizing.

Fourth, a clear puzzle remains over the sources of coercive assurance in brinkmanship interactions.[3] As Schelling reminded us of threat-making in general, "To be coercive, violence has to be…avoidable by accommodation.…The pain and suffering have to appear *contingent* on his [the target's] behavior."[4] Rovner elevates this important internal contradiction in the mad-man model that we only touched on in our article and have been puzzling over since. As he writes, "Mad men cannot provide reassurance…because this requires getting people to believe that they will return to sanity."[5] In other words, once one leader convinces another that he is crazy, why would that adversary then believe that if he does as the leader asks, he will not be destroyed anyway? This raises the following questions: Where does credible coercive assurance come from in nuclear brinkmanship? Or, if coercive assurance is anathema to a strategy of losing control, then why does brinkmanship ever work? Or perhaps it does not work for reasons having little to do with assurance. These puzzles are critically important to investigate.

Applications to the War in Ukraine

Whether or not Russia's invasion of Ukraine is a strong case of nuclear brinkmanship or a good test of any theory is still up for debate. Nonetheless, passing the lens of brinkmanship theory over the case brings much into focus, including the strategies of both sides and the risks of undesirable escalation. Russian President Vladimir Putin has clearly used his nuclear arsenal as a shield behind which to engage in conventional aggression and deter direct intervention by the United States or NATO. He also seemed to think that nuclear saber rattling would instill such fear as to prevent or at least limit military aid provided to Ukraine. Some of these threats have worked and others have not.

[2] For the full argument, see Janice Gross Stein, "Escalation Management in Ukraine: 'Learning by Doing' in Response to the 'Threat that Leaves Something to Chance,'" *Texas National Security Review*, Vol. 6, Iss. 3 (Summer 2023), https://tnsr.org/2023/06/escalation-management-in-ukraine-learning-by-doing-in-response-to-the-threat-that-leaves-something-to-chance/.

[3] One of us also investigated this idea here: Reid B.C. Pauly, "Stop or I'll Shoot, Comply and I Won't: Coercive Assurance in International Politics," (Ph.D. diss., Massachusetts Institute of Technology, 2019).

[4] Schelling, *Arms and Influence*, 2-4. See also chapter two of *The Strategy of Conflict* (1960), based on his 1956 paper "An Essay on Bargaining," *American Economic Review* 46,3 (1956).

[5] See also Roseanne W. McManus, "Revisiting the Madman Theory: Evaluating the Impact of Different Forms of Perceived Madness in Coercive Bargaining," *Security Studies* 28, 5 (2019): 976–1009, https://doi.org/10.1080/09636412.2019.1662482.

As noted above, Stein dubs the Western approach to countering Russian brinkmanship in Ukraine a strategy of "learning by doing."[6] It has a quality of Schelling's "salami tactics." Whether intentionally or not, the Biden Administration's cautious approach of providing aid with gradually increasing lethality is a strategy that has circumvented Russian deterrence—body armor and ammunition, then Javelin and Stinger missiles, then HIMARS, then tanks, then a decision on F-16s, cluster munitions, etc. The idea is that each step is small enough to remain below the threshold of triggering Russian retaliation. Viewed through the lens of risk, each of these choices has boiled down to a question familiar to brinkmanship: how much risk of escalation are you willing to accept to support the deserving Ukrainian defenders? US policymakers have reasonably come up with different answers to this question at different times.

Pragmatic "learning by doing," however, has a significant downside in theory. Over the course of the war, officials have grown more confident that Putin's nuclear threats are bluffs.[7] Our psychological approach, echoed by all the reviewers but especially Stein, advises some caution about such confidence. One can cross a red line without intending to, or without even being aware that the line is being traversed. "Preferences are not stable over time and strategies can change preferences in several different ways," writes Stein. Both political as well as military strategies can influence the nature of preferences. And "the risks of overlearning and overconfidence are significant because leaders do not know their adversary's preferences." This insight connects the nature of personality with that of preference in an important way. Narcissistic leaders, or those who are surrounded by sycophants, may not receive accurate information about their opponents, from capabilities and intentions to preferences. Moreover, as Stein also points out, leaders may remain a mystery to themselves and may not know what they would do in a given situation until it arises. Putin might not himself know what his red lines are until they are crossed.

In this way, uncertainty is doubled, as leaders remain ignorant of their own, as well as their adversaries' preferences and intentions. Moreover, as noted, leaders may be blind to their lack of awareness about their own intentions. Leaders are navigating escalation management in a world of uncertainty, not calculable risks. "There is no underlying probability distribution for the use of nuclear weapons," Stein observes, so analysts or advisors making assessments "have converted uncertainty into subjective estimates of risk." They may then assume that such judgments are in fact more objective and realistic than they are.

Another theoretical puzzle springs from the interaction of Russian and Western strategies in Ukraine. Stein writes that together these two strategies—"Moscow's threat that left something to chance and Washington's disciplined pragmatism"—have helped to define the "the outer limits of indirect military conflict among two nuclear peers." Yet Schelling did not advise signaling your disciplined pragmatism in nuclear crisis bargaining. When the other driver throws the steering wheel out the window, would a pragmatist not swerve? If one of your policy objectives is to discourage Putin from engaging in brinkmanship, why would signaling your monopoly on the responsible high-ground be the best way to do that? And yet it seems to have led to some stability in Ukraine about the expectations of either side—no direct NATO intervention, no Russian attacks on NATO targets, a hurting equilibrium of Western military assistance and Russian willingness to bear high costs. Perhaps Russia is not doing a good job of leaving things to chance (certainly a possible answer), or maybe pragmatists are better at playing (or avoiding) the game of chicken than we think. Or perhaps luck has played a bigger role than anyone is comfortable acknowledging.

---

[6] See also Janice Gross Stein, "Escalation Management in Ukraine: 'Learning by Doing' in Response to the 'Threat that Leaves Something to Chance,'" *Texas National Security Review*, Vol. 6, Iss. 3 (Summer 2023), https://tnsr.org/2023/06/escalation-management-in-ukraine-learning-by-doing-in-response-to-the-threat-that-leaves-something-to-chance/.

[7] See Helene Cooper, Eric Schmitt, and Julian E. Barnes, "U.S. Warms to Helping Ukraine Target Crimea," *New York Times*, January 18, 2023, https://www.nytimes.com/2023/01/18/us/politics/ukraine-crimea-military.html, which included this claim: "fears that the Kremlin would retaliate using a tactical nuclear weapon have dimmed, U.S. officials and experts said."

One of the other most surprising and interesting aspects of the war in Ukraine revolves around the internal fights among the various divisions of the Russian military establishment. The military mutiny undertaken on 23-24 June 2023 by Yevgeny Prigozhin as head of the Wagner group against Putin and the Russian Ministry of Defense was surprising to many observers. Prigozhin claimed he was not rebelling against Putin, but he clearly took issue with other members of the Russian Defense establishment, especially Minister of Defense Sergei Shoigu and Chief of the General Staff Valery Gerasimov. A country divided against itself, particularly in the realm of its coercive fighting forces, will have a harder time conducting an effective campaign against any external enemy. These are the precise circumstances under which Stein notes that the risk of escalation may increase. Leaders who feel cornered can have flight or fight responses, depending on the circumstances, including whether losses occur rapidly or over a long grinding attrition. But the psychological literature includes two other common mammalian responses in the face of threat: freeze and another "f" (pro-social behavior). On a battlefield, freeze may look like an acceptance of paralysis with entrenched front lines. But the pro-social option may encourage peace, bargaining, and negotiation. This response tends to occur most often among animals who know each other well, or humans who are culturally or linguistically similar.

In a recent article for *Foreign Affairs*, we teamed up with eminent psychologist Paul Slovic to examine Vladimir Putin's psychology.[8] The results are not comforting. We identified four specific cognitive biases that may be operating in Putin's mind. First, *security prominence* explains how people generally struggle to weigh conflicting risks, including those involving nuclear weapons. Humans, when faced with complexity, simplify and narrow their focus until a clear choice emerges. Rather than weighing diverse values and objectives in a compensatory manner, we order our goals by priority and focus on achieving the one that holds the highest value for us. In the case of leaders, security tends to be prioritized above all else. This bias can heighten the perceived near-term benefits of force while diminishing the longer term and murkier human, social, cultural, and political consequences.

Second, *psychic numbing* is a phenomenon that explains how people become inured to large losses of life. Most people are familiar with the aphorism that "the death of one man is a tragedy; the death of a million is a statistic." As the number of lives at risk increases, psychic numbing begins to desensitize us. In some cases, the more who die, the less we care. Third, perpetrators of harm almost always believe themselves to be engaged in *virtuous violence* in defense of their in-group. When violence toward an enemy appears to us not only justified but virtuous, the threshold for withholding the use of force, or for avoiding escalation, diminishes greatly.

Finally, each of these psychological pathologies is exacerbated by narcissism, which Putin displays in spades. Of course, it is impossible to precisely assess the odds that Putin will use nuclear weapons in Ukraine. But uncertainty and imprecision are not the same as ignorance. Psychological theorizing can help us identify some of these sources of risk.

Future Empirical Work

Each commentator notes both the importance and challenge of developing empirical demonstrations and tests for various aspects of our theoretical models. We summarize several important prospects here for going beyond our deductive approach and investigating the coercive value of brinkmanship, and specifically its

[8] Rose McDermott, Reid Pauly, and Paul Slovic, "Putin and the Psychology of Nuclear Brinkmanship," *Foreign Affairs*, May 30, 2023, https://www.foreignaffairs.com/ukraine/putin-and-psychology-nuclear-brinkmanship.

psychological pathways. More broadly, all the reviewers seem to agree that the field must continue to integrate human psychology into the study of coercion in general.[9]

First, Rovner writes that "Individuals who are relatively tolerant of uncertainty, or resigned to the limits of their own control over events, may not be sensitive to such [brinkmanship] threats. Because they are comfortable with chance, they will not feel pressure to make concessions and end the crisis." This suggests that the next step is to capture variation in the psychology of targets as well as senders of brinkmanship threats and further explore how each side perceives risk, intentionally or unintentionally manipulated. Different kinds of leaders may perceive brinkmanship differently. It is also possible that it is not just individuals but also regime types that may vary in their ability to both recognize as well as respond to brinkmanship threats. Specifically, democracies, for structural reasons, may have a lower threshold for the recognition and response to external, if not internal, threats. Autocracies may prove more insulated from citizen political pressure and thus can appear less sensitive to threats, unless those threats prove useful for their own internal political dynamics or those threats pose challenges to their regime itself.

Second, Stein rightly pushes the field to tie emotional responses to behavioral ones, and indeed this is a very important both theoretical as well as empirical point. Landau-Wells relatedly argues that we missed "an opportunity to directly explicate the psychological mechanisms that generate and sustain the adversarial relationship at the heart of brinkmanship." If particular emotions are more likely to trigger specific behavioral responses than others, we should be able to predict these outcomes. Is fear more likely to generate submission? Is anger more likely to lead to escalation? Such suppositions should be amenable to experimental testing using mood manipulation techniques in particular. Providing further explication of the relationship between emotion and behavior under conditions of threat and uncertainty offers an important next step in further clarifying this model.

Indeed, this point relates to the one Stein raised earlier regarding the role of history as well as the point Landau-Wells makes about the power of memory. The origin of hostility often casts a long shadow over current conflicts. Many relationships that are infused and defined by hatred are sustained because each side may have some ostensible justification for their hostility, often long buried in the past, which may or may not have anything to do with the actual source of their current antagonism. Clearly, providing greater theoretical specificity, as well as empirical demonstration, of the way that psychological factors both inspire and sustain such hostility should be the next step in the further development of this model.

Third, a methodological point necessarily follows. If we want to know something about the probability of nuclear war, we have only to observe from history: one nuclear war, two uses of nuclear weapons, ten nuclear-armed states, and depending on how you count, about twenty nuclear crises. So our field does its best with limited data, confident that the questions are still important. We can look at this history or model rational decisionmaking. If we want to do better, we have to generate the data ourselves—with experiments, structured interviews, analyses of wargames, and so on.[10]

Moreover, even our limited evidence is difficult to interpret. Is the non-use of nuclear weapons since 1945[11] or the limited observed utility of nuclear coercion[12] evidence against a theory of brinkmanship? As Rovner

[9] For another recent example, see Kathleen E. Powers and Dan Altman, "The Psychology of Coercion Failure: How Reactance Explains Resistance to Threats," *American Journal of Political Science* (June 2022), https://doi.org/10.1111/ajps.12711.

[10] On the challenges of nuclear weapons research, see "What We Talk About When We Talk About Nuclear Weapons," H-Diplo/ISSF Forum, June 15, 2014, http://issforum.org/ISSF/PDF/ISSF-Forum-2.pdf

[11] Reid B.C. Pauly, "Would U.S. Leaders Push the Button? Wargames and the Sources of Nuclear Restraint," *International Security*, Vol. 43, No. 2 (Fall 2018), pp. 151–192, https://doi.org/10.1162/isec_a_00333 .

[12] Todd S. Sechser and Matthew Fuhrmann, *Nuclear Weapons and Coercive Diplomacy* (New York, NY: Cambridge University Press, 2017).

writes, invoking a core paradox of the nuclear age, "Threats that leave something to chance create uncertainty, but in so doing they also convey a lack of resolve…. Clever attempts to manipulate uncertainty are a reminder that nuclear threats are inherently incredible." Yet nuclear war is not impossible, and things that have rarely happened happen all the time.[13] The end of the Soviet Union, the reunification of Germany, and the end of the Cold War were all unthinkable until they were not. Things can change rapidly once momentum builds, and tipping points are not always clear in advance. Complicating empirical brinkmanship research further, Landau-Wells astutely points to the centrality of belief in explaining the motives and behavior of all participants. Even if brinkmanship is exceedingly rare, it is still possible if decisionmakers believe it is possible—fear itself, as they say.

Methodologically, Landau-Wells also warns against abstraction and analogical reasoning. In critiquing Schelling but building on his analogies, we may have indeed tied our hands by debating on his terms. Perhaps that approach is inherently limited when it comes to psychology. Still, Schelling's abstraction may also be a source of his staying power in the field and his influence on policy. The work is approachable and makes complex strategic concepts tangible to quotidian experience. Whether we, building on Schelling, have taken us as far as we can go in the direction of analogical reasoning is for future scholarship to decide.

Nuclear weapons are special, politically unique, and laden with risk. Nuclear coercion, however empirically uncommon, is possible. Many governments, the United States included, express no shortage of fears about it. And some governments, like Russia today, do not seem to shy away from nuclear threats. Our role as scholars is to understand how nuclear brinkmanship works, why it succeeds or fails, and how to distinguish between useful and unnecessary risk, among many other important questions about the nuclear age. We are so grateful to be doing this work in a field with such careful, impressive, and supportive colleagues.

[13] Paraphrasing Scott D. Sagan, *The Limits of Safety* (Princeton, NJ: Princeton University Press, 1993).