“Computable Rationality, NUTS, and the Nuclear Leviathan,” in The Decisionist Imagination: Democracy, Sovereignty and Social Science in the 20th Century,” ed. by Daniel Bessner and Nicolas Guilhot. New York: Bergahn Books, 2018.

<CN>Chapter 6</CN>

<CT>Computable Rationality, NUTS, and the Nuclear Leviathan</CT>

<CA>S. M. Amadae</CA>

<EPI>Liberal democracy was saved only by nuclear weapons. . . . NATO adopted the doctrine of MAD (mutual assured destruction), according to which even conventional Soviet attacks would be answered by an all-out nuclear strike. . . . Without nukes, there would have been no Woodstock, no Beatles and no overflowing supermarkets.

—Yuval Noah Harari, *Homo Deus: A Brief History of Tomorrow*</EPI>

<EPI>A *rational decision process* will be understood . . . to refer to the *entire* reasoning activity that intervenes between the receipt of a decision stimulus and the ultimate decisions. . . . Such an approach forces rational behavior to be thought of as essentially algorithmic. This makes it natural to model a rational player as a suitably programmed computing machine.

—Ken Binmore, *Essays on the Foundations of Game Theory*</EPI>

<EPI>As the case of nuclear strategy makes clear, there was no straightforward way of adding the mind back into questions of rationality. The appeal of the algorithmic definition of rationality was precisely that it avoided the messes that ensued when one attempted to account for the fact that decision makers had personalities, histories, and prejudices.

—Paul Erickson et al., *How Reason Almost Lost Its Mind*</EPI>

<FL>Recent Cold War historiography has demonstrated how military security concerns drove the development of decision technologies such as game theory, bounded rationality, operations research, and systems analysis.[1] Recollecting such history is important because it enables us to see how these decision theories—game theory in particular—were well-suited to the logistics and strategy of conflict. In their reliance on computational techniques that abandon the need for a conscious subject, prominent decision-theoretic tools sharply broke with previous understandings of human reasoning. Under these new frameworks, neither the agents modeled nor the analysts using the models need an intelligible grasp of the problems they are solving in order to identify solutions. The *mindless* strategic rationality of game theory came to define the nuclear security dilemma and, as this chapter argues, is profoundly entangled with the strategic posture it recommends.</FL>

Nuclear strategy is ostensibly more effective the less intelligible it is to observers because deterrent threats achieve credibility at the price of absurdly endangering constituents with apocalyptic terror. Far from the popular conception of the public and journalists, the nuclear strategy sanctioned by game theory is not MAD (Mutual Assured Destruction), but rather NUTS (Nuclear Utilization Targeting Selection), referring to a nuclear war-fighting posture.[2] Whereas MAD is structured to symmetrically hold nuclear weapons in reserve to multilaterally counter a nuclear attack, NUTS entails developing asymmetric advantage based on coercive bargaining, threatening to introduce nuclear warheads into conflict, and preparing to achieve escalation dominance at all levels of engagement. The mindless quality of game theory is useful in nuclear security because, according to strategic rationality, credible deterrence depends on preparing and intending to wage an omnicidal nuclear war, an action that would strike the casual subject as not only recklessly immoral but also pointless. Thus, the fact that strategic rationality jettisons

intelligibility, or understanding of the problem it aims to solve, enables it to rationalize mobilizing and maintaining resources for ends that defy both moral reasoning and purposive action. This alienating logic of unintentional, and not necessarily intelligible, choice informing nuclear deterrence has been elevated to the standard understanding of classical western instrumental rationality, with detrimental consequences. Such strategic rationality now informs modeling and decision-making spanning from nonhuman actors to individuals' choices and collective action, including the exercise of national sovereignty.

Section one introduces recent historiography to show how Cold War strategic rationality—also referred to as rational deterrence, game theory, and rational choice—erased the role of understanding, or a thinking subject, in rational decision-making. This reconceptualization of rational choice foreclosed on the idea of informed citizens who participate in a process of democratic will formation to oversee executive decisions. Section two investigates the entanglement of the content of game theory with the nuclear security posture it recommends. The substance of game theory grounds nuclear strategy, finding that the most effective form of deterrence is to prepare to fight and win a nuclear war. Section three investigates how prominent social theorists use game theory as a way to unify our understanding of decision-making across actors, encompassing biological organisms, persons, and collectives. Corporate agency, the artificial product of collective action, has been subject to theorization at least since Thomas Hobbes's *Leviathan*. Game theory, which served dual Cold War roles of grounding nuclear deterrence and rethinking the intellectual bases of democratic governance and free markets, offers a way to model and enact corporate agency without the need for actors to be aware of their role in realizing inferred joint goals.[3] Given the limited view of agency and choice offered by strategic rationality, this general acceptance of game theory as a tool for institutional design and public policy formation treats individuals as intentionless subjects and governance as

a technocratic exercise of steering equilibrium outcomes through the introduction of appropriate incentives.[4] The conclusion counters contemporary complacency over game theory and atomic warfare, emphasizing that both strategic rationality and the US nuclear weapons program are more entrenched than ever.[5] As this paper shows, game theory's demotion of decision-making, whether individual or collective, to intentionless action makes it theoretically consistent for nations to prepare omnicidal measures to engage in nuclear combat in which they can only achieve their end by perpetually threatening, or actualizing, the complete destruction of humanity. The alternative is to invest in intentional decision-procedures that sustain individual rationality and interpersonal social commitments that insist on intelligibility as a criterion of purposive agency.

<HDA>How Rationality Lost Its Mind</HDA>

<FL>Nuclear deterrence, as a topic of analysis, is typically left to experts. During the early years of the Cold War, the ground for strategic expertise shifted from seasoned military commanders to blackboard consultants with mastery of abstract decision theory.[6] Although their decision-theoretic tools encompassed operations research, linear programming, and cost-benefit analysis, the most articulated coherent decision theory was game theory, formalized by John von Neumann and Oskar Morgenstern in their 1944 *Theory of Games and Economic Behavior*.[7] Game theory spans both strategic rationality, which refers to individualistically vying against others to achieve one's aims, and a theory of expected utility, which provides a technique to build exhaustive single-scale rankings of outcomes.[8] A rational actor must know how to consistently exhibit preferences over whether, for example, she prefers to travel to London, Singapore, or Cairo in all pairwise comparisons. Her choices must also be consistent over lottery tickets, such as making a rational choice between visiting Singapore for sure, or receiving a 20

percent chance of going to London and an 80 percent chance of traveling to Cairo. Game theory puts forward a comprehensive explanatory, normative, and prescriptive theory of rational choice that could be integrated into systems analysis and operations research, as well as a distinctive and well-codified body of mathematics.[9]</FL>

The embrace of game theory had a defining impact on Cold War strategy and social science writ large. First, without knowledge of rational decision theory, or rational deterrence theory, it has been difficult to contribute to or influence debates over nuclear strategy. By the 1970s, nuclear deterrence was synonymous with strategic rationality, which included the calculated use of probabilistic decision-making.[10] Second, the concept of action proposed by game theory, mandatory individualistic strategic competition, became the widely accepted standard for purposive action throughout American social science and the professional programs of law, public policy, and business by the 1980s.[11] There is no widely accepted alternative for formalizing instrumental rationality, and strategic rationality is notorious for not leaving room for other approaches, such as the deliberative approach of Jürgen Habermas or various forms of collective intention.[12] The reconceptualization of theories of markets and governance according to game theory undermined essential pieces of once-conventional wisdom, such as the rationality of voting and voluntary collective action.[13] Third, rational deterrence theory, which is at best amoral, and potentially even immoral, provided the means to "think the unthinkable," to use the phrase of Herman Kahn.[14] This meant matter-of-factly and routinely contemplating the exercise of nuclear threats in the form of flexible-response, lower-yield, first-strike weapons or second-strike counterforce without regard for the intrinsic value of human life.[15] Thus, fourth, the sense of urgency which rationalizes nuclear deterrence and the command-and-control structure it necessitates supported a theory of agency and intelligence that in principle leaves no distinction between human beings and artificial intelligences.[16] Philip Mirowski makes this point repeatedly:

"Because the computer so readily trespasses upon the self-image of man as the thinking animal, it has become equally commonplace to believe that the mind is nothing more than a machine; that is, it operates like a computer."[17]

Although the first three points have been addressed to varying degrees in the literature, the last point, and its particular poignancy with respect to nuclear deterrence, remains unexplored. The oblivious character of rational choice permeates the nuclear security state at all levels of function, from justifying nuclear doctrine and theories of collective choice to normalizing a specific form of rational action and rendering politics the unintended outcome of individuals' preference satisfaction. Hence, blind steering predicated on formulaic maximization of expected utility rationalizes nuclear strategy despite its unintelligibility to the population it allegedly, although not actually, secures from harm.

Game theorists pursue "the quest for algorithmic rationality . . . the complete and consistent calculation of the strategies of the opponent."[18] A benefit of this seamless transition from human to machine actors is that the maintenance of command and control could be decentralized, and in principle could be distilled to "a complete set of instructions that tells every individual what to do in every conceivable circumstance" that could be executed even without a consciously present human decision-maker.[19] Thus, in erasing the line demarcating a consciously present decider with Kantian autonomy, and in postulating that insofar as people think, strategize, and calculate, so can computers, it then became possible to build a complex, diversified, and extended command-and-control network that would carry the burden of prosecuting modern warfare.[20] The newly minted strategic actor obeys a structure of agency limited by the consistency conditions characterizing rational choice theory.[21] Rationality becomes algorithmic. Computer simulations of action and causal implications replaced

experimentation, and military command-and-control needs led to the "diffusion of the computer throughout all levels of military command structure."[22]

Despite the plethora of research on the entanglement of game theory with Cold War nuclear strategy, mainstream economics and social science more broadly continue to rely on rational choice theory without examining this rich intellectual and contextual heritage.[23] This oversight neglects the potential synergy between game theory and conflict, thus proliferating a model for action that is best suited to antagonistic encounters. However, the development of and fascination with strategic rationality is inseparable from the continued US embrace of nuclear weapons at the apex of its military and strategy of full-spectrum dominance.[24] Strategic rationality loses the quality of *mind*, or the characteristic of intelligible grasp of the problem it is harnessed to solve. The very credibility of nuclear deterrence depends on demonstrating the intention and capability to fight and win a nuclear war among superpowers, even though such victory is impossible due to the cataclysmic destructive power of these weapons. Indeed, the United States and Russian Federation have thousands of thermonuclear bombs when a war fought with even only one hundred could, or even much less is likely to, end known civilization. This provocative stance furthermore entails privileging preparing for, and hence rendering more plausible, nuclear war rather than countering the risks of accidental or intentional nuclear war and pursing means to rescind the use of thermonuclear bombs.[25] Game theory, in other words, helps to normalize a nuclear security state in which a "nuclear eternity" is preferable to actively negating the historically demonstrated tendency to, sooner or later, employ deadly technologies on hand. Hence rational choice rationalizes that we all live under, and strategists contribute to, a regime that permanently equates security to living with the doomsday clock at under three minutes to midnight.[26] Thousands of nuclear weapons remain on an unceasing alert status that

will only be abrogated by either the launching of these weapons or a comprehensive rethinking of the logic underlying this exercise of national power.[27]

By discounting the role of anxiety-ridden nuclear nightmares as a primary background to, and motive underlying, strategic rationality, we fail to come to terms with how profoundly the practice of nuclear deterrence, informed by and justifying game theory, continues to shape the cognitive landscape of human geopolitical and economic systems.[28] Thus with every passing year, we further embrace a concept of intelligence that accepts mindless computation and automates social interactions alongside the pedestrian normalization of nuclear weapons with unfathomable destructive potential.[29] Accommodating nuclear weapons relies on algorithmic decision technologies for rationalizing deterrence, and on hybrid AI-human agency to maintain national sovereignty through command-and-control systems during nuclear war wherein casualties likely will interrupt human command chains and communication channels.[30] Both computable rationality and hybrid AI-human decision-making systems deviate from privileging a sovereign human subject, instead turning to algorithmic rule-following punctuated by randomized number generation to ground strategy and carry out commands. So important did defense analysts estimate the significance of command-and-control that it received an additional 90 billion US dollars over and above the 100 billion that were concurrently spent on military hardware throughout the Cold War years.[31]

Erickson et al.'s volume documents the far-reaching transformation during the Cold War of what had formerly been human reason, which connoted a role for consciousness and possibly even a soul in animating decision-making. The authors observe:

> <EXT>In the two decades following World War II, human reason was reconceptualized as rationality. Philosophers, mathematicians, economists,

> political scientists, military strategists, computer scientists, and psychologists sought, defined, and debated new norms for "rational actors," a deliberately capacious category that included business firms, chess players, the mafia, computers, parents and children, and nuclear superpowers.[32]</EXT>

<FL>Rationality, in other words, became synonymous with what had formerly been deemed the lowest level of cognition: routine calculation. Whereas during the Enlightenment and up until the Cold War, reason had referred to evaluative judgments and the formation of ideas, with the innovation of game theory, rationality merely came to connote calculation subject to rules.[33] These rules, also referred to as algorithms, could be followed by low-level human workers or even machines. Initially, it seemed to 1940s contemporaries that even permitting calculation to be conducted by low-paid laborers denigrated "calculation from a mindful to a mindless exercise."[34] However, with the embrace of game theory, following a rule, with definitive and predictable machine-like exactness, came to be the hallmark of rationality. There is no need for a subject with intelligible grasp of the significance of the instructions or their legitimacy. Moreover, the instructions are self-executing in the sense that they ideally specify a singular (or randomized) outcome independent of the faculties of the computor.[35] Unaccounted for and rendered obsolete is the sense of cognizance that could invent rules, understand them, and apply them to diverse and yet unknown circumstances. This ingenuity typified, for example, Immanuel Kant's attempt to establish rules to differentiate between art and technique, or Isaac Newton's discovery of the rules of motion that then could serve as models for understanding mass, momentum, and energy.[36] Thus, rule-following becomes the mindless and exacting execution of a set of instructions, with the rule itself reduced to an algorithm.[37]</FL>

Rationality loses mindfulness, or a conscious subject with intelligible and existential grasp of the problems it solves.[38] This foreclosure on the merits of understanding represents one side of an enduring philosophical divide over whether mind and intelligibility play any causal role in actualizing behavior. Whereas some theorists argue that humans exhibit freedom of will in making deliberate choices, game theorists put forward a theory of rationality that is wholly determined by a set of instructions, or an algorithm, that is enacted as causal process.[39] This renders intelligence, or purposive agency, in principle subject to automated computation in carbon- or silicon-based systems. This has the added benefit, as game theory textbook authors Duncan Luce and Howard Raiffa observe, of making rational decision-making achievable by human or artificial intelligence.[40]

The definitive aspect of such algorithmic rule-following is its exacting production, over and again, of precisely the same outcome for identical sets of input data. This is the opposite of Ludwig Wittgenstein's approach to rule-following, according to which rules do not by themselves specify the outcome of their application.[41] John Searle, who explicitly acknowledges his philosophical affinity to Wittgenstein, challenges both the classical model of rationality encompassing game theory and the view that AI performs intelligence on par with human agents.[42] In differentiating between human and machine intelligence, he sides with Wittgenstein in pointing out the importance of mindful judgment, based on understanding of intelligibility and recognition of veracity. In contrast, promoters of rational choice and game theory view rationality as one concept, whether exhibited by a human, a nonhuman organism, or a machine. Insofar as game theory represents the orthodox statement of instrumental rationality, Alan Turing's conceptualization of intelligence has prevailed.[43] According to the Church-Turing thesis, there are four criteria of programmable rationality.[44] First, it can be stipulated by a finite set of precise instructions stated in a finite set of symbols. Second, if executed without error, it

always produces the same result in a limited number of steps. Third, it can be completed by a human without machine assistance. Fourth, and most significantly, no intelligible grasp or understanding of the instructions is necessary for the one who calculates.[45] Procedures meeting these criteria are deemed to be logically and mechanically computable.[46]

Cold War rationality was integrally connected to military problem solving. George Dantzig's linear programming, which was mathematically equivalent to von Neumann and Morgenstern's two-person, zero-sum game theory, was developed to solve military logistics problems such as the Berlin Airlift operation.[47] Von Neumann and Morgenstern's game theory was used to provide solutions for military duels and other strategic problems, including nuclear brinkmanship and escalation in the Vietnam War.[48] Herbert Simon's bounded rationality was applied to resource allocation problems.[49] According to Erickson et al., "[Herman] Kahn [author of *On Thermonuclear War*] believed that everything about nuclear war could be understood using the core principles of rational-choice theory."[50] Not only did Thomas Schelling rely on the Prisoner's Dilemma game to develop his theoretical defense of mutual assured destruction, but other Cold War intellectuals working at the Arms Control and Disarmament Agency also used the Prisoner's Dilemma to model Nixon's dilemma of whether or not to escalate in Vietnam, and to structure bargaining during conflict and over weapons in arms control.[51]

Herman Kahn, whose NUTS position on nuclear deterrence prevailed by 1980 and still governs US policy, believed that "axiomatic, formalized rationality was the only way to prepare for the vicissitudes of conflict."[52] Even though RAND's staff grew disenchanted with game theory for its failure to generate a science of warfare by the 1960s, the Pentagon still established the Studies Analysis and Gaming Agency dedicated to gaming simulations to analyze how conflicts may be prosecuted.[53] In 1977, RAND analyst Jack Snyder wrote a report identifying the American national security state as a quintessential rational actor.[54] Nuclear strategists modeled

deterrence itself using game theory, and the body of thought comprising it was and remains strategic rationality.[55] Game theory thus provided decision makers with a means of making onerous policy decisions without emotion, context, and even moral judgment or ethical principle and understanding.[56] Without the military decision-making context to sponsor the initial development of game theory to solve logistical and strategic problems, it is unlikely that rational choice theory would have eventually won over economists who, although skeptical at first, were by the 1980s among its greatest proponents, contributing to its preeminence in the American social sciences by the 1980s.[57] Possibly their enthusiasm followed from their propensity to believe that market coordination arises without actors' intentional effort to achieve mutual prosperity or equilibrium.

Von Neumann and Morgenstern formalized game theory to be an all-encompassing science of choice, applicable to every decision in all conceivable circumstances throughout the lifetime of rational agents.[58] Every consideration of worth, insofar as it impinges on choice, can be, the authors insist, encapsulated into this formalism. Strategic rationality is perfect to capture "calmly aggressive selfishness."[59] Game theory, as a decision technology, only recognizes that consequences matter and demands that only individualistic competition is rational.[60] Therefore, only outcomes, and not principled or ethical means, matter to decision-makers.[61] Even if individuals can form coalitions to compete against other actors, once spoils are obtained, every group must disintegrate into individuals who compete among each other. Although game theory promises to deliver a "crystalline definiteness, generality, and conclusiveness" with which to define rational action, it can only guarantee single solutions in contexts of zero-sum competition.[62] In a non-zero-sum game with multiple possible equilibria, rational actors would no longer have a clear template for action that results in predictable outcomes. To address this obvious limitation, game theorists consider repeating interactions in which a history of

engagement may enable actors to reach a Nash equilibrium: a single point outcome from which no single actor has any incentive to diverge given the choices that all other actors made.

Wittingly or not, by reducing agency to mindless action, game theorists end up endorsing these characteristic limitations. Rules and norms are not deemed to guide action, but rather to merely describe behavioral regularities.[63] We need to ask *how* our dominant form of social science and public policy analysis has so roundly accepted that purposive agency can dispense with the meaning of action as a primary motivator.[64] The fact that we use this theory of rationality to inform our nuclear security policy and to solve the challenge of credible deterrence is inseparable from social scientists' and policy analysts' embrace of its potential to explain all human action, individual and collective. It is no coincidence that the theory of rationality governing the US nuclear arsenal under the nuclear war fighting protocol is also the template for rational action used to inform individual and collective action, as well as theories of sovereign decision-making. As a unified theory of action, strategic rationality accounts for the most prosaic choices and the most consequential: to engage in nuclear combat.

<HDA>MAD, NUTS, and Game Theory</HDA>

<FL>Whereas strategic rationality elides the need for an agent who deliberately makes decisions, nuclear war raises numerous existential quandaries. In particular, the now ever-present potential for nuclear war rewrites the social contract between citizens and government. Daniel Deudney argues that after World War II, the modern social contract predicated on the state's ability to take its citizens out of harm's way was only able to maintain the *appearance* of legitimacy given the impossibility of realizing this promise.[65] The state achieves such an appearance by downplaying the role of nuclear weapons in projecting state power and giving citizens the false impression that its nuclear program is oriented toward an anti-nuclear goal.[66] Whereas most sectors of even

US civil society are under the impression that nuclear policy has been and is structured to the end of deterring nuclear war consistent with a policy, and the reality of, mutual assured destruction (MAD), its actual policy stance is, in fact, NUTS. Deudney notices that which has otherwise gone unrecognized: "the doctrine of mutual assured destruction gradually became supplemental and then supplanted by nuclear utilization targeting theory and strategies—known among nuclear strategists as . . . NUTS."[67] The reality of NUTS—i.e., the US preparedness to engage in all levels of nuclear combat regardless of the risks of error, accident, cyber hacking, proliferation, and escalation—is not made manifestly clear to the public because this would erode the remaining legitimacy from the social contract.[68] Not only does government hold citizens hostage as the potential casualties of the failure of nuclear brinkmanship, but the momentum of the state's nuclear program also proliferates the risk of nuclear engagement far beyond the outcome that would result from jointly and progressively reducing nuclear capabilities and attenuating their alert status.</FL>

There are four bodies of literature that provide evidence that nuclear deterrence and game theory are coextensive with each other, that rational deterrence theory is the same as rational decision theory. One is the historical record of practice and engagement.[69] The second is the internal perspective of international relations theory that makes clear on the one hand that the abstract formal theory of rational deterrence is game theory and on the other that the war gaming simulations used to make arguments for different strategies were game-based models and simulations that were indistinguishable from how actual scenarios would be handled.[70] Third is the literature of strategic practice that provides retrospective vision, making clear the extent to which the problem of deterrence was viewed in terms of strategic rationality.[71] Fourth, and finally, strategic rationality developed in conjunction with theoretically exploring problems of nuclear deterrence, most importantly the credibility problem.[72]

Schelling's *Strategy of Conflict* is important to all four bodies of literature. Those who follow his lead modeled the nuclear security dilemma and arms race using the recalcitrant Prisoner's Dilemma game.[73] Within the context of strategic arms control, as addressed by the US Arms Control and Disarmament Agency, "the prisoner's dilemma *did* become a key theoretical framework for thinking about 'the problem of the bomb,' and the game matrix could leap from the mathematics of optimization to psychological laboratories to problems of war and peace writ large."[74] The significance of this is twofold. On the one hand, the Prisoner's Dilemma game, which was originally developed by Merrill Flood and Melvin Drescher at RAND to reflect two prisoners who are each given the choice by a jailor to confess or remain silent, seemed apt to capture the problem of the "Reciprocal Fear of Surprise Attack," and an arms race.[75] However, on the other hand, the fact that hard-nosed strategic rationality led to an unambiguously suboptimal result in this game signified a deep puzzle at the core of game theory.[76] Given their inability to intentionally collaborate, rational actors achieve suboptimal, and therefore inscrutable, results.

In the Prisoner's Dilemma, a district attorney gives two inmates suspected of having committed a crime each two choices: to remain silent, or to confess. Four possible outcomes obtain, depending on the inmates' individual choices of what to do. If both remain silent, they both serve a short sentence. If both confess, they both serve a long sentence. If one confesses and the other remains silent, then the confessor goes free while the one who remains silent serves a lifetime sentence. In orthodox game theory, the moral of the Prisoner's Dilemma is that each individual is better off defecting, regardless of what the other decides to do. Thus both achieve a worse outcome (a long jail sentence) than if they had been able to cooperate (shorter sentences). If the other remains silent, then it is better to confess, thereby achieving freedom. If the other

confesses, then it is still better to confess to save oneself from the worst outcome of lifetime imprisonment (Table 6.1).

		Joe: Silent	Joe: Talk
Bob	Silent	1 year, 1 year	life in jail, freedom
	Talk	freedom, life in jail	10 years, 10 years

Table 6.1: Standard Prisoner's Dilemma

Applied to the Cold War, either in terms of the preemptive fear of surprise attack (Table 6.2) or a nuclear arms race (Table 6.3), each country is better off pursuing unilateral success, because this secures at best dominance and at worst mutual ruin rather than singular defeat. As the analysis goes, in an arms race, regardless of what the USSR does, the US is better off building arms without limit, because this will either grant it supremacy or prevent it from being dominated. In the Prisoner's Dilemma both actors seek unilateral advantage, securing the worst possible outcome for the other agent. These preferences are consistent with an aggressor or revisionist state.[77]

		USSR: Defend	USSR: Attack
US	Defend	security, security	surrender, victory
	Attack	victory, surrender	destruction, destruction

Table 6.2: Reciprocal Fear of Surprise Attack (see, e.g., Schelling, *Strategy of Conflict*)

		USSR: Disarm	USSR: Arm
US	Disarm	peace, peace	submission, dominance
	Arm	dominance, submission	precarity, precarity

Table 6.3: Nuclear Arms Race (see, e.g., Campbell and Sowden, *Paradoxes of Rationality and Cooperation*)

<FL>However, few seem to have asked whether an endless nuclear arms race, such as what the US has pursued even after the Cold War ended, makes sense, unless it is both possible to achieve military supremacy and meaningful to exercise it. Similarly, with respect to actually waging war, experts note that the US strategic policy meets the specifics of striving to realize being "a nuclear-armed state . . . planning to have the capacity to fight and win a nuclear war by disarming enemies [including Russia with its seven thousand nuclear warheads] with a surprise first strike."[78]</FL>

Game theorists originally argued that one possible way out of the reciprocal fear of surprise attack is to ensure a second-strike counterattack that would neutralize the adversary's advantage with what would end up being mutual assured destruction.[79] The hope is that issuing a threat of counterattack, or the certainty of mutual assured destruction, will be sufficient to secure safety from nuclear attack. However, given the severity of the nuclear threat and the fact that all that would remain for the preemptively attacked power is to issue a counterattack amid the ashes of destruction, game theorists worried that issuing a credible threat of retaliation is as impossible

as counting on the two prisoners to follow through on promises to remain silent. In each case, actors' actual interests diverge from their promises or threats, rendering both incredible: I promise to remain silent, but, according to game theory's model of the Prisoner's Dilemma, I really prefer to go free. I seek to deter nuclear attack and hold my arsenal in reserve to prevent the enemy's attack, yet once the attack has occurred, I no longer have a reason to counterattack, according the game theoretic modeling. Thus, the wrenching problem of nuclear deterrence—how to issue a credible threat of mutual destruction after deterrence has failed—was mapped into the conceptual space of the irresolvable Prisoner's Dilemma paradox and explored in the related problems of the Toxin puzzle.[80] The Prisoner's Dilemma modeling of nuclear deterrence renders a nuclear war-fighting posture necessary because the only way to shore up the credibility of deterrence is to perpetually threaten to wage nuclear war with the aim to win.[81]

In this thought experiment designed by Gregory Kavka, a being with acute predictive powers offers the protagonist one million dollars if she can form an intention today to drink a vial of toxin tomorrow that will cause no long-term harm but would make her feel terribly ill for one day. The supernatural being offers a large sum of money if she can promise, and follow through on that promise, to drink nonlethal poison, even if the money is already in her bank account prior to the act of drinking. No matter how many times theorists investigated this puzzle, they could not identify a means by which the rational person could form an intention at time t_1 to perform an action at time t_2 that would be costly at that time and simultaneously would serve no function in promoting the protagonist's goals. Why would the rational agent actually follow through on drinking the toxin *after* the million dollars is already in the bank? Kavka initially reached that conclusion, and David Gauthier similarly reached the same conclusion a few years later.[82] Both saw this thought experiment as equivalent to that of nuclear deterrence in the context of issuing an "apocalyptic threat."[83]

The toxin puzzle has a similar logical structure to the Prisoner's Dilemma as it relates to nuclear deterrence. Forming a credible threat to retaliate with a devastating second strike is an intention that at the time of acting would both be repugnant and serve no instrumental role in furthering the agent's goal. Thus, in terms of game theoretic analysis, no threat of mutual assured destruction is credible: What could motivate a counterstrike that would only wreak further damage after deterrence has already failed? Following this logic, analysts argued that the only way to ensure that deterrence is credible is to integrate it into a plan to prevail in nuclear war at all levels of engagement.[84] This is consistent with the position of escalation dominance in which the US seeks to have sufficient nuclear capability to win on any rung on the escalation ladder of potential conflict. The fact that game theory was designed to be computable and demands that its instructions for action in principle lack intelligibility to agents carrying them out is well-suited to legitimize nuclear strategy without any need for the actual comprehension of that strategy. This possibly underlies the reality that citizens, who may resist endorsing an unintelligible nuclear procurement and deployment posture, are largely left in the dark about US nuclear policy, naively believing that MAD prevails—i.e., that nuclear weapons are on hand to deter attack, and not to vie for dominance.[85]

It is not yet clear why the policy of flexible response, or being the first to deploy nuclear weapons *prior* to their introduction into arenas of conflict by hostile parties, need be invoked given the US's dominance in conventional forces.[86] According to NUTS strategic doctrine, however, in flagrant disregard for the "nuclear taboo," holding that nuclear weapons are categorically different from conventional arms and should remain off the table on a first-use basis, the US should build small tactical nuclear weapons to be launched in lesser conflicts, thus introducing nuclear weapons in an otherwise conventional war.[87] This strategic stance of flexible response, although consistent with Kahn's escalation dominance and preparedness to fight to win

nuclear combat of any scale conceivable, outright contradicts the position of maintaining nuclear weapons only for the sake of deterring a nuclear strike. Proponents of flexible response perceive the need to thoroughly integrate nuclear arms into the US military's standard operating procedure.[88] Thus, the aim of making nuclear threats credible at the highest level of nuclear attack and counterattack merges with assuring other nations that much less provocation could be met with a nuclear strike, including efforts to achieve their own deterrent posture. Hence, against the idea that nuclear weapons are useless, apart from hopefully serving to avoid nuclear war, flexible response seeks to keep them as an option perpetually on the table either to directly strike terror into, or to signal the US's willingness to engage in nuclear war to, potential adversaries, with the overall aim of compelling them to acquiesce to US demands.[89]

Political maneuvering to appease domestic political supporters by projecting strength, and business interests associated with the nuclear weapons industry, may both contribute to the US's reliance on nuclear weapons in its security protocol. However, rational deterrence theory is at least equally complicit. There are, of course, exceptions: Thomas Schelling, Robert Jervis, and Steven Brams deserve recognition as experts who were fluent in game theory but supported a deterrent position consistent with respecting the nuclear taboo and hoping to gradually reduce nuclear arsenals in staged rounds of reciprocal arms reduction under conditions of mutual surveillance.[90] However, from the perspective of pure strategic rationality, the position of using nuclear threats only to counter nuclear aggression and the aim of reducing arms to a reciprocated minimum deterrent stance have not been successfully sustained within rational choice theory.

The position which has prevailed is unmistakably NUTS. Deudney captures the essence of this view, which he calls nuclear strategism, in his 2007 book *Bounding Power*:

> <EXT>This view takes states as given, conflict as endemic, and holds that the quintessential state activity of preparing for and making war defines world politics regardless of the type of weaponry prevalent. Nuclear strategism postulates that states seeking security in a nuclear world will—and should—prepare themselves to exercise a full range of nuclear use options and seek to gain political advantage from relatively small differences in nuclear force levels.[91]</EXT>

<FL>He sketches out four possible deterrence positions: disarmament, automatic deterrence, institutional deterrence, and the assertive stance of escalation dominance, which he calls nuclear strategism. According to strategic rationality, which is presented as a comprehensive and all-inclusive theory of rational choice, the lack of credibility of following through on a deterrent threat requires demonstrating the capability and intention to wage nuclear war. Additional features of nuclear strategism, also known as NUTS, include manipulating opponents' perception of the risk of nuclear war for various actions, and limited nuclear options, both of which violate a no first use pledge. Originally, forgoing the assurance of renouncing first use served to deter a Soviet land invasion into Europe. However, now that US conventional arms are deemed sufficient to the task, the US still reserves the prerogative of first use because doing otherwise would maintain the categorical separation of nuclear weapons from conventional weapons.[92] This would result in the same problem: without constant announcement of the preparedness and capability to resort to nuclear warfare, deterrence—or, similarly, the power to compel others to comply with the state's demands—would lack credibility. Similarly, with the goals of achieving escalation dominance in any conflict or engaging in coercive bargaining to negotiate settlements, the state relies on its ability to threaten harm. Actors are motivated to avoid harm because it is

directly contrary to their interests. The overall strategy of developing and maintaining the physical and institutional infrastructure to perpetuate nuclear threats itself need not be and is not intelligible because it interminably risks the existence of all human life. However, demanding intelligibility requires exiting the framework of strategic rationality and heeding that meaning, understanding, and recognition can provide reasons and causes for action.</FL>

<HDA>From a Unified Social Ontology to a Unified World Order</HDA>

<FL>Although game theory's roots of application lie in the nuclear dilemma, in the subsequent decades, theorists have used strategic rationality to explain all levels of agency, from the biological and human to the corporate and sovereign. Theorists posit that *all* coherent purposive agency must obey the dictates of rational choice. The state, originally analyzed by Thomas Hobbes as an artificial person or Leviathan, is now formally modeled as the outcome of rational self-interest expressed in markets and politics.[93] In particular, the contemporary American state becomes a nuclearized Leviathan whose greatest purview of decision-making, that exerting the most singular impact on life planetwide, is how and when to project power using nuclear arms. Power, in such a situation, can be understood as the ability to influence others' actions by issuing devastating and credible threats of harm. The rational choice approach to politics, collective action, and warfare is realist in the sense that it shuns moralizing justification of action, rather deferring to the unbridled strategic pursuit of ends.[94] This realism permeates the analytic decision technology of strategic rationality itself because it promotes an understanding and practice of communication that denies intention, meaning, and intelligibility. Instead, communication becomes a science of signaling asymmetric information of what world states obtain, and calculating whether costly signals or cheap talk, true or false, best achieves actors' goals.[95]</FL>

In their conclusion to *How Reason Almost Lost Its Mind*, Erickson et al. propose that worries about nuclear security have receded hand-in-hand with the abandonment of mindless rationality.[96] However, just as the nuclear security dilemma remains as pressing as during any time of the Cold War, rationality has never been as mindless as in the first two decades of the twentieth-first century.[97] Now strategic rationality not only provides a tool for decision-making in national security and logistical planning, as well as a paradigm for rational action, but also serves as a means to provide a unifying logic of action that spans across all levels and types of agency. Thus, rational decision theory is applied to evolving organisms, cognition, language, individual agency, collective agency, hybrid systems of human and robotic actors, formal and informal institutions, and arriving at sovereign decisions. The theorist Fancesco Guala refers to this broad application of rational choice as a "unified social ontology."[98] Economist Herbert Gintis claims that game theory unifies the behavioral sciences, which encompass both animal and human behavior.[99] Rather than being a relic of the past, robotic rationality has become so normal that those promoting alternative forms of reason—Wittgensteinian rule-following, commitment to principles as opposed to only being incentivized by outcomes, solidarity and team reasoning, shared intention, and virtue-inspired action, for example—must persistently advocate to maintain these heterodox actions' academic relevance. The greatest prospective casualties are mind, intention, and intelligibility.

The paradigm of computational rationality comfortably sustains the position that mind is an illusion that will gradually fall away.[100] Some urge us to grasp that we are in the midst of a fourth revolution. Just as humans' beliefs that the earth is in the center of the universe, and that humans are of a different order of creation than animals, have given way, so too will our view that we are categorically distinct from artificial intelligence.[101] However, as soon as we settled on using strategic rationality as a template for human intelligence, we *already* conceded that the

conscious presence of a decision-maker with comprehension of the meaning of actions is irrelevant to rational choice.[102] Furthermore, insofar as this model of rationality presents a normative standard, it accepts the premise that subjects need have no comprehension of how following the rules they are incentivized to follow serves the purpose of coordinating their actions for mutual benefit.[103] Rules, alternatively called equilibria by game theorists, are no longer motivating reasons for action. Instead, they reflect regularities of action caused by individuals' automatic preference satisfaction.

The titles of two recent papers convey the trend toward viewing the brain as an organ functioning on par with a computer: "Computational Rationality: Linking Mechanism and Behavior through Bounded Utility Maximization" and "Computational Rationality: A Converging Paradigm for Intelligence in Brains, Minds, and Machines."[104] This branch of research, referred to as computational rationality, uses expected utility theory as the basis for how organisms, including people, register external states of the world in brain states.[105] Life forms are further postulated to have probability estimates of the likelihood of various action-dependent outcomes and act as "information-processing mechanisms by selecting an optimal program for a bounded machine that maximizes utility in some environment."[106] According to a broad review of this new field published in *Science*, "the fields of artificial intelligence (AI), cognitive science, and neuroscience are reconverging on a shared view of the computational foundations of intelligence."[107] Along the same lines, economist Herbert Gintis argues that game theory can explain action all the way down to the way that cognition functions. He observes that "expected utility maximization is not simply an 'as if' story" because "neuroscientists increasingly find that an aggregate decision making process in the brain synthesizes all available information into a single unitary value."[108] These scientific results, and interpretations of them, remain tentative, but at the same time reflect how the Cold War view of rationality maintains its

currency and continues to blur artificial and human intelligence, relegating subjective understanding of the meaning of acts and events to insignificance.[109]

Game theory has become ubiquitous as an explanatory tool. Game theoretic modeling has been applied extensively in evolutionary biology, and lessons learned from these models have been used to draw conclusions about the possibility of and basis for human cooperation.[110] Game theoretic accounts of individual human and collective action are integrated into public policy.[111] Rational choice theory informs the applied schools of law and economics, public choice, and institutional design.[112] Game theory has also been extensively applied in international relations theory spanning nuclear deterrence and hegemonic stability theory, to analyzing treaties as cheap talk insofar as they are not enforceable.[113]

As a method, game theory simply offers a means to model, and possibly predict, political and economic phenomena. However, from the time of its founding, game theory also stood as a normative theory of rational choice with prescriptive implications for individual and collective decision-making.[114] Thus, if, as its advocates propose, game theory is both a normatively valid theory of instrumental action and a descriptively informative account of agency, then it is not surprising that game theorists find it possible to conceive of all levels of agency, from the nonhuman biological to individual and collective action to superordinate action of organized groups including companies and states, and hybrid human-AI systems, as each embodying the same tenets of rational action.[115] However, the question is to what extent descriptive modeling of agency and normative theories of agency could influence the performative dimension of action. This question can be addressed in depth by focusing on various levels of agency: (1) individual action; (2) large-scale collective action; (3) human agency within institutions designed using (a) analytic rational choice and (b) behavioral experiments that rely on orthodox game theory to

provide a template for perfectly rational conduct; and (4) sovereign bodies which are structured to conform to the theoretical results learned from rational choice modeling.

Game theorists address all these expressions of agency with their singular model for action.[116] Modeling individual competition, collective action, and institutions as games is staple in rational choice. More interesting for purposes here is the application of rational choice theory to the sovereign function of governance, particularly by way of the Prisoner's Dilemma.[117] Revisiting the conceptualization of the state as Leviathan, contemporary political theorists have used game theory to provide an analysis of the state of nature and the maintenance of social order offered by Hobbes.[118] According to this new assessment, strategic rationality governs all purposive conduct and must account for the exercise of collective sovereignty from the micromotives of individual choice.[119] Game theorists view Hobbes as a realist, finding common ground with him in their approach, which associates motivation with rational self-interest void of any types of deontological commitment or political obligation. However, Jeremy Bentham is a more fitting forbearer than Hobbes since the former viewed talk of rights and obligations as nonsense, and the latter had a complex approach to natural law, agents' duty to abide by agreements made, and forging the Artificial Man or commonwealth.[120]

A more in-depth inquiry is necessary to discern the nuances between Hobbes's theory of sovereignty and that offered by game theory.[121] Here, it is sufficient to point out that Hobbes's theory bases the sovereign's authoritative rule on individuals' covenants with one another to lay down their rights to all things. Whereas Hobbes identifies gain, safety, and reputation as individuals' primary, and unlikely rational, motivating forces, game theorists impute to individuals the all-encompassing aim of maximizing expected utility.[122] Where Hobbes's *Leviathan* depends on both subjects' sense of allegiance in exchange for the sovereign's promise of security, and the sword to threaten rule-breakers, game theorists assess individuals' calculated

gain against the cost of rule-breaking to identify the impetus underlying stable governance. For Hobbes, achieving the threshold of security will secure the stability of a sovereign. For rational choice theorists, subjects may vigilantly engage in cost-benefit analysis to determine whether to obey or break laws on a case-by-case basis.[123]

Leaving aside individuals' motives and choices to comply with or challenge the rule of law, rational choice also offers the means to both analyze and rationalize government policies. This is because strategic rationality is useful for both assessing what outcomes will obtain from individual choices, and for formalizing the conditions that must be met for collective political will formation to be legitimate in reflecting individuals' interests. As is by now common lore, Kenneth J. Arrow's generation of his impossibility theorem, proving the irrationality of all collective decision-procedures for more than two individuals selecting among three or more alternatives, was at least in part inspired by his early Cold War assignment of formulating an expected utility function conveying Soviet citizens' preferences over nuclear war.[124] The rich body of social choice theory that followed Arrow's 1951 proof remains central to the design of voting procedures, and has been incorporated into search engines such as Google that rely on ranking systems to reflect multiple-criteria decision problems.[125] Constitutional design has also been methodically studied using game theory to ascertain which fundamental principles will structure a society that upholds attractive conditions, such as citizens' sovereignty, efficient use of resources, and egalitarian commitments.[126] Thus, with respect to institutional design of governing bodies, rational choice theory has played a significant role in providing a means to analyze how to achieve desired outcomes, assuming that strategic rationality best reflects purposive agency.

The particular view of collective agency offered by game theory was perfectly attuned to the early Cold War hostility toward organic collective action and collectivism.[127] Recent

literature has drawn attention to the distrust of intentionally organized collective agency, now referred to as conspiracy. Alfred Moore, for example, argues that Friedrich Hayek opposed social theories which accept the possibility of either collective rationality or intentional collective action.[128] Orthodox game theory, which demands thoroughly individualistic interaction and unintended consequences, was thus a fitting social scientific methodology for followers of Hayek and those who sought to help maintain a liberal market order.

Rational choice offers a comprehensive theory for legitimation. Once we accept orthodox game theory as the best statement of the conditions that individual and collective rational agency must obey, if individuals' preferences are complete and consistent, then the only question about strategic interactions is what equilibrium will obtain.[129] In other words, it comes down to whether an equilibrium outcome is Pareto optimal—i.e., a state of affairs in which no single individual can be made better off without making anyone worse off. Individuals can be compelled to accept less through processes of coercive bargaining in which preemptive threats can incentivize actors to make concessions. Rational choice models generally accept that scarce tangible physical resources, as in natural social dilemmas, are interpersonally salient sources of value that individuals alike seek to amass or consume in competition with others.[130] Orthodox game theory thus resonates with realism by anchoring the source of value underlying individuals' preferences in concrete goods, and accepting as necessary de facto outcomes.

Game theory likewise offers a realist approach to politics. Insofar as it claims to be an all-encompassing theory of rational choice, it invalidates types of agency and collective intentionality that are inconsistent with its assumptions, such as reflexive governance based on voluntary rule-following to realize joint goals informed by shared meanings.[131] Nuclear security policy is ratified by game theory three times over. First, game theory offers no means to theorize about legitimate collective agency other than to provide accounts of how collective will

formation reflects individuals' choices. Nuclear security policy is no different. Achieving a US strategic policy different from NUTS would require a different institutional arrangement, one which must privilege intelligibility, commitment, and reflexive governance over computable rationality and rational choice. Second, given that nuclear war has been modeled as a strategic game, this exercise offers clear guidance for action: specifically, it privileges NUTS over MAD as the surest means to achieve credible deterrence. This guidance follows as a computable solution to a "game" that requires no intelligible grasp of its validity, implications, or assumptions. Third, the entire command-and-control structure, which integrates human decision-makers into a hybrid world of telecommunications, computation, and military hardware, invites agents to operate on a level commensurate with computable rationality. Even with a keen sense of oversight, anthropomorphic qualities are easily lost in humans' contribution to a platform of action in which sustainability is defined as the propagation of a war plan and control over military operations.

The realism underlying the sheer destructive power of existing nuclear arms is inseparable from the realism inherent in the game theoretic assessment of value, typically as fungible, interpersonally transferable utility, and its estimation of power with respect to actors' ability to either directly achieve ends, or achieve ends by manipulating others' choices with incentives.[132] Nuclear security expert Daniel Deudney argues that the US nuclear security state cannot meet its basic legitimation demand, and thus it shies away from openly proclaiming its readiness to engage in nuclear war attendant with its bid for global nuclear supremacy.[133] This may seem to be a distinct consideration from examining how rational choice is used to assess legitimacy; however, it is tightly aligned because the US security state purchases its license to act recklessly abroad and to violate civil liberties at home by threatening "the core of the core, the limited government constitution of the United States itself."[134]

Realism can be constructive in demystifying power relations. However, in the case of rational choice, strategic rationality is restrictive because it denies the coherence of alternative means of acting, such as commitment, intrinsic interests, and joint instrumental agency.[135] This, in effect, ends up rationalizing the de facto exercise of power without due process that must be upheld by commitment to constitutional principles.[136] The exigencies of nuclear deterrence and the nuclearized Leviathan who must channel authority to exercise a planetary death warrant and interface with AI systems rendering intelligibility superfluous, facilitate an expression of agency that must, by design, be complicit with this exercise of authority. The ground for legitimation for this NUTS strategy inherently forfeits guaranteeing humans security from nuclear annihilation. This renders denial and unintelligibility the core ingredients of maintaining the sprawling, resource-engulfing, and perpetually endangering nuclear security complex.

The triumvirate of nuclear arms, strategic rationality, and hybrid command-and-control systems threatens constitutional order.[137] One reason that the realism underlying game theory is insufficient to unmask this unsettling fact is that it coopts our theory of linguistic communication, eviscerating it of intelligibility and meaning. Instead, communication is solely based on a strategic platform of signaling, with calculated decisions over whether to be truthful or deceptive. Language itself, when serving the programmed needs of AI and information technologies, becomes purely syntactic with a minimal semantic bridge that associates a physically detectable world state, such as temperature, with a symbol, in this case a numeric temperature scale. This theory of language is consistent with computational rationality and its application to brains, as well as to integrated information processing theory.[138] Runciman recognizes the role played by Cold War military initiatives in generating the "true roots of the digital revolution."[139] Also, much to the point, he quotes political theorist John Gray who notes, "cyberspace is a site of unceasing warfare." Gray likewise muses that "while they are being used

as weapons, electronic technologies may also be creating a terrain on which intelligent life forms could evolve independent of human control."[140] As game theory is developed to accommodate agent-based modeling in contexts of conflict and economic transactions, the transmission of information as the primary means of communication increases in salience. Warfare is described in terms of "the achievement of Information Superiority with characteristics of gross asymmetries and a diversity of 'players.'"[141] Strategic rationality and information theory merge to be pivotal in the projection of power in twenty-first century warfare.

We must wake up to a potential threat that the rational choice revolution, wedded at the hip to the nuclear security state, poses in its theory of rationality that is individualistic, anchors utility in scarce, competitive resources, and, most importantly, denies intelligibility a role in judgment. A constitutional order that rests on game theory, which necessarily neglects the fact that "intrinsic interests" and commitments may animate actors, must by its analytic structure posit "rationalist, interest-based mechanisms of political behavior."[142] The realism of game theory, which presupposes a common world by imputing that environments afford value independent from and prior to sociability, ignores that human communication is predicated on intention and intelligibility, and that a joint social order regulated by a legitimate rule of law "is both a *mechanism* of political commitment and *itself* a political commitment."[143]

This treatment of communication as thoroughly strategic, which represents a recent trend in research into treating communication as signaling and cheap talk without implications for outcomes, views human actors as on par with computing devices that send signals and calculate whether received signals are credible or not.[144] Jürgen Habermas worried that nondialogic market transactions threaten the life-world of human inhabitants, but even his analysis is insufficient to grasp the full implications of accepting a theory of agency and communication that is wholly strategic.[145] Envisioning a vibrant potential for cultural reproduction through

authentic dialogue can open our eyes to the significance of viewing rationality as mindless and computable, and simultaneously treating communication as reducible to strategic signaling that facilitates strategic action. Thus, we can be sensitive to forms of AI that "automate the social," drawing humans into complexes of interaction and communication as though they are governed by either computable rationality or predictable habituated behavioral patterns.[146] We can also recognize, against the reductionist realism of orthodox game theory, that an intelligible grasp of decision-making, which depends on recognition of validity conditions underlying truth claims and the acknowledgment of other like actors, could be positioned to challenge the cogency of a permanent nuclear Leviathan and rule of law reduced to compelling compliance with incentives.[147] Resonating with Habermas's critical stance, even the legal positivist H. L. A. Hart held that "legal validity ultimately rests on a social practice among officials (if not citizens more broadly) of recognizing and accepting certain rules or practices as obligatory."[148]

A unified ontology consistent with orthodox game theory legitimates a nuclear security stance that holds all of humanity hostage.[149] This theory of action indiscriminately applies to animal action, human action, collective action, governance, and hybrid action systems with human and artificial agents. Throughout, intelligibility, intention, and commitment are jettisoned in favor of a simplistic protocol for action: maximizing expected utility. Its convenience for modeling and managing social institutions, in addition to its eschewal of collective intentionality, make it attractive to policy analysts and useful for structuring environments that can force people to reduce their own expressions of selfhood to conform to the prerogatives of AI.[150] As yet, it remains unknown what possibilities lie in a collective rationality that embraces intelligibility, commitment, and shared intention. However, it is clear that strategic rational action makes it possible to mindlessly contribute to systems that neither place an intrinsic value on life, nor have any concern for whether human civilization is obliterated in a miscarriage of every individual's

natural right "to use his own power, as he will himselfe, for the preservation of his own . . . Life; and consequently of doing any thing, which in his own Judgement, and Reason, hee shall conceive to be the aptest means thereunto."[151]

<HDA>Conclusion</HDA>

<FL>The unified social ontology of game theory, its normative clout as the dominant standard of rational choice, and its long-established role in policy-making underlie the contemporary world order in which the United States pursues escalation dominance, negates a no-first-use pledge, and develops flexible nuclear alternatives in the form of limited nuclear options.[152] This is not to say that critiques of this unified social ontology and unified world order do not exist. Nye's *Nuclear Ethics* and Deudney's *Bounding Power* present views of agency and collective security that are not limited to strategic rationality. Nye contrasts strategic rationality with moral judgment. Deudney emphasizes the universal pursuit of "security from predatory violence and for political liberty."[153] Finding "human corporeal vulnerability and the fundamental value of life as prerequisite for all other ends" as a basis for common ground, he calls for a federated international solution to the nuclear security dilemma.[154] These points all resonate as well with Elaine Scarry's *Thermonuclear Hegemony*, which argues that legitimate government arises from the visceral and shared recognition of what it means to be harmed. The social contract must both be intelligible to its constituents and protect its members from wrongful harm to be legitimate.</FL>

Thus, it shades from ludicrous to alarming that the decision technologies that were first developed to address nuclear security, then applied to modeling all manner of individual and collective decisions, and developed to construct public policies and design institutions, presuppose at the outset that there is no comprehending subject in the decision-maker's seat.[155]

The fact of mindless rationality is consistent with its design. Game theory's prominence is inseparable from its generative application to the nuclear security dilemma that selected the strategy of NUTS—preparing to fight and win a nuclear war—as the only way to solve the supposed paradox of MAD deriving from the lack of credibility of a mutually suicidal retaliatory strike. But any fact, including the reality of mutual assured destruction, cannot be a paradox. The paradox of deterrence was generated internal to rational decision theory, which, having jettisoned intelligibility as a distinguishing feature of human reasoning, then proposed an incredible and omnicidal solution to its logical reconstruction of a security dilemma. As Deudney observes, given how rapid the decision to pull the nuclear trigger would be, strategic rationality not only takes us from MAD to NUTS, but further to DEAD: Destructive-Entrusted Automation Devices.[156] There is no pretense of a conscious decision-maker in strategic games. Moreover, under US strategic protocol and also that of the Russians, potentially and even necessarily the decision to actualize a nuclear confrontation will be assisted with, if not implemented by, algorithmic decision machines.[1] This current situation invites us to question whether we should leave the future of humanity to strategic and computable rationality, which presents the greatest challenges as those of maintaining credibility and strategic parity in a world of overkill nuclear strike capability. We are thus poised to pursue a security posture that is intelligible and serves humans worldwide in affording security by minimizing the chances of intentional and accidental nuclear conflict.

<FL>S. M. Amadae is the author of *Prisoners of Reason* (Cambridge University Press, 2016) and the award winning *Rationalizing Capitalist Democracy* (University of Chicago Press, 2003).

---

[1] Geist and Lohn, "How Might Artificial Intelligence Affect the Risk of Nuclear War?," 18.

This research trajectory investigates the political theoretical implications of rational decision theory for classical liberal and neoliberal theories and practices. Amadae holds appointments as a University Lecturer of World Politics in the Dept. of Economic and Political Studies, University of Helsinki; Associate Professor of international political economy in Swansea University's Department of Politics and International Relations and Research Affiliate in Science, Technology, and Society Studies, Massachusetts Institute of Technology.<FL>

<HDA>Bibliography</HDA>

Amadae, S. M. "Arrow's Impossibility Theorem and the National Security State." *Studies in History and Philosophy of Science Part A* 36, no. 4 (2005): 734–43.

_____. "Game Theory, Cheap Talk and Post-Truth Politics," *Journal for the Theory of Social Behavior,* May 2018, 1-24.

———. "The Computability of Rational Choice." *What Robots Can and Should Do*. Proceedings of Robo-Philosophy: IOS Press, 2016.

———. *Prisoners of Reason: Game Theory and Neoliberal Political Economy*. New York: Cambridge University Press, 2016.

———. *Rationalizing Capitalist Democracy: Cold War Origins of Rational Choice Liberalism*. Chicago, IL: University of Chicago Press, 2003.

Arrow, Kenneth. *Social Choice and Individual* Values. New Haven: Yale University Press, 1951.

Axelrod, Robert. *Evolution of Cooperation*. New York: Basic Books, 1984.

Axelrod, Robert, and R. O. Keohane. "Achieving Cooperation." In *Neorealism and Neoliberalism: The Contemporary Debate*, edited by David A. Baldwin, 85–115. New York: Columbia University Press, 1993.

Bacharach, Michael, Natalie Gold, and Robert Sugden. *Beyond Individual Choice: Teams and Frames in Game Theory*. Princeton, NJ: Princeton University Press, 2006.

Ball, Desmond. "US Strategic Forces: How Would They Be Used?" *International Security* 7, no. 3 (1982): 31–60.

Ball, Desmond, and Robert C. Toth. "Revising the SIOP: Taking War-Fighting to Dangerous Extremes." *International Security* 14, no. 4 (1990): 65–92.

Becker, Gary S. *The Economic Approach to Human Behavior*. Chicago, IL: University of Chicago Press, 1978.

Bentham, Jeremy. *A Fragment on Government*. New York: The Lawbook Exchange, Ltd., 2001.

Berry K, Lewis P, Pélopidas B, Sokov N, Wilson W. Delegitimizing Nuclear Weapons: Examining the Validity of Nuclear Weapons. Monterey, CA: Monterey Institute for International Studies. 2010 May..

Betts, Richard. "Should Strategic Studies Survive?" *World Politics* 50, no. 1 (1997): 9–11.

Binmore, Ken. *Essays on the Foundations of Game Theory.* Oxford: Blackwell, 1990.

———. *Natural Justice*. Oxford: Oxford University Press, 2005.

Binmore, Ken, and Hyun Song Shin. "Algorithmic Knowledge and Game Theory." *Knowledge, Belief, and Strategic Interaction* 141 (1992): 154.

Brams, Steven J. *Superpower Games: Applying Game Theory to Superpower Conflict*. New Haven, CT: Yale University Press, 1985.

Brandt, Felix, Vincent Conitzer, and Ulle Endriss. "Computational Social Choice." In *Multiagent Systems,* edited by Gerhard Weiss, 213-283 Cambridge, MA: MIT Press, 2012.

Bringsjord, Selmer, Naveen Sundar Govindarajulu, Simon Ellis, Evan McCarty, and John Licato. "Nuclear Deterrence and the Logic of Deliberative Mindreading." *Cognitive Systems Research* 28 (2014): 20–43.

Buchanan, James M. *The Limits of Liberty: Between Anarchy and Leviathan*. Chicago, IL: University of Chicago Press, 1975.

Buchanan, James M., and Gordon Tullock. *The Calculus of Consent*. Ann Arbor: University of Michigan Press, 1962.

Campbell, Richmond, and Lanning Sowden, eds. *Paradoxes of Rationality and Cooperation: Prisoner's Dilemma and Newcomb's Problem*. British Columbia: UBC Press, 1985.

Cebrowski, A. "Network Centric Warfare and Information Superiority." Keynote address from the proceedings, Royal United Services Institute (RUSI) conference "C4ISTAR Achieving Information Superiority." Whitehall, London: RUSI, 2000.

Church, Alfred. "A Set of Postulates for the Foundation of Logic." *Annals of Mathematics*, second series 33 (1936): 346–66.

Churchland, Paul M. *Matter and Consciousness*. Rev. ed. Cambridge, MA: MIT Press, 1988.

Coeckelbergh, Mark. "The Automation of the Social: What Robots Can Teach Us about the Social." In *Robo-Philosophy: Philosophy of, for, and by Social Robotics*, edited by Johanna Seibt, Raul Hakli, and Marco Nørskov, ##–##. Cambridge, MA: MIT Press, forthcoming 2018.

Crowther-Heyck, Hunter. *Age of System: Understanding the Development of Modern Social Science*. Baltimore, MD: Johns Hopkins University Press, 2015.

———. *Herbert A. Simon: The Bounds of Reason in Modern America*. Baltimore, MD: JHU Press, 2005.

Deudney, Daniel. *Bounding Power: Republican Security Theory from the Polis to the Global Village*. Princeton, NJ: Princeton University Press, 2007.

———. "Omniviolence, Arms Control, and Limited Government," in Jeffrey Tulis and Stephen Macedo, eds., *The Limits of Constitutionalism* (Princeton University Press, 2010), pp.297-316..

———. "Political Fission: State Structure, Civil Society, and Nuclear Security Politics in the United States." In *On Security*, edited by Ronnie Lipschutz, 87–123. New York: Columbia University Press, 1995.

———. "Whole Earth Security: A Geopolitics of Peace." *Worldwatch Paper 55*. Washington, DC: Worldwatch Institute, 1983.

Erickson, Paul. *The World the Game Theorists Made*. Chicago, IL: University of Chicago Press, 2015.

Erickson, Paul, Judy L. Klein, Lorraine Daston, Rebecca Lemov, Thomas Sturm, and Michael D. Gordin. *How Reason Almost Lost Its Mind: The Strange Career of Cold War Rationality*. Chicago, IL: University of Chicago Press, 2013.

Ellsberg, Daniel. "Theory of the Reluctant Duelist." *The American Economic Review* 46, no. 5 (1956): 909–23.

Espinola-Arredondo, Ana, and Felix Munoz-Garcia. "Free-Riding in International Environmental Agreements: A Signaling Approach to Non-enforceable Treaties." *Journal of Theoretical Politics* 23, no. 1 (2011): 111–34.

Farrell, Joseph, and Matthew Rabin. "Cheap Talk." *The Journal of Economic Perspectives* 10, no. 3 (1996): 103–18.

Floridi, Luciano. *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality*. Oxford: Oxford University Press, 2014.

Freedman, Lawrence. *The Evolution of Nuclear Strategy*. New York: St. Martin's Press, 1981.

Galison, Peter. "The Ontology of the Enemy: Norbert Wiener and the Cybernetic Vision." *Critical Inquiry* 21, no. 1 (1994): 228–66.

Gartzke, Erik, and Jon R. Lindsay. "Thermonuclear Cyberwar." *Journal of Cyber Security* 3, no. 1 (March 2017): 37–48. doi: 10.1093/cybsec/tyw017.

Gauthier, David. "Assure and Threaten." *Ethics* 104, no. 4 (1994): 690–721.

———. "Deterrence, Maximization, and Rationality." *Ethics* 94, no. 3 (1984): 474–95.

———. *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*. Oxford: Oxford University Press, 1969.

———. "Rethinking the Toxin Puzzle." In *Rational Commitment and Social Justice: Essays for Gregory Kavka*, edited by Jules L. Coleman, Christopher W. Morris, and Gregory S. Kavka, 47–58. Cambridge: Cambridge University Press, 1998.

Gavin, Francis J. *Nuclear Statecraft*. Ithaca, NY: Cornell University Press, 2012.

Geist, Edward and Andrew J. Lohn. "How Might Artificial Intelligence Affect the Risk of Nuclear War?" Santa Monica, CA: RAND Corporation, 2018. https://www.rand.org/pubs/perspectives/PE296.html.

Gerson, Michael S. "No First Use: The Next Step for US Nuclear Policy." *International Security* 35, no. 2 (2010): 7–47.

Gershman, Samuel J., Eric J. Horvitz, and Joshua B. Tenenbaum. "Computational Rationality: A Converging Paradigm for Intelligence in Brains, Minds, and Machines." *Science* 349, no. 6245 (2015): 273–78.

Ghamari-Tabrizi, Sharon. *The Worlds of Herman Kahn. The Intuitive Science of Thermonuclear War.* Cambridge, MA: Harvard University Press, 2005.

Gilbert, Margaret. *Joint Commitment: How We Make the Social World.* Oxford: Oxford University Press, 2015.

Gintis, Herbert. *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*. Princeton, NJ: Princeton University Press, 2009.

Giocoli, Nicola. "Do Prudent Agents Play Lotteries? Von Neumann's Contribution to the Theory of Rational Behavior." *Journal of the History of Economic Thought* 28, no. 1 (2006): 95–109.

Gray, John. *The Soul of the Marionette: A Short Enquiry into Human Freedom.* London: Allen Lane, 2015.

Guala, Francesco. *Understanding Institutions: The Science and Philosophy of Living Together*. Princeton, NJ: Princeton University Press, 2016.

Guala, Francesco, and Frank Hindricks. "A Unified Social Ontology." *The Philosophical Quarterly* 65, no. 259 (2015): 177–201.

Habermas, Jürgen. *Between Facts and Norm, Contribution to Discourse Theory of Law and Democracy*. Translated by William Rehg. Cambridge, MA: MIT Press, 1996.

———. *The Theory of Communicative Action.* Vol. 1, *Reason and the Rationalization of Society*, translated by Thomas McCarthy. Boston, MA: Beacon, 1984.

———. *The Theory of Communicative Action.* Vol. 2, *Lifeworld and System: A Critique of Functionalist Reason*, translated by Thomas McCarthy. Cambridge: MIT Press, 1987.

Harari, Yuval Noah. *Homo Deus: A Brief History of Tomorrow*. New York: Random House, 2016.

Hardin, Russell. *Collective Action*. Baltimore, MD: Johns Hopkins University Press, 1982.

Hargreaves Heap, Shaun, and Yanis Varoufakis. *Game Theory: A Critical Introduction*. Abingdon: Routledge, 2004.

Hausman, Daniel H. *Preference, Value, Choice and Welfare*. Cambridge, UK: Cambridge University Press, 2011.

Heath, Joseph. *Communicative Action and Rational Choice*. Cambridge, MA: MIT Press, 2001.

Herken, Gregg. *Counsels of War*. New York: Alfred A. Knopf, 1985.

Hobbes, Thomas. *Leviathan.* Rev. ed. Edited by Richard Tuck. Cambridge, UK: Cambridge University Press, 1996.

Jardini, David R. "Out of the Blue Yonder: The RAND Corporation's Diversification into Social Welfare Research, 1946–1968." PhD diss., Carnegie Mellon University, 1996.

Jervis, Robert. "Deterrence Theory Revisited." *World Politics* 31, no. 2 (1979): 289–324.

———. *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon*. Ithaca, NY: Cornell University Press, 1989.

Jones, Todd. "Uncovering 'Cultural Meaning': Problems and Solutions." *Behavior and Philosophy* (2004): 247–68.

Kahn, Herman. *Thinking the Unthinkable*. New York: Avon, 1962.

Kaplan, Fred. *The Wizards of Armageddon*. Palo Alto, CA: Stanford University Press, 1991.

Kavka, Gregory S. "Deterrence, Utility, and Rational Choice." *Theory and Decision* 12, no. 1 (1980): 41–60.

———. *Hobbesian Moral and Political Theory*. Princeton, NJ: Princeton University Press, 1986.

———. *Moral Paradoxes of Nuclear Deterrence*. Cambridge, UK: Cambridge University Press, 1987.

———. "Some Paradoxes of Deterrence." *The Journal of Philosophy* 75, no. 6 (1978): 285–302.

———. "The Toxin Puzzle." *Analysis* 43, no. 1 (1983): 33–36.

Krause, Keith. "Rationality and Deterrence in Theory and Practice." In *Contemporary Security and Strategy*, edited by Craig A. Synder, 120–49. New York: Routledge, 1999.

Kristensen Hans M., Matthew McKinzie, and Theodore A. Postol "How US Nuclear Force Modernization Is Undermining Strategic Stability: The Burst-Height Compensating Super-Fuze." *Bulletin of the Atomic Scientists,* vol. 1(2017) Available online at https://thebulletin.org/2017/03/how-us-nuclear-force-modernization-is-undermining-strategic-stability-the-burst-height-compensating-super-fuze/.

Lackey, Douglas P. "The American Debate on Nuclear Weapons Policy." *Analyse & Kritik* 9, no. 1–2 (1987): 7–46.

Lebow, Ned. *Coercion, Cooperation and Ethics in International Relations.* London: Routledge, 2007.

Leonard, Robert J. "Creating a Context for Game Theory." In *Toward a History of Game Theory*, edited by E. Roy Weintraub, 29–76. Durham, NC: Duke University Press, 1992.

Levinson, Daryl J. "Parchment and Politics: The Positive Puzzle of Constitutional Commitment." *Harvard Law Review* (2011): 657–746.

Lewis, Richard L., Andrew Howes, and Satinder Singh. “Computational Rationality: Linking Mechanism and Behavior through Bounded Utility Maximization.” *Topics in Cognitive Science* 6, no. 2 (2014): 279–311.

Luce, R. Duncan, and Howard Raiffa. *Games and Decisions: Introduction and Critical Survey*. New York: Courier Corporation, 1957.

Markey-Towler, Brendan. “I, Roboticus Oeconomicus: The Philosophy of Mind in Economics, and Why It Matters.” *Cambridge Journal of Economics* 41, no. 1 (2017): 203–37.

Mayberry, John P. “The Notion of Threat.” In *Models of Gradual Reduction of Arms*, co-authored by Robert J. Aumann, John C. Harsanyi, John P. Mayberry, Michael Maschler, Herbert E. Scarf, Reinhard Selten, and Richard Stearns, 29-46, Washington D.C.: Arms Control and Disarmament Agency, ACDA/ST-116, 1967.

Maynard Smith, John. *Evolution and the Theory of Games*. Cambridge, UK: Cambridge University Press, 1982.

McCloskey, Deirdre N. *Bourgeois Equality: How Ideas, not Capital or Institutions, Enriched the World*. Chicago, IL: University of Chicago Press, 2016.

Mele, Alfred R. *Motivation and Agency*. Oxford: Oxford University Press, 2003.

Mirowski, Philip. *Machine Dreams: Economics Becomes a Cyborg Science*. Cambridge, UK: Cambridge University Press, 2002.

Moffat, J. *Command and Control in the Information Age: Representing Impact*. London: The Stationary Office, 2002.

———. *Complexity Theory and Network Centric Warfare.* CCRP Publications Series, 2003. http://internationalc2institute.org/ccrp-books/.

Moore, Alfred. “Hayek, Conspiracy, and Democracy.” *Critical Review* 28, no 1 (2016): 44–62. DOI: 10.1080/08913811.2016.1167405.

Mueller, Dennis, C. “Public Choice III.” New York: Cambridge University Press, 2003.

Myerson, Roger B. *Game Theory: Analysis of Conflict*. Cambridge, MA: Harvard University Press, 1991.

Nash, John. “Non-cooperative Games.” *Annals of Mathematics* (1951): 286–95.

New York Times Editorial Board. “The Finger on the Nuclear Button.” *New York Times*. 5 February 2017. https://www.nytimes.com/2017/02/05/opinion/the-finger-on-the-nuclear-button.html?_r=0.

North, Douglass C. *Institutions, Institutional Change and Economic Performance*. Cambridge, UK: Cambridge University Press, 1990.

Nye, Joseph S. *Nuclear Ethics*. New York City: Simon and Schuster, 1988.

Pelopidas, Benoît. “A Bet Portrayed as a Certainty: Reassessing the Added Deterrent Value of Nuclear Weapons.” *The War That Must Never Be Fought: Dilemmas of Nuclear Deterrence*, ed. by Shultz, George P., and James E. Goodby, eds. Stanford, CA: Hoover Press, 2015.

, 2015.

———. *“*The Birth of Nuclear Eternity,*” The Future: interdisciplinary perspectives*, ed. by Jenny Andersson and Sandra Kemp. Oxford: Oxford University Press, forthcoming *2019*.

———. “The Oracles of Proliferation: How Experts Maintain a Biased Historical Reading That Limits Policy Innovation.” *Nonproliferation Review* 18, no. 1 (2011): 297–314.

———. “The Unbearable Lightness of Luck: Three Sources of Overconfidence in the Manageability of Nuclear Crises.” *European Journal of International Security* 2, no. 2 (July 2017): 240–62.

Pettit Philip. “Deliberative Democracy and the Discursive Dilemma.” *Noûs*. 2001 Oct;35:268-99.

———. *Rules, Reasons, and Norms*. Oxford: Clarendon Press, 2002.

Rapoport, Anatol, and Albert M. Chammah. *Prisoner’s Dilemma: A Study of Conflict and Cooperation*. Ann Arbor, MI: University of Michigan Press, 1965.

Rawls, John. *Theory of Justice*. Cambridge, MA: Belknap Press, 1970.

Reynolds, Andrew, ed. *The Architecture of Democracy: Constitutional Design, Conflict Management, and Democracy*. Oxford: Oxford University Press, 2002.

Roth, Alvin E. “The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics.” *Econometrica* 70, no. 4 (2002): 1341–78.

Runciman, David. “Political Theory and Real Politics in the Age of the Internet.” *The Journal of Political Philosophy* 25, no. 1 (2016): 3–21.

Scarry, Elaine. *Thermonuclear Monarchy: Choosing between Democracy and Doom*. New York: W. W. Norton & Company, 2014.

Schelling, Thomas C. *Arms and Influence*. New Haven, CT: Yale University Press, 1966.

_____. “The Life You Save May Be Your Own.” In *Problems in Public Expenditure Analysis*, ed. by Samuel B. Chase, Jr. Washington D.C.: Brookings Institution 1968.

———. “An Astonishing Sixty Years: The Legacy of Hiroshima.” *The American Economic Review* 96, no. 4 (2006): 929–37.

———. *Micromotives and Macrobehavior*. New York: W. W. Norton & Company, 2006.

———. *Strategy of Conflict*. Cambridge, MA: Harvard University Press, 1960.

Schweller, Randall L. “Neorealism’s Status Quo Bias: What Security Dilemma?” *Security Studies* 5, no. 3 (1996): 90–121.

Scowcroft, Brent. "C3 Systems for the President and Military Commanders." In *National Security Issues 1981 Symposium: Strategic Nuclear Policies, Weapons, and the C3 Connection*. Ed. by Drexel M. Ace. Bedford, MA: Mitre Corporation, Mitre Document M82-30, pp. 93-98.

Searle, John R. "Insight and Error in Wittgenstein." *Philosophy of the Social Sciences* 46, no. 6 (2016): 527–47.

———. *Making the Social World: The Structure of Human Civilization.* Oxford: Oxford University Press, 2010.

———. "Minds, Brains and Programs." *Behavioral and Brain Sciences* 3 (1980): 417–57.

———. *Rationality in Action*. Cambridge, MA: MIT Press, 2001.

———. "Wittgenstein and the Background." *American Philosophical Quarterly* 48, no. 2 (2011): 119–28.

Sen, Amartya K. "Goals, Commitment, and Identity." *Journal of Law, Economics and Organization* 1 (1985): 2.

Shetterly, Margot Lee. *Hidden Figures: The Untold Story of the African American Women Who Helped Win the Space Race*. London: Harper Collins UK, 2016.

Shlapak, David A., and Michael Johnson. *Reinforcing Deterrence on NATO's Eastern Flank: Wargaming the Defense of the Baltics.* Santa Monica, CA: RAND Corporation, 2016. https://www.rand.org/pubs/research_reports/RR1253.html.

Shoham, Yoav, and Kevin Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge, UK: Cambridge University Press, 2008.

Snidal, Duncan. “The Limits of Hegemonic Stability Theory.” *International Organization* 39, no. 4 (1985): 579–614.

Snyder, Glenn H. “‘Prisoner’s Dilemma’ and ‘Chicken’ Models in International Politics.” *International Studies Quarterly* 15, no.1 (1971): 66–103.

Snyder, Jack L. *The Soviet Strategic Culture: Implications for Limited Nuclear Options*. Santa Monica, CA: RAND Corporation, 1977.

Stern, Nicholas Herbert. *The Economics of Climate Change: The Stern Review*. Cambridge, UK: Cambridge University Press, 2007.

Thomas, William. *Rational Action: The Sciences of Policy in Britain and America, 1940–1960*. Cambridge, MA: MIT Press, 2015.

Tononi, Giulio, Boly, M., Massimini, M., & Koch, C. “Integrated Information Theory: From Consciousness to Its Physical Substrate.” *Nature Reviews Neuroscience* 17, no. 7 (2016): 450–61.

Trachtenberg, Marc. “Social Scientists and National Security Policymaking.” Paper presented at Notre Dame University, 22–23 April 2010. http://www.sscnet.ucla.edu/polisci/faculty/trachtenberg/cv/notre%20dame(2010).pdf.

Trivers, Robert L. “The Evolution of Reciprocal Altruism.” *The Quarterly Review of Biology* 46, no. 1 (1971): 35–57.

Tuck, Richard. *Free Riding*. Cambridge, MA: Harvard University Press, 2008.

Tuomela, Raimo. *Social Ontology: Collective Intentionality and Group Agents*. Oxford: Oxford University Press, 2013.

Turing, Alan. “Intelligent Machinery.” National Physical Laboratory Report. In *Machine Intelligence 5*, edited by B. Meltzer and D. Michie. Edinburgh: Edinburgh University

Press, 1969. Digital facsimile viewable at http://www.AlanTuring.net/intelligent_machinery.

U.S. Department of Defense. *Joint Vision 2020*. Washington, DC, 2000. https://web.archive.org/web/20011129104507/.

Volmar, Daniel. "The Power of the Atom: US Command Control and Communications, 1945–1965." PhD diss., Harvard University, 2017.

von Neumann, John, and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 1947.

Wittgenstein, Ludwig. *Philosophical Investigations*. Trans. GEM Anscombe. New York: The Macmillan Company, 1999.

Yann, Chevaleyre, Ulle Endriss, Jérôme Lang, and Nicolas Maudet. "A Short Introduction to Computational Social Choice." *International Conference on Current Trends in Theory and Practice of Computer Science*. Berlin, Heidelberg: Springer, 2007.

Zagare, Frank C. "Classical Deterrence Theory: A Critical Assessment." *International Interactions* 21, no. 4 (1996): 365–87.

Zagare, Frank C., and D. Marc Kilgour. *Perfect Deterrence*. Cambridge, UK: Cambridge University Press, 2000

Notes

---

1. Paul Erickson, Judy L. Klein, Lorraine Daston, Rebecca Lemov, Thomas Sturm, and Michael D. Gordin, *How Reason Almost Lost Its Mind: The Strange Career of Cold War Rationality* (Chicago, IL: University of Chicago Press, 2013); Paul Erickson, *The World the Game Theorists*

Made (Chicago, IL: University of Chicago Press, 2015); William Thomas, *Rational Action: The Sciences of Policy in Britain and America, 1940–1960* (Cambridge, MA: MIT Press, 2015).

2. Yuval Noah Harari, *Homo Deus: A Brief History of Tomorrow* (New York: Publisher, 2016).

3. On the former, see S. M. Amadae, *Rationalizing Capitalist Democracy: Cold War Origins of Rational Choice Liberalism* (Chicago, IL: University of Chicago Press, 2003); Alvin E. Roth, "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics," *Econometrica* 70, no. 4 (2002): 1341–78.

4. Philip R. Cohen and Hector J. Levesque, "Intention Is Choice with Commitment," *Artificial Intelligence* 42, no. 2–3 (1990): 213–61, see esp. 217, 257.

5. On the former, see, for example, Erickson et al., *How Reason Almost Lost its Mind.*

6. Amadae, *Prisoners of Reason*, 76–79; Thomas, *Rational Action*, 135.

7. On the former, see Thomas, *Rational Action*, 201, 203*;* Philip Mirowski, *Machine Dreams: Economics Becomes a Cyborg Science* (New York: Cambridge University Press, 2002); Amadae, *Rationalizing Capitalist Democracy;* Erickson, *The World the Game Theorists Made*.

8. John von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior* (Princeton, NJ: Princeton University Press, 1947).

9. Keith Krause, "Rationality and Deterrence in Theory and Practice," in *Contemporary Security and Strategy*, ed. Craig A. Snyder, 120–49 (New York: Routledge, 1999).

10. See, for example, Thomas Schelling, *Strategy of Conflict* (Cambridge, 1960). The most optimistic statement of Schelling's influence apart from Schelling himself comes from Richard Ned Lebow who writes that "his ideas have also had considerable impact in the policy community, where they continue to shape the way in which policymakers seek to influence adversaries." In *Coercion, Cooperation and Ethics in International Relations* (London: Routledge, 2007), 255. [The essay was originally published in 1996]. Marc Trachtenberg,

Richard Betts, and Francis J. Gavin are much more sceptical of Schelling's influence and critical of his overstatement on this issue. See Marc Trachtenberg, "Social Scientists and National Security Policymaking," paper presented at Notre Dame University, 22–23 April 2010, 8 available at http://www.sscnet.ucla.edu/polisci/faculty/trachtenberg/cv/notre%20dame(2010).pdf; Richard Betts, "Should Strategic Studies Survive?" *World Politics* 50, no. 1 (1997): 9–11; Francis J. Gavin, *Nuclear Statecraft* (Ithaca, NY: Cornell University Press, 2012), 4 and chapter 6 on the limited impact of game theory on American policymakers. My argument is compatible with this caution about overemphasizing the role of game theory in constructing nuclear policy because, as Schelling originally observed, game theory has impact as more of a mentality and approach than the demand to provide a formal payoff matrix and decision tree for every choice. Thus, so long as game theory provides the widely accepted standard for instrumental rationality, then its overarching approach to rational choice remains implicit in mainstream strategic policies.

11. Daniel Ellsberg, "Theory of the Reluctant Duelist," *The American Economic Review* 46, no. 5 (1956); Amadae, *Rationalizing Capitalist Democracy*; Mirowski, *Machine Dreams*, 114–15. In Mirowski's close reading of the development of game theory within the context of military purpose and purview, game theory—along the lines of what would become the orthodox Nash equilibrium approach of mutual-best-reply, which postulated a calculable (machine representable) agent with hallmark consistent preferences—increasingly yielded ground to John von Neumann's post-1940s view of mechanism design as itself a form of algorithmic governance that spans from military command and control to markets (94–152, 536–45).

12. Jürgen Habermas, *The Theory of Communicative Action,* vol. 1, *Reason and the Rationalization of Society*, trans. Thomas McCarthy (Boston, MA: Beacon, 1984); Jürgen Habermas, *The Theory of Communicative Action*, vol. 2, *Lifeworld and System: A Critique of*

Functionalist Reason, trans. Thomas McCarthy (Cambridge: MIT Press, 1987). On Habermas, see Joseph Heath, *Communicative Action and Rational Choice* (Cambridge, MA, 2001); Margaret Gilbert, *How We Together Make the Social* World (New York, 2015); John R. Searle, *Making the Social World: The Structure of Human Civilization* (Oxford, UK, 2010); Raimo Tuomela, *Social Ontology: Collective Intentionality and Group Agents* (New York, 2013). Alternatives within the rationality paradigm may include Herbert Simon's bounded rationality (Mirowski, *Machine Dreams*, 529–32; Crowther-Heyck, *Herbert A. Simon: The Bounds of Reason in Modern America*. Baltimore, MD: JHU Press, 2005) and Norbert Wiener's cybernetics (Peter Galison, "The Ontology of the Enemy: Norbert Wiener and the Cybernetic Vision," *Critical Inquiry* 21, no. 1 [1994]: 228–66), but Mirowski makes clear that not only were these programs similarly inspired by military aims, but also that they share with game theory the view of agency as programmable or algorithmic.

13. Amadae, *Rationalizing Capitalist Democracy*; Richard Tuck, *Free Riding* (Cambridge, MA, 2008).

14. Joseph S. Nye, *Nuclear Ethics* (New York, 1988); Herman Kahn, *Thinking the Unthinkable*, (New York: Avon, 1962).

15. Daniel Deudney, "Whole Earth Security: A Geopolitics of Peace," *Worldwatch Paper 55* (Washington, DC: Worldwatch Institute, 1983); on the economic value of a statistical human life, see Thomas C. Schelling, "The Life You Save May Be Your Own," in *Problems in Public Expenditure Analysis*, ed. Samuel B. Chase, Jr. (Washington D.C.: Brookings Institution, 1968).

16. On command and control, see Erik Gartzke and Jon R. Lindsay, "Thermonuclear Cyberwar," *Journal of Cyber Security* 1–2 (2017); on the principle computability of strategic rationality, see Amadae, "The Computability of Rational Choice."

17. Mirowski, *Machine Dreams*, 532.

18. Mirowski, *Machine Dreams*, 512.

19. Von Neumann and Morgenstern, *Theory of Games and Economic Behavior*, 31.

20. Desmond Ball, "US Strategic Forces: How Would They Be Used?" *International Security* 7, no. 3 (1982): 39, 56–58.

21. Mirowski, *Machine Dreams*, 440.

22. Mirowski, *Machine Dreams*, 190; see also Gartzke and Lindsay, "Thermonuclear Cyberwar."

23. See, for example, Ken Binmore, *Natural Justice* (Oxford, UK, 2005); Herbert Gintis, *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences* (Princeton, NJ, 2009); Daniel H. Hausman, *Preference, Value, Choice and Welfare* (Cambridge, UK, 2011); Francesco Guala, *Understanding Institutions: The Science and Philosophy of Living Together* (Princeton, NJ, 2016).

24. U.S. Department of Defense, *Joint Vision 2020*, (Washington, DC: Publisher, 2000), https://web.archive.org/web/20011129104507/. Here I share ground with Erickson et al.'s recent assessment that nuclear security and other military concerns set the stage for economizing decision-making using game theory, and other means of algorithmic calculation, during the early Cold War, Erickson et al., *How Reason Almost Lost Its Mind.*

25. Ken Berry et al., *Delegitimizing Nuclear Weapons: Examining the Validity of Nuclear Deterrence* (Monterey, CA, 2010); Benoît Pelopidas, "The Oracles of Proliferation: How Experts Maintain a Biased Historical Reading That Limits Policy Innovation," *Nonproliferation Review* 18, no. 1 (2011); Benoît Pelopidas, "A Bet Portrayed as a Certainty: Reassessing the Added Deterrent Value of Nuclear Weapons," in *The War That Must Never Be Fought: Dilemmas of Nuclear Deterrence*, ed. J. Goodby and G. P. Shultz (Stanford, CA, 2015); Benoît Pelopidas, "The Unbearable Lightness of Luck: Three Sources of Overconfidence in the Manageability of Nuclear Crises," *European Journal of International Security* 2, no. 2 (July 2017): 240–62.

26. Benoît Pelopidas, *"*The Birth of Nuclear Eternity,*" The Future: interdisciplinary perspectives*, ed. by Jenny Andersson and Sandra Kemp (Oxford: Oxford University Press, forthcoming *2019*.

27. Daniel Deudney, *Bounding Power: Republican Security Theory from the Polis to the Global Village* (Princeton, NJ: 2007).

28. Frank C. Zagare and D. Marc Kilgour, *Perfect Deterrence* (Cambridge: Cambridge University Press, 2000), "classical deterrence theory remains a potent intellectual force shaping the policy debates in the US . . . and elsewhere" (36).

29. Mark Coeckelbergh, "The Automation of the Social: What Robots Can Teach Us about the Social," in *Robo-Philosophy: Philosophy of, for, and by Social Robotics*, ed. J. Seibt, R. Hakli, and M. Nørskov (Cambridge, MA, 2018 forthcoming); Elaine Scarry, *Thermonuclear Monarchy: Choosing between Democracy and Doom* (New York, 2014).

30. Desmond Ball and Robert C. Toth, "Revising the SIOP: Taking War-Fighting to Dangerous Extremes," *International Security* 14, no. 4 (1990): 65–92; Gartzke and Lindsay, "Thermonuclear Cyberwar." For a frank discussion of this potential reliance on AI to maintain the chain of command see Brent Scowcroft, "C3 Systems for the President and Military Commanders," In *National Security Issues 1981 Symposium: Strategic Nuclear Policies, Weapons, and the C3 Connection*, d. by Drexel M. Ace (Bedford, MA: Mitre Corporation, Mitre Document M82-30), pp. 93-98.

31. Daniel Volmar, "The Power of the Atom: US Command Control and Communications, 1945–1965" (PhD dissertation, Harvard University, 2017).

32. Erickson et al., *How Reason Almost Lost Its Mind,* 29.

33. Erickson et al., *How Reason Almost Lost Its Mind,* 37; Amadae, "The Computability of Rational Choice."

34. Erickson et al., *How Reason Almost Lost Its Mind,* 43.

35. Although it is a fictionalized account of the historical book *Hidden Figures* by Margot Lee Shetterly (New York, 2016), the film by the same name (2016) accurately captures the difference between a computerized calculation without the demand of intelligibility on the part of a computing machine and a human "computor" who understands and can vouch for the meaning, purpose, and validity of a calculation, in this case astronaut John Glenn's reentry trajectory and landing coordinates.

36. Erickson et al., *How Reason Almost Lost Its Mind,* 40–41.

37. Erickson et al., *How Reason Almost Lost Its Mind,* 39.

38. Rationality itself is postulated to have been programmed via evolution into living beings so that they can survive and propagate: R. L. Trivers, "The Evolution of Reciprocal Altruism," *The Quarterly Review of Biology* 46, no. 1 (1971): 35–57; John Maynard Smith, *Evolution and the Theory of Games* (Cambridge, UK, 1982); Philip Pettit, *Rules, Reasons, and Norms* (Oxford, UK, 2002).

39. For discussion, see John R. Searle, *Rationality in Action* (Cambridge, MA, 2001); Alfred R. Mele, *Motivation and Agency* (Oxford, UK, 2003); Amadae, "The Computability of Rational Choice"; Brendan Markey-Towler, "I, Roboticus Oeconomicus: The Philosophy of Mind in Economics, and Why It Matters," *Cambridge Journal of Economics* 41, no. 1 (2017): 203–37.

40. R. Duncan Luce and Howard Raiffa, *Games and Decisions: Introduction and Critical Survey* (New York, 1957), 6.

41. Ludwig Wittgenstein, *Philosophical Investigations*, trans. G. E. M. Anscombe (New York, 1999).

42. See John Searle, *Rationality in Action*, as well as the following by Searle: "Minds, Brains and Programs," *Behavioral and Brain Sciences* 3 (1980): 417–57; "Wittgenstein and the Background," *American Philosophical Quarterly* 48, no. 2 (2011): 119–28; "Insight and Error in Wittgenstein," *Philosophy of the Social Sciences* 46, no. 6 (2016): 527–47.

43. Erickson et al., *How Reason Almost Lost Its Mind,* 30.

44. Amadae, "The Computability of Rational Choice."

45. Alfred Church, "A Set of Postulates for the Foundation of Logic," *Annals of Mathematics*, second series, 33 (1936): 346–66.

46. Alan Turing, "Intelligent Machinery" (London, 1948) in *Machine Intelligence 5*, ed. Bernard Meltzer and Donald Michie (Edinburgh, 1969), 7.

47. Erickson et al., *How Reason Almost Lost Its Mind.*

48. Erickson et al., *How Reason Almost Lost Its Mind,* 136; Amadae, *Prisoners of Reason*, 73–93.

49. This vision of algorithmic rule-following is sufficiently comprehensive that Herbert Simon argued that scientific discovery itself can be modeled, and hence has a determinist quality, in *Models of Discovery*, 1977. For a comprehensive analysis of Simon's contributions to cybernetics and bounded rationality, see Hunter Crowther-Heyck, *Herbert A. Simon*. See also Heyck's *Age of System: Understanding the Development of Modern Social Science* (Baltimore, MD, 2015).

50. Erickson et al., *How Reason Almost Lost Its Mind,* 87. See also Amadae, *Prisoners of Reason*, 76–84

51. Schelling, *Strategy of Conflict*; Amadae, *Prisoners of Reason*; Erickson et al., *How Reason Almost Lost Its Mind*, 135; Robert J. Aumann et al., *Models of Gradual Reduction of Arms*, Arms

Control and Disarmament Agency, ACDA/ST-116, Sept. 1, 1967; Amadae, *Prisoners of Reason*, 49–58.

52. Sharon Ghamari-Tabrizi, *The Worlds of Herman Kahn: The Intuitive Science of Thermonuclear War* (Cambridge, MA, 2005); Deudney, *Whole Earth Security*; Amadae, *Prisoners of Reason*, 65–140; quote from Erickson et al., *How Reason Almost Lost Its Mind,* 88.

53. Amadae, *Prisoners of Reason*, 106 fn19; Mirowski, *Machine Dreams*.

54. Jack L. Snyder, *The Soviet Strategic Culture: Implications for Limited Nuclear Options* (Santa Monica, CA, 1977). In a similar vein, see Glenn H. Snyder, "'Prisoner's Dilemma' and 'Chicken' Models in International Politics," *International Studies Quarterly* 15, no. 1 (1971): 66–103.

55. Amadae, *Prisoners of Reason*; Selmer Bringsjord et al., "Nuclear Deterrence and the Logic of Deliberative Mindreading," *Cognitive Systems Research* 28 (2014): 20–43; David A. Shlapak and Michael Johnson, *Reinforcing Deterrence on NATO's Eastern Flank: Wargaming the Defense of the Baltics* (Santa Monica, CA, 2016).

56. Nye, *Nuclear Ethics*.

57. Robert J. Leonard, "Creating a Context for Game Theory," in *Toward a History of Game Theory*, ed. E. Roy Weintraub (Durham, NC, 1992); Mirowski, *Machine Dreams*; Amadae, *Rationalizing Capitalist Democracy* and *Prisoners of Reason*; Erickson, *The World the Game Theorists Made*.

58. Von Neumann and Morgenstern, *Theory of Games and Economic Behavior*, 30–33.

59. Erickson et al., *How Reason Almost Lost Its Mind,* 29.

60. Amartya K. Sen, "Goals, Commitment, and Identity," *Journal of Law, Economics and Organization* 1 (1985): 2; Daniel H. Hausman, *Preference, Value, Choice and Welfare*,

(Cambridge, UK, 2011); John Nash, "Non-cooperative Games," *Annals of Mathematics* (1951): 286–95.

61. See, for example, Deirdre N. McCloskey, *Bourgeois Equality: How Ideas, not Capital or Institutions, Enriched the World* (Chicago, IL: 2016).

62. Erickson et al., *How Reason Almost Lost Its Mind,* 30

63. Francesco Guala and Frank Hindricks, "A Unified Social Ontology," *The Philosophical Quarterly* 65, no. 259 (2015): 177–201.

64. See, for example, McCloskey, *Bourgeois Equality*.

65. Daniel Deudney, "Political Fission: State Structure, Civil Society, and Nuclear Security Politics in the United States," in *On Security*, ed. Ronnie Lipschutz (New York: 1995), 87–123.

66. Deudney, "Political Fission," 102–9.

67. Deudney, *Whole Earth Security*, 33.

68. Ball and Toth, "Revising the SIOP: Taking War-Fighting to Dangerous Extremes"; Hans M. Kristensen, Matthew McKinzie, and Theodore A. Postol, "How US Nuclear Force Modernization Is Undermining Strategic Stability: The Burst-Height Compensating Super-Fuze." *Bulletin of the Atomic Scientists,* vol. 1(2017) Available online at https://thebulletin.org/2017/03/how-us-nuclear-force-modernization-is-undermining-strategic-stability-the-burst-height-compensating-super-fuze/. Gartzke and Lindsay, "Thermonuclear Cyberwar"; Deudney, "Political Fission."

69. Fred Kaplan, *The Wizards of Armageddon* (Palo Alto, CA, 1991); Gregg Herken, *Counsels of War* (New York, 1985); David R. Jardini, "Out of the Blue Yonder: The RAND Corporation's Diversification into Social Welfare Research, 1946–1968" (PhD dissertation, Carnegie Mellon University, 1996); Mirowski, *Machine Dreams*; Amadae, *Rationalizing Capitalist Democracy* and *Prisoners of Reason*; Erickson et al., *How Reason Almost Lost Its Mind*.

70. Robert Jervis, "Deterrence Theory Revisited," *World Politics* 31, no. 2 (1979): 289–324; Frank C. Zagare, "Classical Deterrence Theory: A Critical Assessment," *International*

Interactions 21, no. 4 (1996): 365–87; Krause, "Rationality and Deterrence in Theory and Practice"; Amadae, *Prisoners of Reason*, 106.

71. Lawrence Freedman, *The Evolution of Nuclear Strategy* (New York, 1981); Douglas P. Lackey, "The American Debate on Nuclear Weapons Policy," *Analyse & Kritik* 9, no. 1–2 (1987): 7–46.

72. Gregory S. Kavka, "Some Paradoxes of Deterrence," *The Journal of Philosophy* 75, no. 6 (1978): 285–302; "Deterrence, Utility, and Rational Choice," *Theory and Decision* 12, no. 1 (1980): 41–60; "The Toxin Puzzle," *Analysis* 43, no. 1 (1983): 33–36; *Hobbesian Moral and Political Theory* (Princeton, NJ, 1986); *Moral Paradoxes of Nuclear Deterrence* (Cambridge, UK, 1987); David Gauthier, "Deterrence, Maximization, and Rationality," *Ethics* 94, no. 3 (1984): 474–95; "Assure and Threaten," *Ethics* 104, no. 4 (1994): 690–721; and Richmond Campbell and Lanning Sowden, eds. *Paradoxes of Rationality and Cooperation: Prisoner's Dilemma and Newcomb's Problem* (Vancouver, 1985).

73. Erickson et al., *How Reason Almost Lost Its Mind,* 134.

74. Erickson et al., *How Reason Almost Lost Its Mind,* 148. See also John Mayberry, "The Notion of Threat," in *Models of Gradual Reduction of Arms*, ed. Robert J. Aumann (Princeton, NJ, 1967).

75. Schelling, *Strategy of Conflict*; Amadae, *Prisoners of Reason*.

76. Anatol Rapoport and Albert M. Chammah, *Prisoner's Dilemma: A Study of Conflict and Cooperation* (Ann Arbor, MI: University of Michigan Press, 1965); Campbell and Sowden, *Paradoxes of Rationality and Cooperation*.

77. Randall L. Schweller, "Neorealism's Status Quo Bias: What Security Dilemma?" *Security Studies* 5, no. 3 (1996): 90–121.

78. Kristensen et al., “How US Nuclear Force Modernization Is Undermining Strategic Stability.”

79. Schelling, *Strategy of Conflict*; Amadae, *Prisoners of Reason*.

80. Campbell and Sowden, *Paradoxes of Rationality and Cooperation*; Kavka, “The Toxin Puzzle”; David Gauthier, “Rethinking the Toxin Puzzle,” in *Rational Commitment and Social Justice: Essays for Gregory Kavka,* ed. Jules L. Coleman, Christopher W. Morris, and Gregory S. Kavka (Cambridge, UK, 1998), 47–58

81. Frank C. Zagare and D. Marc Kilgour argue for updating classical deterrence theory to perfect deterrence based on game theory’s perfect Bayesian equilibrium. Their case rests on embracing the PD preferences characterizing an aggressor state, *Perfect Deterrence* (Cambridge: Cambridge University Press, 2000).

82. Kavka, *Moral Paradoxes of Nuclear Deterrence*; Gauthier, “Assure and Threaten.”

83. Gauthier, “Assure and Threaten,” 719.

84. Ball and Toth, “Revising the SIOP,” 66; Kristensen et al., “How US Nuclear Force Modernization Is Undermining Strategic Stability,” 1.

85. New York Times Editorial Board, “The Finger on the Nuclear Button,” *New York Times*, 5 February 2017, https://www.nytimes.com/2017/02/05/opinion/the-finger-on-the-nuclear-button.html?_r=0.

86. Michael S. Gerson, “No First Use: The Next Step for US Nuclear Policy,” *International Security* 35, no. 2 (2010): 7–47.

87. Thomas C. Schelling, “An Astonishing Sixty Years: The Legacy of Hiroshima,” *The American economic review* 96, no. 4 (2006): 929–37.

88. The term for this is SIOP, or single integrated operational plan. For discussion, see Ball, “US Strategic Forces: How Would They Be Used?”; Ball and Toth, “Revising the SIOP.”

89. Thomas C. Schelling, *Arms and Influence* (New Haven, CT, 1966). For a recent statement of this concern see Edward Geist and Andrew J. Lohn, "How Might Artificial Intelligence Affect the Risk of Nuclear War?" (Santa Monica, CA: RAND Corporation, 2018), p. 11. https://www.rand.org/pubs/perspectives/PE296.html.

90. Schelling, *Strategy of Conflict*; Schelling, *Arms and Influence*; Robert Jervis, *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon* (Ithaca, NY, 1989); Steven J. Brams, *Superpower Games: Applying Game Theory to Superpower Conflict* (New Haven, CT, 1985).

91. Deudney, *Bounding Power*, 246.

92. Gerson, "No First Use."

93. Thomas Hobbes, *Leviathan*, rev. ed., ed. by Richard Tuck (Cambridge, UK: Cambridge University Press, 1996).

94. Mayberry, "The Notion of Threat"; Amadae, *Prisoners of Reason*, 57.

95. This insight resonates with David Runciman's exploration of the implications of new information technologies for schools of political realism and idealism. See David Runciman, "Political Theory and Real Politics in the Age of the Internet," *The Journal of Political Philosophy* 25, no. 1 (2016): 3–21.

96. Erickson et al., *How Reason Almost Lost Its Mind*, 159–88.

97. On the nuclear security dilemma, see Kristensen et al., "How US Nuclear Force Modernization Is Undermining Strategic Stability."

98. Guala, *Understanding Institutions*.

99. Gintis, *Bounds of Reason.*

100. S. J. Gershman et al., "Computational Rationality: A Converging Paradigm for Intelligence in Brains, Minds, and Machines," *Science* 349, no. 6245 (2015): 273–78.; Paul M. Churchland, *Matter and Consciousness*, rev. ed. (Cambridge, MA, 1988), 43–49; 156–65

101. See, for example, Luciano Floridi, *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality* (Oxford, UK, 2014)

102. Amadae, "The Computability of Rational Choice."

103. On the normative status of rational choice, see Roger B. Myerson, *Game Theory: Analysis of Conflict* (Cambridge, MA, 1991); and Shaun Hargreaves Heap and Yanis Varoufakis, *Game Theory: A Critical Introduction* (Abingdon, 2004). On institutions, see Douglass C. North, *Institutions, Institutional Change and Economic Performance* (Cambridge, UK, 1990); and Roth, "The Economist as Engineer." On the nonreflexive quality of norms and rule-following, see Guala, *Understanding Institutions*.

104. R. L. Lewis et al., "Computational Rationality: Linking Mechanism and Behavior through Bounded Utility Maximization," *Topics in Cognitive Science* 6, no. 2 (2014): 279–311; Gershman et al., "Computational Rationality," 273–78.

105. Herbert Simon's bounded rationality is used because it is computable in biological systems.

106. Lewis et al., "Computational Rationality: Linking Mechanism and Behavior through Bounded Utility Maximization," 281.

107. Gershman et al., "Computational Rationality," 273.

108. Gintis, *Bounds of Reason*, 3.

109. See also Todd Jones, "Uncovering 'Cultural Meaning': Problems and Solutions," *Behavior and Philosophy* 32, no. 2 (2004): 247–68.

110. Trivers, "The Evolution of Reciprocal Altruism"; John Maynard Smith, *Evolution and the Theory of Games*; Robert Axelrod, *Evolution of Cooperation* (New York, 1984); Robert Axelrod

and R. O. Keohane, "Achieving Cooperation," in *Neorealism and Neoliberalism: The Contemporary Debate*, ed. David A. Baldwin (New York, 1993), 85–115.

111. Russell Hardin, *Collective Action* (Baltimore, MD, 1982); Dennis C. Mueller, *Public Choice III* (Cambridge, UK, 2003). See also, for example, Nicholas Stern, *The Economics of Climate Change: The Stern Review* (Cambridge, UK, 2007).

112. Gary S. Becker, *The Economic Approach to Human Behavior* (Chicago, IL, 1978); Mueller, *Public Choice III*; North, *Institutions, Institutional Change and Economic Performance*; and Roth, "The Economist as Engineer."

113. Duncan Snidal, "The Limits of Hegemonic Stability Theory," *International Organization* 39, no. 4 (1985): 579–614; Ana Espinola-Arredondo and Félix Munoz-Garcia, "Free-Riding in International Environmental Agreements: A Signaling Approach to Non-enforceable Treaties," *Journal of Theoretical Politics* 23, no. 1 (2011): 111–34.

114. Myerson, *Game Theory*, 22.

115. On systems with AI actors, see Yoav Shoham and Kevin Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations* (Cambridge, UK, 2008).

116. Pettit, *Rules, Reasons, and Norms*; Binmore, *Natural Justice*; Gintis, *The Bounds of Reason*; Guala, *Understanding Institutions*.

117. See, example, Hargreaves, Heap, and Varoufakis, *Game Theory.*

118. David P. Gauthier, *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes* (Oxford, UK, 1969); Pettit, *Rules, Reasons, and Norms*, 308–43; Amadae, *Prisoners of Reason*, 153–74.

119. Thomas C. Schelling, *Micromotives and Macrobehavior* (New York, 2006).

120. Jeremy Bentham, *A Fragment on Government* (New York, 2001); Hobbes, *Leviathan*, 191–200, chap. 14. For discussion, see Runciman, "Political Theory and Real Politics in the Age of the Internet," 16.

121. Amadae, *Prisoners of Reason*.

122. Hobbes, *Leviathan*, 88; Myerson, *Game Theory*, 3.

123. See, for example, Becker, *Economic Approach to Human Behavior.*

124. Kenneth J. Arrow, *Individual Values and Social Choice* (New Haven, CT, 1951); S. M. Amadae, "Arrow's Impossibility Theorem and the National Security State," *Studies in History and Philosophy of Science Part A* 36, no. 4 (2005): 734–43.

125. Chevaleyre Yann et al., "A Short Introduction to Computational Social Choice," *International Conference on Current Trends in Theory and Practice of Computer Science* (Heidelberg, 2007), 51–69; Felix Brandt et al., "Computational Social Choice," in *Multiagent Systems*, ed. Gerhard Weiss (Cambridge, MA, 2012).

126. James M. Buchanan and Gordon Tullock, *The Calculus of Consent* (Ann Arbor, 1962); John Rawls, *Theory of Justice* (Cambridge, MA, 1970); James M. Buchanan, *The Limits of Liberty: Between Anarchy and Leviathan* (Chicago, IL, 1975); Andrew Reynolds, ed., *The Architecture of Democracy: Constitutional Design, Conflict Management, and Democracy* (Oxford, UK, 2002).

127. Amadae, *Rationalizing Capitalist Democracy.*

128. Alfred Moore, "Hayek, Conspiracy, and Democracy," *Critical Review* 28, no. 1 (2016): 44–62.

129. On the conditions rational action must obey, see von Neumann and Morgenstern, *Theory of Games and Economic Behavior*, 30–33; Myerson, *Game Theory*, 1–33; and Hargreaves Heap and Varoufakis. *Game Theory*, 8–12.

130. Myerson, *Game Theory*, 1–5

131. On rational choice theory's "all things considered preferences," see Hausman, *Preference, Value, Choice and Welfare*.

132. On fungible interpersonally transferable utility, see Nicola Giocoli, "Do Prudent Agents Play Lotteries? Von Neumann's Contribution to the Theory of Rational Behavior," *Journal of the History of Economic Thought* 28, no. 1 (2006): 95–109.

133. Deudney, "Political Fission." "Basic legitimation demand" is Runciman's vocabulary—see "Political Theory and Real Politics in the Age of the Internet," 10.

134. On acting recklessly abroad, see Runciman, "Political Theory and Real Politics in the Age of the Internet," 10; Daniel Deudney, "Losing the Constitution? Omniviolence, Arms Control, and Limited Government," in his *On Security* (Princeton, NJ, 2009), 298.

135. On commitment, see Heath, *Communicative Action and Rational Choice*, 86–92; Hausman, *Preference, Value, Choice and Welfare*, 57–62. On intrinsic interests, see Daryl J. Levinson, "Parchment and Politics: The Positive Puzzle of Constitutional Commitment," *Harvard Law Review* (2011): 657–746. On joint instrumental action, see Michael Bacharach, Natalie Gold, and Robert Sugden, *Beyond Individual Choice: Teams and Frames in Game Theory* (Princeton, NJ, 2006).

136. Levinson, "Parchment and Politics," 729.

137. Deudney, "Losing the Constitution?"; Deudney, "Political Fission"; Scarry, *Thermonuclear Monarchy*.

138. Gershman et al., "Computational Rationality"; Giulio Tononi et al., "Integrated Information Theory: From Consciousness to Its Physical Substrate," *Nature Reviews Neuroscience* 17, no. 7 (2016): 450–61.

139. Runciman, "Political Theory and Real Politics in the Age of the Internet," 16.

140. John Gray, *The Soul of the Marionette: A Short Enquiry into Human Freedom* (London, 2015).

141. J. Moffat, *Complexity Theory and Network Centric Warfare* (CCRP Publications Series, 2003), 55; A. Cebrowski, "Network Centric Warfare and Information Superiority," Keynote address from the proceedings, Royal United Services Institute (RUSI) conference "C4ISTAR Achieving Information Superiority," (Whitehall, London, UK: 2000); J. Moffat, *Command and Control in the Information Age: Representing Impact* (London: 2002).

142. Levinson, "Parchment and Politics," 693, 690.

143. On the strategic account of sociability, see Amadae, *Prisoners of Reason*, 125–35; Levinson, "Parchment and Politics," 697.

144. Joseph Farrell and Matthew Rabin, "Cheap Talk," *The Journal of Economic Perspectives* 10, no. 3 (1996): 103–18; S.M. Amadae, "Game Theory, Cheap Talk and Post-Truth Politics," *Journal for the Theory of Social Behavior,* May 2018, 1-24, https://doi.org/10.1111/jtsb.12169.

145. Jürgen Habermas, *Between Facts and Norm: Contribution to Discourse Theory of Law and Democracy*, trans. William Rehg (Cambridge, MA, 1996).

146. Coeckelbergh, "Automation of the Social."

147. Searle, *Rationality in Action*.

148. Levinson, "Parchment and Politics," 698.

149. Amadae, *Prisoners of Reason*, 69–140.

150. On game theory and policy analysis, see Stern, *Economics of Climate Change*; Coeckelbergh, "Automation of the Social."

151. Hobbes, *Leviathan*, 88.

152. See, for example, the Congressional Commission on the Strategic Posture of the United States (2009), available at http://media.usip/reports/strat_posture_report.pdf. See Frank C.

Zagare and D. Marc Kilgour's *Perfect Deterrence* or the statement of the long dure hold of rational deterrence theory over mainstream nuclear policy, as well as their effort to update the role of game theory in deterrence policy using the concept of the subperfect equilibrium.

153. Deudney, *Bounding Power*, xii.

154. Deudney, *Bounding Power*, 27.

155. On public policy, see Mueller, *Public Choice III*. On institutional design, see North, *Institutions, Institutional Change and Economic Performance*.

156. Deudney, *Whole Earth Security*, 36–37.