Penn State Journal of Law & International Affairs

Volume 7
Issue 3 *Symposium Issue*

April 2020

Lessons From Game Theory About Humanizing Next-Generation Weapons

Richard Jordan

ISSN: 2168-7951

Penn State
Journal of Law & International Affairs

2020 SYMPOSIUM ISSUE

LESSONS FROM GAME THEORY ABOUT HUMANIZING NEXT-GENERATION WEAPONS

Richard Jordan[*]

ABSTRACT

This article draws a parallel between nuclear weapons and the next generation of military technology, autonomous systems. It outlines some legal and ethical dilemmas the latter pose, and in particular aspects of the technology that make it dehumanizing. Autonomous systems share all of these attributes with nuclear weapons. This fact should be encouraging, because the dehumanizing effects of nuclear systems have been overcome. Drawing on the evolution of nuclear strategy and the nuclear taboo, I argue that, in negotiating the legal and ethical dilemmas posed by autonomous weapons systems, the role of international law and of normative entrepreneurs will be primarily one of imagination, not regulation. The first and most important task is to create focal points in popular and elite consciousness. To this end, I make three modest suggestions for normative entrepreneurs: to take political incentives seriously, including the impossibility of abolition or non-use; to first establish simple, guiding ideas accessible to a broad population before turning to finer points of law; and to focus on interstate, rather than transnational, cooperation.

TABLE OF CONTENTS

ABSTRACT 1
I. AUTONOMOUS WEAPONS AND DEHUMANIZED VIOLENCE 2
A. How Autonomous Systems Dehumanize Violence 3
1. Civilian Casualties 4
2. Distance 5
3. Abstraction 6
4. Automation 8
B. Comparing Nuclear and Autonomous Weapons 9

[*] Richard Jordan is an Assistant Professor of Political Science at Baylor University.

II. GAME THEORY, NUCLEAR WEAPONS, AND DEHUMANIZED VIOLENCE 10
A. Should Nuclear Strategy Use Game Theory? 11
B. The Problem of Multiple Equilibria 15
C. Focal Points as a Solution to Multiple Equilibria 17
D. Focal Points as Re-Humanizing 19
III. FOCAL POINTS, IMAGINATIVE COOPERATION, AND AUTONOMOUS SYSTEMS 23
A. Examples of Focal Points in International Law 24
B. Creating Focal Points for Autonomous Systems 26
1. Begin with Incentives 27
2. Keep it Simple 29
3. Start at the Top 31
IV. CONCLUSION 33

I. AUTONOMOUS WEAPONS AND DEHUMANIZED VIOLENCE

The central ethical problem of autonomous weapons is dehumanized violence. The threats they present to innocent lives, the gaps they interpose between decisionmakers and actual violence, and the cold rationality of their deployment, are all moral dilemmas on a scale militaries have rarely seen. So pronounced is this problem, that some scholars and public intellectuals have advocated for the total abolition of these weapons.

Autonomous systems are a product of modern technology, yet the focus on their strangeness or novelty obscures the most important part of the problem they present: we have faced it before. Nuclear weapons and nuclear strategy share all the problems that make autonomous systems dehumanizing. In fact, they confronted modern societies with the same problems on an even greater scale. Encouragingly, these problems were, by and large, overcome.

If the history of nuclear strategy is any guide, the problem of dehumanization will *not* be met by changing the weapons themselves, nor through their abolition. Technology is not the answer. Instead, the re-humanization of autonomous systems will come through imagination. Through devices like narrative journalism, speculative

fiction, and public awareness campaigns, normative and legal entrepreneurs can restore a human dimension to their moral calculus *without* undercutting the strategic logic necessary for their successful use. This approach has clear precedent in international law and cooperation. It is essentially a solution of *imagination*.

The history of nuclear strategy teaches a second lesson: the role of imagination must proceed from, and not replace, the same rationality, distance, and abstraction that dehumanized violence in the first place. Put another way, the dehumanizing aspects of autonomous systems are inescapable. Humanity is retained (or restored) to these systems, not by doing away with their dehumanizing character, but by pairing it with something more.

Both lessons derive from the application of game theory to human conflict. Before developing these lessons, though, this article first reviews the reasons scholars, philosophers, policymakers, and journalists are worried by autonomous weapons.

A. How Autonomous Systems Dehumanize Violence

Dehumanized violence is a particularly common problem in the modern world. In a famous passage, Hannah Arendt captures its danger while describing the trial of Adolf Eichmann, a Nazi officer in the SS:

> Eichmann was not Iago and not Macbeth . . . Except for an extraordinary diligence in looking out for his personal advancement, he had no motives at all . . . He merely, to put the matter colloquially, never realized what he was doing . . . such remoteness from reality and such thoughtlessness can wreak more havoc than all the evil instincts taken together which, perhaps, are inherent in man.[1]

[1] HANNAH ARENDT, EICHMANN IN JERUSALEM: A REPORT ON THE BANALITY OF EVIL (1963).

The problem, Arendt relates, was not that Eichmann thought of others as monsters, but that he did not think of them at all. His world, and his victims, had been thoroughly dehumanized.

1. Civilian Casualties

At least four qualities of autonomous systems, when used in war, tend to dehumanize their violence: harm to noncombatants; distance from actual killing; abstraction of human targets; and automation of life-and-death decisions. The most common critique of autonomous weapons, especially drones, focuses on civilian casualties. Autonomous weapons' role in noncombatant deaths in Afghanistan, Pakistan, and especially Yemen attract significant controversy, and non-profit groups expend substantial resources drawing attention to the cost in innocent lives. A small minority of scholars have even argued that, because of their harm to noncombatants, manufacturing drones is unethical and tantamount to abetting murder.[2]

When tallying drones' civilian toll, some think tanks, such as the *New America Foundation*, have set the number of non-militant deaths quite high, between 747 and 1076;[3]Other organizations, such as the Bureau for Investigative Journalism, put the figures even higher,[4] but official figures are much lower.[5] Ultimately, the opacity with which the United States conducts these overseas operations make reliable figures difficult to cite with any degree of certainty.

---

[2] Edmund F. Byrne, *Making Drones to Kill Civilians: Is It Ethical?*, 147 J. of Bus. Ethics 81, 81-93 (2018).

[3] NEW AMERICA FOUNDATION, AMERICA'S COUNTERTERRORISM WARS: TRACKING THE UNITED STATES DRONE STRIKES AND OTHER OPERATIONS IN PAKISTAN, YEMEN, AND SOMALIA (June 2019), https://www.newamerica.org/in-depth/americas-counterterrorism-wars/.

[4] THE BUREAU OF INVESTIGATIVE JOURNALISM, DRONE WARFARE (June 2019), https://www.thebureauinvestig ates.com/projects/drone-war.

[5] OFFICE OF THE DIRECTOR OF NATIONAL INTELLIGENCE, SUMMARY OF 2016 INFORMATION REGARDING UNITED STATES COUNTERTERRORISM STRIKES OUTSIDE AREAS OF ACTIVE HOSTILITIES (Jan. 19, 2017), https://www.dni.gov/index.php/newsroom/reports-publications/reports-publications-2017/item/1741-summary-of-information-regarding-u-s-counterterrorism-strikes-outside-areas-of-active-hostilities.

Still, there is substantial evidence that drones do *not* kill civilians at disproportionate rates from similar weapons.[6] Indeed, the rapid technological advance of these systems has rendered them less dangerous to innocent lives than more traditional forms of air power. Nonetheless, because autonomous systems tend to be low-cost, they could still increase civilian casualties by making decisionmakers more willing to take violent action.[7] Whatever their effects on individual encounters, it does seem certain that, by decreasing the costs of engagements, autonomous weapons increase the utility of violent force, and thus the likelihood policymakers use it.

2. Distance

Second, autonomous systems often place humans at a significant distance from actual killing. This distance can make it difficult for remote pilots (or for the machines themselves) to distinguish combatants from non-combatants.[8] Consequently, some philosophers and legal theorists have argued for deliberately reducing this distance, even if that means endangering one's own soldiers.[9] As well, the relative safety of one side can blur the moral distinction between warfare and policing. If soldiers are not at risk, are they really soldiers anymore?[10] Less obviously, but no less dangerously, this distance can make human decisionmakers less aware of others' humanity. Removed from the circumstances of a decision, decisionmakers may not fully recognize or appreciate the decision's imperatives. This distance can weaken human agents' ability to employ moral reasoning or even undermine their sense of responsibility.[11]

---

[6] Vivek Sehrawat, *Legal Status of Drones Under LOAC and International Law*, 5 PENN ST. J.L. & INT'L AFF. 164, 168-70 (2017).

[7] *The Constitutional and Counterterrorism Implications of Targeted Killing: Hearing Before the S. Judiciary Subcomm. on the Constitution, Civil Rights, and Human Rights*, 113th Cong. 12-14 (2013) (statement of Prof. Rosa Brooks).

[8] ALEX J. BELLAMY, JUST WARS: FROM CICERO TO IRAQ 182 (2006).

[9] MICHAEL WALZER, JUST AND UNJUST WARS: A MORAL ARGUMENT WITH HISTORICAL ILLUSTRATIONS 156 (2015).

[10] Paul W. Kahn, *War and Sacrifice in Kosovo*, 19 PHIL. AND PUB. POL'Y 1, 2 (1999).

[11] P. M. Asaro, *Remote-Control Crimes*, 18 IEEE ROBOTICS AUTOMATION MAG. 68, 68-69 (2011).

As above, it is not clear that the distance autonomous systems create between human operators and military targets actually imperils civilians, at least not any more than similar weapons. In theory, drones and related technology can actually create much more accurate depictions of military targets than traditional artillery or manned bombers,[12] though whether practice lives up to this potential is unclear. Yet even if it does, the *moral* distance still exists. At root, the dehumanizing power of distance stems from its impairment of moral intuition, and this seems inherent in the technology.

3. Abstraction

Similar to distance, autonomous systems make human targets abstract. At best, the human image may be filtered through screens; at worst, it is never observed at all. Popular-press articles often translate this by saying war becomes "like a video game,"[13] but the problem does not need to be so stylized.[14] By removing the human face, or by encountering the human image in a virtual or other strange environment,[15] abstracting from actual human beings increases an

---

[12] Christian Enemark, *Unmanned Drones and the Ethics of War*, *in* ROUTLEDGE HANDBOOK OF ETHICS AND WAR: JUST WAR IN THE TWENTY-FIRST CENTURY 332 (Nicholas G. Evans, Fritz Allhof, and Adam Henscke eds., 2013).

[13] Noah Shachtman, *Drone School: A Ground's-Eye View*, WIRED MAG. (May 27, 2005), https://www.wired.com/2005/05/drone-school-a-grounds-eye-view/; Chris Cole, Mary Dobbing, and Amy Hallwood, *Convenient Killing: Armed Drones and the 'Playstation' Mentality* (Sept. 2010), https://dronewarsuk.files.wordpress.com/2010/10/conv-killing-final.pdf.

[14] In fact, some journalists have challenged this idea, arguing that drone pilots experience *more* graphic violence than other soldiers. Eyal Press, *The Wounds of the Drone Warrior*, N.Y. TIMES (June 13, 2018), https://www.nytimes.com/2018/06/13/magazine/veterans-ptsd-drone-warrior-wounds.html; Sarah McCammon, *The Warfare May Be Remote But the Trauma is Real*, NAT'L PUB. RADIO (Apr. 24, 2017, 2:40 PM), https://www.npr.org/2017/04/24/525413427/for-drone-pilots-warfare-may-be-remote-but-the-trauma-is-real.

[15] A related point is developed in the controversial book *Barbed Wire*, which argues the widespread use of barbed wire before WWII helped Nazi concentration camps see their victims as "mere biological objects." Reviel Netz, Barbed Wire: An Ecology of Modernity 153 (2004).

individual's willingness to commit violence and impairs their moral judgment.[16]

From the perspective of Christian ethics, this tendency is particularly troubling. A human being can recognize another's dignity because he can recognize the image of God.[17]The more abstractly he sees another human being, the less he perceives this image, and so the less he perceives another's inherent worth.[18]

In addition to soldiers, this abstraction afflicts the wider public. It can make war more likely: "People are more likely to support the use of force as long as they view it as costless."[19] It can also incur unseen strategic costs, especially if its victims respond against their dehumanization.[20] The dehumanization of targets can foment terrorism, instability, and general animosity—in short, autonomous systems can create as many enemies as they kill. Indeed, the abstract, dehumanizing effects on the public may be the most important costs of all. Scholars have repeatedly shown that the experience of war disinclines a decisionmaker from using any kind of force.[21] The most

---

16 "By removing warriors completely from risk and fear, unmanned systems create the first complete break in the ancient connection that defines warriors and their soldierly values . . . toward 'virtueless war,' a result of remote soldiers' no longer having any 'emotional connectivity with the battlespace.'" Peter Warren Singer, Wired For War: The Robotics Revolution and Conflict in the 21st Century 332 (2009).

17 David H. Calhoun, *Human Exceptionalism and the Imago Dei, in* HUMAN DIGNITY IN BIOETHICS: FROM WORLDVIEWS TO THE PUBLIC SQUARE 20 (Stephen Dilley and Nathan J. Palpant eds. 2013).

18 Perversely, at the same time that human operators are less perceptive of another's person, the machine itself may be hyper-personalized. Charles J. Dunlap, Jr., *The Hyper-Personalization of War: Cyber, Big Data, and the Changing Face of Conflict*, 15 GEO. J. OF INT'L AFF. 108, 112-13 (2014).

19 Singer, *supra* note 17, at 316; Stephan Sonnenberg, *Why Drones Are Different, in* PREVENTIVE FORCE: DRONES, TARGETED KILLING, AND THE TRANSFORMATION OF WARFARE 123-29 (Kerstin Fisk and Jennifer M. Ramos eds. 2016).

20 Michael J. Boyle, *The Costs and Consequences of Drone Warfare*, 89 INT'L AFF. 1, 14-16 (2013).

21 . "It may surprise some, but those in the armed forces-especially those who have seen the horrific consequences of war firsthand-are often the ones most opposed to the use of force." Charles Dunlap, Jr., *Clever or Clueless? Observations About Bombing Norm Debates, in* THE AMERICAN WAY OF BOMBING: CHANGING ETHICAL

dangerous dehumanization, then, might be occurring in the minds of a public that never sees, hears, or reads about violence inflicted by autonomous weapons—a trend exacerbated by the growing disconnect between soldiers and citizens.[22]

4. Automation

Finally, by definition autonomous systems remove decisions from human control. Indeed, it seems likely, even inevitable, that future systems will leave computers with the final decision to use lethal force. The increasing speed at which autonomous military instruments need to make decisions will compel greater and greater leeway given to machines. Yet even before that bridge is crossed, autonomous systems pose an important question: who is responsible for their decisions?

In terms of both international law and just war theory, some scholars fear that autonomous weapons, especially ones making life-and-death decisions, will force a retreat from *jus in bello* (justice during war) to *jus ad bellum* (justice in going to war) when evaluating the justice or injustice of a past action.[23] The former encompasses the evolution of norms and laws to govern which actions soldiers can (and cannot) take during armed conflict; the latter only describes those circumstances in which an actor may legitimately go to war. For many reasons dangerous, this retreat from *jus in bello* to *jus ad bello* would be especially troubling because the legal and ethical protections of noncombatants flow primarily through the former, not the latter. In short, the fear is that autonomous systems will push legal and ethical norms towards a *laissez-faire* approach in which the primary criterion for just warfare is simply whether the war was begun for a just reason.

AND LEGAL NORMS, FROM FLYING FORTRESSES TO DRONES 126 (Matthew A. Evangelista and Henry Shue eds. 2014). Beyond anecdote, this finding is robustly confirmed by, for instance, Michael C. Horowitz and Allan C. Stam, *How Prior Military Experience Influences the Future Militarized Behavior of Leaders*, 68 INT'L ORGS. 527, 527-59 (2014).

[22] WARRIORS AND CITIZENS: AMERICAN VIEWS OF OUR MILITARY 1-5 (Jim Mattis and Kori Schake eds. 2016).

[23] Heather M. Roff, *Killing in War, in* ROUTLEDGE HANDBOOK OF ETHICS AND WAR: JUST WAR THEORY IN THE TWENTY-FIRST CENTURY 353 (Nicholas G. Evans, Fritz Allhoff, and Adam Henschke eds. 2013).

B. Comparing Nuclear and Autonomous Weapons

In some ways, humans have encountered these problems since ancient times. Distance inheres in any sort of ranged weapon, as does abstraction; crude traps and urban sieges involve a form of automatic death; and civilian casualties are as old as war. Perhaps the inventor of the sling and stone wrestled with the same moral quandaries we are debating now.[24] Yet autonomous systems, especially drones, seem to take the problem a step further. While the difference may be one of degree, not of kind, it is still enormous.

But there is another field where these problems loomed even larger: nuclear strategy. Nuclear weapons are distant: intercontinental ballistic missiles (ICBMs) long predate the first drone pilots operating behind air-conditioned desks in Nevada. Civilian casualties remained an inescapable part of nuclear weaponry from its inception. However deadly autonomous systems become, and however remotely controlled, they will never rival the intercontinental extinction of unarmed millions.

Still more similar is the way violence becomes automatic. Nuclear strategists considered automatic responses a feature, not a bug. Policymakers sought "trip-wires"[25] that would automatically commit them to military, even nuclear, responses.[26] They coined new jargon to describe this desired attribute: *automaticity*.[27] In nuclear bargaining, a leader who could credibly surrender his freedom of choice gained enormous leverage, because a nuclear system that would

---

[24] Medieval longbowmen "need have no sense of initiating an act of killing, therefore; it was probably their technical and professional sense which was most actively engaged." JOHN KEEGAN, THE FACE OF BATTLE: A STUDY OF AGINCOURT, WATERLOO, AND THE SOMME 93 (1976).

[25] THOMAS C. SCHELLING, ARMS AND INFLUENCE 47 (1966).

[26] American policy still relies on trip-wires today. For a recent study on their continued role, see A. Lanoszka and M.A. Hunzeker, *Landpower and American Credibility*, 45 PARAMETERS: THE U. S. ARMY'S SENIOR PROF. J. 17, 17-26 (2016).

[27] *See* Schelling, *supra* note 26, at 50 n. 9.

respond automatically to an enemy's aggression was the strongest possible deterrent.[28]

Perhaps most of all, nuclear strategy was dehumanized because of its *abstraction*. Except for the bombings of Japan, potential civilian deaths from nuclear weapons were and remain hypothetical. Anodyne words like "counter-value" concealed strategies that targeted noncombatants in the millions. And the whole subject was studied by economists and civilian strategists, often with mathematical tools that reduced human beings to numbers on a chalkboard.

Because we have encountered the same problems before, and on a much more frightening scale, it will be instructive to examine how we solved them in the past. While they may not translate precisely, these solutions, one hopes, can help us navigate the coming era of remote-controlled warfare and "killer robots."

II. GAME THEORY, NUCLEAR WEAPONS, AND DEHUMANIZED VIOLENCE

The novelty of nuclear weapons demanded a new kind of military strategy. Bernard Brodie, a military thinker later celebrated for his commentary on Clausewitz, became the first to theorize the bomb in *The Absolute Weapon*.[29] A bevy of academics soon joined him; they became known as "defense intellectuals." As these strategists gained prominence among policymakers and in the public eye, they came under increasing fire. Their critics, aghast at their willingness to debate

---

[28] Thomas C. Schelling, *An Essay on Bargaining*, 46 THE AM. ECON. REV. 281, 281-306 (1956). This ideal was famously satirized in Stanley Kubrick's *Dr. Strangelove*, which imagined a Soviet weapon, the Doomsday Machine, that would automatically destroy the entire world if any atomic device were detonated. DR. STRANGELOVE, OR, HOW I LEARNED TO STOP WORRYING AND LOVE THE BOMB (Columbia TriStar Home Entertainment 1964). *See also* Charles Maland, *Dr. Strangelove (1964): Nightmare Comedy and the Ideology of Liberal Consensus*, 31 AM. Q. 697, 697-717 (1979); PETER DANIEL SMITH, DOOMSDAY MEN: THE REAL DR. STRANGELOVE AND THE DREAM OF THE SUPERWEAPON (2007).

[29] BERNARD BRODIE ET AL., THE ABSOLUTE WEAPON: ATOMIC POWER AND WORLD ORDER (1964); BERNARD BRODIE, STRATEGY IN THE MISSILE AGE (1959).

the pros and cons of mass death, recoiled in horror. Herman Kahn, whose manner and appearance could give the impression of an almost gleefully cynical detachment, attracted the most ire, but all experienced a similar opprobrium.[30] Academics and peace activists condemned the cool, abstract way these civilian intellectuals tallied human casualties. Detached analysis, argued academics and activists, was no way to study the prospect of nuclear annihilation.

A microcosm of this conflict is that between Thomas Schelling and Anatol Rapaport. Their disagreement is especially telling, since they focused on the most abstract and dehumanized of all the nuclear strategist's tools: game theory. In particular, they disagreed about whether this mathematical tool, in which both were experts, should be applied to nuclear strategy.

A. Should Nuclear Strategy Use Game Theory?

Game theory uses formal, mathematical logic to study problems of strategic interaction. It typically assumes some sort of rationality among actors. As a subfield of the larger rational-choice tradition, game theory's picture of the world is a limited one, but its practitioners hope (they place a "methodological bet")[31] that what it ignores will be less important than what it reveals.

In the 1950s and 1960s, game theory achieved prominence for its applications to nuclear bargaining. Thomas Schelling, who went on to win the Nobel Prize, exemplified this approach. A Harvard economist, Schelling began his career working on foreign aid (he helped administer the Marshall Plan), and his exposure to international politics quickly led him to the study of nuclear strategy. In major

---

[30] Kahn's influence and demeanor can still repel strategic thinkers (Andrew Bacevich calls him "creepy" in the popular press). One scholarly monograph made this judgment on his *Thermonuclear War*: it "is a massive window into a warped mind." Andrew Bacevich, *Rationalizing Lunacy: The Intellectual as Servant of the State*, Huffington Post (Dec. 6, 2017), https://www.huffpost.com/entry/rationalizing-lunacy_b_6828460; David A. Baldwin, *Andrew J. Bacevich, ed., The Long War: A New History of U.S. National Security Policy Since World War II*, 10 J. of Cold War Stud. 149, 149-151 (2008).

[31] DAVID A. LAKE AND ROBERT POWELL, STRATEGIC CHOICE AND INTERNATIONAL RELATIONS 16 (1999).

contributions to the budding field, he showed how the logic of nuclear bargaining did not differ substantially from that of mafia protection rackets; he elaborated the idea of commitment devices (like trip-wires) to stabilize international politics between the superpowers; and he analyzed "salami tactics" and explained how the United States could prevent their abuse. Perhaps most importantly, Schelling was the first strategist to recognize and explain how nuclear weapons could increase American bargaining leverage even when the States could not credibly threaten to use them. By engaging in a "competition in risk-taking," the United States could compel an adversary to back down by increasing the risk of mutual disaster—even though neither side actually wanted the disaster to happen.[32] Schelling articulated all of these ideas in readable, analytic prose, feeling as free to discuss mass death as coercive parenting strategies. A famous quip epitomizes this style: "Against defenseless people there is not much that nuclear weapons can do that cannot be done with an ice pick."[33]

Schelling's cold approach met with heated criticism. Anatol Rapaport, himself a game theorist,[34] denounced Schelling for applying math to nuclear strategy. Rapaport felt that, as a tool, game theory should be applied only with great caution to human behavior, and never to the Cold War arms race. He pointed out that, as a branch of mathematics, game theory necessarily dehumanized its subjects.[35] Between strategy and conscience, there is an "essential incompatibility," he argued:

> Seduction lurks also in the mental habit of rational analysis. For this analysis requires detachment. While detachment is a source of supreme strength in the investigation of nature, it may be debilitating if it is

---

32 Schelling, *supra* note 26, at 91.

33 *Id.* at 19.

34 Rapaport is perhaps most famous for his victories in Robert Axelrod's "tournaments," which pitted competitors' strategies against each other in repeated Prisoner's Dilemmas. Rapaport won with the simple strategy of tit-for-tat. ROBERT AXELROD, THE EVOLUTION OF COOPERATION 42 (1984).

35 "Mathematics is a great leveler. When a problem is mathematically formulated, its *content* has disappeared and only the form has remained. To the strategists 'targets' are indeed only circles on maps; overkill is a coefficient." ANATOL RAPAPORT, STRATEGY AND CONSCIENCE 192 (1964).

> carried over bodily from natural science . . . One cannot play chess if one becomes aware of the pieces as living souls.[36]

In a review of Rapaport's book, Schelling objected to this line of reasoning which alleged "strategic thinking is bad no matter how good it is."[37] He criticized Rapaport for disagreeing, not with his logic, but with its coldness. In a matter as cataclysmic as nuclear armageddon, cool heads would seem to be in demand. Nonetheless, Schelling recognized that his disagreement with Rapaport was a normative one, not a scientific one, and he left it to others to decide who was right.

By any measure, Rapaport and figures like him lost the debate about game theory and nuclear strategy. Schelling remains the classic text on the subject, assigned in policy schools around the country, while Rapaport's objections have passed out of print. One scholar wrote that Schelling "has contributed as much as and perhaps more than any other thinker of the scientific genre to the theory of international relations."[38] It is no overstatement to say that Schelling gave strategists and policymakers the intellectual framework they still use today to think about nuclear weapons.[39]

Perhaps the best explanation for Schelling's victory comes from Hedley Bull. Bull was an acerbic critic of modern social science, and his works came to found the so-called English School of international relations against the American combination of positivism, mathematics, and hypothesis-testing. Yet, despite his antipathy for applying science to human beings, Bull nonetheless sided with the defense intellectuals against Rapaport. He observed, "[there are] certain conflicts in the world that simply have to be taken as

---

[36] *Id.* at 109, 195.

[37] Thomas C. Schelling, Book Review, 54 THE AM. ECON. REV. 1082, 1088 (1964) (reviewing ANATOL RAPAPORT, STRATEGY AND CONSCIENCE (1964)).

[38] Hedley Bull, *International Theory: The Case for a Classical Approach*, 18 WORLD POL. 361, 368 (1966).

[39] Robert Ayson, Thomas Schelling and the Nuclear Age: Strategy as Social Science (2004). The only full-length treatment of Schelling, his thought, and its influence. While unfairly critical in places, and overly obsessive with the concept of stability throughout, it is a helpful companion to Schelling's corpus.

given,"[40] and he faulted Rapaport for never showing where a less rational, more "conscientious" approach to nuclear strategy produced better outcomes.[41] Whatever the objections to applying mathematics to human beings, analytic detachment is not one of them. Bull concluded:

> When one asks oneself what the history of strategic policy in the West might have been in the last ten years had this influence not been brought to bear, or when one contemplates the moral and intellectual poverty of the debate about nuclear affairs (or of that part of it we are able to see) in the Soviet Union where in fact no such influence exists, it is difficult to escape the conclusion that even though the civilian strategists have sometimes committed the errors I have been exploring, they have served us well.[42]

Rapaport claimed that strategists cannot play chess with living souls. And yet, of course, they do. Strategists do because they must.[43] He was right that abstract analysis was dehumanizing, in the same way that autonomous systems are dehumanizing today. Yet he was wrong to see this objection as final. The answer to the question, "is abstraction dehumanizing," must be an unequivocal *yes*. But that answer does not imply it should not be done. A lesson emerges from the history of nuclear strategy: when the stakes are high, rational

---

[40] Hedley Bull, *Strategic Studies and Its Critics*, 20 World Pol. 593, 604 (1968).

[41] As another reviewer pointed out, "Ideologies *are* sometimes filled out with a complement of new and liberating conceptions, but rarely by 'ideological disarmament.'" Arthur Lee Burns, *Must Strategy and Conscience Be Disjoined?*, 17 World Pol. 687, 688 (1965) (reviewing Anatol Rapaport, Strategy and Conscience (1964)).

[42] Bull, *supra* note 41, at 605

[43] Kahn opens the first chapter of *Thinking About the Unthinkable* with this observation: Seventy-five years ago white slavery was rampant in England. Each year thousands of young girls were forced into brothels and kept there against their will . . . One reason why this lasted as long as it did was that it could not be talked about openly in Victorian England; moral standards as to subjects of discussion made it difficult to arouse the community to necessary action." HERMAN KAHN, THINKING ABOUT THE UNTHINKABLE 17 (1962).

analysis—no matter how cold and detached—ought not and will not be discarded.

Raymond Aron, in his introduction to Kahn's *Thinking About the Unthinkable*, summarizes the issue well. He writes:

> The analyst who calculates in millions or tens of millions of deaths resulting in a matter of a few minutes or a few hours resulting from thermonuclear exchanges does indeed forget the human significance of these figures . . . without the ability to neutralize his feelings, the analyst's profession would become impossible. But the acquired capacity to coldly examine possible horrors does not prove the analyst has lost his humanity.[44]

Aron is correct, yet he leaves the most important question unanswered. If calculating in the abstract millions *is* dehumanizing, but the analyst has not lost his humanity, then what rehumanized the process? The answer to this question also arises from game theory: by navigating multiple equilibria with the moral imagination.

B. The Problem of Multiple Equilibria

According to Harrison Wagner, there are two fundamental insights of game theory. First is the problem of strategic interdependence, i.e. that one actor's interests depend on other actors' choices. Second is the problem of multiple equilibria, i.e. that many different outcomes are rationally possible.[45]

Consider an example. Scholars of international politics often discuss a simple game called a Stag Hunt.[46] In this game, two hunters

---

[44] Raymond Aron, *Introduction* to HERMAN KAHN, THINKING ABOUT THE UNTHINKABLE 10 (1962).

[45] R. HARRISON WAGNER, WAR AND THE STATE: THE THEORY OF INTERNATIONAL POLITICS 101 (2007).

[46] Kenneth Waltz seems to have introduced this game to the study of international relations, drawing on a famous vignette by Jean-Jacques Rousseau. KENNETH N. WALTZ, MAN, THE STATE AND WAR: A THEORETICAl ANALYSIS

go into the woods in pursuit of the evening's dinner. Their best outcome is to hunt a stag together, in which case they will feast on venison. However, either one of them can also choose to chase a rabbit on his own, in which case he will catch the rabbit and have a satisfying but meager meal, while the other huntsman (if he still pursues the stag) will go hungry. This game is summarized in the table below.

		Hunter 2	
		Stag	Rabbit
Hunter 1	Stag	3,3	0,1
	Rabbit	1,0	1,1

The first insight of game theory, strategic interdependence, highlights how one hunter's decision hinges on what he expects the other hunter will do. It further guarantees that there are only two possible outcomes: either both hunt the stag, or both hunt rabbits. Anything else is not stable—not an "equilibrium," in the parlance of game theory—because at least one hunter would prefer to change his behavior to achieve a better outcome.[47] For instance, if Hunter 1 chased a rabbit, but Hunter 2 chased the stag, then Hunter 2 would deviate to chasing rabbits. Thus, strategic interdependence helps eliminate a variety of outcomes as implausible because reasonable agents would not behave in such a way, at least not for any sustained period of time.[48]

(1954); For an important recent application of this game to state formation, *see* BRIAN SKYRMS, THE STAG HUNT AND THE EVOLUTION OF SOCIAL STRUCTURE (2004).

[47] A third, mixed-strategy equilibrium exists, in which each hunter pursues the stag or the rabbit probabilistically. For ease of exposition, I ignore this possibility here.

[48] Thus, game theory refuses the postmodern or constructivist idea that the world can be whatever "states make of it." ALEXANDER WENDT, SOCIAL THEORY OF INTERNATIONAL POLITICS (1999).

The second fundamental insight, multiple equilibria, stresses that, once implausible outcomes are eliminated, there remain many ways for the world to turn out. The Stag Hunt, though a simple game, has many possible outcomes. This multiplicity confronts both the theorist and the practitioner with a problem. Because the theorist seeks to predict outcomes, he must find ways to refine the number of possible equilibria; otherwise, he can offer no clear hypotheses to test. The theorist therefore turns to cultural norms and expectations in order to discern which outcomes are likely to occur in his own circumstances. To return to the Stag Hunt: if there are two possible equilibria, either both hunting the stag or both chasing rabbits, then the theorist must discern whether his society is one in which the hunters trust each other enough to hunt the stag. The practitioner faces a similar problem: because she likely finds herself in an unsatisfying equilibrium, she must devise ways to move from a less-desirable outcome to a more-desirable one. To put all this more simply, the problem of multiple equilibria is identical to the classic questions, "which of the possible worlds do we inhabit, which do we wish to inhabit, and how can we move from one to the other?" Game theory does not erase these idealistic questions; its importance is in stressing that imagination and idealism must be framed within the incentive structures of the relevant agents.

In the Stag Hunt, the problem of multiple equilibria is easily solved. Even if the two hunters are used to chasing rabbits, they can coordinate with relative ease to start hunting stags. But imagine a far more complex hunt, with thousands of hunters, hundreds of options, and uncountably many competing interests. How could human beings navigate this complexity? The answer is surprisingly simple: they look for a focal point.

C. Focal Points as a Solution to Multiple Equilibria

Thomas Schelling was among the first to think seriously about the problem of multiple equilibria. He addressed it with a novel solution: the idea of a *focal point*.[49] A focal point is an outcome which

[49] Thomas C. Schelling, *Bargaining, Communication, and Limited War*, 1 CONFLICT RESOL. 19, 21 (1957).

"stands out," which attracts the actors' attention, and by so doing makes coordination easier. Schelling popularized this idea with an intuitive example. Imagine two strangers: each is given a photograph of the other and told to meet that person at noon in New York City. They have no other information. Can they do it? Schelling's surprising answer was, *yes*: both will probably go to Grand Central Station, or maybe Times Square.[50] These hubs are focal points: popularized by movies, books, and television, they stand out in the public imagination as spots unlike the others.

Other examples could include avoiding bicycle collisions by intuitively swerving right[51]or (if they have droll senses of humor) a husband and wife, separated in a department store, looking for each other at the lost-and-found. In these examples, thousands of possible equilibria exist. Cooperation emerges because both actors intuit a pattern of behavior that, for reasons either inherent or cultural, stands out from all the others—even though others could be perfectly "rational" or "reasonable."

The use of focal points is particularly important with tacit bargaining, where much must go unsaid. Tacit bargaining happens on a vast scale, as among millions of citizens who cannot possibly all talk to each other, and on a small scale, as among world leaders who cannot say or communicate everything they might wish. Drivers negotiating a poorly marked intersection are engaging in tacit bargaining (especially if one is more aggressive than the other); world leaders maneuvering carrier groups into a strategic waterway are doing the same thing. Indeed, an emerging consensus in political science sees state development as a kind of massive, unspoken bargain between rulers and ruled.[52] In these situations, because it is impossible to explicitly coordinate action, the implicit appeal of focal points becomes decisively important.

---

[50] *Id.*

[51] PRESH TALWALKAR, THE JOY OF GAME THEORY: AN INTRODUCTION TO STRATEGIC THINKING 26-8 (2014).

[52] Wagner, *supra* note 46; Charles Tilley, *Coercion, Capital, and European States, AD 990-1990 in* COLLECTIVE VIOLENCE, CONTENTIOUS POLITICS, AND SOCIAL CHANGE 140-54 (2017).

Since Schelling, the study of focal points has progressed, though *not* primarily within game theory.[53] (He once observed, "Focal points have done more for the theory of games, than game theory has done for the theory of focal points.")[54] Instead, the concept opened an avenue for normative and sociological approaches to intersect rational, economic thinking:[55] "here ideas contribute to outcomes in the absence of a unique equilibrium."[56] By studying which ideas gain salience, a social scientist can distinguish between multiple equilibria. This continued a research agenda Schelling envisioned, one in which game theory would illuminate which social conventions can (and cannot) come about and how durable they might be once they emerged.[57]

D. Focal Points as Re-Humanizing

The practice of identifying focal points is empathetic and deeply humane. It requires an observer to put himself, not just in someone else's rational position, but in their emotional, cultural, and religious shoes, as well. He must get inside their heads *and* their hearts. Identifying focal points is an exercise of moral imagination.

Unlike a rationalistic approach that would try to trace everything to first principles, moral imagination begins by situating oneself in particular cultural and historical circumstances. (The phrase "moral imagination" is due to Edmund Burke, who was arguing against the excessive rationalism of the French Revolution.)[58] In this, it

---

53 For an exception, see the work of Robert Sugden, who studies focal points rigorously within mathematical game theory. Robert Sugden, *A Theory of Focal Points*, 105 THE ECON. J. 533, 533-50 (1995).

54 Jean-Paul Carvalho, *An Interview with Thomas Schelling*, 2 OXONOMICS 1,4 (2007).

55 Albert S. Yee, *Thick Rationality and the Missing Brute Fact: The Limits of Rationalist Incorporations of Norms and Ideas*, 59 THE J. OF POL. 1001, 1001-39 (1997).

56 IDEAS AND FOREIGN POLICY: BELIEFS, INSTITUTIONS, AND POLITICAL CHANGE 12 (Judith Goldstein and Robert O. Keohane eds., 1993).

57 In fact, Schelling believed that David Lewis, the eminent philosopher of conventions, took his main idea from Schelling's course on bargaining, where Lewis had been his student. N. Emrah Aydinonat, *An Interview with Thomas C. Schelling: Interpretation of Game Theory and the Checkerboard Model*, 2 ECON. BULL. 1, 3 (2005).

58 "By this 'moral imagination,' Burke signifies that power of ethical perception which strides beyond the barriers of private experience and momentary events." RUSSELL KIRK, REDEEMING THE TIME 71 (1996).

resembles the moral philosophy of Adam Smith and his modern heirs, who emphasize a kind of empathy as requisite to just reasoning.[59] Thus, through focal points, strategic reasoning not only involves but comes to *demand* a moral vision, one shared with other human beings.

Schelling himself emphasized the role of ethical imagination. Finding a focal point, he argued, may "depend on imagination more than on logic; it may depend on analogy, precedent, accidental arrangement, symmetry, aesthetic or geometric configuration . . . Poets may do better than logicians at this game."[60] In fact, Schelling valued the role of imaginative boundaries so highly that he favored them even when they might cut against short-term bargaining power.[61] Because states and culture are themselves a kind of "tacit bargain," the ethical norms that evolve in them encapsulate how individuals are coordinating between multiple equilibria.[62]

Through this evolutionary process of moral coordination around a focal point, nuclear weapons became re-humanized. That is, not the technology nor its deployment, but the imaginative ways the larger culture came to conceive it, restored a degree of humanity to nuclear strategy. So complete has this ethical evolution been, that some scholars argue the memory of Hiroshima has been "sanctified" in the public imagination.[63] This process is most clearly seen in the nuclear taboo.

---

[59] ADAM SMITH, THE THEORY OF MORAL SENTIMENTS (D.D. Raphael and A.L. Macfie eds. 1976). It is worth noting that Smith is tracking a divide between *reason* and *sentiment* and inquiring which one offers the wellspring of moral judgment. "For Smith, the key mechanism of sympathy is imaginatively placing oneself in another's position, or what would now be called *simulation*." Antti Kauppinen, MORAL SENTIMENTALISM, IN THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY 1 (Edward Zalta ed. 2018).

[60] Thomas C. Schelling, *War Without Pain, and Other Models*, 15 WORLD POL. 465 (1963).

[61] Alexander Field, *Schelling, von Neumann, and the Event that Didn't Occur*, 5 GAMES 53, 70 (2014).

[62] Thomas C. Schelling, *Game Theory and the Study of Ethical Systems*, 12 J. OF CONFLICT RESOL. 34, 34-44 (1968).

[63] Alvin M. Weinberg, *The Bell and the Bomb*, COSMOS J. (1997).

The nuclear taboo is the international norm against the use of any nuclear weapon for any purpose. The taboo draws a distinction between nuclear and other weapons, and its prohibition on the former is total. How it arose through an interplay of strategic thought and moral imagination is well summarized by Nina Tannenwald:

> Game theorists hold that norms can serve as focal points, thus contributing to stable outcomes in the absence of a unique equilibrium. The analysis here helps to explain why one equilibrium was chosen over another. The development of the taboo has been the result of both self-interested and normative concerns, and has depended importantly on discursive strategies——how nuclear weapons became categorized, interpreted, and politicized.[64]

We take for granted that nuclear weapons and conventional weapons are qualitatively different from each other. Yet this distinction is arbitrary: "the line between conventional and nuclear weapons did not always exist but had to be created."[65] After all, today some nuclear weapons are less powerful than some conventional ones, and yet while the United States has periodically used the latter, it has never again used the former—despite strong temptation otherwise.[66] Moreover, the taboo is not a matter of law, since the use of nuclear weapons is *not* legally prohibited.[67] Furthermore, the distinction certainly was not inevitable: many early Cold Warriors, including Secretary Dulles, expected they would eventually by treated as ordinary weapons not unlike any others.[68]

---

[64] Nina Tannenwald, *Stigmatizing the Bomb: Origins of the Nuclear Taboo*, 29 INT'L SECURITY 5, 41 (2005).

[65] *Id.* at 12.

[66] Thomas C. Schelling, *An Astonishing Sixty Years: The Legal of Hiroshima*, 96 AM. ECON. REV. 929, 931-2 (2006).

[67] The United States has never agreed to any general prohibition, and in fact has vetoed such proposals. *See also* Michael N. Schmitt, *The International Court of Justice and the Use of Nuclear Weapons*, 51 NAVAL WAR C. REV. 91, 91-116 (1998); Dale Stephens, *Human Rights and Armed Conflict: The Advisory Opinion of the International Court of Justice in the Nuclear Weapons Case*, 4 YALE HUM. RTS. AND DEV. L. J. 1, 1 (2001).

[68] Tannenwald, *supra* note 65, at 5.

It seems clear that, without this artificial convention, nuclear weapons likely would have been used sometime in the past sixty years.[69] Yet, "[a]lthough rationalist variables are important, the taboo cannot be explained simply as the straightforward result of rational adaptation to strategic circumstances."[70] Instead, it emerged through a normative evolution propagated by activists, novelists, and public discourse. This evolution exerted increasing pressure on policymakers, constraining them within the moral dimensions in which the public saw nuclear force.[71] In fact, many advances in military technology were explicitly passed over by policymakers, for fear of eroding the distinction between nuclear and other weapons. (Similarly, PNEs—peaceful nuclear explosions, whose use was proposed for large-scale construction projects like canals—also became taboo.)[72]

Thomas Schelling dedicated his Nobel acceptance speech to this theme. He said:

> A large part of the credit for nuclear weapons not having been used must be due to the "taboo" that Secretary of State John Foster Dulles perceived to have been attached to these weapons as early as 1953 . . . These weapons are unique, and a large part of their uniqueness derives from their being *perceived* as unique. We call most other weapons "conventional," and that word has two distinct senses. One is "ordinary, familiar, traditional," words that can be applied to food, clothing, or housing. The more interesting sense of "conventional" is something that arises as if by compact, by agreement, by *convention*. It is simply an established convention that nuclear weapons are different.[73]

---

69 Nina Tannenwald, *The Nuclear Taboo: The United States and the Normative Basis of Nuclear Non-Use*, 53 Int'l Org. 433, 463 (1999).

70 Tannenwald, *supra* note 65, at 7.

71 *Id.* at 23-7.

72 Schelling, *supra* note 67, at 932.

73 *Id.* at 929.

It is striking that Schelling, a game theorist and an economist, privileges an imaginative convention over rule-making, institutionalization, or regulation:

> We depend on nonproliferation efforts to restrain the production and deployment of weapons by more and more countries; we may depend even more on universally shared inhibitions on nuclear use. Preserving those inhibitions and extending them, if we know how to preserve and extend them, to cultures and national interests that may not currently share those inhibitions will be a crucial part of our nuclear policy.[74]

In short, nuclear weapons became rehumanized through a normative evolution *outside* of nuclear strategy. This evolution did not overwhelm strategic incentives—the campaign for nuclear abolition never came close to succeeding, and the United States never adopted a no-first-use policy—but it channeled these strategic decisions within a moral vision of the world. As a result, international cooperation emerged within a deeply satisfying ethical framework despite the fundamentally dehumanizing nature of nuclear weapons, a nature which to this day remains basically unchanged.

III. FOCAL POINTS, IMAGINATIVE COOPERATION, AND AUTONOMOUS SYSTEMS

If we take the nuclear taboo as a guide, to develop international norms around autonomous systems, we should take their dehumanizing aspects as given, and instead seek to rehumanize nuclear strategy through the popular imagination. This prescription is not far from that suggested by Sir John Keegan's seminal *The Face of Battle*. There, he stresses the value of a "de-sensitized," even a "dehumanizing," approach to war, one which has helped make Western militaries the most powerful in the world: "the deliberate

---

[74] *Id.* at 935.

injection of emotion into an already highly emotive subject will seriously hinder, if not indeed altogether defeat, the aim of officer-training . . . [given] that battles are going to happen, it is powerfully beneficial."[75] Dehumanization is a necessary part of military power. But Keegan goes on to outline an important role for the military historian, which is to recapture the 'realism' of battle by imbuing its necessarily abstract components with something of their original life; to re-humanize it: hence his title, the *face* of battle.[76]

What Keegan did for battle, John Hersey's *Hiroshima* did for nuclear weapons.[77] His son relates: "He told me about getting the idea of using novelistic devices to structure his reporting. He wanted to put faces and names to the story . . . He wanted to show their humanity."[78] His approach sets a valuable precedent to follow. If we want to understand how to talk and write about the effects of autonomous weapons, his example is a good place to begin.[79]

A. Examples of Focal Points in International Law

One of the most important facts about the nuclear taboo is that it is entirely arbitrary. The distinction upon which it rests does not actually have a basis in any first principles or pure reason. Rather, it is entirely the result of an ethical evolution. Its moral force, which is entirely real, flows from norms that are not inherent but have been

---

75 Keegan, *supra* note 25, at 18-21.

76 In one instructive comparison, he contrasts the historiography of Julius Caesar, whose "subordinate figures are cardboard," with that of Thucydides, whose figures "are individuals, with wills of their own . . .
Thucydides' army . . . [is] the product of human conduct and character at every level." *Id.* at 68.

77 JOHN HERSEY, HIROSHIMA (1999).

78 Russell Shorto, *John Hersey, the Writer Who Let Hiroshima Speak for Itself*, THE NEW YORKER (Aug. 31, 2016), https://www.newyorker.com/culture/culture-desk/john-hersey-the-writer-who-let-hiroshima-speak-for-itself.

79 Some scholarly efforts have recently been exerted in this direction, especially by scholars from the liberal arts, but their ponderous execution and academese have limited their appeal. FROM ABOVE: WAR, VIOLENCE, AND VERTICALITY (Mark Whitehead, Peter Adey, and Alison Williams eds. 2013). Charles J. Dunlap, Jr., *Book Review: From Above: War, Violence, and Verticality*, 44 PARAMETERS 187, 187-89 (2014).

created. Early in the Cold War, Schelling and Hedley Bull converged toward the idea that "'a common language, a common epistemology and understanding of the universe, a common religion, a common ethical code" is necessary for stable cooperation; they differed in that Schelling believed "norms were much more created than made."[80] Schelling turned out to be right. Before returning to autonomous systems, it will be useful to briefly consider two other examples where focal points emerged through the efforts of normative entrepreneurs to shape international law: the campaign to save the whales, and the regulation of outer space.

Similar to the nuclear taboo is the legal moratorium on whaling. The modern norm against whaling is novel; in fact, even in the 1960s U.S. pilots still used whales for target practice.[81] While their killing was regulated by international bodies, it was seen as an entirely legitimate practice. Just as nuclear weapons once were considered just another kind of bomb, so too were whales just another kind of marine life. During the 1960s and 70s, increasing disagreement about the proper bounds on whaling made cooperation difficult. Environmentalists exploited this fissure to create a new focal point for international norms: zero whaling. The clarity of this idea made it easy to coordinate the relevant actors, while their opponents could never create a similarly imaginative focus.[82] Once these actors came to dominate the International Whaling Commission, international law swiftly followed suit. "They [environmentalists] mobilize for political action best when an issue can be framed in fairly simple terms indicating a clear policy preference without the need for a highly detailed explanation of why the chosen policy is better than others."[83] The moratorium has been in force ever since.

---

[80] Robert Ayson, *A Common Interest in Common Interest: Hedley Bull, Thomas Schelling, and Collaboration in International Politics*, *in* REMEMBERING HEDLEY 56, 61 (Coral Bell and Meredith Thatcher eds. 2008).

[81] Mark J. Peterson, *Whalers, Cetologists, Environmentalists, and the International Management of Whaling*, 46 INT'L ORG. 147, 159 n. 23 (1992).

[82] DAVID A. LAKE, JEFFRY A. FRIEDEN, AND KENNETH A. SCHULTZ, WORLD POLITICS: INTERESTS, INTERACTIONS, INSTITUTIONS 501 (2nd Ed. 2013).

[83] Peterson, *supra* note 82, at 155.

Perhaps the best example of creating a new focal point is outer space. Americans take it for granted that space is a sort of commons, something belonging to all humanity, but this was not obvious when its exploration first became possible. There were multiple equilibria. Two of these potential equilibria came to stand out: in one, space would be treated like airspace, a kind of property divided among sovereign nations; in the other, it would be treated like the high seas, navigable to all. Western lawyers, eager for an internationalist vision of the next frontier, began advancing the analogy of the high seas, and this became the foundation of international law on the matter.[84] To be sure, there were traditional balance-of-power concerns in the negotiations; but the legal norms that emerged were essentially determined by imaginative metaphors, *not* power politics. We owe the freedom of space, now a cornerstone of the global economy, to the imaginative analogies drawn in the early years of the Cold War.[85]

B. Creating Focal Points for Autonomous Systems

The major obstacle to cooperation around autonomous systems is the absence of a clear focal point. At present, no coordination point stands out besides non-use, and unlike with nuclear weapons or whaling this outcome is not an equilibrium. The problem is especially pressing, since action needs to be taken before suboptimal patterns evolve on their own.[86]The pressure is especially acute for the United States, who "sets the standard for bombing practices and remains the focus of efforts to change those practices."[87] The technology itself, though, is not the problem, but its dehumanizing effects.[88] As with nuclear weapons, we should look for their solution, not in the technology, but in the way our culture morally imagines it.

---

[84] To settle disputes over celestial objects and property rights, this analogy was later combined with one comparing space to Antarctica, whose legal status had been settled in 1959.

[85] M.J. Peterson, *The Use of Analogies in Developing Outer Space Law*, 51 INT'L ORG. 245, 245-74 (1997).

[86] Brooks, *supra* note 7, at 14.

[87] THE AMERICAN WAY OF BOMBING: CHANGING ETHICAL AND LEGAL NORMS, FROM FLYING FORTRESSES TO DRONES 2 (Matthew A. Evangelista and Henry Shue eds. 2014).

[88] Patrick Eberle, *To UAV or Not to UAV: That is the Question; Here is One Answer*, AIR AND SPACE POWER J. (2001).

In this section, I want to draw three lessons from the previous discussion of nuclear strategy: begin with incentives; keep it simple; and start at the top.

1. Begin with Incentives

Cooperation must begin by identifying potential equilibria—or, put another way, it must begin by studying the relevant incentives. Once again, it is useful to turn to Thomas Schelling:

> As a starting point for legal analysis or moral judgment, it may be helpful to draw the matrix of choices . . . Game theory would be most pertinent to those constraints that affect people's expectations about each other . . . And in that process, the difference between ethics and law, or ethics and instinct, may be less important than the similarities.[89]

Imagination is not an excuse to escape reality. Rather, imagination must begin with what is probable.[90] The simplest proposal for dealing with autonomous weapons is to abolish them or, at the least, to ban their use.[91] This proposal will not succeed. Immediately, the enormous pressures on democratic leaders to avoid casualties mean that nonuse[92] or abolition is not an option.[93] Public support for drone strikes is strong and, by some polls, overwhelming, especially when American lives are not at risk; moreover, this support is

---

89 Schelling, *supra* note 63, at 35-6.

90 Here, my argument tracks that of the philosopher Irving Babbitt. Babbitt contrasts this approach to imagination with that of Rousseau, who said in his "Second Discourse," "Let us begin by setting aside all the facts." IRVING BABBITT, DEMOCRACY AND LEADERSHIP 79 (1924).

91 Robert Sparrow, *Robots and Respect: Assessing the Case Against Autonomous Weapon Systems*, 30 ETHICS AND INT'L AFF. 93, 93-116 (2016).

92 The nuclear taboo was opposed by major political figures in the 1950s (including Dwight Eisenhower); but quickly politicians of all parties came to recognize that, whatever its short-run costs, preventing escalation to nuclear war was worth the price. This is clearly seen in Lyndon Johnson's insistence on the qualitative difference between nuclear and other weapons. Unlike nuclear weapons, with autonomous ones the public *supports* their use.

93 Though, in terms of sheer creativity, it's hard to top the droll. *Campaign to Stop Killer Robots*, https://www.stopkillerrobots.org/about-us/.

bipartisan.[94] So long as capital-rich democracies prefer to lose expensive machines rather than people—a preference that seems universal and enduring—autonomous systems will be a fact of modern war.[95]

Moreover, given their increasing accuracy and precision, many observers argue that armed forces have a *responsibility* to use these new technologies, especially unmanned vehicles.[96] In fact, some legal scholars argue that international law ought to make their use a positive duty.[97] Far from abolishing them, autonomous systems may be ethically required of 21st century war.

Another common proposal is Meaningful Human Control (MHC).[98] Unlike abolition, this proposal does not seek to eliminate the technology but to circumvent the legal and moral dilemmas it poses. It would require human beings to make the most ethically and legally-fraught decisions, taking these responsibilities away from "killer

---

[94] *The Ap-GfK Poll: A Survey of the American General Population (ages 18+)*, GFK PUBLIC AFF. AND CORPORATE COMM. (Apr. 2015), http://surveys.ap.org/data/GfK/AP-GfK Poll April 2015 Topline drones.pdf; Alyssa Brown and Frank Newport, *In US, 65% Support Drone Attacks on Terrorists Abroad*, GALLUP (Mar. 25, 2013), https://news.gallup.com/poll/161474/support-drone-attacks-terrorists-abroad.aspx; Jacquelyn Schneider and Julia Macdonald, *U.S. Public Support For Drone Strikes: When Do American Prefer Unmanned Over Manned Platforms?*, CENT. FOR A NEW AM. SECURITY (Sept. 20, 2016), https://www.cnas.org/publications/reports/u-s-public-support-for- drone-strikes.

[95] Tellingly, the most recent strategic documents by the U.S. military all emphasize autonomous systems as an integral part of emerging doctrine. U. S. MARINE CORPS, *Commandant's Planning Guidance: 38th Commandant of the Marine Corps* 13; *See also* Matthew Fuhrmann and Michael C. Horowitz, *Droning On: Explaining the Proliferation of Unmanned Aerial Vehicles*, 71 INT'L ORG. 397, 397-418 (2017).

[96] Bradley Jay Strawser, *Moral Predators: The Duty to Employ Uninhabited Aerial Vehicles*, 9 J. OF MIL. ETHICS 342, 342-368 (2010).

[97] Oren Gross, *The New Way of War: Is There a Duty to Use Drones?*, 67 FLA. L. REV. 1, 60 (2015).

[98] K. Vignard, *The Weaponization of Increasingly Autonomous Technologies: Considering How Meaningful Human Control Might Move Discussion Forward*, 2 UNIDIR RESOURCES 1, 1-9 (2014).

robots."[99] Though a proposal still in its infancy,[100] there are reasons to be skeptical of its viability or staying power. The rapidity with which a decision must be made as the technology advances will make effective human control impossible. In fact, even so minimal a requirement as "eyes on target" will become not only strategically but morally untenable, as it will make collateral deaths *more* likely.[101] Finally, the relative ease with which nonstate actors can acquire the technology means that state actors will struggle to limit their own use of such instruments.

Simply put, these proposals are not equilibria. Actors face too many pressures to deviate, to cheat, and to change the rules. Keith Abney observes, "given 'ought implies can,' tactical and technical changes [in warfare have] . . . led to changes in just war theory."[102] We should recognize that international law will not succeed in prohibiting this technology, and such efforts may prove very damaging along the way.[103] The technology will progress; we must use our imaginations so that our laws and norms keep pace.[104]

2. Keep it Simple

When debating the ethics of drones, there is a tendency among policymakers and especially among scholars to move too quickly into

---

[99] Mary Ellen O'Connell, *Banning Autonomous Killing: The Legal and Ethical Requirement That Human Make Near-Time Lethal Decisions*, *in* THE AMERICAN WAY OF BOMBING: CHANGING ETHICAL AND LEGAL NORMS, FROM FLYING FORTRESSES TO DRONES 224-36 (Matthew A. Evangelista and Henry Shue eds. 2014).

[100] Marc C. Canellas and Rachel A. Haga, *Toward Meaningful Human Control of Autonomous Weapons Systems Through Function Allocation, in* 2015 IEEE INT'L SYMP. ON TECH. AND SOC'Y (ISTAS) 1-7 (2015).

[101] Keith Abney, Autonomous Robots and the Future of Just War Theory, *in* ROUTLEDGE HANDBOOK OF ETHICS AND WAR: JUST WAR IN THE TWENTY-FIRST CENTURY 343-45 (Nicholas G. Evans, Fritz Allhof, and Adam Henscke eds., 2013).

[102] *Id.* at 338.

[103] Kenneth Anderson and Matthew C. Waxman, *Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How Laws of War Can* (2013).

[104] Here, the failure of the campaign to abolish nuclear weapons, including the recent and ongoing embarrassment of the Treaty on the Prohibition of Nuclear Weapons (which no nuclear state has joined), is instructive: failing to account for incentive structures will doom idealistic schemes for cooperation to years of wasted effort.

the weeds. For instance, Rosa Brooks, a law professor at Georgetown, in testifying to Congress, moved quickly from "encourag[ing] transparency" to forming "a non-partisan blue ribbon commission" and concluding with "creating a [new] judicial mechanism, perhaps similar to the existing Foreign Intelligence Surveillance Court . . . "[105] It is perhaps unfair to criticize a lawyer in front of Congress for focusing on legal remedies; but her testimony typifies a common response. Often, discussion turns to debating legal and moral intricacies accessible only to a highly educated audience.[106] This contrasts strikingly with the development of the nuclear taboo, which began by drawing a distinction comprehensible to the ordinary man; with the campaign to save the whales; and even with the imagination of space as a commons rather than as private property. All of these, while eventually highly technical and recondite, began with simple ideas and simple illustrations. These ideas then informed the imaginations of the publics and politicians whose beliefs would structure subsequent cooperation.

Samuel Taylor Coleridge, the English Romantic, famously compared imagination to a kind of healthy mania. This mania relates everything to a central, organizing passion. He observed: "He that knows the state of the human mind in deep passion must know, that it approaches to that condition of madness, which is not absolute frenzy or delirium, but which models all things to one reigning idea"[107] Thomas Schelling's work followed just this pattern: "extract the central lessons (in this case from international strategy), capture them in simple formulations, and describe them in language that is both lucid and vivid."[108] He became the greatest of the nuclear strategists because he understood the need to order complex phenomena around a simple

---

105 Brooks, *supra* note 7, at 16-18.

106 For a useful overview of constitutional debates around drone strikes, *see* Jordan Cash and Dave Bridge, *Attack of the Drones: National Security, Due Process, and the Constitutionality of Unmanned Strikes, in* PRESIDENTIAL LEADERSHIP AND NATIONAL SECURITY 127-141 (2017).

107 SAMUEL TAYLOR COLERIDGE, SEVEN LECTURES ON SHAKESPEARE AND MILTON 132 (1856). *See also* BASIL WILLEY, NINETEENTH CENTURY STUDIES: COLERIDGE TO MATTHEW ARNOLD 13-14 (1949).

108 Richard Zeckhauser, *Distinguished Fellow: Reflections on Thomas Schelling*, 3 J. of ECON. PERSP. 153, 157 (1989).

center: "If you're going to do theory and develop concepts that you hope will help people have a deeper sense of the structure of some part of the world, then give them concepts they can use easily."[109] Like his books, his lectures at Harvard were celebrated for their illuminating analogies, like comparing nuclear blackmail to child-rearing. Almost alone among the defense intellectuals, Schelling was and remains inescapable: "it is difficult to recall how the world looked before we first saw it with the aid of Thomas Schelling's vision."[110] We should strive for the same clarity of insight. The time for details and minutiae will come, but not until after the overarching ideas are in place. A focal point, to facilitate cooperation, must *stand out*, and prominence demands a kind of simplicity—a simplicity whose essence can be quickly grasped, no matter how complex its details.

3. Start at the Top

Anne-Marie Slaughter, later the Director of Policy Planning under Secretary Clinton, outlined in a 1997 *Foreign Affairs* article a vision of "the real new world order."[111] In this vision, most international cooperation occurs, not between heads of government, but within transnational networks of bureaucrats, judges, regulators, etc, and their counterparts abroad. For instance, "[j]udges are building a global community of law," she says, such as the Organization of the Supreme Courts of the Americas.[112] These networks bypass traditional actors and treaty-making, focusing on extending cooperation at alternative levels of the state. A similar argument is made by legal scholars studying "global administrative law."[113] They emphasize the progress of international legal cooperation among networks of low-level actors and organizations, especially when inter-state cooperation is ineffective or slow-moving, or when existing treaties are inadequate

---

109 *Id.* at 162.

110 Vincent P. Crawford, *Thomas Schelling and the Analysis of Strategic Behavior, in* STRATEGY AND CHOICE 265 (Richard Zeckhauser ed. 1991).

111 Anne-Marie Slaughter, *The Real New World Order*, 75 FOREIGN AFF. 183-197 (1997).

112 *Id.* at 186, 189.

113 Sabino Cassese, *Global Administrative Law: The State of the Art*, 13 INT'L J. OF CONST. L. 465, 465-468 (2015).

to the problems at hand.[114] This transnationalism comes to resemble a new kind of constitutional regime in international law,[115] and it is unsurprising that it presents certain challenges to the norm of sovereignty.[116]

Whether or not Slaughter correctly describes the emerging "new world order," it would be a mistake to begin with such midlevel actors. The proliferation of low-tech drone technology[117]—which scholarship on legal cooperation often ignores[118]—makes coordination among the relevant actors daunting. These technologies disproportionately benefit nonstate actors, especially terrorist groups, and their usage will disrupt efforts at regulation by transnational bureaucrats. Coordination would be better advised to begin at the highest levels, where autonomous technologies are out of reach of all but a few state actors.

Besides technological proliferation, there is a still more important reason to begin with the big issues. The nuclear taboo was created and sustained, not by transnational cooperation, but by the moral imaginations of domestic and international publics. A transnational approach would not appreciate "the importance of democratizing domestic policymaking on nuclear weapons . . . [including] civilian nuclear analysts and arms control groups, and other groups in civil society, as well as public education about nuclear weapons."[119] Whatever norms ultimately guide the use of autonomous

---

114 Benedict Kingsbury and Lorenzo Casini, *Global Administrative Law Dimensions of International Organizations Law*, 6 INT'l ORG. L. REV. 319, 319-358 (2009).

115 David P. Fidler, *Constitutional Outlines of Public Health's New World Order*, 77 TEMP. L. REV. 247 (2004).

116 Anne-Marie Slaughter, *Sovereignty and Power in a Networked World Order*, 40 STAN. J. INT'L L. 283 (2004); Sabino Cassese, *The Globalization of Law*, 37 NYUJ INT'L L. & POL. 973 (2004).

117 One sensationalist article draws attention, not without reason, to the "Kalashnikov drone." Liz Sly, *The Kalashnikov assault rifle changed the world. Now there's a Kalashnikov drone.*, THE WASH. POST (Feb. 23, 2019), https://www.washingtonpost.com/world/2019/02/23/kalashnikov-assault-rifle-changed-world-now-theres-kalashnikov-kamikaze-drone/.

118 Kerry Chavez and Ori Swed, *Low-Tech, High-Yield: Assessing the Marginal Utility of Violent Non-State Actors' Drone Acquisition* (2019).

119 Tannenwald, *supra* note 65, at 48.

systems, they must be ingrained in the popular consciousness before they can constrain international actors. Bypassing the people will not create a sustainable norm.

In the same way that the shaping of the public's moral imagination around a nuclear taboo was more important than the NPT or arms control treaties, so too constraining autonomous systems will depend on shaping the public imagination. Successfully confining drones and "killer robots" will depend on the form these technologies take in the public imagination. A better approach is that of normative entrepreneurship. An extensive literature traces how entrepreneurial actors construct normative frame.[120] These figures focus on public argumentation in order to frame elite-public discourse, which then determines what sort of cooperation occurs.[121] The insight of this constructivist scholarship, in contrast to that of more institutionalist approaches, is the need to create a social idea before mobilizing institutional bodies. If there is no shared conception, there cannot be a focal point of cooperation. Transnationalism, while perhaps an appropriate tool down the road, must proceed from, and not precede, the creation of a common moral axis around which the details of future cooperation can revolve.

IV. CONCLUSION

To some extent, the dehumanizing influence of autonomous systems cannot be escaped, but that does not mean it cannot be counteracted. Nuclear weapons were far more dehumanizing than drones or killer robots will ever be, and their inhumanity was overcome. Their dehumanization was overcome by creating a focal point that could stand out in the popular moral imagination. In some ways, this point was arbitrary—in fact, there is no qualitative difference

---

120 Martha Finnemore and Kathryn Sikkink, *International Norm Dynamics and Political Change*, 52 INT'L ORG. 887, 887-917 (1998); Rodger A. Payne, *Persuasion, Frames, and Norm Construction*, 7 EUR. J. OF INT'L REL. 37, 37-61 (2001).

121 NETA CRAWFORD, ARGUMENT AND CHANGE IN WORLD POLITICS: ETHICS, DECOLONIZATION, AND HUMANITARIAN INTERVENTION (Vol. 81 2002).

between nuclear and other weapons—but its arbitrariness does not make it less ethically salient.

We can draw manifold lessons from nuclear strategy in the moral imagination, but I briefly developed only three. I suggest the way forward should begin by analyzing the incentives of the relevant actors. Such an analysis requires us to look elsewhere than non-use or abolition. It should then proceed to imaginative construction of focal points around which cooperation might emerge. These foci will necessarily be simple: experts who move too quickly to minutiaea misunderstand the state of the problems we face. Finally, to construct these imaginative foci, I suggest normative entrepreneurs begin at the top and work their way down. By doing so, they will both bypass the thorniest areas of cooperation (where the multiplicity of actors will make coordination extremely difficult) and, more importantly, they will form the moral imaginations of the most important actors: the democratic publics and politicians who alone can sustain long-term cooperation.

The dehumanizing effects of weapons like drones and nuclear warheads cannot be counteracted in the moment of their use, nor in adjustments to the technology itself, nor in the way strategy is formulated; rather, the weapons are humanized by the ways the larger culture comes to talk about them. So, here is the question facing the next generation of weaponry: can we draw clear moral distinctions that align with our incentives, focus our cooperation, and, by capturing our imaginations, re-humanize a dehumanizing technology? It is the same question that faced nuclear weapons in the early Cold War. Given our success then, we need not fear a little optimism now.