Benchmarks for Reducing Civilian Harm in Armed Conflict: Learning Feasible Lessons About Systemic Change

Peter Margulies*

INTRODUCTION

While the U.S. drone strike that killed Al Qaeda leader Dr. Ayman al-Zawahri on July 31, 2022 apparently did not harm any civilians, sharp questions have arisen recently about the U.S. commitment to reducing such harm in armed conflicts.[1] Other countries engaged in armed conflicts, such as Russia in Ukraine, have triggered more severe impacts.[2] However, the U.S. military has admitted that several operations that may have cost civilian lives stemmed from inaccurate or incomplete information.[3] In addition, the *New York Times* documented multiple errors in U.S. strikes against the Islamic State

* Professor of Law, Roger Williams University School of Law; B.A., Colgate; J.D., Columbia Law School.

1. Before approving the strike on Dr. Zawahri, President Biden asked about the risk of collateral damage. *See* Julian E. Barnes & Eric Schmitt, *How the C.I.A. Tracked the Leader of Al Qaeda*, N.Y. TIMES (Aug. 2, 2022), https://www.nytimes.com/2022/08/02/us/politics/cia-qaeda-al-zawahri.html.

2. *See* Beth Van Schaack, Ambassador-at-Large for Global Criminal Justice, U.S. Dep't of State, Foreign Press Centers Briefing, *War Crimes and Accountability in Ukraine* (June 15, 2022), https://www.state.gov/briefings-foreign-press-centers/war-crimes-and-accountability-in-ukraine (describing Russia's apparent "deliberate and indiscriminate attacks against . . . civilian infrastructure within Ukraine.").

3. David Vergun, *Air Force Official Briefs Media on Deadly Drone Strike in Kabul*, DOD NEWS (Nov. 3, 2021), https://www.defense.gov/News/News-Stories/Article/Article/2831896/air-force-official-briefs-media-on-deadly-drone-strike-in-kabul/ (reporting on briefing by U.S. Air Force Inspector General, Lt. Gen. Sami D. Said, discussing "inaccurate" interpretation of facts and incomplete information in the Kabul strike in the waning days of U.S. involvement in Afghanistan); Michael X. Garrett, Dep't of the Army, Memorandum for Secretary of Defense, Executive Summary: Review of the Civilian Casualty Incident that Occurred on 18 March 2019 at Baghuz, Syria (May 11, 2022), https://int.nyt.com/data/documenttools/mfr-for-secdef-v/21afedb925372e3b/full.pdf [hereinafter Garrett Executive Summary] (acknowledging flaws in information known to U.S. targeters at the time of air strike during a 2019 battle between U.S.-supported Kurdish militia and ISIS fighters).

(ISIS) organization.[4] Questions about U.S. airstrikes led the U.S. Department of Defense to offer its Civilian Harm Mitigation Response Action Plan in August 2022.[5] Although there have been ebbs and flows in civilian harm attributable to U.S. actions,[6] a more consistent approach is needed. This Article argues that systemic efforts to reduce harm fall within states' legal duty to use "constant care" to spare civilians.[7]

Four key barriers to effective civilian harm-reduction require systemic remedies: access to technology, cognitive flaws, procedural deficits, and inadequate training and institutions. Access to improved technology, such as high-resolution video and artificial intelligence (AI), is needed.[8] Cognitive barriers that adversely affect targeting

---

4. *See, e.g.* Azmat Khan, *Hidden Pentagon Records Reveal Patterns of Failure in Deadly Airstrikes,* N.Y. TIMES (Dec. 18, 2021), www.nytimes.com/interactive/2021/12/18/us/airstrikes-pentagon-records-civilian-deaths.html; Azmat Khan et al., *Documents Reveal Basic Flaws in Pentagon Dismissals of Civilian Casualty Claims*, N.Y. TIMES (Jan. 6, 2022), https://www.nytimes.com/2021/12/31/us/pentagon-airstrikes-syria-iraq.html (documenting flaws in U.S. military investigations of alleged civilian casualties); Dave Philipps, Eric Schmitt & Mark Mazzetti, *Civilian Deaths Mounted as Secret Unit Pounded ISIS*, N.Y. TIMES (Dec. 12, 2021), https://www.nytimes.com/2021/12/12/us/civilian-deaths-war-isis.html (focusing on Special Operations Talon Anvil unit).

5. *See* U.S. Dep't of Defense, *Civilian Harm Mitigation Response Action Plan (CHMR-AP)* 6 (Aug. 25, 2022) (hereinafter DoD Civilian-Harm Mitigation Plan), https://media.defense.gov/2022/Aug/25/2003064740/-1/-1/1/CIVILIAN-HARM-MITIGATION-AND-RESPONSE-ACTION-PLAN.PDF.

6. *See* MITT REGAN, DRONE STRIKE: ANALYZING THE IMPACTS OF TARGETED KILLING 227 (2022).

7. *See* Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), June 8, 1977, 1125 U.N.T.S. 3. [hereinafter Additional Protocol I]. Scholars have suggested that such a systemic approach is required by international law. *See* Peter Margulies, *The Other Side of Autonomous Weapons: Using Artificial Intelligence to Enhance IHL Compliance*, *in* THE IMPACT OF EMERGING TECHNOLOGIES ON THE LAW OF ARMED CONFLICT (Ronald T.P. Alcala & Eric Talbot Jensen eds., 2019) (arguing that "constant care" duty includes efforts to use new technology, such as artificial intelligence (AI), to provide attack planners with more information and ensure more effective deliberation and implementation regarding targeting); Asaf Lubin, *The Reasonable Intelligence Agency*, 47 YALE J. INT'L L. 119, 140–47 (2022) (discussing duty of intelligence agencies to provide sound and comprehensive information that attack planners utilize in planning strikes).

8. *See* Vergun, *supra* note 3 (noting the conclusion of the U.S. Air Force Inspector General that lack of high-resolution video played a role in the mistaken August 2021 Kabul strike); LARRY LEWIS & ANDREW ILACHINSKI, CENTER FOR NAVAL ANALYSES, LEVERAGING AI TO MITIGATE CIVILIAN HARM 20–28 (2022), https://www.cna.org/archive/CNA_Files/pdf/leveraging-ai-to-mitigate-civilian-harm.pdf [hereinafter CNA REPORT] (noting that AI could provide more information about civilian patterns of life that attack planners could use to reduce civilian harm); Margulies, *supra* note 7 (discussing the role of technology in flagging the unanticipated presence of civilians at target site). Users of AI must take care to

include confirmation bias: the human tendency to read all evidence as confirming the decisionmaker's current thesis, even if new evidence is either neutral or inconsistent with that thesis.[9] In U.S. targeting, flawed compliance with procedures has played a role; certain targeting units relied excessively on claims that unit self-defense against an imminent threat justified the use of lethal force. Invoking these claims of exigency sidestepped rules of engagement (ROE) and processes for deliberate targeting that senior commanders and civilian officials had imposed to limit civilian harm.[10] Inadequate training and an absence of institutions also have an adverse effect. According to the independent RAND Corporation, training in avoiding civilian harm is "negligible" in many U.S. combatant commands.[11] Widening this gap, RAND also critiqued the absence in the U.S. military of "structures and capabilities for... analyzing and monitoring civilian-harm trends over time...."[12] Indeed, each of the barriers described here has an institutional component; large organizations that operate in high-risk settings often discount small errors that can snowball into crises or ignore wide swings in decisions on similar

address racial, gender, socio-economic, and cultural biases that can emerge from information that developers input to AI agents. *See* Peter Margulies, *Autonomous Weapons in the Cyber Domain: Balancing Proportionality and the Need for Speed*, 96 INT'L L. STUD. 394, 408–09 (2020) (cautioning about the risk of bias in AI agents); *cf.* Kristin N. Johnson, *Automating the Risk of Bias*, 87 GEO. WASH. L. REV. 1214, 1239–42 (2019) (discussing the risk of AI bias in the consumer and financial sector); Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 PROCEEDINGS MACHINE LEARNING RSCH. 1 (2018) (discussing problems of AI in identifying women of color); David S. Rubenstein, *Acquiring Ethical AI*, 73 FLA. L. REV. 747, 775–77 (2021) (discussing ways in which government procurement policies can address the risk of bias).

9. *See* DANIEL KAHNEMAN ET AL., NOISE: A FLAW IN HUMAN JUDGMENT 169, 172 (2021); Mark S. Martins, *Rules of Engagement for Land Forces: A Matter of Training, Not Lawyering*, 143 MIL. L. REV. 1, 43–45 (1994) (discussing the impact of confirmation bias on targeting decisions); *see also* Vergun, *supra* note 3 (noting that the U.S. Air Force Inspector General found that confirmation bias played a role in the mistaken 2021 Kabul strike).

10. *See* Khan, *Hidden Pentagon Records Reveal Patterns of Failure in Deadly Airstrikes*, *supra* note 4; *cf.* Gary P. Corn, *Should the Best Offense Ever Be a Good Defense? The Public Authority to Use Force in Military Operations: Recalibrating the Use of Force Rules in the Standing Rules of Engagement*, 49 VAND. J. TRANSNAT'L L. 1, 9–12 (2016) (discussing rationale for carefully tailored unit self-defense rules and arguing that on occasion units interpreted rules unduly aggressively).

11. MICHAEL J. MCNERNEY, GABRIELLE TARINI, KAREN M. SUDKAMP, LARRY LEWIS, MICHELLE GRISÉ & PAULINE MOORE, RAND CORP., U.S. DEPARTMENT OF DEFENSE CIVILIAN CASUALTY POLICIES AND PROCEDURES 56 (2022), https://www.rand.org/pubs/research_reports/RRA418-1.html [hereinafter RAND REPORT]; *see* Garrett Executive Summary, *supra* note 3, at 2 (acknowledging the need for more training on the protection of civilians).

12. RAND Report, *supra* note 11, at 59.

facts by organization personnel.[13]

The law of armed conflict (LOAC)—sometimes called international humanitarian law (IHL)—applicable to individual attack planners does not address these obstacles to civilian-harm reduction.[14] Derived from the balance of military necessity and humanity, LOAC centers on a guiding principle and two key rules. The principle of distinction bars the targeting of civilians in an armed conflict.[15] The rule of proportionality prohibits attacks that a planner expects will prompt "excessive" collateral harm when the planner measures harm against the military advantage that an attack will yield.[16] Under the rule of precautions in attack, planners must take "feasible" measures to minimize the harm to civilians expected from an attack.[17] Because of express and implicit qualifiers on the IHL duties of attack planners, flawed technology, cognition, and intelligence will sometimes result in avoidable civilian harm that does not rise to the level of an IHL violation by an individual planner.

To show the gap that may arise between avoidable civilian harm and IHL provisions governing individual planners, consider the limitation on required precautions to safeguards that are "feasible." Suppose high-resolution video would facilitate the targeting cell's ability to discern civilians at a target site. Although even in sophisticated militaries such as that of the United States the

---

13. *See* Peter M. Madsen, *Organizational Learning as Reliability Enhancement, in* ORGANIZING FOR RELIABILITY: A GUIDE FOR RESEARCH AND PRACTICE 143, 147 (Rangaraj Ramanujam & Karlene H. Roberts eds., 2018) (observing that managers may be lulled into complacency and come to mistakenly believe that "resources devoted to safety may be reduced"); *see also* KAHNEMAN, ET AL., *supra* note 9, at 24–27, 248–53 (noting that organizations either tolerate or fail to detect wide variations, which authors attribute to irrelevant factors or "noise," among decisionmakers within the entity); *see also id.* at 17 (noting that variations in rulings by juvenile court judges in cases with similar facts correlate with the performance of the local football team: if the team loses, the judge issues harsher decisions).

14. This Article uses these terms interchangeably.

15. Additional Protocol I, *supra* note 7, art. 51(2); *see* TALLINN MANUAL 2.0 ON THE INTERNATIONAL LAW APPLICABLE TO CYBER OPERATIONS 422–23 (Michael N. Schmitt ed., 2d ed. 2017) [hereinafter TALLINN MANUAL 2.0]; YORAM DINSTEIN, THE CONDUCT OF HOSTILITIES UNDER THE LAW OF ARMED CONFLICT 102 (3d ed. 2016).

16. Additional Protocol I, supra note 7, art. 51(5)(b), 57(2)(a)(iii); TALLINN MANUAL 2.0, supra note 15, at 471; *see also* Isabel Robinson & Ellen Nohle, *Proportionality and Precautions in Attack: The Reverberating Effects of Using Explosive Weapons in Populated Areas*, 98 INT'L REV. RED CROSS 107, 112 (2016) (arguing that the rule of proportionality covers secondary effects of attacks, such as reasonably foreseeable damage to civilian infrastructure, including sewer systems, that stems from an attack on a structure or individuals located on ground above the system).

17. Additional Protocol I, supra note 7, art. 57(2)(a)(ii); *see also* Geoffrey S. Corn, *War, Law, and the Oft Overlooked Value of Process as a Precautionary Measure*, 42 PEPP. L. REV. 419, 459 (2015).

exigencies of armed conflict can reduce any given unit's access to technology, diminished access can also result from poor supply-chain management that was entirely preventable.[18] The attack planner is not responsible for flaws in supply-chain management, and it is obviously not feasible for a targeting cell to overhaul a state force's logistics while it is planning an attack. A planner who relies on lower-resolution video in preparing for an attack has not committed an IHL violation unless a reasonable commander would view the video feed available as too vague to distinguish combatants from protected civilians. Similarly, avoidable flaws in intelligence collection and analysis may classify an individual as the operational leader of an active terrorist group in an armed conflict with a state force. The appearance of a leader such as Al Qaeda's Dr. Zawahiri was known throughout the world.[19] However, that degree of notoriety is rare.[20] A commander who receives information from a state intelligence service classifying an individual as an operational head of a terrorist group must rely on "reasonably available" information.[21] The commander in the field cannot conduct a fresh research project on each person that a state intelligence agency designates as a senior operational figure.[22] Therefore knowledge of the mistaken identification is not "reasonably available" to the commander, although the mistake was preventable. Addressing such gaps requires proactive state policies, not merely compliance by individual commanders.

This Article argues that states have a *lex lata* duty—a duty that is binding under current law—to adopt a systemic approach. The law of armed conflict's duty to make systemic efforts starts with the

---

18. *See* Christoph Bode, Stephan M. Wagner, Kenneth J. Petersen & Lisa M. Ellram, *Understanding Responses to Supply Chain Disruptions: Insights from Information Processing and Resource Dependence Perspectives*, 54 ACAD. MGMT. J. 833, 834–40 (2011) (explaining how experience, cognition, and interpersonal relationships affect firms' approaches to supply-chain management).

19. *See Ayman Al-Zawahiri: Who Was Al-Qaeda Leader Killed by US*, BBC (Aug. 2, 2022), https://www.bbc.com/news/world-middle-east-13789286.

20. Al-Zawahiri's killing raises questions under LOAC, although intelligence failures do not come into play. Those questions arise because of the passage of time since 9/11, Al Qaeda's current limited activity, and the apparent limits on Al-Zawahiri's recent operational role. Questions of this kind involve the interpretation of facts, not disagreements about the accuracy of intelligence reporting those facts. *See* Robert Chesney, *On the Legality of the Strike that Killed Ayman al-Zawahiri*, LAWFARE (Aug. 3, 2022), https://www.lawfareblog.com/legality-strike-killed-ayman-al-zawahiri.

21. *See* TALLINN MANUAL 2.0, *supra* note 15, at 424 (referring to United Kingdom law of war manual); *see id.* at 432 (noting consensus of international experts contributing to TALLINN MANUAL).

22. *See* Lubin, *supra* note 7, at 134–36.

constant-care obligation.[23] While most analysts address the place of constant care in the deliberation of individual attack planners,[24] a broader role for this duty is required by the term itself, as well as its logic. The adjective "constant" indicates that care applies at all levels and at all times. On this view, constant care occurs in both individual and systemic contexts. The two contexts are interdependent: inattention to systemic issues, including the dissemination of technology and knowledge about LOAC itself, would limit the utility of individual choices by attack planners. Similarly, systemic improvements lack meaning and impact if they fail to reduce civilian harm caused by individual attacks. Ensuring that attack planners' choices are not idle gestures, but instead actually matter in reducing civilian harm, is one aspect of a state's duty under international law to comply with LOAC in good faith.[25] The systemic view attracts additional support from the International Court of Justice's view in the *Nuclear Weapons* case that compliance with IHL indicates that a state has followed human rights law's ban on arbitrary deprivations of life.[26] Lethal targeting would be arbitrary if it stemmed from the obstacles described in this Article, such as haphazard logistics, flawed cognition, inadequate training, and failure to follow a state's own procedures. Therefore, a state's compliance with IHL should include systemic efforts to address these issues.

While a state's duty to make systemic efforts is *lex lata*, good faith is a relaxed standard. The particular form of systemic efforts will vary widely with each state's capabilities and resources. As one *lex ferenda* suggestion, this Article advances a benchmarking approach. Benchmarking would entail setting goals and seeking to preserve or restore past superior performance. It would focus on best practices, not quantitative measures, although the latter may be helpful as one index of performance. To aid in this endeavor, the benchmarking approach adopts mainstays of administrative law:[27] an assessment of

23. Additional Protocol I, *supra* note 7, art. 57(1).

24. *See generally* Corn, *supra* note 17.

25. *See* Michael N. Schmitt & Sean Watts, *Common Article I and the Duty to "Ensure Respect"*, 96 INT'L L. STUD. 674, 684–85 (2020) (discussing duty of good faith, sometimes known as *pacta sunt servanda* (adhere to the purposes of the agreement) in treaty and customary law).

26. Legality of Threat or Use of Nuclear Weapons, Advisory Opinion, 1996 I.C.J. 226, ¶ 25 (July 8) [hereinafter *Nuclear Weapons*] (observing that the "protection of the International Covenant of Civil and Political Rights [including the prohibition on arbitrary deprivations of life] does not cease in times of war" and that compliance with IHL will generally comport with this prohibition).

27. Professor Lubin has also discussed the importance of administrative law concepts, although his approach does not rely on the specific methods and doctrines outlined above, and also owes much to tort law. *See* Lubin, *supra* note 7, at 143–48.

a policy's effects, here called a civilian impact statement;[28] reasoned explanation of policy choices;[29] and notice and comment procedures that include both internal stakeholders like other government units and external stakeholders, such as nongovernmental organizations (NGOs).[30] These methods will improve deliberation, refine priorities, and check arbitrary actions and omissions regarding civilian-harm reduction.

Professor Chachko has discussed the relevance of administrative law to national security. Elena Chachko, *Administrative National Security*, 108 GEO. L.J. 1063 (2020). Professor Chachko's work has focused more on the relationship in U.S. law between presidential control over operational national security decisions and the influence of agency officials. *Id.* at 1115–22.

28. Pulp Mills on River Uruguay (Arg. v. Uru.), Judgment, 2010 I.C.J. 14, ¶ 204 (Apr. 20) (asserting international law duty to "undertake an environmental impact assessment where there is a risk that the proposed industrial activity may have a significant adverse impact in a transboundary context"); Eric Talbot Jensen & Sean Watts, *Cyber Due Diligence*, 73 OKLA. L. REV. 645, 676–78 (2021) (discussing *Pulp Mills* decision); Tseming Yang, *The Emergence of the Environmental Impact Assessment Duty as a Global Legal Norm and General Principle of Law*, 70 HASTINGS L.J. 525 (2019) (linking emerging international law requiring EIS with examples from states' municipal (domestic) law).

29. *See* Motor Vehicle Mfrs. Ass'n v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 43 (1983) (holding that agency action—in that case a decision to abandon the requirement that car manufacturers install air bags—is arbitrary and capricious under U.S. Administrative Procedure Act if the action lacks a reasoned explanation); Dep't of Homeland Sec. v. Regents of Univ. of California, 140 S. Ct. 1891, 1912–13 (2020) (holding that the Trump administration had failed to provide reasoned explanation of its effort to rescind the Deferred Action for Childhood Arrivals (DACA) program); *see also* Peter Margulies, *The DACA Case: Agencies' "Square Corner" and Reliance Interests in Immigration Law*, 2019 CATO SUP. CT. REV. 127, 149–51 (2020) (discussing the application of reasoned explanation requirement in *Regents*); Benjamin Eidelson, *Reasoned Explanation and Political Accountability in the Roberts Court*, 130 YALE L.J. 1748, 1773–85 (2021) (discussing the reasoned explanation requirement as a method for encouraging executive candor); *cf.* Jonathan H. Adler, *Super Deference and Heightened Scrutiny*, 74 FLA. L. REV. 267, 272–74 (2022) (suggesting that the reasoned explanation requirement encourages agency deliberation but leaves substantial room for agency policy choices).

30. Elec. Privacy Info. Ctr. v. Dep't of Homeland Sec., 653 F.3d 1, 6 (D.C. Cir. 2011) (discussing the rationale for notice and comment procedures under APA); Nina A. Mendelson, *Regulatory Beneficiaries and Informal Agency Policymaking*, 92 CORNELL L. REV. 397, 408–409 (2007) (arguing that notice-and-comment procedure remedies the "lack of procedural discipline [that] can raise the risk of agency action that . . . does not properly engage public preferences"). In one of the initiatives discussed in this Article, the U.S. DoD has recognized the need to "[b]uild enduring engagements and collaboration across industry, academia, and civil society to promote development, adoption, and implementation" of responsible AI practices. *See* U.S. DEP'T OF DEFENSE RESPONSIBLE AI WORKING COUNCIL, U.S. DEPARTMENT OF DEFENSE RESPONSIBLE ARTIFICIAL INTELLIGENCE STRATEGY AND IMPLEMENTATION PATHWAY 30 (June 2022), https://media.defense.gov/2022/Jun/22/2003022604/-1/-1/0/Department-of-Defense-Responsible-Artificial-Intelligence-Strategy-and-Implementation-Pathway.PDF [hereinafter RAI PATHWAY].

The Article is in five Parts. Part I distills IHL rules and principles. Part II discusses obstacles to civilian-harm reduction. Part III outlines the legal foundation for the systemic approach. Part IV outlines the benchmarking method and applies that method to three examples: the relationship between President Obama's Presidential Policy Guidance (PPG) on drone strikes outside active battlefields and President Trump's Principles, Standards, and Procedures (PSP) revising President Obama's guidance;[31] the recent multilateral Political Declaration on protecting civilians in urban settings from explosive weapons;[32] the U.S. DoD's pathway on responsible AI;[33] and the U.S. DoD's Civilian Harm Mitigation Response Action Plan. Part V addresses possible alternatives to the benchmarking method.

I. THE LOAC FRAMEWORK

Analysis of civilian harm requires an understanding of underlying LOAC rules and principles. The law of armed conflict—in Latin, the *jus in bello*—governs the actual conduct of hostilities between states or between a state and a nonstate actor, such as an internal rebel force or terrorist group. The coverage of LOAC thus distinguishes it from international law on the use of force—the *jus ad bellum*—which governs the justifications for initiating an armed conflict.[34] The rules and principles of LOAC reflect customary

---

31. *See* PRESIDENT BARACK OBAMA, PROCEDURES FOR APPROVING DIRECT ACTION AGAINST TERRORIST TARGETS LOCATED OUTSIDE THE UNITED STATES AND AREAS OF ACTIVE HOSTILITIES (May 22, 2013), https://www.justice.gov/oip/foia-library/procedures_for_approving_direct_action_against_terrorist_targets/download [hereinafter PPG]; PRESIDENT DONALD TRUMP, PRINCIPLES, STANDARDS, AND PROCEDURES FOR U.S. DIRECT ACTION AGAINST TERRORIST TARGETS (2017), https://int.nyt.com/data/documenttools/trump-psp-drone-strike-rules-foia/52f4a4baf5fc54c5/full.pdf [hereinafter PSP].

32. *See* Political Declaration on Strengthening the Protection of Civilians from the Humanitarian Consequences Arising from the Use of Explosive Weapons in Populated Areas, Nov. 18, 2022, https://www.dfa.ie/media/dfa/ourrolepolicies/peaceandsecurity/ewipa/EWIPA-Political-Declaration-Final-Rev-25052022.pdf [hereinafter Political Declaration]; *cf.* Michael W. Meier, *A New Political Declaration on Civilian Harm: Progress or Mythical Panacea?*, LIEBER INST. (July 20, 2022), https://lieber.westpoint.edu/political-declaration-civilian-harm/ (discussing Political Declaration).

33. DoD Civilian Harm Mitigation Response Action Plan, *supra* note 5. This section also addresses the DOD blueprint for artificial intelligence. *See* RAI PATHWAY, *supra* note 30.

34. *See* Charter of the United Nations and Statute of the International Court of Justice arts. 2(4), 51, June 26, 1945, 59 Stat. 1031; T.S. No. 933; 3 Bevans 1153 (respectively barring the use of force generally and allowing states to respond in self-defense to an armed attack). Scholars and practitioners strive to keep the *jus in bello* and *jus ad bellum* separate, since conflating them could confuse commanders and impose liability on ordinary soldiers for acts of aggression over which the soldiers

international law (CIL). In addition, many states have approved Additional Protocol I (AP I) of the Geneva Conventions, which outlines these rules and principles.[35]

A. LOAC'S BALANCING ACT

The laws of armed conflict turn on the balance between humanity and military necessity.[36] Military necessity is a crucial component, because LOAC assumes that armed conflicts will occur for a range of reasons, including aggression by one state against another, such as Russia's recent intervention in Ukraine. While international law bars the use of force in all instances except self-defense against an armed attack[37], LOAC provides guidance on *how* a state should use force, whether or not that force is permissible in the first instance. The military-necessity value in LOAC is both enabling and limiting. It is enabling because it acknowledges that states have a legitimate interest in pursuing military advantage in an armed conflict. Since states do not want to wage endless wars or leave themselves vulnerable to an adversary, the military-necessity value gives states leeway in planning attacks or other military actions that provide an edge over an adversary.

As we shall see when we move beyond these general values and discuss more specific rules and principles, IHL recognizes that actions that are necessary militarily will sometimes entail incidental harm to civilians and civilian objects. Whether an armed conflict is kinetic or virtual, civilians and civilian objects such as dwellings, hospitals, and cultural sites are often close to the site of hostilities. All civilian harm is tragic, but because of that proximity, no state or other force in an

---

had no control. Robert D. Sloane, *The Cost of Conflation: Preserving the Dualism of* Jus ad Bellum *and* Jus in Bello *in the Contemporary Law of War*, 34 Yale J. Int'l L. 47, 48–49 (2009).

35. The United States has not ratified Additional Protocol I, but regards most of its provisions as customary law. *See* Michael J. Matheson, *The United States Position on the Relation of Customary International Law to the 1977 Protocols Additional to the 1949 Geneva Conventions*, 2 Am. U. J. Int'l L. & Pol'y 419 (1987).

36. Michael N. Schmitt, *Military Necessity and Humanity in International Humanitarian Law: Preserving the Delicate Balance*, 50 VA. J. INT'L L. 795, 796 (2010); see also David Wallace, Shane Reeves & Trent Powell, *Direct Participation in Hostilities in the Age of Cyber: Exploring the Fault Lines*, 12 HARV. NAT'L SEC. J. 164, 183–84 (2021) (arguing that an unduly narrow definition of lawful targets can disrupt the balance of military necessity and humanity and therefore impede states' acceptance of LOAC).

37. *See* Matthew C. Waxman, *Regulating Resort to Force: Form and Substance of the UN Charter Regime*, 24 Eur. J. Int'l L. 151, 151 (2013).

armed conflict, whether that force is aggressive or acting in self-defense, can avoid causing *some* harm to civilians. At the same time, military-necessity entails tethering attacks to a specific, concrete military purpose, such as gaining territory currently held by an adversary, countering an attack, disrupting the adversary's chain of command, or impairing the adversary's ability to supply its military in the field. Attacks that are indiscriminate or random do not fit the military-necessity value. Moreover, even attacks that comport with military necessity are still subject to the constraints of humanity. The principle of humanity prohibits needless harm to civilians or those, such as wounded or detained combatants, who are outside the field of combat. While conjoining humanity and the inherent violence of war may seem oxymoronic,[38] LOAC assumes that integrating these overarching values will both recognize the ongoing existence of armed conflict and regulate armed conflict's excesses.

B. THE PRINCIPLE OF DISTINCTION

The law of armed conflict's core principle of distinction requires that combatants refrain from targeting civilian persons or objects.[39] To implement this principle, an attack planner should gather all

---

38. One scholar has forcefully articulated this view in a recent work. *See* SAMUEL MOYN, HUMANE: HOW THE UNITED STATES ABANDONED PEACE AND REINVENTED WAR (2021).

39. See Additional Protocol I, supra note 7, art. 48 (noting obligation to distinguish between civilians and combatants); *id.* art. 51(2) (mandating that "[t]he civilian population as such, as well as individual civilians, shall not be the object of attack"); TALLINN MANUAL 2.0, *supra* note 15, at 422. In NIACs between states, whose armed forces typically wear uniforms, and nonstate actors, whose forces often do not, it can be difficult for a state military force to determine whether a possible target is a civilian outside of combat or a person directly participating in hostilities (DPH) on behalf of a nonstate actor such as ISIS. *See* Kenneth Watkin, *Fighting at the Legal Boundaries: Controlling the Use of Force in Contemporary Armed Conflict*, 42 N.Y.U. J. INT'L L. & POL. 307–11 (2016) (discussing debate); Rachel E VanLandingham, *Meaningful Membership: Making War a Bit More Criminal*, 35 CARDOZO L. REV. 79, 101–11 (2013) (discussing targeting criteria); Wallace, et al., *supra* note 36, at 183–84 (outlining view of United States and its allies that narrow DPH definition will undermine state acceptance of LOAC); *see also* Michael N. Schmitt, *Deconstructing Direct Participation in Hostilities: The Constitutive Elements*, 42 N.Y.U. J. INT'L L. & POL. 697, 699 (2010) (discussing disagreement between United States and allied states, who argued for a functional definition of DPH resting on the operational role in armed group, and other contributors to study by International Committee of the Red Cross (ICRC), who sought a narrower definition that the United States and its allies asserted created a "revolving door" for nonstate actors); Kenneth Watkin, *Opportunity Lost: Organized Armed Groups and the ICRC "Direct Participation in Hostilities" Interpretive Guidance*, 42 N.Y.U. J. Int'l L. & Pol. 641, 643–44 (2010) (same).

"reasonably available" information to ascertain whether a possible target is a lawful objective.[40] As with the other rules and principles of IHL, reasonableness is key. The planners of an attack need not seek to gather *all* information that might possibly be available. That would impose too high a standard, since collection of all data is a never-ending task and some residue of uncertainty is present in most human decisions.

C. THE RULE OF PROPORTIONALITY

Reasonableness is also the hallmark of the rule of proportionality, which bars attacks that a commander expects will cause "excessive" collateral civilian harm when that harm is weighed against the anticipated military advantage.[41] There is no rigid, quantitative formula for gauging excessive harm in a particular operational context. The term "excessive" is by definition highly context-dependent. For example, killing a single enemy foot soldier will entail a lower threshold for excessive civilian harm than killing a senior commander, such as Osama bin Laden or the Russian generals killed by Ukrainian forces in the current conflict in Ukraine.[42] But for any legitimate target, there is a dividing line, however blurry at the margins, between expected harm that—while tragic—fits within the rule and harm that violates it. Moreover, as this explanation indicates, the assessment of what counts as excessive turns on the information reasonably available to the planner of an attack *before* the attack, when planning is occurring.[43] In some cases, to provide more information to the planner of an attack, the chain of command for a state military force may act proactively, by compiling a "no-strike list" (NSL) of protected sites, including "historical, archaeological,

---

40. See Tallinn Manual 2.0, *supra* note 15, at 424 (citing United Kingdom law of war manual); *cf.* William H. Boothby, The Law of Targeting 71 (2012) (explaining that attack planner should assess character of a proposed target in "good faith" and based on "available information," taking all "feasible precautions" to avoid mistakes).

41. Additional Protocol I, supra note 7, arts. 51(5)(b), 57(2)(a)(iii). Additional Protocol I, along with CIL, defines excessive harm by weighing that harm against the military advantage that a commander anticipates. Id. art. 57(2)(a)(iii); Tallinn Manual 2.0, supra note 15, at 471–72; *cf.* Michael Newton & Larry May, Proportionality in International Law 4 (2014).

42. Tallinn Manual 2.0, supra note 15, at 473 (noting that "extensive collateral damage [to civilians or civilian objects] may be legal if the anticipated concrete and direct military advantage is sufficiently great").

43. *Id.* at 474 (observing that the rule of proportionality "requires an assessment of the reasonableness of the determination at the time the attack in question was planned, approved, or executed"; at that time, "all apparently reliable information that is reasonably available must be considered").

economic, and politically sensitive" locations.[44]

D. THE RULE OF PRECAUTIONS IN ATTACK

The rule of precautions in attack mandates that "constant care . . . be taken to spare" civilians.[45] The term, "constant," indicates that the duty is of a "continuing nature" that does not turn on arbitrary starting or stopping points.[46] The passive voice of the language of Additional Protocol I—"constant care *shall be taken*"—leaves some interpretive leeway for imposing this duty not merely on individual commanders planning an attack, but on all those in the chain of command who play a role in armed conflict, including senior officials; this duty applies "[i]n the conduct of military operations."[47] While that language might suggest that the duty is limited to specific military activities, not to overall planning and supply of a military campaign, the next subsection of Article 57 starts with more particular phrasing that flags precautions "[w]ith respect to attacks."[48] The resort to particularized language on "attacks" suggests that the earlier language on "conduct of military operations" represents a deliberate choice to use more

---

44. U.S Dep't of Def. Law of War Manual, § 5.10.3 n. 325 (2016); Peter Margulies, *At War with Itself: The DoD Law of War Manual's Tension Between Doctrine and Practice on Target Verification and Precautions in Attack*, *in* THE UNITED STATES DEPARTMENT OF DEFENSE LAW OF WAR MANUAL: COMMENTARY AND CRITIQUE 201, 209–10 (Michael A. Newton, ed. 2018); *see also* Ronald T.P. Alcala, *Babylon Revisited: Establishing a Corps of Specialists for the Protection of Cultural Property in Armed Conflict*, 6 HARV. NAT'L SEC. J. 206, 233–39 (2015) (discussing special legal protections for cultural property, including multilateral treaties, and suggesting that this protection is also part of binding customary international law; asserting that the United States failed to fully observe its international obligations in this regard during the Second Gulf War; and recommending creation of a trained unit within the U.S. military to ensure future compliance).

45. Additional Protocol I, supra note 7, art. 57(1); Tallinn Manual 2.0, supra note 15, at 477 (interpreting "constant care" language as requiring "commanders and all others involved in the operations to be continuously sensitive to the effects of their activities" on civilians and to "seek to avoid any unnecessary effects" on civilian persons or objects); *cf. id.* at 476 (indicating that this duty is part of CIL, as well as part of a multilateral treaty); 1 International Committee of the Red Cross, Customary International Humanitarian Law 51 (Jean-Marie Henckaerts & Louise Doswald-Beck eds., 2005) [hereinafter ICRC Customary IHL Study]; 2 ICRC Customary IHL Study 337–39 (outlining state practice that fits "constant care" phrase in Additional Protocol I); *see* International Committee of the Red Cross, Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949 677, 712 (Yves Sandoz *et al.* eds., 1987) (acknowledging "constant care" standard).

46. Tallinn Manual 2.0, supra note 15, at 477 (reading language as imposing duty of "situational awareness at all times").

47. Additional Protocol I, supra note 7, art. 57(1) (emphasis added).

48. *Id.* art. 57(2).

general terms with broader applicability.[49]

As one feature of constant care, those planning an attack must take "feasible" measures to minimize the harm to civilians expected from an attack.[50] Traditional precautions include launching an attack on an industrial site at night to minimize civilian presence; providing warnings to the civilian population; and consulting with more senior officers regarding the timing of an attack or the explosive impact of weapons.[51] Here, again, the touchstone is reasonableness. A feasible precaution does not include all *possible* precautions. Given the need for expeditious military action, the presence of budgetary constraints, and limits on technology, requiring employment of all possible precautions would impose an unduly onerous burden on attack planners.[52] Take the example of consulting with more senior commanders about the military advantage of an attack at a particular time, given a particular expected level of civilian harm.[53] That consultation, endorsed by the DoD Law of War Manual, may suggest alternatives that reduce civilian harm, such as the use of lower-impact explosive weapons or a different—but equally effective—mode of attack.[54]

---

49. *See also* Michael N. Schmitt & Michael Schauss, *Uncertainty in the Law of Targeting: Toward a Cognitive Framework*, 10 HARV. NAT'L SEC. J. 148, 179–80 (2019); Jean-Francois Queguiner, *Precautions Under the Law Governing the Conduct of Hostilities*, 88 Int'l. Rev. Red Cross 793, 796 (2006).

50. Additional Protocol I, supra note 7, art. 57(2)(a)(ii); Tallinn Manual 2.0, supra note 15, at 479–80; Corn, *supra* note 17, at 459; Geoffrey Corn & James A. Schoettler Jr., *Targeting and Civilian Risk Mitigation: The Essential Role of Precautionary Measures*, 223 Mil. L. Rev. 785, 837 (2015).

51. Schmitt & Schauss, *supra* note 49, at 186–90; Corn & Schoettler, *supra* note 50, at 837; Corn, *supra* note 17, at 459.

52. Schmitt & Schauss, *supra* note 49, at 187–88.

53. Corn & Schoettler, *supra* note 50, at 837–39; U.S Dep't of Def. Manual, *supra* note 44, § 5.10.3 & n. 327; Margulies, *supra* note 44, at 214–15.

54. Corn & Schoettler, *supra* note 50, at 838–39. A precaution is not "feasible" under IHL if it reduces the military advantage yielded by an attack. Schmitt & Schauss, *supra* note 49 at 186. The focus here is on the feasibility of a different mode of attack on a specific military objective. Language in Article 57(3) of Additional Protocol I also appears to require an attack planner to choose between several legitimate military objectives. Under this subsection, an attack planner must determine whether an attack on one of several possible military objectives—such as different locations of an adversary's forces—1) yields a military advantage that is "similar" to an attack on any of the other possible objectives, and, 2) poses a lower risk to civilians. *See* Additional Protocol I, *supra* note 7, art. 57(3). This is an accurate statement of customary law, if the interpretation of the duty also recognizes that a precise comparison of different military objectives can be difficult in the exigency of combat. *See* Corn & Schoettler, *supra* note 50, at 820–23. Given this difficulty, an attack planner must act reasonably and in good faith under the circumstances. *Id.*

E. THE GAP BETWEEN ATTACK PLANNERS' INDIVIDUAL DUTIES AND SYSTEMIC REDUCTION OF CIVILIAN HARM

The law of armed conflict provides vital safeguards for civilians. However, that protection is not absolute, given LOAC's effort to balance humanity with military necessity. Some harm to civilians is inevitable in an armed conflict. Other harm to civilians is avoidable in two subsets of cases. In some cases, harm will result from violation of LOAC's specific rules and principles, including the principle of distinction and the rules of proportionality and precautions in attack. In addition, civilian harm can result when—on a systemic level—a military force's chain of command has failed to give individual attack planners the tools that a higher-functioning system would provide. That systemic gap may reflect a given state's modest resources, in comparison with other states that are richer and more technologically sophisticated. The law of armed conflict recognizes that such variations in resources will necessarily occur in an imperfect world.[55] However, LOAC and human rights law may have more to say about systemic gaps that stem from institutional dysfunctions in training, procedure, and access to available technology. Unfortunately, recent accounts indicate that these two subsets of cases account for a material number of recent episodes in which U.S. attacks have caused civilian harm.[56] The next subsection analyzes cause of that harm in greater depth.

II. FACTORS DRIVING AVOIDABLE CIVILIAN HARM

A material portion of recent U.S. attacks reveal violations of specific LOAC rules and/or systemic failure to reduce avoidable civilian harm.[57] No evidence, even in critical accounts, suggests that U.S. forces intentionally and deliberately attacked civilians. Rather,

---

55. *See* Michael N. Schmitt, *Precision Attack and International Humanitarian Law*, 87 Int'l Rev. Red Cross 445, 460 (2005) (noting that because of concerns about disparate resources among states and shortages that inevitably occur during wartime, IHL does not require states to use precision-guided munitions).

56. *See* Khan, *Hidden Pentagon Records Reveal Patterns of Failure in Deadly Airstrikes*, *supra* note 4; RAND REPORT, *supra* note 11, at 21–59.

57. *See* Khan, *Hidden Pentagon Records Reveal Patterns of Failure in Deadly Airstrikes*, *supra* note 4; RAND REPORT, *supra* note 11, at 21–42. Assessing the full number of civilian casualties of U.S. air strikes is difficult. Moreover, the *New York Times*' reports (by Khan), while useful and informative, do not purport to assess all U.S. strikes since the United States intervened in Afghanistan after the Sept. 11, 2001 attacks. In compiling a useful, but incomplete, "numerator" of strikes that may be problematic under LOAC, the *Times* series does not present a "denominator" that represents all U.S. strikes, including those that manifestly follow LOAC requirements.

possible U.S. LOAC violations by individual attack planners show a lack of due care in gathering reasonably available information, a failure to question initial theories in light of contrary evidence, and a disregard for set procedures. Systemic failures undermining performance of the United States' "constant care" duty include the lack of methodical programs to improve access to technology, compliance with procedure, and exercise of attack planners' judgment.

Some nations have track records that are demonstrably worse; for instance, Russia's unlawful aggression in Ukraine has resulted in thousands of needless civilian deaths and injuries, and massive harm to civilian objects, including residences and hospitals.[58] This Article focuses on problems with the U.S. approach precisely because the United States has made substantial efforts to reduce harm to civilians.[59] An assessment of the patterns that prompt avoidable and U.S.-caused civilian harm will also be relevant to many other states.

Moreover, Congress has mandated that the DoD report on civilian harm and the annual reports issued by DoD indicate an awareness of areas where there is room for improvement. A recent DoD report indicates progress in deploying technology to reduce civilian harm.[60] However, more remains to be done.

A. How Organizations Fail

A systemic focus on organizations is a helpful prelude to this discussion, because organizations tend to make errors that form recognizable patterns.[61] Participants in organizations, including personnel with extensive professional training, are susceptible to errors, including: undue haste in making high-risk judgments; failure to follow procedures; extreme variations in assessments of identical

---

58. *See* Van Schaack, *supra* note 2.

59. *See* Jennifer M. O'Connor, Gen. Couns., U.S. Dep't of Def., Applying the Law of Targeting to the Modern Battlefield 3-8, Address at N.Y.U. School of Law (Nov. 28, 2016), *published at* https://dod.defense.gov/Portals/1/Documents/pubs/Applying-the-Law-of-Targeting-to-the-Modern-Battlefield.pdf (discussing U.S. approach to targeting, including gathering data and seeking legal advice).

60. *See* U.S. Dep't of Defense Ann. Rep. on Civilian Casualties in Connection with U.S. Mil. Operations in 2020, at 19 (2021), https://int.nyt.com/data/documenttools/annual-report-civilian-casualties-2020/7d258e324d84d499/full.pdf.

61. *See* Madsen, *supra* note 13, at 147; Rangaraj Ramanujam & Paul S. Goodman, *Latent Errors and Adverse Organizational Consequences: A Conceptualization*, 24 J. Organizational Behav. 815, 819 (2003) (analyzing types of errors); *see also* Tara Lamont & Justin Waring, *Safety Lessons: Shifting Paradigms and New Directions for Patient Safety Research*, 20 J. Health Servs. Rsch. & Pol'y 1, 2–3 (2015) (discussing the role of organizational dynamics and context).

data; and a tendency to discount contrary evidence.[62] Organizations as a whole often fail to plan effectively; they underestimate the likelihood of future systemic failures, the resources required to deal with future challenges, and the time involved to complete projects.[63] Many errors are latent—they are hidden in the distracting environment of organizational routine, where they multiply to undermine the organization from within. A pattern of small errors in a large organization can snowball into "highly preventable major failures."[64]

Complicating quality control in large organizations, such as state military forces, subgroups within such entities often fail to predict, monitor, or understand the effects of their decisions on other subgroups.[65] Concepts or processes that work well on the test bench or on paper may not function in the field without significant adjustments. Hierarchies within such entities can further impede the learning process. Moreover, since many errors are latent, accountability for errors is rare.[66] Those personnel who make errors lack incentives to perform more effectively. State militaries are hardly immune from these organizational woes.

B. FACTORS IN FAILING TO REDUCE CIVILIAN HARM

This subsection breaks down existing problems into four areas:

---

62. *See, e.g.,* KAHNEMAN, ET AL., *supra* note 9, at 14–17 (describing studies that showed large variation in federal judges' sentencing of similar defendants); *id.* at 275–79 (describing wide variations and common flaws in elite physicians' reading of X-rays).

63. *See* Madsen, *supra* note 13, at 147 (describing organizations that discount the risk of failure and the need for resources to ensure safety); Bode, et al., *supra* note 18, at 850 (noting that "[i]nsufficient attention to the risk of supply chain disruption is a constant threat" in large business organizations); Sang-Hyun Kim & Brian Tomlin, *Guilt by Association: Strategic Failure Prevention and Recovery Capacity Investments*, 59 MGMT. SCI. 1631, 1632–33 (2013) (suggesting that organization form and structure influence decisions on investing in failure-prevention for industrial systems); Tobin E. Porterfield, John R. Macdonald & Stanley E. Griffis, *An Exploration of the Relational Effects of Supply Chain Disruptions*, 51 TRANSP. J. 399, 404–05, 410–14 (2012) (observing that personal relationships, effort, and performance influenced supply chain issues and recovery from shortages); KAHNEMAN, ET AL., *supra* note 9, at 162 (recounting evidence of "planning fallacy" that leads to underestimating time needed to perform projects).

64. *See* Mark D. Cannon & Amy C. Edmondson, *Confronting Failure: Antecedents and Consequences of Shared Beliefs about Failure in Organizational Work Groups*, 22 J. ORGANIZATIONAL BEHAV. 161, 162–63 (2001).

65. Melissa A. Valentine, *Renegotiating Spheres of Obligation: The Role of Hierarchy in Organizational Learning*, 63 ADMIN. SCI. Q. 570, 573 (2018).

66. *See* Madsen, *supra* note 13, at 145, 147, 152–53.

deficits in technology, cognition, procedure, and training and institutions. I address each in turn.

1. Technology and Logistics

Access to technology can reduce civilian harm by ensuring that attack planners have the information they need to identify proposed targets, determine the expected amount of collateral damage, and tailor the mode and means of attack. Further development of advanced technology, including AI, could provide additional benefits.[67] Improved logistics would enhance access to currently available technology, including high-definition video.

In many civilian-harm incidents involving the U.S. military, imperfect information is a major cause.[68] Targeters lack access to information that would identify a proposed target as a protected person or site, or information that would fully account for civilians whose presence near a target would figure into the proportionality calculus.

In the mistaken air strike, in August 2021, that killed ten people, including seven children, in Kabul, Afghanistan, it appears that poor video quality contributed to the targeting cell's decision.[69] The U.S. and independent journalists reviewing the footage agreed that the video showed children in a courtyard running to the vehicle of Zemari Ahmadi, the U.S. military's target, who was killed in the strike but was later found to have no terrorist ties.[70] Unfortunately, the video was blurry and shot from above, both characteristics that would make it difficult to ascertain individuals' heights or other cues that could have

67. *See* CNA REPORT, *supra* note 8, at 20–40.

68. *See* RAND REPORT, *supra* note 11, at 21; CNA REPORT, *supra* note 8, at 20; *see also* Khan, *Hidden Pentagon Records Reveal Patterns of Failure in Deadly Airstrikes*, *supra* note 4 (discussing inconsistent dissemination of high-definition video, as part of *New York Times* study); Vergun, *supra* note 3 (discussing briefing by U.S. Air Force Inspector General acknowledging that "communication breakdowns" played a role in the failure to recognize the presence of civilians at the scene of the mistaken Kabul air strike in final days of U.S. military presence in Afghanistan); Garrett Executive Summary, *supra* note 3, at 1 (acknowledging that commanding officer of targeting cell that ordered Baghuz strike, "through no fault of his own . . . .relied on data that was not fully accurate.").

69. *See* Charlie Savage, Eric Schmitt, Azmat Khan, Evan Hill & Christoph Koettl, *Newly Declassified Video Shows U.S. Killing of 10 Civilians in Drone Strike*, N.Y. TIMES (Jan. 19, 2022), https://www.nytimes.com/2022/01/19/us/politics/afghanistan-drone-strike-video.html.

70. *See* Vergun, *supra* note 3. In fact, Ahmadi worked for a California-based aid group, Nutrition and Education International. *See* Savage, et al., *supra* note 69.

identified the figures in the video as children.[71] Poor video quality may have also played a role in a 2019 air strike, in Baghuz, Syria, that reportedly killed dozens of civilians, according to the *New York Times*.[72]

The United States military does not have a monopoly on air strikes in which lack of access to technology contributes to tragic outcomes. In the 2014 Protective Edge campaign against Hamas in the Gaza Strip, an Israeli Defense Force (IDF) pilot with no direct visual contact with the target zone conducted a strike that killed four Palestinian boys who had been playing in the vicinity of a waterfront structure that IDF intelligence believed was operated by Hamas police. The pilot lacked a high-definition video feed or other data that would have allowed targeters to determine that the visible figures were children. Making a split-second decision without clear video and informed by an intelligence report that Hamas operatives were about to launch an attack on IDF personnel, the pilot launched this lethal strike. In fact, the boys, who ranged from 9-11 years old, had no role in any Hamas action. Onlookers on the beach or at a widely used dock within 100 yards of the strike, including journalists, had earlier taken note of the children playing. More accurate and complete video or other methods for determining the height of the figures in the video feed could have averted this tragedy.[73]

---

71. Savage, et al., *supra* note 69.

72. Garrett Executive Summary, *supra* note 3, at 1 (acknowledging that U.S. targeting cell "relied on data that was not fully accurate."); Eric Schmitt & Dave Philipps, *Pentagon Faults Review of Deadly Airstrike but Finds No Wrongdoing*, N.Y. TIMES (May 17, 2022), https://www.nytimes.com/2022/05/17/us/politics/us-airstrike-civilian-deaths.html#:~:text=WASHINGTON%20%E2%80%94%20A%20Pentagon%20investigation%20into,reporting%20delays%20and%20information%20gaps. (reporting on U.S. military findings); *see also* Dave Philipps & Eric Schmitt, *How the U.S. Hid an Airstrike That Killed Dozens of Civilians in Syria*, N.Y. TIMES (Nov. 15, 2021), https://www.nytimes.com/2021/11/13/us/us-airstrikes-civilian-deaths.html (reporting that targeting cell used standard-definition video while Central Command operations center in Qatar had real-time access to high-definition video of the same site; high-definition footage showed a few armed men near a large group of women and children). Michael X. Garrett, the U.S. Army general officer tasked with investigating the Baghuz strike, found no violation of LOAC and contested some of the allegations in the *New York Times* report. *See* Garrett Executive Summary, *supra* note 3, at 2 (asserting that the *New York Times* reporting, and the claims of its sources & government whistleblowers, were incorrect).

73. Lubin, *supra* note 7, at 159–61; IDF Military Advocate General (MAG), *Decisions Regarding Exceptional Incidents that Allegedly Occurred During Operation 'Protective Edge' – Update No. 4,* at 7 (June 11, 2015), http://www.law.idf.il/163-7353-en/Patzar.aspx (Isr.); *see also* Chen Maanit & Assoc. Press, *Israeli High Court Rejects Petition to Reopen Probe into Deadly Gaza Strike*, Haaretz (Apr. 24, 2022), https://www.haaretz.com/israel-news/2022-04-24/ty-article/.premium/israeli-high-court-rejects-petition-to-reopen-probe-into-deadly-gaza-strike/00000180-

Lack of access to technology can also exacerbate the harm produced by the unanticipated presence of civilians at a targeting site.[74] Armed conflict is dynamic at the strategic and tactical level. Conditions "on the ground" can be volatile, making plans for a strike look less like textbook exercises and more like drawing stick figures on sand.

Consider the soda-straw problem.[75] The problem arises because attack planners, particularly those who rely on unmanned aerial vehicles (drones) often zero in tightly on a projected target in the seconds before an attack. That tight focus resembles looking at an object in a dynamic landscape through a soda-straw; the focus is narrow, and the viewer fails to see any object outside of the small frame. This issue first rose to prominence during the NATO campaign in Kosovo in 1999, in which a NATO pilot tasked with destroying a bridge used by Serbian forces failed to see an oncoming train carrying civilian passengers. The missiles fired by the aircraft hit the train, resulting in at least 10 killed or missing civilians and at least 15 injured civilians.[76] In other situations, U.S. drone pilots focusing on a moving target, such as a suspected Al Qaeda operative in a ground vehicle, have failed to see oncoming vehicles carrying civilians.[77] Here, too, improved dissemination of technology that is currently available could ease the soda-straw problem.

655b-d824-ad9e-e77f596d0000 (reporting that Israel High Court of Justice had denied a request to resume the IDF investigation of the incident). In a report on investigation of the incident finding no violation of LOAC, IDF personnel recommended technological changes to reduce the risk of future mistakes of this type.

74. *See* RAND Report, *supra* note 11, at 16.

75. *See* PETER W. SINGER, WIRED FOR WAR: THE ROBOTICS REVOLUTION AND CONFLICT IN THE 21ST Century 75 (1st ed. 2009).

76. *See* Aaron Schwabach, *NATO's War in Kosovo and the Final Report to the Prosecutor of the International Criminal Tribunal for the Former Yugoslavia*, 9 TUL. J. INT'L & COMP. L. 167, 178 (2001) (discussing NATO bombing of Grdelica Gorge bridge); *see also* Int'l Crim. Trib. for the Former Yugoslavia, Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia, ¶ 59, (2000) (citing explanation of incident by NATO Supreme Allied Commander General Wesley Clark that pilot, whose aircraft was miles away from the intended target, saw an oncoming train in the seconds before impact of the missile; the train appeared as a small image on his 5-inch navigation screen; the pilot fired a second missile seconds later, when the pilot assumed that the train had stopped short of the bridge, although the train had in fact proceeded onto the bridge); *cf. id.* at ¶ 62 (finding that incident did not merit investigation as a war crime in light of short period between appearance of train and firing of both weapons and limited indication of train's presence provided by instruments available to the pilot and crew of the NATO aircraft).

77. CNA Report, *supra* note 8, at 15 (finding that "unanticipated presence" of civilians contributed to civilian harm in U.S. air strikes).

In another setting where the U.S. military has already recommended corrective action, lack of access to technology has hampered prompt identification of protected sites. In the 2015 Kunduz attack by U.S. aircraft on a Medicins Sans Frontières (MSF or Doctors Without Borders) facility, one problem in a parade of errors arose because the U.S. aircraft took off hurriedly in a fire-fight without loading an electronic No-Strike List (NSL) that included the GPS coordinates of all protected sites in the area, including the MSF facility.[78] Operating at night during a battle, United States forces mistakenly identified the MSF building as another building in the area, which housed Taliban fighters. The U.S. aircraft ultimately fired missiles at the MSF facility, killing over 40 people, including medical personnel and patients. Access to an electronic NSL would have warned the U.S. air crew that their identification was mistaken, thus sparing the MSF facility and the people inside. In the aftermath of this horrific episode, investigators recommended pre-loading electronic NSLs in aircraft to avoid omissions that occurred in the heat of battle.[79] A more comprehensive, deliberative approach to reducing civilian harm might have recommended this straightforward safeguard in advance of the Kunduz strike, thereby averting the tragedy.

At any given time, a large entity like a state military force will exhibit some internal variation in technical capabilities. However, as of August 2022, there was no publicly available evidence that the U.S. military has developed a workable plan for disseminating high-definition video to all targeting cells and drone assets. High-definition video and other technology that could reduce civilian harm, such as GPS-linked site maps, are commercially available features present in vast numbers of homes, workplaces, and vehicles. Nevertheless, the U.S. Defense Department, more than twenty years after the start of armed conflict with Al Qaeda, had not formulated, issued, and

78. Kristina Daugirdas & Julian Davis Mortenson, *Contemporary Practice of the United States Relating to International Law*, 110 AM. J. INT'L L. 554, 581–82 (2016); U.S. FORCES-AFG., INVESTIGATION REPORT ON THE AIRSTRIKE ON THE MEDICINS SANS FRONTIERES/DOCTORS WITHOUT BORDERS TRAUMA CENTER IN KUNDUZ, AFGHANISTAN ON 3 OCTOBER 2015, at 52 (2016) [hereinafter U.S. FORCES, KUNDUZ AIRSTRIKE REPORT); Rachel E. VanLandingham & Geoffrey S. Corn, *Why No Courts-Martial over Kunduz?*, USA TODAY (May 8, 2016, 9:02 AM), https://www.usatoday.com/story/opinion/2016/05/08/kunduz-doctors-without-borders-msf-court-martial-afghanistan-bomb-strike-column/84040012; *cf.* Alcala, *supra* note 44 (documenting U.S. forces' failure to safeguard cultural and archaeological sites in Iraq during Second Gulf War).

79. U.S. CENT. COMMAND, RELEASE NO. 20160429-10, SUMMARY OF THE AIRSTRIKE ON THE MSF TRAUMA CENTER IN KUNDUZ, AFGHANISTAN ON OCTOBER 3, 2015: INVESTIGATION AND FOLLOW-ON ACTIONS (Nov. 21, 2015) [hereinafter U.S. CENT. COMMAND, MSF AIRSTRIKE INVESTIGATION AND FOLLOW-ON ACTIONS).

implemented a plan for using this common technology to reduce civilian harm.[80]

Indeed, evidence suggests that the U.S. military by the mid teens had backslid from compliance efforts earlier in the post-9/11 era. While the United States' drone strategy has always been controversial,[81] earlier targeting efforts revealed substantially greater care. For example, drone pilots tried to limit the soda-straw effect by taking a wide video sweep of the target area.[82] The U.S. military has pivoted from these practices without an explanation.

2. Cognition and Deliberation

Just as links between civilian harm and inconsistent access to technology are longstanding, time has cemented harm's link to cognitive errors. State military forces are subject to flaws that mar overall human judgment and inference.[83] These flaws affect targeting, as the U.S. experience shows in the Kabul, Kunduz MSF, and Syria strikes. Two related problems of human inference that plague targeting are confirmation bias[84] and base-rate neglect.[85] Here, as with access to technology, the problem is not isolated; it is systemic.

a. Confirmation Bias

Confirmation bias entails viewing all new information as reinforcing a thesis the targeter has already formed.[86] Ideally, a

80. *See* RAND Report, *supra* note 11, at 20–21 (observing that U.S. military has devoted substantial resources to promoting interdisciplinary situational awareness of adversaries' behavior in post-9/11 period, but that this focus has caused "information gaps" in understanding and addressing behavior of civilians swept up in armed conflict).

81. *See generally* Craig Martin, *The Means-Methods Paradox and the Legality of Drone Strikes in Armed Conflict*, 19 Int'l. J. HUM. RTS. 3 (2015).

82. *See* CNA Report, *supra* note 8, at 20.

83. Nobel Prize winner Daniel Kahneman and his colleagues and co-authors have outlined these errors most accessibly for a general audience. *See* DANIEL KAHNEMAN, THINKING FAST AND SLOW 80–88 (2011); *see also* KAHNEMAN, ET AL., NOISE, *supra* note 9, at 170–74.

84. KAHNEMAN, ET AL., NOISE, *supra* note 9, at 172 (noting that confirmation bias is a "tendency that leads us, when we have a prejudgment . . . to disregard conflicting evidence").

85. *Id.* at 166–67.

86. Simone Galperti, *Persuasion: The Art of Changing Worldviews*, 109 Am. Econ. Rev. 996, 1016 (2019) (citation omitted) (explaining that in social science studies, "[e]xperimental subjects show 'a clear tendency to resist evidence' inconsistent with their hypotheses").

rational decisionmaker, including a member of a targeting cell, will be open to interpreting new information in one of three ways: as bolstering the decision-maker's initial view; as neutral—neither proving nor disproving that first thesis; or as rebutting the view previously held. Unfortunately, most individual decision-makers, including experts, fall short of this ideal and instead attach disproportionate weight to initial interpretations.[87] In an armed conflict, where targeting necessarily involves acting on incomplete information, drawing reasonable inferences from evidence is essential.[88] Equally important is acknowledging that the evidence at hand is inconclusive, requiring the attack planner to seek further evidence to confirm a hypothesis.[89] A judgment that evidence is inconclusive may prompt the attack planner to consult with a more senior officer, who may suggest alternative scenarios or ways of conducting an attack.[90] Confirmation bias, sometimes called "overconfidence," hinders inference, acknowledgment, and consultation.[91] It, therefore, compromises the accuracy and reliability of targeting decisions.

The U.S. military has acknowledged for decades that confirmation bias adversely affects targeting. Illustrating this awareness, U.S. Department of Defense investigators used a related

---

87. *Id.* at 1016 (reporting that, "once people have adopted a hypothesis as the single explanation of something, they often preclude the possibility of interpreting data as supporting any other explanations"); Michael A. Bruno et al., *Understanding and Confronting Our Mistakes: The Epidemiology of Error in Radiology and Strategies for Error Reduction*, 35 RadioGraphics 1668, 1671–72 (2015) (analyzing studies showing that trained physicians specializing in radiology regularly commit X-ray reading error called "satisfaction of search," in which the doctor overlooks a key abnormality "because of a failure to continue to search" after spotting an initial anomaly that is less serious or reflects a different medical cause; the analyst tacitly and incorrectly infers that any further abnormalities will merely buttress preliminary diagnosis instead of materially altering it).

88. Corn & Schoettler, *supra* note 50, at 837–39.

89. Corn, *Process as a Precautionary Measure*, *supra* note 17, at 459.

90. DoD Manual, *supra* note 44, at ¶ 5.10.3 & n.327.

91. *See* Kathrin E. Maki, Matthew K. Burns & Amanda L. Sullivan, *School Psychologists' Confidence in Learning Disability Identification Decisions*, 41 Learning Disability Q. 243, 251 (2018) (noting that for professionals, the ability to "self-monitor" one's own practices and make appropriate adjustments is vital, and that overconfidence impedes this key attribute). Economic and political analysts commit similar errors and fail to develop means to address those mistakes. *See* Philip E. Tetlock, Expert Political Judgment 65 (2005) (noting that in political forecasting, experts with extensive training and education performed only marginally better than laws of chance); Simon Gervais et al., *Overconfidence, Compensation Contracts, and Capital Budgeting*, 66 J. Fin. 1735, 1743–46 (2011) (discussing overconfidence in setting compensation); *see also* Kahneman, et al., *supra* note 9, at 140–42 (discussing overconfidence in forecasting).

term—"scenario fulfillment"—to describe the 1988 strike in the Persian Gulf by the *U.S.S. Vincennes* on an aircraft that turned out to be an Iranian passenger jet.[92] Officers on the *Vincennes* came to believe that the Iranian aircraft was hostile because the jet's trajectory, plotted on a horizontal axis, appeared to be bringing it closer to the U.S. vessel, although this reading was an incomplete assessment of the data available to the *Vincennes*'s officers.[93] The Iranian aircraft failed to respond to repeated warnings, although until moments before launching its missiles the *Vincennes* had issued on frequencies used solely by military aircraft, which a civilian aircrew would be unlikely to monitor. In addition, the threat scenario that the *Vincennes'* officers constructed also reflected a 1987 Persian Gulf episode in which an Iraqi military aircraft had attacked a U.S. naval vessel, the *U.S.S. Stark*.[94]

With this narrative of a hostile aircraft front and center, officers on board the *Vincennes* failed to properly interpret another data point the shipboard instruments revealed: while the Iranian aircraft was drawing closer to the U.S. vessel along a horizontal axis, on a vertical plane it was moving further away, rapidly gaining altitude. This account of forward horizontal movement and a steep vertical climb precisely fits the profile of a passenger jet on take-off.[95] Unfortunately, influenced by confirmation bias, the officers on the *Vincennes* discounted this data point or failed to process it because it did not fit their preconceived narrative. That undue cognitive discount resulted in a strike that destroyed the passenger jet and killed 290 people on board.[96]

---

92. U.S. Department of Defense, *Formal Investigation in the Circumstances Surrounding the Downing of Iran Air Flight 655 on 3 July 1988*, §IV(A)(12) (1988), https://en.wikisource.org/wiki/Formal_Investigation_into_the_Circumstances_Surrounding_the_Downing_of_Iran_Air_Flight_655_on_3_July_1988/Formal_Report [hereinafter DoD *Vincennes* Report]; Martins, *supra* note 9, at 43-45 (analyzing *Vincennes* episode); Michael N Schmitt & Jeffrey S. Thurnher, *"Out of the Loop": Autonomous Weapons Systems and the Law of Armed Conflict*, 4 Harv. Nat'l Sec. J. 231, 248–49 (2013) (citing *Vincennes* incident as illustrating human error in targeting).

93. DoD *Vincennes* Report, *supra* note 92, at § II(C)(1); Andreas F. Lowenfeld, *The Downing of Iran Air Flight 655: Looking Back and Looking Ahead*, 83 Am. J. Int'l L. 318, 336–37 (1989); Margulies, *supra* note 7; Martins, *supra* note 9, at 43–45.

94. Corn, *supra* note 10, at 14 n.53.

95. Lowenfeld, *supra* note 93, at 336.

96. After Iran initiated a lawsuit in the International Court of Justice, the United States agreed to pay Iran compensation for the tragedy. However, U.S. officials asserted that the payments were *ex gratia*, reflecting concern for the victims and their families, not binding legal obligation. Charles P. Trumbull IV, *Autonomous Weapons: How Existing Law Can Regulate Future Weapons*, 34 Emory Int'l L.J. 533, 592 n.366 (2020).

b. Base-Rate Neglect

On occasion, targeting accuracy can suffer from confirmation bias and base-rate neglect. The latter flaw occurs in everyday thinking when an individual makes probability judgments about an event or condition based on a distinctive, memorable fact without considering the overall probability of that event or condition's occurrence. Because of base-rate neglect—sometimes called probability neglect—some events or conditions seem far more likely than a rational cognitive process would find them to be.[97]

Consider a question linking personality traits with employment. Targeting in a NIAC—where potential targets rarely wear uniforms—involves analysis of personal behavior. Therefore, this question has some relevance to the kinds of decisions that targeting cells make. In our hypo, we ponder the case of Steve, whom we describe as shy and fastidious.[98] What are the respective probabilities that Kurt is either: 1) a librarian or 2) an employee of Walmart, Amazon, or FedEx? In controlled experiments with similar hypotheticals, most respondents peg Steve as a librarian because Steve's personality seems to echo traits we associate with that group.[99] However, librarians are a relatively small occupational group, while Walmart, Amazon, and FedEx are three of the five largest employers in the United States, with a total of approximately 5 million employees. Assuming that a numerically significant group of U.S. persons—although not necessarily a majority—are both shy and fastidious, the sheer number of workers for the huge firms mentioned makes Steve's employment at one of those companies far more likely than his service as a librarian. The higher base rate of employees of major corporations compared with the much lower percentage of librarians in the general population has a substantial effect on probability. However, due to the cognitive flaw of base-rate neglect, most people ignore this effect.

The combination of base-rate neglect and confirmation bias contributed to the mistaken 2015 U.S. air strike on the MSF facility in Kunduz, Afghanistan. Targeting personnel relied excessively on two structural features shared by the MSF facility and the Taliban stronghold that the U.S. team had sought to locate and attack: an arch and an outer perimeter wall.[100] In identifying the MSF facility as their

97. *See* Kahneman, *supra* note 83, at 88, 146–51; Kahneman, et al., *supra* note 9, at 166–67.

98. *See* Kahneman, *supra* note 83, at 88.

99. *Id.*

100. U.S. Air Forces, *Kunduz Airstrike Report*, *supra* note 78, at 034 (noting that Taliban headquarters that targeting team sought to identify had an "arched-shaped

intended target based on these shared structural features, the targeting team failed to acknowledge that arches and perimeter walls are common. The high incidence of these features suggested an alternative scenario that the targeting team did not consider sufficiently: the MSF facility that they were observing was not their intended target—the Taliban stronghold—but was instead a protected structure housing civilians.[101]

Confirmation bias exacerbated the effects of base-rate neglect in the MSF incident. In assessing the probative value of structural features such as an arch that the MSF facility shared with the team's intended target, the team should have regarded that evidence as neutral or mildly probative, at best. However, under the sway of confirmation bias, the aircrew viewed this evidence as compelling. Without the cognitive space for an alternative scenario, the crew settled on the MSF facility as its target, with tragic consequences.

More recently, the combination of base-rate neglect and confirmation bias influenced the August 29, 2021, mistaken strike in Kabul. The U.S. targeting cell, in the wake of a bombing at the Kabul airport, received intelligence that ISIS elements were planning another bombing. Intelligence, which may have been faulty, directed U.S. personnel to an individual who turned out to work for a U.S. aid agency.[102] Tracking this individual through the course of a working day, the U.S. targeting cell perceived similarities between the actions

---

gate" and also had an "outer perimeter wall, with multiple buildings inside of it"); Daugirdas & Mortenson, *supra* note 78, at 581–82 (discussing shared presence of outer perimeter wall in both the sought-after target and the MSF facility, respectively); Margulies, *supra* note 7 (discussing presence of an arch).

101. The MSF facility's coordinates were included in the electronic NSL that the U.S. military had compiled for the area, but that source was not available to the aircrew, although it was available to the ground-force commander coordinating targeting. *See* U.S. Air Forces, *Kunduz Airstrike Report*, *supra* note 78, at 081 (noting that the ground-force commander (GFC) "did not know, but should have been aware of the MSF Trauma Center's location"; staff under GFC's direction "were provided with the location of the MSF Trauma Center prior to the GFC's decision to engage"). The facility had MSF flags prominently displayed on the roof, the building's front and sides, and the courtyard. *Id.* at 082. The facility did not display internationally recognized insignia such as a red cross, red crescent, or red letter, "H." *Id.* Display of one of those insignia would have reduced the risk of attack. *Id.* According to MSF, the government of Afghanistan, whose approval would have been required for display of these insignia, informed MSF that exhibiting the group's own logos was sufficient. *See* Francoise Bouchet-Saunier & Jonathan Whitfall, *An environment conducive to mistakes? Lessons learnt from the attack on the Médecins Sans Frontières hospital in Kunduz, Afghanistan*, 100 Int'l Rev. Red Cross 337, 362–64 (2018).

102. Pentagon Press Secretary John F. Kirby and Air Force Lt. Gen. Sami D. Said Hold a Press Briefing (Nov. 3, 2021), https://www.defense.gov/News/Transcripts/Transcript/Article/2832634/pentagon-press-secretary-john-f-kirby-and-air-force-lt-gen-sami-d-said-hold-a-p/.

of the earlier bomber and the proposed target. For example, the bomber had participated in the transfer of a computer bag in the course of his preparations.[103] The mistaken target on August 29, who actually worked for a U.S. aid group, had also received and handed off bags as he drove through the city. Focusing on this detail was a classic example of base-rate neglect. As the U.S. Air Force inspector general conceded after an investigation of the strike, in a large city such as Kabul, thousands of people every day pass bags back and forth for myriad reasons.[104] Most of these transfers are thoroughly mundane. The transfer of bags by the mistaken target on August 29 did not confirm that the target was a would-be bomber; instead, the evidence was neutral. However, as the inspector general explained, confirmation bias framed the bag transfer as bolstering the targeting team's prejudgment that the proposed target was another terrorist bomber.[105] As in the MSF attack several years earlier, the combination of base-rate neglect and confirmation bias paved the way for a tragedy.[106]

3. Procedural Deficits

Flaws in procedure are also a major cause of civilian harm. In some cases involving the U.S. military, these flaws involve failure to comply with established procedural rules, including rules limiting

103. *See* Azmat Khan, *U.S. Military Inquiry Reveals Errors in Botched Kabul Drone Strike*, N.Y. Times, Jan. 7, 2023.

104. *Id.*

105. *Id.*

106. It is also fair to ask whether ethnic and religious bias played a role in these targeting decisions and in overall U.S. policy since 9/11. Stereotypes influence overall perceptions. The portrayal of Muslims as devious, fanatical, and intolerant has been a fixture of Western discourse since the dawn of the Enlightenment. *See* Sahar Aziz, The Racial Muslim: When Racism Quashes Religious Freedom 91 (2022); Shirin Sinnar, *Separate and Unequal: The Law of "Domestic" and "International" Terrorism*, 117 Mich. L. Rev. 1333, 1338 (2019) (describing inequality in law enforcement); Maryam Jamshidi, *Bringing Abolition to National Security*, Just Security (Aug. 27, 2020), https://www.justsecurity.org/72160/bringing-abolition-to-national-security/ (outlining argument for abolition of counterterrorism laws that yield disproportionate targeting of Muslim-Americans, Arabs, and South Asians); *see also* Kahneman, et al., *supra* note 9, at 303 (noting in employment context that, "[i]nterviews are a minefield of psychological biases . . . interviewers tend, often unintentionally, to favor candidates who are culturally similar to them or with whom they have something in common, including gender, race, and educational background"). Perhaps a simulated study of targeters' attitudes toward white, Judeo-Christian Europeans and Americans would have eliminated ethnic and religious bias as a factor. In the absence of such a definitive study, the question is both legitimate and vexing.

resort to a unit self-defense rationale.[107] In other cases, flaws may entail omissions of simple tasks, including failing to 1) coordinate with other units in the U.S. military forces in targeting and subsequent investigations, and 2) consult with nongovernmental organizations (NGOs) regarding past civilian harm and strategies for harm reduction. Patterns of noncompliance and omissions cause harm that military forces can feasibly avoid.

a. Excessive Reliance on Unit Self-Defense Claims

One salient procedural flaw is unduly broad use of the unit self-defense standard.[108] The self-defense standard is significant because restrictive Rules of Engagement (ROE) often limit permissible collateral damage to very low rates (such as one civilian in any given attack, with departures hinging on approval by senior officers). Moreover, the deliberative targeting process involves safeguards, such as consultation with military lawyers and other officials.[109] Due to its exigent nature, a claim of unit self-defense will preempt more restrictive ROE and the safeguards of the deliberative targeting process. A unit, when at risk for imminent attack, can respond immediately with all necessary and proportionate lethal force, even if that force exceeds ROE.[110] Unit self-defense is a necessary complement to restrictive ROE, preventing catastrophic attacks and giving a unit authority to protect itself in exigent situations. But that complement comes with a cost.

Unit self-defense is also a potential loophole, authorizing far more extensive lethal force with fewer safeguards than ROE or the

---

107. Khan, *supra* note 4.

108. The unit self-defense standard permits an armed force to respond to imminent threats to life. *See* Corn, *supra* note 10, at 9–12.

109. Lubin, *supra* note 7, at 145 (noting that sophisticated military forces distinguish between planned strikes, which entail an array of safeguards, and unplanned or "dynamic" targeting, which may stem from exigent situations in which application of full array of safeguards is impracticable); O'Connor, *supra* note 59, at 3–8 (discussing, in talk by then-General Counsel of U.S. Department of Defense, the targeting process used by U.S. forces in planned strikes, including discussing of weaponeering, timing, and other factors to reduce collateral damage); Gregory S. McNeal, *Targeted Killing and Accountability*, 102 Geo. L.J. 681, 701–20 (2014) (discussing multi-step process used prior to drone strikes targeting specific individuals participating in hostilities on behalf of Al Qaeda or affiliated groups); VanLandingham, *supra* note 39 (same); Michael N. Schmitt & John J. Merriam, *The Tyranny of Context: Israeli Targeting Practices in Legal Perspective*, 37 U. Pa. J. Int'l L. 53, 70–99 (2015) (discussing steps and criteria in targeting by the Israeli Defense Force (IDF)).

110. Corn, *supra* note 10.

deliberative targeting process would permit. Because of the lowered guard rails for unit self-defense, U.S. attack planners have sometimes invoked unit self-defense improperly, as a basis for *offensive* military operations that lack a self-defense justification.[111] This end-run around restrictive ROE and the deliberative targeting process can imperil compliance with the IHL principle of distinction and rules of proportionality and precautions in attack. Multiple reliable sources suggest that in the U.S. armed conflict with ISIS, U.S. high-volume targeting units in Syria regularly interpreted unit self-defense too broadly.[112]

b. Deficits in Internal and External Collaboration

Both the targeting process and subsequent investigations exhibit a failure to collaborate. That flaw in procedure is important, because collaboration can ease the cognitive flaws discussed in the previous subsection. Both internal and—where feasible—external collaboration have benefits that the U.S. military has sometimes failed to recognize.

In targeting, U.S. planners have not always tasked intelligence personnel with gathering current information on civilian movements in the battlespace. Such areas are "active and dynamic."[113] Adversaries in the battlespace can move; civilians can, too. However, U.S. attack planners have too often handled information about civilians as a static backdrop for operations, instead of treating it as an evolving series of data points subject to rapid change.[114] In the heat of battle, getting fully up-to-date reports of civilian movements can be challenging. However, it appears that sometimes U.S. military protocols fail to encourage reasonable efforts.[115]

The August 2021 mistaken Kabul strike provides a related

111. *Id.* at 10–11 (observing that U.S. attack planners in post-9/11 armed conflicts have sometimes used unit self-defense pretextually to "conduct hasty, tactical targeting" in offensive operations); *id.* at 11 (explaining that attack planners on occasion have "improperly leveraged [the unit self-defense justification] . . . to draw insurgents out and thereby trigger self-defense authorities," in a hack of restrictive ROE that is referred to as "baited self-defense").

112. *See* Khan, *supra* note 4; Schmitt & Philipps, *Pentagon Faults Review of Deadly Airstrike*, *supra* note 72; Philipps, Schmitt & Mazzetti, *supra* note 4; *see also* CNA Report, *supra* note 8, at 11 (noting role of self-defense claims in sidelining restrictions on collateral damage); Corn, *supra* note 10, at 11 (suggesting that invocation of unit self-defense rationale has "expanded beyond legally permissible limits").

113. *See* RAND Report, *supra* note 11, at 21.

114. *Id.* at 21–22.

115. *Id.*

example of a targeting cell failing to communicate with intelligence personnel. The inspector general's report on the Kabul strike indicated that the targeting cell did not communicate with intelligence and other support personnel regarding the timing of the proposed strike.[116] In particular, the targeting cell failed to inform intelligence personnel of the very short time left before the planned attack. That lack of communication about timing dulled the urgency that intelligence personnel felt about sharing information about civilians at the target site. Timely sharing of information about civilians might have altered the proportionality calculus and led to consultation with commanders and intelligence personnel that would have identified the putative target as a worker for a U.S. aid agency. Failing to share information about the timing of targeting with intelligence personnel thus helped pave the way for the mistaken attack and the human toll that resulted.

The U.S. military has also on occasion failed to consult internal sources in its investigation of allegations of civilian harm. Under IHL customary and treaty law, investigations of claims of civilian harm must be independent and effective.[117] To be effective, an investigation must consult reasonably available sources of data, including data that the military itself has developed or compiled before, during, and after the strike at issue. Indeed, U.S. military procedural rules impose this requirement on investigations.[118] Unfortunately, personnel conducting U.S. military investigations have sometimes failed to consult internal sources. In some cases, investigative personnel have determined that reports of civilian harm lacked credibility without consulting the military's *own data*, including video and imaging sources that show the location of civilians and the effects of a strike.[119] An investigation that reaches conclusions prematurely without examination of the military's own data cannot be effective under IHL.

In addition, in at least one prominent case, the U.S. military failed to follow its own procedures on timely investigation of allegations of civilian harm resulting from attacks. An investigation that suffers from undue delays cannot benefit as much as it should from either internal or external outputs. Delays reflect unwillingness and inability to interact with both internal and external players. Consider the 2019 U.S. strike on an ISIS position in Baghuz, Syria which may have resulted in the deaths of scores of civilians, including a substantial

116. Vergun, *supra* note 3.

117. Michael N. Schmitt, *Investigating Violations of International Law in Armed Conflict*, 2 Harv. Nat'l Sec. J. 31, 80 (2011).

118. RAND Report, *supra* note 11, at 21–22.

119. *Id.* at 23–24.

number of women and children.[120] Disputes continue regarding the lawfulness of the strike and the precise harm, if any, that the strike caused to individuals that IHL would classify as protected civilians.[121] Given concerns about the lawfulness of the strike and the extent of civilian harm, it would have been reasonable to expect that military personnel responsible for initiating and completing an investigation would have carefully followed the procedures of the Department of Defense and made every effort to address those pressing questions in a reasonably expeditious manner, given the complexity of the task. However, in May 2022, more than three years after the incident, General Michael X. Garrett, who conducted a review of the matter, found that "[n]umerous policy compliance deficiencies at multiple levels of command led directly to numerous delays in reporting" the Baghuz incident.[122]

As General Garrett concluded, such failures to follow the military's own procedures erode the confidence of outside observers in the military's ability to conduct an independent and effective

120. See Schmitt & Philipps, *Pentagon Faults Review of Deadly Airstrike But Finds No Wrongdoing*, *supra* note 72.

121. A review by U.S. Army general Michael X. Garrett concluded, based—as IHL requires—on the attack planners' knowledge at the time of the attack, that the attack was proportionate. *Garrett Executive Summary*, *supra* note 3, at 1. However, General Garrett acknowledged that the targeting cell relied on information that was not "fully accurate." *Id.* The lack of such information would not affect the lawfulness of the attack if more fully accurate information was not "reasonably available" to the targeting team and the team had no reasonable basis to believe at the time that the information it relied on was inconclusive in material ways. *See* Tallinn Manual 2.0, *supra* note 15 at 424, Rule 95, ¶ 3 (citing U.K. law of war manual); *see also id.* at 432, Rule 97, ¶ 13 (citing consensus of experts who contributed to Tallinn Manual). However, General Garrett did not use the reasonableness standard recommended by the Tallinn Manual; instead, he used a lower good-faith standard, in which an investigator focuses on what information was "known at the time" by the attack planner. *See Garrett Executive Summary*, *supra* note 3, at 1. Under that standard, an attack planner would behave lawfully in relying on information in her possession, even if other information prompting materially different inferences was reasonably available. This gap between the standard used by General Garrett and the higher standard preferred by the United Kingdom and a group of distinguished international experts raises questions about the Baghuz attack's lawfulness. *Cf.* Ryan Goodman, *CentCom's Full Statement on Baghuz Strike: Annotated*, Just Security (Nov. 22, 2021), https://www.justsecurity.org/79304/centcoms-full-statement-on-baghuz-strike-annotated/ (discussing apparent anomalies between IHL rules and earlier statements by U.S. military on the Baghuz strike). In addition, General Garrett's executive summary did not conclusively resolve questions, independent of the lawfulness of the strike, about the number of civilians—if any—harmed in attack. General Garrett noted that the U.S. military's inquiry into the extent of civilian harm caused by the attack "currently remains open." *Garrett Executive Summary*, *supra* note 3, at 2. The military's decision not to close the investigation suggests that investigators have not yet definitely resolved the issue of civilian harm.

122. *See Garrett Executive Summary*, *supra* note 3, at 2.

investigation. General Garrett concluded that these procedural flaws encouraged outside observers to suspect that the military was "not treating this . . . incident seriously."[123] According to General Garrett, procedural violations also fostered the sentiment that the military was not being "transparent" about civilian harm.[124] The harm to the U.S. military's standing and strategic goals is self-evident.

While the Baghuz incident was the most high-profile example of U.S. military personnel's failure to follow established procedures on investigations, independent sources have found a substantial number of other instances. The New York *Times* study of civilian harm found that personnel tasked with deciding whether reports of civilian harm were credible in connection with specific strikes sometimes failed to adequately consider all of the available evidence.[125] In a related finding, researchers contributed to a RAND Corporation report found that investigators actually resorted to a standard higher than the preponderance of the evidence standard that the military asserts it uses to determine the credibility of civilian-harm reports.[126] Clear guidance on this standard is crucial, since applying the standard determines whether a full investigation will commence on an attack's lawfulness and the precise extent of civilian harm.[127] The RAND researchers concluded that, in practice, investigators applied a higher standard, such as clear and convincing evidence, for commencing an investigation.[128] Yet, at this stage, military investigators relied on incomplete evidence, such as aerial video footage, which cannot show the extent of civilian harm within structures that an attack damaged or destroyed.[129] This rigid standard deterred investigations even when a reasonable assessment would have found a basis for further inquiry under the U.S. military's own procedures.[130]

---

123. *Id.*

124. *Id.* A more recent strike in Syria has also raised questions that the Department of Defense is investigating as of May 2023. *See* Omar Nezhat, Meg Kelly, Alex Horton & Imogen Piper, *U.S. officials walk back claim drone strike killed senior al-Qaeda leader*, Wash. Post, May 18, 2023 (reporting on questions that have arisen on whether Lotfi Hassan Misto, the target of a May 3, 2023 drone strike in a region of northwest Syria controlled by an organizational rival of Al Qaeda, was an Al Qaeda leader or, contrary to initial U.S. claims, was merely a farmer); Brianna Rosen, *Death by Drones: Does the Pentagon Always Know Who it is Killing?*, Just Security (May 23, 2023), https://www.justsecurity.org/86678/death-by-drones-does-the-pentagon-always-know-who-it-is-killing/.

125. *See* Khan, *supra* note 4.

126. *See* RAND Report, *supra* note 11, at 22–24.

127. *Id.*

128. *Id.*

129. *Id.*

130. *Id.*

In conducting investigations, the U.S. military has also failed on occasion failed to consult external sources with useful information, including the media and NGOs. The record here is mixed. On the one hand, a majority of episodes of admitted civilian harm flow from initial reports by outside groups.[131] On the other had, the U.S. military has sometimes discouraged input from external sources.[132] Here, again, the inadequate U.S. response to civilian harm allegations regarding the 2019 Baghuz, Syria attack is a case in point. In his report in May, 2022, General Garrett singled out the New York *Times* for criticism, stating that there "no evidence" to support the allegations in the three extensively documented *Times* articles on Baghuz.[133] General Garrett also stated that whistleblower reports that comprise part of the *Times* pieces' sourcing were "unsubstantiated" and "based on inaccurate or false information."[134] On the merits, General Garrett's conclusory verdict lacked any explanation of why the *Times* pieces were unreliable. That lack of explanation was particularly glaring because the *Times* based its reporting on respective complaints by a whistleblower who viewed a high-resolution video feed of the Baghuz attack in real-time, another whistleblower who investigated the incident on behalf of the Pentagon's inspector general, and an Air Force military lawyer.[135]

While disputes continue about the lawfulness and civilian toll of the attack and a carefully documented account of the strike by the U.S. military would be welcome,[136] General Garrett's unsupported disparagement of well-sourced journalistic accounts was counterproductive. Given General Garrett's acknowledgment of the "numerous policy compliance deficiencies ... [and] delays" that afflicted the Baghuz investigation,[137] it seems reasonable to infer that renewed commitment to investigation of the incident stemmed from pressure by the media and NGOs, fueled by whistleblower complaints. Since General Garrett acknowledged that the Baghuz targeting team acted without "fully accurate" information and that the U.S. military had not yet concluded its inquiry into civilian harm caused by the attack, a less adversarial approach to media reports would have

131. *Id.* at 32.

132. *Id.* at 22–24.

133. *Garrett Executive Summary*, *supra* note 3, at 2.

134. *Id.*

135. *See* Philipps & Schmitt, *How the U.S. Hid an Airstrike That Killed Dozens of Civilians in Syria*, *supra* note 72.

136. *See* Goodman, *supra* note 121.

137. *Garrett Executive Summary*, *supra* note 3, at 2.

recognized the media's helpful role.[138]

The media's role seemed even more salient in light of General Garrett's recommendation of improvements in the U.S. military's approach to civilian harm, such as better training and development of "common understanding and practice" across combat commands; "increased situational awareness" at all levels on the risk of civilian harm; and more "refined processes" at the Mission Command level.[139] General Garrett's recommendations, apparently discussed in detail in his full report, were a helpful complement to the Department of Defense Directive on civilian harm, which was still awaiting public issuance in early August, 2022. A complete picture of the contributions of the media and NGOs would affirm the RAND Corporation's observation that these players provide "important perspectives that may challenge conventional wisdom and reduce the risk of groupthink."[140] Unfortunately, General Garrett's summary failed to acknowledge these contributions, instead adopting a needlessly adversarial tone. A pivot to broader engagement with civil society is urgently needed.

The comprehensiveness of investigations—at least in public reports—has also experienced backsliding in recent years. The 2016 report on the mistaken Kunduz MSF facility strike[141] was exceptionally detailed and thorough about the chain of errors that led to that tragic incident.[142] In contrast, recent public reports about the Baghuz and Kabul strikes have provided much less information. General Garrett's executive summary of his Baghuz strike review was conclusory and often adversarial. It provided no in-depth analysis of the time-line leading up to the attack or the targeting cell's deliberations. Perhaps the Defense Department will ultimately release a suitably redacted version of the complete Garrett report; in the meantime, both U.S. military personnel and outside players lack sufficient information or analysis to learn the right lessons. In contrast with General Garrett's conclusory summary regarding Baghuz, U.S. Air Force Inspector General Sami D. Said's briefing on the mistaken Kabul strike provided more detail on that attack and acknowledged that "confirmation bias" appeared to play a role in the targeting cell's

138. *See id.* at 1 (discussing information "known at the time" by the targeting team); *id.* at 2 (noting that civilian-harm investigation remained open at the time of submission of Garrett's report).

139. *Id.* at 2.

140. RAND Report, *supra* note 11, at 32.

141. *See* Daugirdas & Mortenson, *supra* note 78.

142. *See, e.g.*, U.S. Air Forces, *Kunduz Airstrike Report*, *supra* note 78, at 052 ¶¶ 48–50 (addressing lack of electronic NSLs); *id.* at 082 (noting ground-force commander's access to information about facility's location).

mindset.[143] Nevertheless, while General Said conducted his briefing in November 2021, there has been no public issuance of a written report.

The trend seems to be less transparent investigations, at a time when the *Times* stories and other media and NGO accounts have highlighted transparency's importance. Moreover, the U.S. military has failed to provide a comprehensive, reasoned explanation for this backsliding in the transparency realm. The unexplained drift toward diminished transparency is one of the most disturbing aspects of recent U.S. military practice regarding targeting.

4. Structure and Training

Flaws in the structure of the targeting process and the training of members of the targeting team also play a role in civilian harm. Each flaw impairs the response to confirmation bias and the other deficits noted in this section. In addition, each impedes learning that might result in civilian-harm reduction.

The U.S. military has conceded that its training on civilian-harm reduction must improve. In his report on the Baghuz episode, General Michael Garrett recommended more and better training on reducing civilian harm.[144] Independent organizations such as the RAND Corporation have come to similar conclusions, arguing that current personnel have "negligible training" in methods to reduce civilian casualties.[145] Moreover, the U.S. military has not devoted sufficient attention and resources to leveraging new technology, such as artificial intelligence (AI), to aid in civilian-harm reduction.[146]

The structure of targeting creates particular problems in so-called dynamic targeting, in which military personnel order a strike because of exigent circumstances. As we have seen in studying the MSF Kunduz and Kabul attacks, members of the targeting cell do not receive arguments that a proposed target was protected under IHL or that the team had incorrectly estimated the number of civilians at the scene. In the deliberative targeting process, where time is on the targeters' side, it is arguably easier to explore counter-arguments. The

143. *See* Vergun, *supra* note 3.

144. *See Garrett Executive Summary*, *supra* note 3, at 2 (also recommending "improvements to CIVCAS [civilian casualty] policy to establish . . . procedures common to all Services").

145. RAND Report, *supra* note 11, at 56.

146. CNA Report, *supra* note 8, at 28–43; *cf.* Margulies, *supra* note 7 (discussing ways to use AI to promote situation awareness, improve deliberation by targeting cells, and increase information available about civilian persons and objects that an attack could adversely affect).

ethos of dynamic targeting, on the other hand, makes cautionary views seem like a luxury that the team cannot afford. But even in the dynamic realm, undue haste can lead to tragedy. In the dynamic MSF Kunduz and Kabul strikes, where the team took one hour and eight hours, respectively, time was available to examine alternative theories.[147] Sadly, targeting personnel frequently did not take advantage of the time available to fully assess alternative scenarios that would have produced better outcomes.

In a broader sense, the U.S. military has—at least until establishment of the new Civilian Harm Center of Excellence—lacked the "structures and capabilities" required for spotting and addressing trends in civilian harm.[148] The U.S. military needs skilled personnel whose primary responsibility lies in this area.[149] As RAND and General Garrett agreed, more clarity and focus are vital.[150]

Here, too, evidence suggests that backsliding has occurred. Earlier in the post-9/11 era, the U.S. military in Afghanistan conducted "periodic reviews" to identify trends and patterns and formulate responses.[151] The U.S. military has backed away from this institutional turn, relying on the flawed procedures described in the preceding section.[152]

III. SYSTEMIC DUTIES UNDER IHL

This section argues that under the IHL duty of "constant care,"[153] state military forces, including those of the United States, have a systemic duty to respond to the system-wide deficits in civilian-harm reduction. To expand the "constant care" duty beyond individual attack planners to the systemic level, this Part relies on several rules, principles, and practices, starting with the international law

147. *See* U.S. Air Forces, *Kunduz Airstrike Report*, *supra* note 78; Daugirdas & Mortenson, *supra* note 78; Vergun, *supra* note 3.

148. RAND Report, *supra* note 11, at 59.

149. *Id.* at 56–57.

150. *Id.* at 57 (noting that officials "move from crisis to crisis" without a "central focal point"); *id.* at 58 (observing that personnel are not "learning lessons" on how to avoid future mistakes and tragedies); *Garrett Executive Summary*, *supra* note 3, at ¶ 4 (noting the importance of developing "common understanding and practice" regarding civilian-harm reduction).

151. RAND Report, *supra* note 11, at 33. This "data-driven approach" led to lower levels of civilian casualties.

152. *Id.* at 33.

153. Additional Protocol I, *supra* note 7; Lubin, *supra* note 7, at 138; *The Other Side of Autonomous Weapons*, *supra* note 7, at 147.

requirement that states implement their obligations in good faith.[154] Further support for this systemic duty stems from the relationship of IHL to international human rights law's prohibition on the arbitrary deprivation of life.[155] Moreover, a systemic duty is consistent with current obligations, such as the duty to perform an independent, effective investigation of colorable IHL abuses[156] and the widespread state practice of drafting law of war manuals.[157]

The systemic duty is *lex lata*—i.e., currently binding on states. As we shall see, states receive deference in building particular pathways to implementing the systemic duty. The following section describes vision that is *lex ferenda*—a suggested set of best practices whose implementation is recommended, but not legally controlling.[158]

A. CONSTANT CARE AND THE REQUIREMENT OF GOOD FAITH

A general principle of international law requires states to implement their obligations in good faith.[159] That principle is of paramount importance because of the weakness of direct enforcement of international law.[160] In private law, such as rules governing property, a wronged party can go to a court for an order enforcing legal rights.[161] In contrast, much of public international law

---

154. Schmitt & Watts, *supra* note 25, at 684–85. In treaty law, this duty to discharge duties in good faith is referred to as *pacta sunt servanda*—the principle that a States Party to a treaty must act in a way that serves the underlying purposes of the treaty, instead of seeking to undermine the agreement. *See* Vienna Convention on the Law of Treaties art. 26, *opened for signature* May 23, 1969, 1155 U.N.T.S. 331 (entered into force Jan. 27, 1980).

155. *Nuclear Weapons*, *supra* note 26, at ¶ 25.

156. *See* Schmitt, *Investigating Violations of International Law in Armed Conflict*, *supra* note 117.

157. *See* W. Michael Reisman & William K. Leitzau, *Moving International Law from Theory to Practice: The Role of Military Manuals in Effectuating the Law of Armed Conflict*, 64 Int'l L. Stud. 1 (1991).

158. *Cf.* Kubo Mačák, *Military Objectives 2.0: The Case for Interpreting Computer Data as Objects Under International Humanitarian Law*, 48 Isr. L. Rev. 55, 59–60 (2015) (discussing distinction between *lex lata* and *lex ferenda*).

159. *See* Statute of the International Court of Justice, art. 38 ¶ 1(c) (Jun. 26, 1945) (discussing International Court of Justice's application of "general principles of law"); *see also* Schmitt & Watts, *supra* note 25, at 684–85; Russell Buchan & Nicholas Tsagourias, *Hacking International Organizations: The Role of Privileges, Immunities, Good Faith and the Principle of State Sovereignty*, 104 Int'l Rev. Red Cross 1171, 1186–87 (2022) (noting that good faith as a general principle of international law stems from the ancient Roman concept of *bona fides*, which connotes "trustworthiness, conscientiousness and honourable conduct").

160. Buchan & Tsagourias, *supra* note 159, at 1187.

161. For example, a mortgage lender can seek to foreclose on a mortgage if the borrower defaults on payments, and a purchaser of real property can seek specific

depends on voluntary compliance.[162] Voluntarism only works if parties trust each other; without trust, states would lack confidence that other states will observe rules of international law. As a result, each state would seek disregard international law, and instead turn to maximizing its own short-term advantage. Imposing a duty of good faith on states provides a check on this short-sighted behavior, and hence promotes the long-term stability of the global system.

The duty of good faith requires that states promote compliance with international law by their own personnel and agents. A state must be consistent and reliable in monitoring the performance of persons who act on its behalf. Merely going through the motions of compliance is insufficient. Instead, a state must be proactive in ensuring compliance. Good faith does not require clairvoyance; a state cannot predict or prevent all violations of international law by its personnel. However, a state must take reasonable measures to ensure that its personnel comply.

This good-faith requirement shapes states' constant-care duty under IHL. Attack planners must show constant care to avoid needless civilian harm, but states must provide planners and other personnel with the tools they need. As we shall see, training and dissemination of knowledge about IHL is one aspect of this good faith requirement regarding IHL.[163] Where a state can predict that cognitive flaws, such as confirmation bias, will adversely affect compliance or cause civilian harm that a state could feasibly prevent, a response to those cognitive flaws is also required. Similarly, when a state can predict that supply-chain issues or haphazard procurement policies impair access to technology and a state can feasibly modify its practices, the requirement of good faith would require those steps.

Another aspect of this duty of good faith is formulating, implementing, and enforcing procedures, including procedures regarding both consultation with senior officers and intelligence agencies in targeting and internal information-sharing in

---

performance if the seller breaches their duty to convey title to the property and other attributes of ownership.

162. Buchan & Tsagourias, *supra* note 159, at 1187; *see also* Anne van Aaken & Betül Simsek, *Rewarding in International Law*, 115 Am. J. Int'l L. 195, 213–15 (2021) (discussing how states that comply voluntarily with international law receive reputational benefits that allow them to achieve their goals in foreign relations and trade); *cf.* Jack Goldsmith & Daryl Levinson, *Law for States: International Law, Constitutional Law, Public Law*, 122 Harv. L. Rev. 1792 (2009) (discussing rationales for compliance with international law); Harold Hongju Koh, *Why Do Nations Obey International Law?*, 106 Yale L.J. 2599 (1997) (discussing explanations for compliance in light of stress on voluntariness).

163. *See infra* Section III.C.3.

investigations. Where consultation with external players such as the International Red Cross is a necessary condition of internal compliance, that outreach is required.

States that provide a good-faith backstop for the compliance efforts of their personnel will not eliminate needless civilian harm in every case. No state has the capacity to reach perfection. States comply with their good-faith duty when they formulate, implement, and review plans for civilian-harm reduction throughout the military and associated groups, including software developers and other contractors, even though those safeguards will not guarantee reductions in civilian harm in every possible case.

B. READING IHL AGAINST THE BACKDROP OF INTERNATIONAL HUMAN RIGHTS LAW

In addition to gaining support from states' good-faith requirement, a systemic view of IHL compliance draws from IHL's relationship with international human rights law (IHRL). Under the International Covenant on Civil and Political Rights, no individual should be "arbitrarily deprived" of life.[164] This Article takes a middle view of the relevance of IHRL to IHL, skirting "maximalist" theories that view IHL as displacing IHRL or view IHRL as preempting IHL. That middle view of the influence of the ICCPR's prohibition on arbitrary deprivation of life bolsters a systemic view of IHL compliance.

Some background may be useful on the complex relationship between IHL and IHRL.[165] Usually, human rights rules governing law

---

164. *See* International Covenant on Civil and Political Rights, art. 6 ¶ 1, Dec. 16, 1966, 999 U.N.T.S. 171 (hereinafter ICCPR); *see also* U.N. Hum. Rts. Comm., *General Comment No. 36 (2018) on Article 6*, Right *to* Life ¶ 4, 11 64 (Oct. 30, 2018) (hereinafter HRC, *GC No. 36*) (providing interpretation of ICCPR's prohibition on arbitrary deprivation of life by U.N. treaty body, including discussion of relationship between right to life and IHL). While the approach taken in this Article agrees with the Human Rights Committee's view that a measure of state transparency is important regarding IHL targeting criteria, the Human Rights Committee should have expressly acknowledged that a state must balance transparency against the need for safeguarding sources and methods and avoiding providing adversaries—particularly those in noninternational armed conflicts (NIACs)—with a "road map" that will provide those adversaries with an unfair advantage.

165. A large body of literature exists on this topic, which is beyond the scope of this Article to canvass comprehensively. *See* Daniel Bethlehem, *The Relationship between International Humanitarian Law and International Human Rights Law in Situations of Armed Conflict* 2 Cambridge J. INT'L & COMPAR. L. 180 (2013); Oona Hathaway et al, *Which Law Governs During Armed Conflict? The Relationship between International Humanitarian Law and Human Rights Law*, 96 96 Minn. L. Rev. 1883 (2012); Noam Lubell, *Parallel Application of International Humanitarian Law and*

enforcement apply, including strict targeting criteria that limit use of deadly force to situations in which a law enforcement officer is responding to an imminent threat to her life or the life of another.[166] In contrast, the distinct rules and principles of LOAC apply to situations where a nonstate actor has sufficient cohesion to be considered an organized armed group and where violence is of sufficient intensity and duration to set it apart from ordinary criminality or even civil unrest.[167] In an armed conflict, a party can target forces of an adversary based on their status as belligerents involved in the conflict, whether or not those particular forces happen to pose an imminent threat.[168]

Those relaxed targeting rules can pose tensions with human rights law's prohibition on arbitrary deprivations of life. According to the International Court of Justice (ICJ), the prohibition applies in armed conflict.[169] Generally, application of LOAC will be consistent

---

International Human Rights Law: An Examination of the Debate, 40 ISR. L. Rev. 648 (2007); Marko Milanovic, *Norm Conflict in International Law: Whither Human Rights?*, 20 Duke J. Comp. & INT'L L.69 (2009); Marco Sassoli & Laura Olson, *The Relationship between International Humanitarian and Human Rights Law Where It Matters: Admissible Killing and Internment of Fighters in Non-International Armed Conflicts*, 90 INT'L Rev. Red Cross 599 (2008).

166. *See* McCann v. United Kingdom, App. No. 18984/91, 21 Eur. H.R. Rep. 97 (1995).

167. Prosecutor v. Tadic, Case No. IT-94-1-I, Decision on Defence Motion for Interlocutory Appeal on Jurisdiction, ¶70 (Int'l Crim. Trib. for the Former Yugoslavia Oct. 2, 1995); *see also* Mitt Regan, *From Protecting Lives to Protecting States: Use of Force Across the Threat Continuum*, 10 J. NAT'L SEC. L. & POL'Y 171, 196–207 (2019) (reviewing Watkin, *Fighting at the Legal Boundaries: Controlling the Use of Force in Contemporary Armed Conflict*, *supra* note 39) (analyzing *Tadic* factors and cautioning against unduly narrow interpretation that will fail to acknowledge presence of armed conflict); Laurie R. Blank & Geoffrey S. Corn, *Losing the Forest for the Trees: Syria, Law, and the Pragmatics of Conflict Recognition*, 46 Vand. J. TRANSNAT'L L.. 693, 725–31 (2013) (suggesting that unduly narrow reading of *Tadic* factors in the context of the violence in Syria undermined accountability for war crimes); Deborah Pearlstein, *Armed Conflict at the Threshold?*, 58 Va. J. INT'L L. 369, 385–87 (2019) (analyzing the view expressed by Blank and Corn); Adil Ahmad Haque, *Whose Armed Conflict? Which Law of Armed Conflict?*, 45 Ga. J. INT'L & Comp. L. 475, 478–79 (2017) (suggesting that threshold for armed conflict should be lower to enhance accountability for war crimes, particularly with respect to nonstate actors such as ISIS, which—unlike states—are not subject to human rights law).

168. Blank & Corn, *supra* note 167, at 712–13; Laurie R. Blank & Benjamin R. Farley, *Identifying the Start of Conflict: Conflict Recognition, Operational Realities and Accountability in the Post-9/11 World*, 36 Mich. J. INT'L L. 467, 525 (2015); *see also* Laura T. Dickinson, *National Security Policymaking in the Shadow of International Law*, 2021 Utah L. Rev. 629, 649–51 (2021) (discussing contrast and overlap between IHL and human rights paradigms).

169. Legality of Threat or Use of Nuclear Weapons, Advisory Opinion, 1996 I.C.J. 226, ¶ 25 (July 8) [hereinafter *Nuclear Weapons*] (observing that the "protection of the International Covenant on Civil and Political Rights [including the prohibition on

with the prohibition. On this view, LOAC, including its relaxed targeting rules, is *lex specialis*—a distinct body of law applicable in armed conflicts.[170] The *lex specialis* model hinges on the practicalities of the relevant context for LOAC. According to this theory, the exigencies of armed conflict require permitting targeting based on status as a belligerent, instead of limiting use of lethal force to individuals whose actual behavior poses an imminent threat. In traditional armed conflicts, involving massed ground, sea, or air forces bearing uniforms or insignia of opposing states, one party cannot wait to inquire whether a given enemy soldier or tank poses an imminent threat. Before the inquiry is complete, the enemy force may have used lethal force against that party. Requiring the equivalent of civil discovery prior to the use of lethal force in an armed conflict would be impracticable.

Furthermore, at least in traditional armed conflicts in which opposing forces bear the uniforms or identifying insignia of their respective states, the problem of false positives—persons or objects targeted by mistake—does not typically arise. However, that problem reasserts itself in noninternational armed conflicts (NIACs) like the U.S. post-9/11 conflicts with Al Qaeda and ISIS, in which opposing forces usually do not wear uniforms. In that context, status-based targeting can present a greater risk of arbitrary or mistaken deprivations of life.[171]

A systemic account of IHL that imposes duties on states, not merely on individual attack planners, best reconciles IHL and the ICCPR. Systematic, methodical state planning can provide attack planners with the maximum feasible information, technology, and deliberative tools. Through attention to training, enhancements to

---

arbitrary deprivations of life] does not cease in times of war"); *see also* HRC, *GC No. 36, supra* note 163, ¶ 64 (agreeing that prohibition is applicable); Tallinn Manual 2.0, *supra* note 15, at 181 n.392 (agreeing that LOAC does not override human rights law and that the relationship between LOAC and ICCPR requires discerning use of "legal methodology and interpretation").

170. *Nuclear Weapons*, 1996 I.C.J. 226, ¶ 25; *see also* U.N. Hum. Rts. Cmte, Gen. Cmt. 36, ¶ 64 (noting that, "in general," use of lethal force "consistent with international human rights law" is "not arbitrary").

171. *See* VanLandingham, *supra* note 39, at 103–09. This concern also extends to detention of members of nonstate adversaries in NIACs. *See* Matthew C. Waxman, *Detention as Targeting: Standards of Certainty and Detention of Suspected Terrorists*, 108 Colum. L. Rev. 1365 (2008); *see also* Ashley S. Deeks, *Predicting Enemies*, 104 Va. L. Rev. 1529, 1532–33 (2018) (discussing role of AI in identifying members of enemy groups in NIACs for purposes of detention and targeting); Margulies, *supra* note 7 (asserting that AI can assist in avoiding false positives); *cf.* Monica Hakimi, *A Functional Approach to Targeting and Detention*, 110 Mich. L. Rev. 1365 (2012) (suggesting streamlined test for detention and targeting that synthesizes LOAC and human rights law).

procurement and distribution, adherence to procedures, and design of institutions, the state can address recurrent problems that individual attack planners cannot solve on their own. International humanitarian law insulates attack planners from state failures beyond their control with qualifiers in IHL rules and principles, such as the principle of distinction's condition that attack planners need only consider "reasonably available" information; the rule of proportionality's assessment of fault based on information reasonably available to the attack planner at the time of attack; and the limiting language in the rule of precautions in attacks that precautions must be "feasible."

These qualifiers, while necessary to avoid undue second-guessing of targeting decisions, nevertheless create a gap in which risks to civilians can proliferate. Risks include the "soda-straw" problem and the risk of unanticipated civilians present at a targeting site; the perils of confirmation bias and base-rate neglect; the excessive use of unit self-defense justifications to circumvent safeguards in the deliberative targeting process; and the flaws in investigating instances of civilian harm. Without a systemic commitment by states, civilian harm may flow from haphazard and avoidable anomalies in distribution of technology; lack of focused training; inability to learn lessons from investigations of past civilian harm; and organizational discounting of small errors regarding civilians that snowball into tragedies such as the mistaken Kunduz MSF or Kabul airstrikes.[172] That haphazard approach to civilian harm would introduce a level of arbitrariness into targeting that conflicts with the ICCPR's protections, even when individual attack planners' actions do not themselves rise to the level of IHL violations.

A systemic approach will not wholly eliminate such problems. But it will marshal the resources of states to minimize them. Since the safeguards in this systemic approach focus on process, not on particular outcomes, it will be feasible for all states to adopt these or related guard-rails. The prohibition on arbitrary deprivations requires this effort, even when the effort does not result in perfection.[173]

---

172. U.S. Air Forces, *Kunduz Airstrike Report*, *supra* note 78, ¶¶ 48–50 (discussing failure to load electronic NSLs onto aircraft used in Kunduz airstrike); *cf.* Madsen, *supra* note 13, at 156 (discussing accumulating costs of latent errors in organizations).

173. The approach taken in this Article addresses a gap between individual planners' decisions and civilian harm that Professor Asaf Lubin has also considered in an important recent piece. *See* Lubin, *supra* note 7, at 137–40. This Article's approach differs from Professor Lubin's approach in the following way: Professor Lubin imposes a tort-based duty of care on intelligence agencies that provide flawed

C. OTHER SYSTEMIC DUTIES UNDER IHL

While most IHL regarding targeting focuses on individual attack planners, certain practices are part of a state's responsibility. Those include the respective duties to investigate, employ reasonable protections regarding technology, and educate state forces about HL. I discuss each in turn.

1. The Duty to Investigate

The duty to investigate, which requires an independent and effective inquiry but otherwise does not mandate a particular approach, stems in part from treaty law but is also likely customary in nature.[174] Article 87 of Additional Protocol I imposes a duty on States Parties and parties to an armed conflict to in turn require commanders to "prevent and . . . suppress" IHL violations and report such violations to appropriate authorities.[175] A state cannot abdicate this responsibility to supervise commanders of its armed forces. That abdication would leave a gap between the "undertakings entered into

---

information to targeting cells. *Id.* at 138 (discussing "duty of care" that Lubin refers to as the "'reasonable intelligence agency' test"); *cf.* Marko Milanovic, *Mistakes of Fact When Using Lethal Force in International Law: Part I*, EJIL TALK! (Jan. 14, 2020), https://www.ejiltalk.org/mistakes-of-fact-when-using-lethal-force-in-international-law-part-i/ (arguing that an honest and reasonable mistake by an attack planner would not violate IHL or human rights law); Milanovic, *Mistakes of Fact When Using Lethal Force in International Law: Part III*, EJIL TALK! (Jan. 15, 2020), https://www.ejiltalk.org/mistakes-of-fact-when-using-lethal-force-in-international-law-part-iii/ (suggesting that honest but unreasonable mistake in use of force, such as Iran's 2020 downing of a Ukraine passenger jet that had just taken off from Tehran, was a violation of the *jus ad bellum* that comprises international norms on the initiation of armed conflicts and limits use of force to instances of self-defense against an armed attack); Rebecca Crootof, *War Torts*, 97 N.Y.U. L. REV. 1063 (arguing for imposing regime of strict liability on states entailing obligation to pay compensation for any and all civilian harm in armed conflict, even when an attack that causes civilian harm stems from a reasonable decision by attack planners who have accurate targeted an adversary and caused harm to civilian persons or objects that is not excessive, either from the standpoint of the attack planner at the time of the attack or even viewed in hindsight). The Article discusses differences between the benchmarking approach and the tort approach in the section just before the Article's Conclusion on alternatives to the benchmarking model.

174. *See* Schmitt, *supra* note 117, at 36; Geneva Convention Relative to the Protection of Civilian Persons in Time of War, art. 146, Aug. 12, 1949, 6 U.S.T. 3516, 75 U.N.T.S 287 (requiring states that have approved Fourth Geneva Convention to apprehend and prosecute or transfer custody to regional or international tribunals those who have committed grave breaches of any of the Geneva conventions). The other three Geneva Conventions contain similar language. Schmitt, *supra* note 117, at 36 n. 18.

175. Additional Protocol I, *supra* note 7, art. 87(1).

by Parties to the conflict and the conduct of individuals" serving in those parties' armed forces.[176] A gap would be evidence that a state was not fulfilling its treaty or customary duties in good faith. To fill that gap, states must require that commanders "act proactively."[177] States and state officials up and down the chain of command bear responsibility for implementing this proactive approach.[178]

2. Safeguards Regarding Technology

States and other authorities also currently highlight the systemic value of certain safeguards. For example, the United States Department of Defense's law of war manual urges the compiling and distribution of a no-strike list (NSL) that designates cultural and other protected sites, presumptively taking them off the list of potential targets.[179] In the wake of the mistaken 2015 airstrike on the MSF Kunduz facility, the U.S. military recommended a further systemic safeguard: pre-loading electronic NSLs, including GPS coordinates of protected sites, onto aircraft that commanders may use in ongoing operations.[180] In the U.S. military, the recommendation of such systemic safeguards reflects and implements the lessons learned through experience, including the MSF Kunduz tragedy.[181]

Manuals on the law of armed conflict and related issues such as countermeasures in the cyber realm have also recommended approaches to particular situations.[182] For example, the Tallinn

176. INT'L COMM. OF THE RED CROSS, COMMENTARY ON THE ADDITIONAL PROTOCOLS OF 8 JUNE 1977 TO THE GENEVA CONVENTIONS OF 12 AUGUST 1949 3550 (Yves Sandor et al. eds., 1987).

177. Schmitt, *supra* note 117, at 41.

178. *Id.* at 42.

179. DoD Manual, *supra* note 44, ¶ 5.18.4 n.507; Margulies, *supra* note 44, at 209–10; *see also* Alcala, *supra* note 44, at 233–39 (discussing specific safeguards for cultural property such as archaeological sites, and suggesting that safeguarding such sites is required under both treaty and customary law).

180. *See* U.S. Central Command, *supra* note 79.

181. *Cf.* RAND Report, *supra* note 11, at 32, 59 (suggesting that U.S. military must develop broader capabilities for learning lessons regarding civilian-harm reduction); *Garrett Executive Summary*, *supra* note 11, at 2 (acknowledging that U.S. military needs to expand its capabilities on civilian-harm reduction and expressing support for pending Department of Defense guidance on this issue).

182. A countermeasure is a state response to another state's breach of an international law duty owed to the victim state. Such breaches, which could include interventions in the victim state's internal governmental affairs, are typically below the threshold of force. For example, a state or its agents might launch a distributed denial of service (DDoS) attack on the website of a governmental unit of the victim state, disabling the website temporarily. Countermeasures by the victim state must be proportional and below the use of force threshold. Nonstate actors are not

Manual recommends that states, where feasible, use experts in cyber to advise them on precautions.[183] The use of experts is important because of the sophistication and interdependence of cyber systems.[184] This recommendation dovetails with the benchmark approach. The recommendation of expert participation goes beyond individual attack plans into the realm of systems. An individual attack planner in an exigent situation lacks the time to consult the Internet and obtain an expert's contact information. Instead, senior officials in the state force's chain of command must plan *systemically* to have experts available, with the understanding that in an armed conflict, access to experts will vary with circumstances.

Similarly, in the cyber realm, a state force should undertake particular categories of activities only with great care. For example, as the Tallinn Manual suggests, extreme caution should accompany the use of thumb drives to place malware on an adversary's putatively closed military network.[185] Given the interdependence of cyber systems, even a supposedly closed military network may have links to civilian computer architecture. The adverse effects on civilian systems will figure in both the proportionality calculus and compliance with precautions in attack.[186] This caution about means and methods of cyber warfare is not merely individual. A particular attack planner cannot reinvent the wheel of IHL compliance for every new engagement. Instead, such guidance is inherently systemic, providing an overall framework to guide planners.

---

permissible targets for countermeasures. *See* Tallinn Manual 2.0, *supra* note 15, at 128; Air Service Agreement of 27 March 1946 between the United States of America and France, Fr.-U.S., Mar. 27, 1946, 18 R.I.A.A. 417; Michael N. Schmitt, *"Below the Threshold" Cyber Operations: The Countermeasures Response Option and International Law*, 54 VA. J. INT'L L. 697, 715 (2014); *see also* Frederick Gilles Sourgens, *Cyber-Nuisance*, 42 U. PA. J. INT'L L. 1005, 1020–34 (2021) (arguing that connections between states and non-state actors on intrusive cyber actions complicate the law of countermeasures, since countermeasures can only target state actors, and attribution of responsibility for intrusive actions can be difficult between state officials and associated nonstate players such as "hacktivists").

183. Tallinn Manual 2.0, *supra* note 15, at 477.

184. *Id.*

185. *Id.* at 480–81.

186. The proportionality analysis will assess whether an attack planner considering use of a thumb drive should reasonably have expected that harm to civilian objects would be excessive, in light of the military advantage to be gained. The rule of precautions in attack would require consideration of whether any alternative approach, such as a different design for the malware, could reduce civilian harm.

3. The Role of State Law of War Manuals

State drafting and issuance of law of war manuals is another systemic step.[187] Manuals distribute ideas, much like military logistics specialists distribute technology; each is part of the attack planner's toolkit. Law of war manuals often contain instructions on systemic measures. Moreover, state military organizations do not draft manuals to sit in long-term storage. To accomplish their role, a state's military must publicize the manual for its intended audience of combat operators and their legal advisors.[188] Education of all stakeholders is a necessary incident of compliance with IHL.[189] Without that education and awareness, compliance will be sporadic and inconsistent, leaving a gap between a state's abstract positions on IHL and its compliance on the ground.[190] As elsewhere, taking steps to fill that gap is necessary for good faith compliance.

Manuals also prompt dialogue within and among states, experts, and advocates that advances the development of international law.[191] The United States has treated its manual as an iterative project that adjusts to comments by external and internal audiences. A related dialogic occurs with all manuals, which are always subject to revision. In addition, states cite the manuals of other states. Moreover, international groups such as the Tallinn Manual's assemblage of experts and the International Committee of the Red Cross use states' manuals as a reference point for their work, which in turn constitutes evidence of international law. This give-and-take is an essential

187. *See* Michael A. Newton, *Framing Thoughts on the DoD Law of War Manual and this Commentary*, in THE UNITED STATES DEPARTMENT OF DEFENSE LAW OF WAR MANUAL: COMMENTARY AND CRITIQUE 3, 4 (Michael A. Newton ed., 2018); Reisman & Lietzau, *supra* note 157.

188. Reisman & Lietzau, *supra* note 157, at 2 (declaring that manuals serve essential goal of "dissemination and effective internalization of authoritative norms . . . 'in the field'").

189. *Id.* at 3 (explaining that dissemination is a "necessary step if law is to be transformed from an exercise in theory to a matter of practice").

190. *Id.* (observing that education in a state's guidance regarding IHL is a prerequisite to effective enforcement, and, without providing education in advance, the investigation, adjudication, and punishment of IHL violations seems "arbitrary, retroactive and *ex post facto* and undermine rather than reinforce the symbol of law").

191. *Id.* at 7 (asserting that "manuals are an important mode for making international law as well as evidencing its existence"); *see also* Charles J. Dunlap, *Practitioners and the Law of War Manual*, in THE UNITED STATES DEPARTMENT OF DEFENSE LAW OF WAR MANUAL: COMMENTARY AND CRITIQUE, 65, 69 (Michael A. Newton ed., 2018) (noting that the U.S. DoD Manual "make[s] a real contribution to norm development").

element in international law's evolution.[192] It thus testifies to the importance of law of war manuals and the integral role of a systemic approach to IHL compliance.

IV. BENCHMARKING AS SYSTEMIC CIVILIAN-HARM REDUCTION: STRUCTURE AND APPLICATION

The *lex lata* duty to systemically address civilian harm may elicit a range of state responses. The standard is flexible: any response that meets the test of good faith and reduces arbitrary results is consistent with this obligation.[193] As a *lex ferenda* alternative, this Article suggests benchmarking. A benchmark on this view is a standard that serves as a reference. It can indicate a particular level of performance that an individual or entity has reached before and aspires to reach again. It can also be aspirational, stating an aim or objective, while leaving the subject of regulation discretion to reach that goal in a variety of ways. Benchmarking is not necessarily quantitative, although measurements such as percentage and absolute reductions in civilian harm will often be useful. Instead of hinging on quantitative results, benchmarking relies on enhancing deliberation.[194] Attention to goals and diagnosing failures to meet those goals will accomplish more than a rigidly prescriptive fixation on results.

To focus on deliberation, benchmarking turns to three mainstays of U.S. administrative law: impact assessments, reasoned explanation, and notice and comment. At least one of those doctrines, the impact statement, has also made substantial inroads in international law.[195] The others are implicit or are beginning to surface in steps like the new Political Declaration on explosive weapons in urban areas.[196] The

192. *See* Marko Milanovic & Sandesh Sivakumaran, *Assessing the Authority of the ICRC Customary IHL Study*, 104 INT'L REV. RED CROSS 1856 (2022).

193. *See* Beth Van Schaack, *Book Review: Sigrid Redse Johansen, The Military Commander's Necessity: The Law of Armed Conflict and Its Limits*, 115 AM. J. INT'L L. 176, 179–80 (2021) (arguing that states and commanders are entitled to measure of deference in targeting decisions, given exigencies of armed conflict).

194. Lubin, *supra* note 7, at 147–48 (reflecting on need for deliberation); *see also* Peter Margulies, *A Moment in Time: Autonomous Cyber Capabilities, Proportionality, and Precautions, in* AUTONOMOUS CYBER CAPABILITIES UNDER INTERNATIONAL LAW 152 (Rain Liivoja & Ann Väljatagam eds., 2021) (discussing importance of review of performance of autonomous agents).

195. Arg. v. Uru., 2010 I.C.J. at ¶ 204; Jensen & Watts, *supra* note 28, at 676–78 (analyzing decision).

196. *See, e.g.*, Political Declaration, *supra* note 32, at § 4.7 (envisioning meetings "in a collaborative spirit" with other nations, the United Nations, the ICRC, and "other relevant international . . . and civil society organizations").

application of these mainstays here is a matter of soft law.[197] While administrative law in the United States turns on judicial review of agency decisionmaking,[198] benchmarking as this Article uses the term refers to more informal and collaborative interactions. That may be a downside of benchmarking: opportunities for deliberation do a lot of work, while avenues for enforcement are limited. On balance, the space for deliberation justifies the approach.

A. THE COMPONENTS OF BENCHMARKING

Benchmarking relies on civilian impact statements, reasoned explanations, and notice and comment.[199] This subsection discusses each in turn, using the example of the interaction of the Obama PPG and Trump PSP on targeting.

1. Civilian Impact Statements

The need for an environmental impact statement (EIS) is well-established under international law.[200] Drafting and issuance of an EIS is one aspect of the due diligence that states owe in supporting actions of private parties, such as mining or navigational projects, that may

197. *See* Ganesh Sitaraman & Ingrid Wuerth, *The Normalization of Foreign Relations Law*, 128 HARV. L. REV. 1897, 1973–74 (2015) (discussing domestic law significance of soft, non-binding law).

198. Chachko, *supra* note 27, at 1130–33 (discussing judicial review in national security cases).

199. Proposed U.S. legislation, the Protection of Civilians in Military Operations (POCIMO) Act, employs a parallel approach, with differences discussed below. Senator Elizabeth Warren was among a group of Democratic senators who introduced the bill. *See* Protection of Civilians in Military Operations Act, S. 4108, 117th Cong. (2022), https://www.warren.senate.gov/imo/media/doc/Protection%20of%20Civilians%20in%20Military%20Operations%20Act1.pdf; *see also* Annie Shiel & Sarah Yager, *Congressional Action on Civilian Harm Resulting from U.S. Military Operations: Part I*, JUST SECURITY (Apr. 28, 2022), https://www.justsecurity.org/81303/congressional-action-on-civilian-harm-resulting-from-u-s-military-operations-part-i/ (analyzing POCIMO); Laura Dickinson, Brianna Rosen & Rachel VanLandingham, *Congressional Action on Civilian Harm Resulting from U.S. Military Operations: Part II*, JUST SECURITY (Apr. 28, 2022), https://www.justsecurity.org/81315/copy-3/ (discussing companion legislation regarding DoD transparency). POCIMO does not adopt the specific administrative law framework that benchmarking utilizes, although it would require methodical DoD disclosures of past harm and remedial action. POCIMO also includes provisions on documenting the basis for all past U.S. strikes. These provisions would be unduly cumbersome and distract from essential reforms. *See* Crootof, *supra* note 173 (discussing alternatives to benchmarking, including strict liability for civilian harm and POCIMO framework).

200. Arg. v. Uru., 2010 I.C.J. at ¶ 204; Jensen & Watts, *supra* note 28, at 676–78.

adversely affect other states' interests.[201] An EIS will outline those impacts. In doing so, it will promote deliberation by the state issuing the EIS on the consequences of its choices and also provide a referent for interactions with the affected state. A civilian impact statement (CIS) would serve the same function, encouraging state officials to be forthright about their decisions and the effects of those decisions on civilian-harm reduction.

A CIS could promote the positive impact of steps to refine military targeting. For example, two distinguished LOAC scholars with substantial military experience have recently suggested that the U.S. military add a civilian risk-mitigation officer to the targeting team. That officer would prioritize civilian protection, countering the groupthink and confirmation bias that sometimes plague targeting. A CIS would discuss the positive impact of such a structural step.[202]

A mix of positive and negative impacts was characteristic of the Obama PPG, issued in 2013 with media mentions and made public with some redactions in 2016.[203] The PPG centralized planning for targeted strikes in these areas.[204] Responding to criticism by NGOs and others, the PPG set a very restrictive condition for permissible harm to civilian persons, requiring "near certainty" that the proposed action would not result in such harm.[205] Continuing the theme of limits, the PPG also required "near certainty" that a target was present at the target site prior to launching an attack.[206] This limit was helpful because a more casual approach to the target's presence might have increased the risk of mistaken identification and harm to civilians at the scene.

A CIS on the Obama PPG would have noted the positive impact of

---

201. The duty of due diligence applies when a state works with private actors whose actions have adverse effects on a neighboring state, through pollution or other factors. *See* Trail Smelter (U.S. v. Can.), 3 R.I.A.A. 1905, 1963 (Conv. of Ottawa Tribunal 1938 & 1941) (asserting that each state "owes at all times a duty to protect other States against injurious acts by individuals from within its jurisdiction").

202. *See* Geoffrey S. Corn & Michael W. Meier, *Enhancing Civilian Risk Mitigation by Expanding the Commander's Information Aperture*, *in* THE GLOBAL COMMUNITY: YEARBOOK OF INTERNATIONAL LAW AND JURISPRUDENCE 159, 183-89 (Giuliana Ziccardi Capaldo ed., 2019).

203. *See* PPG, *supra* note 31. The Biden administration issued guidance in October 2022 that largely echoed the Obama administration's approach. *See* Charlie Savage, *White House Tightens Rules on Counterterrorism Drone Strikes*, N.Y. Times, Oct. 10, 2022 (requiring senior officials' approval for certain strikes, in contrast to greater leeway for commanders that Trump guidelines provided).

204. *See* PPG, *supra* note 31 at 2–5 (including approval of "operational plans" through interagency process, culminating in presidential sign-off).

205. *Id.* at 1.

206. *Id.* at 3.

these restrictions. However, to be useful in IHL's axiomatic balance of humanity and military necessity, the CIS would also have noted that an *ex ante* announcement of such strict limits could cede the tactical and strategic initiative to the United States' adversaries, such as Al Qaeda and ISIS. Perhaps the CIS would then have noted that this concern was overstated. The PPG also permitted departure from the guidelines in extraordinary circumstances.[207] Moreover, experience since the PPG's 2013 issuance has indicated that targeters can be effective even within restrictive parameters on civilian harm, as the strike on Al Qaeda's Dr. Zawahri showed.[208]

A complete CIS on the Obama PPG would have had to acknowledge that a limit on the PPG's scope circumscribed its utility in driving civilian risk-reduction. Consider the PPG's express limitation to targeting outside active combat zones.[209] This curb on the PPG's coverage also confined its impact on civilian harm reduction, since civilians in active combat zones, including Afghanistan, were still at risk. A complete CIS would have clearly acknowledged the effects of this limitation.

Finally, the Obama PPG contained an enigmatic exception. It expressly disclaimed coverage of "otherwise lawful and properly authorized activities that may have lethal effects, which are incidental to the primary purpose of the operation."[210] The average reader of this disclaimer would not understand the disclaimer's subject or how it could affect civilian harm. A more knowledgeable reader might construe this oblique reference as a recognition of a unit's right to act in self-defense based on an imminent threat it encountered in the course of a mission.[211] Under a benchmarking conception, a CIS would have to assess the impact of this disclaimer.

Viewed as a CIS, the Trump PSP would raise even more serious questions than the Obama PPG did.[212] The Trump PSP continued the Obama PPG's general requirement of near certainty that a campaign or individual strike would not result in civilian harm.[213] However, the overall trajectory of the Trump PSP entailed greater delegation of discretion to attack planners, with less oversight by the executive

207. *Id.* at 17.

208. Barnes & Schmitt, *supra* note 1.

209. PPG, *supra* note 31, at 1.

210. *Id.* at 2 n.1.

211. *See* Corn, *supra* note 10. It is possible that further internal guidance on the PPG explained the meaning of this disclaimer.

212. *See* Chachko, *supra* note 27, at 1085 (asserting that the Trump PSP loosened restrictions in the Obama PPG).

213. PSP, *supra* note 31, at 4.

branch.[214] Yet the Trump PPG contained no assessment of whether civilian harm would increase as a result, and, if so, the extent of the increase. Moreover, the Trump PSP contained an express exception for unit self-defense.[215] As noted in an earlier section, unduly expansive unit self-defense claims have been a major source of hasty targeting decisions.[216] Despite this concern, which was discernible based on the records of drone strikes in Syria and elsewhere, the Trump PSP did not indicate that an express exception for unit self-defense claims could encourage attack planners to frame an excessive number of strikes in this fashion. Nor did the Trump PSP address how sidestepping restrictions in rules of engagement or deliberate targeting could increase civilian harm.

2. Reasonable Explanation

The reasonable explanation prong of benchmarking complements the CIS. It requires concrete, express consideration of material issues concerning the policy. The requirement of a reasonable explanation, like the CIS, has beneficial effects *ex ante*; it encourages decision-makers to consider the substance of a policy in light of the respective ease or difficulty that the decision-makers will experience in seeking to explain it.[217] Ideally, an unsound policy proposal will not survive this rigorous test, including a targeting policy proposal that tilts too far toward military necessity and therefore unduly discounts the humanity principle that also drives IHL.

Applying the reasoned explanation requirement to both the Obama PPG and Trump PSP sharpens the questions that the CIS factor triggered. Once again, despite triggering questions of its own, the Obama PPG comes closer than the Trump PSP to meeting the criteria. Consider first the gaps in the Obama PPG's explanations. To see a crucial gap, juxtapose the Obama PPG's limit on geographic scope to places away from active combat areas and the counterinsurgency focus of the conflict *within* those combat areas.[218] Suppose that in a

214. *Id.* at 5 (expressly allowing authorization of persistent campaign of U.S. "direct action" without concrete oversight after officials approved operational criteria for the campaign).

215. *Id.* at 1 n.1.

216. Corn, *supra* note 10, at 11.

217. Motor Vehicle Mfrs. Ass'n v. *State Farm* Mut. Auto. Ins. Co., 463 U.S. 43.

218. *See* Jamie A. Williamson, *Using humanitarian aid to 'win hearts and minds': a costly failure?*, 93 INT'L REV. RED CROSS 1035 (2011); *cf.* Regan, *supra* note 6 (discussing this limitation).

counterinsurgency effort, reduction in civilian harm was a strategic imperative, necessary for winning hearts and minds. If that was so, the PPG's restrictions should have been applied across the board. There may have been adequate explanations for the limitation on the PPG's geographic scope, such as the need to delegate more discretion to individual commanders within areas where combat raged. However, the PPG failed to provide those answers.

In addition, the PPG failed to acknowledge the benchmark provided by the useful periodic assessments that the U.S. military had conducted in Afghanistan from 2009 to 2012.[219] Those assessments had allowed the United States to "tailor adaptations to reduce risks to civilians."[220] Yet there was no mention in the PPG of why the United States had failed to continue this "data-driven" approach[221] or why it was not part of the PPG itself.

Despite these concerns, the Obama PPG did prod targeting cells to provide reasonable explanations. For example, the PPG mandated that dossiers compiled for proposed targets include all relevant information and flag "gaps in existing intelligence, as well as inconclusive and contradictory intelligence reports."[222] This requirement encouraged reflection and a sober second look at sources, countering confirmation bias.

The Trump PSP raised even more questions than the Obama PPG along the reasoned-explanation axis. Although the Trump PSP delegated more control to commanders in the field, it did not explain why this expanded delegation was necessary. Moreover, the PSP and related documents diluted the focus of the Obama PPG on the nomination of specific high-value targets such as Dr. Zawahri. Stating that a proposal from a combat agency "need not distinguish between identified high-value terrorist targets ... and other targets" encouraged a turn to less identifiable targets whose status in the conflict with the United States may have been less clear. That tendency posed tensions with the principle of distinction. The Trump administration failed to explain why creating this tension was necessary or whether it had considered alternatives.

3. Notice and Comment

In administrative law, notice and comment procedures serve to

219. *See* RAND Report, *supra* note 11, at 33.
220. *Id.*
221. *Id.*
222. PPG, *supra* note 31, at 11-12.

keep bureaucrats honest. A proposed policy may attract thousands of comments from interested parties, including NGOs and advocates. An agency must respond to those comments. If an agency makes arbitrary choices, stakeholders can highlight the consequences of those choices.[223] That feedback can also help set targeting policy, since individual perceptions are biased and a crowdsourcing approach will often detect cognitive flaws.[224]

For notice and comment, the pattern noted above continues regarding the Obama PPG and the Trump PSP: there are questions for the former, but questions for the latter are more numerous. The Obama PPG were part of a process that also included ongoing engagement with NGOs.[225] However, the PPG failed to establish a comprehensive reporting mechanism or a formal channel for communication with civil society groups. As a result, NGOs had to guess whether the U.S. government had followed the PPG's procedures and, if so, what benefits had accrued. The notice and comment component of benchmarking required more.

Despite these questions, the Obama PPG was closer to the spirit of notice and comment than the Trump PSP. Civil society groups have virtually no input into the Trump policy. They critiqued the policy, but there is no evidence that Trump administration officials responded to their concerns. In this sense, the Trump PSP represented a notable instance of backsliding from President Obama's approach, violating one of benchmarking's central tenets.

B. THE POLITICAL DECLARATION ON EXPLOSIVE WEAPONS IN URBAN AREAS

A recent Political Declaration by the United States and many other countries on the use of explosive weapons in populated areas convey a broader commitment than either the PPG or PSP to

223. *See* Hoctor v. U.S. Dep't of Agric., 82 F.3d 165, 170–71 (7th Cir. 1996) (discussing rationale for requiring formal rulemaking process by agency, including notice and comment).

224. *See* Kahneman et al., *supra* note 9, at 85.

225. *See* Jonathan Hafetz, *A Problem of Standards?: Another Perspective on Secret Law*, 57 Wm. & Mary L. Rev. 2141, 2163-64 (2016) (noting that PPG were part of a process that also had included public discussion of the legal authority for drone strikes by senior officials such as Harold Koh, then Legal Adviser to the State Department); *see also* Jonathan Chait, *Five Days That Shaped a Presidency*, N.Y. Mag. (Aug. 25, 2016) (quoting President Obama as acknowledging in an interview, "I'm glad the left pushes me [on criteria for drone strikes]" while also quoting President Obama's view that PPG stemmed largely from the President's own sense that more precise guidance was necessary).

principles underlying benchmarking.[226] The Declaration leaves unanswered questions on the depth of states' commitment to reducing civilian harm from such weapons and the level of engagement with NGOs.[227] Moreover, the Declaration is not a substitute for each state's individual benchmarking on this topic. However, the Declaration sends a positive signal about openness to collaboration and change.

In the Declaration, Ireland—which led the effort—and other participants such as the United States and the United Kingdom called on states to "[i]mplement, and, where necessary, review, develop or improve national policy and practice" on the use of explosive weapons—including missiles, bombs, and artillery fire—in urban areas.[228] The Declaration also expresses the intent of state signatories to "[e]nsure that . . . armed forces adopt and implement a range of policies and practices to help avoid civilian harm."[229] In addition, the Declaration expresses a commitment to "[c]ollect, share, and make publicly available . . . data on the direct and indirect effects on civilians" of military actions.[230]

These steps are not outlined in a rigidly prescriptive way. For example, under the Declaration, states should "take into account the direct and indirect effects on civilians" of military actions and conduct "damage assessments," where the latter is "feasible."[231] The feasibility test, which echoes the language in the rule of precautions in attack, recognizes the challenges of conducting detailed assessments in a dynamic battlespace. Nevertheless, the commitment to developing concrete policies and examining outcomes is consistent with the CIS component of benchmarking.

The Declaration acknowledges the impact of the use of explosive weapons but falls short of the kind of specific findings that a Civilian Impact Statement (CIS) would include. For example, the Declaration

---

226. *See* Political Declaration, *supra* note 32; *see also* Robinson & Nohle, *supra* note 16, at 115–16 (discussing impact); *see generally* Int'l Committee of the Red Cross, *International humanitarian law and the challenges of contemporary armed conflicts: Recommitting to protection in armed conflict on the 70th anniversary of the Geneva Conventions*, 101 INT'L REV. RED CROSS 869 (2019).

227. *See* Meier, *supra* note 32 (discussing risk that Declaration will do little to change state practice); *see also* Simon Bagshaw, *Protecting Civilians from Explosive Weapons in Populated Areas: A New Political Declaration*, Just Security (July 6, 2022), https://www.justsecurity.org/82220/protecting-civilians-from-explosive-weapons-in-populated-areas-a-new-political-declaration/ (discussing deliberations that contributed to Declaration's final wording).

228. Political Declaration, *supra* note 32, § 3.1.

229. *Id.* § 3.3.

230. *Id.* § 4.2.

231. *Id.* § 3.4.

admits that explosive weapons in urban areas can have a "devastating" impact.[232] It concedes that these effects, including indirect effects on civilian infrastructure such as sewer systems, can cause injury and death and cause long-term health effects, as well as disrupt key services such as education and health care. [233] However, the Declaration would have been even more convincing on this score if it had taken the next step of offering concrete examples of harm to civilians, such as the harm from the campaign of the United States and its allies against ISIS and the pivotal assault in 2017 on the ISIS stronghold of Raqqa. According to the RAND Corporation, the United States did not seek adequate information about the movements of civilians within Raqqa prior to the use of explosive weapons.[234]

Acknowledgment of this omission would have shown that the United States, which has been a principal participant in armed conflicts in the post-9/11 period, has "skin in the game." That showing would have been helpful as a signal that the United States was on board with efforts to reduce the toll on civilians. Without such specific, concrete acknowledgment of past impacts, there is a risk that the principles and values in the Declaration, while commendable, will be seen as vacuous generalities rather than a guide to action.[235]

Another flaw in the Declaration is its insufficient attention to technological fixes for civilian harm. The Declaration does mention the need for states to exchange "technical and tactical expertise."[236] However, the Declaration fails to expand on the importance of video and AI technology in reducing civilian harm. For example, consider the soda-straw problem of unanticipated civilian presence discussed earlier in this Article.[237] Rather than relying on one drone, one camera, and one pilot, the military could use two drones or equip each drone with two cameras. In addition, a force could use an AI agent to quickly toggle between the narrow and broad feeds to detect oncoming civilians. The AI agent, capable of working far more rapidly than a human, could send a sensor alert regarding the unanticipated presence of civilians or temporarily halt the attack autonomously.[238] A detailed explanation of this kind might well have been inappropriate

232. *Id.* § 1.2.

233. *Id.* §§ 1.3–1.5.

234. RAND Report, *supra* note 11, at 21–22.

235. *See* Kishanthi Parella, *Improving Human Rights Compliance in Supply Chains*, 95 NOTRE DAME L. REV. 727, 759 (2019) (discussing how certain state efforts at compliance with international law become suspect as "cheap talk" with little substance).

236. Political Declaration, *supra* note 32, § 4.1.

237. *See supra* text accompanying notes 75–82.

238. *See* Margulies, *supra* note 7, at 159.

in a statement of broad principles such as the Declaration. However, the Declaration should have devoted more space to the potential of technology, even if it framed that discussion in general terms.

For similar reasons, the Declaration offers little that is specific in terms of reasoned explanations. The generalities of the Declaration on the humanitarian cost of explosive weapons in urban areas are manifestly accurate, but contribute little but positive sentiment to addressing the problem. The Declaration could have explained why it had opted to explore modifying the use of explosive weapons in urban contexts, rather than categorically barring their use in such environments. The answer may have been that nonstate actors such as Hezbollah have embedded their own weapons in urban areas, making effective targeting impossible without some recourse to explosive weapons.[239] Stakeholders could have accepted this explanation or urged states to enact a broader prohibition. Without that explanation, debate on the merits is more difficult, leaving generalities to occupy the field.

Similarly, the Declaration includes notice and comment processes, although those are vaguely defined. It notes that state signatories should "[f]acilitate the work of the United Nations, the ICRC, other relevant international organizations and civil society organizations aimed at protecting and assisting civilian populations and addressing" the effects of explosive weapons in urban areas.[240] Civil society groups can also participate in meetings "review[ing] in a collaborative spirit the implementation" of the Declaration.[241] The Declaration could have done more to integrate civil society groups into the formulation of approaches to reducing the impact of the use of explosive weapons. For example, it could have set up temporal benchmarks for meetings, committing state signatories to meet with civil society groups periodically in advance of state meetings and share tentative proposals. Perhaps follow-up activity will take a more concrete turn. Methodical follow-up will be necessary to move the Declaration from a promising start to a blueprint for action.

C. BENCHMARKS AND RESPONSIBLE AI

Like the Political Declaration on explosive weapons, the U.S. Defense Department's recent guidance on AI sounds the themes

239. Scholars have discussed how such nonstate groups violate IHL by imposing such risks on civilians. *See* Eric Talbot Jensen, *Precautions against the effects of attacks in urban areas*, 98 INT'L. REV. RED CROSS 147, 156–57 (2016).

240. Political Declaration, *supra* note 32, § 4.6.

241. *Id.* § 4.7.

featured in the benchmarking approach.[242] Much of the language in the responsible AI (RAI) DoD documents seems open and collaborative. Moreover, the documents acknowledge AI's flaws.[243] Nevertheless, beneath the gleaming surface, questions arise, particularly on the scope of collaboration with civil society groups.

1. A Brief Introduction to AI

AI is an imprecise term for software, often in algorithmic form, that analyzes data. Around the world, a range of sectors, including finance, law enforcement, housing, and health care use AI to screen applicants, devise and implement selection criteria, and produce documents.[244] AI agents have vast memories and can spot patterns in a blizzard of variables that could overwhelm human analysts. In doing work with blinding speed, AI agents also have the potential to remain immune from human traits such as deception, anger, hatred, and fear. However, along with these virtues, AI agents also have flaws, which an impact assessment should identify and analyze.[245]

AI's flaws appear whether AI agents work autonomously—without prior specific human approval—or collaborate with human personnel. Because AI agents rely on the data that developers provide, their outputs can be "brittle"—AI agents' lack of contextual understanding can prompt sharp swings in the evaluation of very similar phenomena.[246] Moreover, because a software developer may

242. *See* RAI PATHWAY, *supra* note 30.

243. *See* JARED DUNNMON ET AL., DEFENSE INNOVATION UNIT, U.S. DEP'T OF DEF., RESPONSIBLE AI GUIDELINES IN PRACTICE 5 (2021), https://assets.ctfassets.net/3nanhbfkr0pc/acoo1Fj5uungnGNPJ3QWy/3a1dafd64f22efcf8f27380aafae9789/2021_RAI_Report-v3.pdf (discussing problem of bias in AI) [hereinafter RAI Guidelines].

244. The literature on this subject is vast; tracing its development is beyond the scope of this article. Even a tiny selection of scholarly contributions shows the complexity of the issues and the extraordinary work being done. *See, e.g.*, Jane R. Bambauer et al., *When a Small Change Makes a Big Difference: Algorithmic Fairness Among Similar Individuals*, 55 U.C. DAVIS L. REV. 2337, 2347–53 (2022) (discussing examples of algorithms causing different results in credit ratings and targeted consumer advertising); Aziz Z. Huq, *Constitutional Rights in the Machine-Learning State*, 105 CORNELL L. REV. 1875, 1881–83 (2020) (discussing algorithms in criminal justice, detection of child abuse, and allocation of government services and benefits); *see also* Margot E. Kaminski & Jennifer M. Urban, *The Right to Contest AI*, 121 COLUM. L. REV. 1957, 2012–40 (2021) (discussing range of models allowing individuals to challenge impact of AI on decisions affecting them).

245. *See* Rubenstein, *supra* note 8, at 788 (describing the challenges of anticipating and integrating social impacts, such as fairness and nondiscrimination, into AI); *see generally* Andrew D. Selbst, *An Institutional View of Algorithmic Impact Assessments*, 35 HARV. J. L. & TECH. 117 (2021).

246. *See* Margulies, *supra* note 8, at 405–08; Katherine J. Strandburg, *Rulemaking*

input an unrepresentative data set to train the AI agent, the AI agent may show bias, accurately identifying people of color at lower rates than it recognizes white people.[247] In addition, the outputs of many AI models, including neural networks, are not readily explainable; AI agents do not provide a conventional written record of how they derived outputs from inputs.[248] That explainability gap can be problematic, particularly when apparent mistakes prompt a need for accountability and correction.[249]

2. AI in the DoD: Military and National Security Applications

Since military forces are often large entities, they can find uses for AI in myriad contexts. For example, the U.S. military provides medical care to active-duty service members and veterans. Using AI to interpret diagnostic tests such as X-rays, CAT-scans, and MRIs could be a massive benefit if AI agents interpreted the tests correctly and efficiently.[250] The healthcare sector overlaps only in a small way with the battlefield activities that are the subject of this Article, but it also has high stakes for life and safety. In addition, DoD is currently testing AI models to ferret out foreign election-influence operations across

---

and Inscrutable Automated Decision Tools, 119 COLUM. L. REV. 1851, 1877–78 (2019) (discussing brittleness in AI outputs); Rubenstein, *supra* note 8, at 772 (noting brittleness of AI and how government procurement officials should address this flaw); Bambauer et al., *supra* note 244, at 2367–72 (discussing how small changes in inputs can yield big differences in AI outcomes, while noting that normative significance of this input-output disparity may vary with nature of input and reasonableness of its link to output). For example, by adding small shapes such as specks to an inputted image, researchers have been able to drastically alter the AI agent's classification of the image. Adding specks to an image of a stop sign can cause an AI agent to view the image as representing something materially different, such as a yield or speed limit sign. Even a fifth grader has enough contextual knowledge of the crucial elements of a stop sign, such as its octagonal shape and red color, to discount minor changes such as the addition of specks. An AI agent lacks this context, unless a developer inputs data that will provide it. *See* Margulies, *supra* note 8, at 406; PAUL SCHARRE, ARMY OF NONE: AUTONOMOUS WEAPONS AND THE FUTURE OF WAR (2018) (discussing adversarial examples in which foes seek to frustrate AI agents by small changes to data that the agent absorbs).

247. Margulies, *supra* note 8, at 408–09; Buolamwini & Gebru, *supra* note 8, at 1.

248. *See* Strandburg, *supra* note 246, at 1877–78.

249. *See* Margulies, *supra* note 8, at 410; Rubenstein, *supra* note 8, at 778–79; Bambauer et al., *supra* note 244, at 2370–71.

250. RAI PATHWAY, *supra* note 30, at 11–12. AI is currently playing a useful role in this space. Experienced physicians often show wide swings in spotting and interpreting anomalies in test results that may indicate a need for treatment. *See* Kahneman et al., *supra* note 9, at 275–79. As a result, thousands of patients each year fail to receive treatment or receive treatment that is not appropriate. Providing more consistent analyses would benefit patients in the military and elsewhere.

social media.[251] This work has profound national security implications, although it is also distinct from battlespace use.[252] Finally, albeit in a limited way, DoD is conducting projects on using AI in battlespaces to find targets and reduce civilian harm.[253]

While more concrete moves toward the deployment of AI agents in the battlespace will require further assessment by scholars, NGOs, and other external stakeholders, commitment to benchmarking is evident in DoD's current AI policy documents. In CIS terms, the documents acknowledge risks, including brittleness, bias, and lack of explainability.[254] The Guidelines also require establishing processes

---

251. RAI Guidelines, *supra* note 243, at 13–15.

252. The U.S. government also uses AI to assist in guiding and analyzing overseas surveillance and intelligence collection. *See* Peter Margulies, *Surveillance by Algorithm: The NSA, Computerized Intelligence Collection, and Human Rights*, 68 FLA. L. REV. 1045, 1063–71 (2016) (explaining how AI can be used to guide overseas intelligence gathering); Emily Berman, *A Government of Laws and Not of Machines*, 98 B.U. L. REV. 1277, 1286–90 (2018) (explaining how AI can be used to ascribe predictive meaning to data already collected); *cf.* David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. Davis L. Rev. 653 (2017) (canvassing the field).

253. The civilian-harm reduction dimension of AI has promise in providing more information to targeting cells and checking hasty actions that cause needless civilian casualties. *See* CNA Report, *supra* note 8, at 20–50; Margulies, *supra* note 7. Using AI as autonomous weapons to find targets and use lethal force against them without specific prior human approval has engendered controversy and substantial literature on both the risks and possible benefits. *See* John Cherry & Durward Johnson, *Maintaining Command and Control (C2) of Lethal Autonomous Weapons System: Legal and Policy Considerations*, 27 SW. J. INT'L L. 1 (2021); Ashley Deeks, Noam Lubell & Daragh Murray, *Machine Learning, Artificial Intelligence, and the Use of Force by States*, 10 J. OF NAT'L SEC'Y L. AND POL'Y 1 (2019); Michael C. Horowitz, *The Ethics and Morality of Robotic Warfare: Assessing the Debate Over Autonomous Weapons*, 145 DEADALUS 25 (Fall 2016); Chris Jenks, *The Gathering Swarm: The Path to Increasingly Autonomous Weapons Systems*, 57 JURIMETRICS J. 341-59 (2017); Shane R. Reeves, Ronald T.P. Alcala & Amy McCarthy, *Challenges in Regulating Lethal Autonomous Weapons Under International Law*, 27 SW. U. J. INT'L L. 101 (2021); Michael N. Schmitt & Jeffrey S. Thurnher, *"Out of the Loop": Autonomous Weapons Systems and the Law of Armed Conflict*, 4 HARV. NAT'L SEC'Y J. 231 (2013); Alan Schuller, *At the Crossroads of Control: The Intersection of Artificial Intelligence in Autonomous Weapons Systems with International Humanitarian Law*, 8 HARV. NAT'L SEC'Y J. 379 (2017); *see also* Peter Margulies, *Making Autonomous Weapons Accountable: Command Responsibility for Computer-Guided Lethal Force in Armed Conflicts*, *in* RESEARCH HANDBOOK ON REMOTE WARFARE 405 (Jens David Ohlin ed., 2017) (discussing modification of command responsibility doctrine to promote accountability for mistakes of AI agents in armed conflict); Charles J. Dunlap Jr., *Accountability and Autonomous Weapons: Much Ado About Nothing?*, 30 TEMP. INT'L. & COMPAR. L. J. 63, 70–71 (2016) (arguing that current law and procedure can accommodate accountability for mistakes of autonomous weapons); Trumbull, *supra* note 96.

254. *See* RAI Guidelines, *supra* note 243, at 6 n.6 (citing Buolamwini & Gebru, *supra* note 8) (flagging bias); *Id.* at 26 (discussing brittleness); RAI PATHWAY, *supra* note 30, at 5 (mandating development of "transparent and auditable

to monitor harm from AI agents and classify harms that pose special concerns.[255]

The DOD RAI documents highlight the importance of reasoned explanations. The ability to assess options and consider a broad range of alternatives is central to the reasoned explanation prong. Illustrating commitment to this value, the Guidelines' drafters recognized that deployment was merely one option for an AI project. Pulling the plug on a project because of errors that were not susceptible to correction was an "acceptable outcome," given the cost of errors in the field.[256]

However, a deeper look at what the Guidelines do *not* say flags issues for the future on the reasoned explanation front. For example, the Guidelines were positive about the vendors with whom DoD officials had partnered on demonstration projects involving health-care and foreign-influence operations.[257] That positive assessment, including the accompanying discussion of the vendor's commitment to critical engagement, may be a testament to the sound path traveled by DoD officials in this fraught area.[258] However, the Guidelines could have done more to acknowledge that vendors have agendas of their own.[259] Furthermore, the Guidelines could have contained at least one

---

methodologies"). The Guidelines cite to a prominent critique of AI natural-language processing. *See* RAI Guidelines, *supra* note 243, at 15 n.11 (citing Emily M. Bender et al., *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?*, 81 PROC. MACH. LEARNING RSCH. 1 (2021)). The State Department has recently taken an important step forward regarding military use of AI that echoes the stress on accountability, review, and fairness in the DoD AI pathway materials. *See* U.S. Dep't of State, *Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy* (Feb. 16, 2023), https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy/#:~:text=Use%20of%20AI%20in%20armed,chain%20of%20command%20and%20control (discussing accountability and related values and noting importance of consultation between like-minded states).

255. RAI Guidelines, *supra* note 243, at 23 (discussing physical injury, psychological trauma, deprivation of civil liberties such as privacy, reduced access to services, and environmental impacts).

256. *Id.* at 17.

257. *Id.* at 13–16.

258. The vendor who collaborated on the foreign-influence project seemed especially attuned to the need to cast a critical eye on possible methodologies. *Id.* at 13–15, n.9 (citing Margaret Mitchell et al., *Model Cards for Model Reporting*, Report from the Conference on Fairness, Accountability, and Transparency (2019), https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf.) (discussing vendor's use of periodic review, solicitation of feedback from end-users, and use of approach that requires analysts to document AI agent's weaknesses as well as strengths).

259. *See generally* Bender et al., *supra* note 254 (discussing conflicts between public interest and commercial developers' bottom line); Margulies, *supra* note 194, at 410–11 (discussing problems with Boeing 737 Max systems that led to air crashes

critical observation about the vendors' performance in demonstration projects. A negative observation may have been jarring and may have spurred pushback from the developer. But sending a signal that implementation is imperfect would have been worth the trouble.

Similarly, regarding notice and comment, the RAI Guidelines include welcoming language, but may not adequately account for needed follow-through procedures. The Guidelines envision exchanges on broad principles with external stakeholders, including those in "industry, academia, and civil society."[260] However, the Guidelines contain no provisions for specific dialogue with external stakeholders on particular projects. The drafters mention open-source data only once.[261] More seriously, the Guidelines do not seriously consider the open-source method as a potential design alternative that will promote inclusion and wider perspectives.

The Guidelines' drafters acknowledge the trade-off between consultation with persons and entities affected and the need for secrecy in development and deployment.[262] Yet, the Guidelines' drafters devote only modest attention to discussing whether such trade-offs are inevitable. In some cases, according to the Guidelines, a "proxy" can provide feedback that will substitute for direct consultation with large groups whose participation could jeopardize secrecy.[263] In theory, a proxy could be a trusted group of academics or an NGO that will preserve operational secrecy while providing candid and comprehensive feedback.[264] Unfortunately, the AI Guidelines do not develop this proxy idea. That gap in provisions for external notice and comment is a flaw in the Guidelines' approach.

In sum, the DoD's RAI documents sound many of the themes that the benchmarking approach includes. However, the development of those themes is still uncertain. In this sense, both the RAI documents and the Political Declaration are works in progress.

---

and how U.S. agencies had failed to address these problems).

260. RAI Guidelines, *supra* note 243, at 4.

261. *Id.* at 13. This mention is irrelevant to dialogue with stakeholders; it describes open-source data as one source of content for the government's AI project on combating foreign government election misinformation.

262. *Id.* at 22.

263. *Id.*

264. This could be a role like that played by *amici curiae* in the U.S. Foreign Intelligence Surveillance Court (FISC), who provide a counter to government requests to maintain or expand surveillance programs. *See* Ira Rubinstein & Peter Margulies, *Risk and Rights in Transatlantic Data Transfers: EU Privacy Law, U.S. Surveillance, and the Search for Common Ground*, 54 CONN. L. Rev. 395, 403 (2022).

D. THE DOD CIVILIAN-HARM MITIGATION PLAN

The Defense Department's civilian-harm mitigation plan[265] is a fitting companion for the Responsible AI pathway and Political Declaration on Explosive Weapons. The civilian-harm plan has gaps and prompts additional questions. However, those gaps and questions are part of the dialogic process that the new plan initiated.

1. The Plan's Pillars

Consider DoD's establishment of a Center of Excellence on reducing civilian harm.[266] The Center should be a home for discerning and distributing lessons learned on civilian-harm reduction. The RAND report indicated that the U.S. military often approached potential lessons in a haphazard way that diminished their impact.[267] As an entity with a defined role, the Center should provide a more methodical approach. In this sense, the new Center answered the RAND report's call for a renewed emphasis on institutions.[268]

Structural innovation also figured in two other steps from the Defense Department plan. First, the Defense Department committed itself to establishing Civilian Environment Teams.[269] The Defense Department planned to place teams of experts in infrastructure and urban systems in combatant commands, such as Central Command, to provide information to targeting cells on issues such as second-order effects of attacks on sewers and water supplies.[270] The plan responded to concerns raised by RAND and others that the military had not paid sufficient attention to this issue.[271]

In addition, the new plan contemplated the formation of so-called "red teams" that would provide an opposing voice on targeting scenarios and assumptions.[272] The red teams will "explore ... alternatives" to a particular attack.[273] Those alternatives could include

---

265. DoD Civilian-Harm Mitigation Plan, *supra* note 5.

266. *Id.* at 6.

267. *See* RAND Report, *supra* note 11, at 57-61.

268. *Id.*

269. DoD Civilian-Harm Mitigation Plan, *supra* note 5, at 9.

270. *Id.*

271. *Id.* at 12-13.

272. *Id.* at 15; *see also* Oona Hathaway, *National Security Lawyering in the Post-War Era: Can Law Constrain Power?*, 68 UCLA L. Rev. 2, 82 (2021) (discussing possible use of red teams to argue against executive branch use of legal strategies that may exceed executive authority).

273. DoD Civilian-Harm Mitigation Plan, *supra* note 5, at 15.

consulting with senior officials who can free up additional resources or devising an attack with different weapons that will pose less serious risks to civilians.[274] Each of these steps will defuse the cognitive bias that has plagued U.S. targeting efforts.[275] These ideas may stem in part from analysis by Professor Geoffrey Corn and Mike Meier, a longtime DoD lawyer, on reconfiguring targeting teams to include civilian-risk mitigation officers.[276]

The Defense Department plan also highlighted the promise of technology. It called for new uses of AI that can develop information about civilian patterns of life.[277] Artificial intelligence and machine learning can have uses besides going on offense against adversaries, as the DoD's Responsible AI pathway suggested in June, 2022 and the Center for Naval Analysis report co-authored by civilian-harm reduction expert Larry Lewis noted earlier.[278] AI's most important use can be in minimizing risks to civilians.[279]

Finally, in yet another structural innovation, the Defense Department plan included provision for Civilian Harm Assessment and Investigation Coordinators who will ensure a uniform approach to investigations.[280] The *Times* reports and RAND noted that investigations were often spotty and relied on a narrow range of information, sometimes not even including information from the military's own sources, let alone the data obtained by civil society groups.[281] A plan for careful and consistent coordination of investigations will build public trust and facilitate the dissemination of lessons learned.[282]

2. Benchmarking the Plan

Analyzed through the prism of benchmarking's factors, the Defense Department plan was a welcome initiative, although it also prompted questions. Consider how the plan fared under the civilian impact assessment factor. On the one hand, the plan urged a "robust understanding" of the civilian environment and the impact military

274. *See* Corn, *supra* note 17, at 459.
275. DoD Civilian-Harm Mitigation Plan, *supra* note 5, at 15.
276. *See* Corn & Meier, *supra* note 202.
277. DoD Civilian-Harm Mitigation Plan, *supra* note 5, at 14.
278. *See* CNA Report, *supra* note 8, at 28-48.
279. *See* Margulies, *supra* note 7.
280. DoD Civilian-Harm Mitigation Plan, *supra* note 5, at 20.
281. *See* RAND Report, *supra* note 11, at 22-24.
282. DoD Civilian-Harm Mitigation Plan, *supra* note 5, at 22-24.

operations would cause.[283] That recognition, which the new plan endeavored to integrate into operational planning, augured well for a more deliberate approach. On the other hand, certain features of this integration remained nebulous. The section on red teams was particularly vague, with little guidance on the number of red teams and their precise operational role. The new plan stipulated that red teams would be an "ad-hoc element at tactical organizations . . . as needed."[284] The stated "ad hoc" and "as-needed" qualifiers of red teams' role raised questions about the role of this entity. Attack planners could request a red team to probe significant doubts about a positive target-identification or the degree of expected collateral harm. However, confirmation bias is a persistent problem precisely because it obscures such doubts. A more robust and comprehensive trigger for red teams' deployment would address these concerns.

While the deliberative aspects of the Defense Department plan were a positive sign for the promise of reasoned explanations, this component also suffered from gaps. For example, the plan did not address past overreliance on unit self-defense. Since that particular procedural pretext led to excesses in targeting,[285] focus on the issue would have been appropriate. More broadly, while the plan offers tacit admissions about the influence of confirmation bias and other factors on targeting, both internal and external stakeholders could have benefited from more specific acknowledgment of past mistakes. Guidance from legal rules is often most effective when it includes examples. Even at this late date, not everyone knows the meaning of confirmation bias. Nor can everyone diagnose the problem when they encounter it. Confirmation bias is insidious precisely because it seems to spring from "natural" and easy turns of thought. Training the mind to recognize the perils of those turns requires examples. The Defense Department plan would have been a good place to start.

In addition, acknowledging past mistakes shows that an entity is fully invested in reform. Institutional change is a formidable task and backsliding is a constant risk. Acknowledging past mistakes supports internal reformers and strips the backsliders of cover.

The Defense Department plan also raised questions about notice and comment. The plan was insufficiently specific about interactions with external stakeholders. The content and tone of the document suggested new transparency and engagement. But a roadmap for future interactions would have been welcome. The U.S. military will

283. *Id.* at 9.

284. *Id.* at 15.

285. *See supra* notes 108-112 and accompanying text; Corn, *supra* note 10.

have to balance the benefits of transparency against the importance of preserving sources and methods. A roadmap would have illustrated that the military is weighing each of these values carefully, without putting a thumb on the scale.

That said, the commitment to deliberation in the DoD plan mirrored other initiatives. The DoD "Responsible AI" pathway asked tough questions about the need for certain AI products and their effect on civil rights and liberties. The multilateral Political Declaration on Explosive Weapons in Populated Areas committed the United States to dialogue with stakeholders on reducing civilian risk.[286] The reflective turn in these measures and the new Defense Department civilian-harm mitigation plan may transform U.S. military practice.[287]

V. ALTERNATIVES TO THE BENCHMARKING APPROACH

This section considers two alternatives to benchmarking: the tort and retrospective approaches. Each approach has merit. However, each approach, if taken to extremes, would stifle innovation and intrude unduly on the discretion that attack planners need in the interests of military necessity.

A. THE TORT APPROACH

The tort approach adapts tort concepts like reasonable care and strict liability to the IHL context. Indeed, there is actually more than one tort approach. As we shall see, the reasonable care approach taken by Professor Asaf Lubin incorporates some of the concepts advanced here.[288] In contrast, the strict liability approach advanced by Professor Rebecca Crootof would modify IHL in ways that states are unlikely to accept.[289]

The reasonable care approach takes as its central text the same "constant care" requirement that both customary and treaty law

286. *See* Meier, *supra* note 32.

287. DoD Civilian-Harm Mitigation Plan, *supra* note 5, at 6.

288. *See* Lubin, *supra* note 7.

289. *See* Crootof, *supra* note 173; *see also* Rebecca Crootof, *War Torts: Accountability for Autonomous Weapons*, 164 U. PA. L. REV. 1347, 1394–98 (2016) (arguing for strict liability for wrongs committed by autonomous weapons). My critique of Professor Crootof's proposal should not obscure the insightful work that Professor Crootof has done on the interaction of humans and AI models. *See, e.g.*, Rebecca Crootof, Margot E. Kaminski & W. Nicholson Price II, *Humans in the Loop*, 76 Vand. L. Rev. 429 (2023) (discussing benefits and pitfalls of human interaction with AI decisions).

upholds. As with the approach taken here, the reasonable care approach seeks to upgrade best practices on a systemic basis. Moreover, Professor Lubin discusses the relevance of administrative law principles to systemic reform,[290] although he does not apply administrative law practices and doctrines, including the requirement of an impact statement; a reasonable explanation modeled after the Supreme Court's *State Farm* decision; and notice to and comment from internal and external stakeholders. Nevertheless, the reasonable care model's stress on "deliberative and participatory processes" meshes with benchmarking's emphasis on consultation.[291]

The more problematic tort model imposes strict liability on states for any collateral damage to civilians or civilian objects, even when there was no individual or systemic fault at work.[292] This approach is an audacious and original reframing of accountability under IHL. It places civilian harm front and center. Moreover, it prioritizes compensation for civilians in a formal mechanism that replaces the informal regime of ad hoc *ex gratia* payments to injured parties or survivors that have long dominated the armed conflict space.[293] That said, strict liability is a powerful medicine for IHL's ills, which might end up causing greater ills than it remedies.

A strict liability regime would upset IHL's careful balance between humanity and military necessity.[294] In the fog of war, even the most effective systemic and individual safeguards, rules, and principles will not prevent all collateral damage. The need to wage a war effectively makes collateral damage inevitable, even with all reasonable safeguards in place.[295] Imposing legal liability on states in such situations creates uncertainty for planners. Imagine a diligent

290. Lubin, *supra* note 7, at 146.

291. *Id.* Lubin uses the term, "benchmarks," although he refers to benchmarking as a consequence of his approach, rather than as a centerpiece. *See id.* at 148 (explaining that "treating intelligence production as a trade with a set of well-defined industry standards" allows observers to rank actors and policies and hence "[o]ver time . . . will result in the creation of actual benchmarks against which we may be able to continuously assess new breaches").

292. *See* Crootof, *supra* note 173; Crootof, *Accountability for Autonomous Weapons*, *supra* note 289, at 1394–98.

293. Readers will note that this Article has not addressed issues of compensation, which merit full discussion on their own.

294. *See* Schmitt, *supra* note 36.

295. *Cf.* Torres v. Dep't of Pub. Safety, 142 S. Ct. 2455, 2465 (2022) (quoting Charles Evan Hughes, War Powers Under the Constitution (Sept. 5, 1917)) (noting, in construing scope of U.S. Congress's war powers and authority to bar state discrimination against military veterans, that "the power to wage war is the power to wage war successfully"); Matthew C. Waxman, *The Power to Wage War Successfully*, 117 COLUM. L. REV. 613, 618–58 (2017) (discussing context of Hughes's speech and its influence).

government lawyer seeking to provide advice to an attack planner on the legality of a strike. The planner may react with understandable confusion if the lawyer explains that an attack would trigger liability even if, 1) the planner's chain of command had studiously followed the benchmarking model and methodically reduced the risk of civilian harm, and, 2) the individual attack that the planner proposed was consistent with the principle of distinction and the rules of proportionality and precautions in attack.

An attack planner hearing such advice might be puzzled that no course of action— apart from not conducting the attack— would avoid liability. The planner's puzzlement would increase if the attack was necessary to relieve other state forces at risk, preclude an adversary from gaining additional territory, or liberate civilians whom an adversary had wrongfully held captive. The planner might also ponder whether the lawyer's counsel was worth seeking in the future.

These consequences do not serve IHL compliance. Yet the strict liability view does not provide sufficient guidance on how to avoid such adverse effects. In contrast, the reasonable care and benchmarking models supplement current individual IHL rules but do not propose to replace them.

B. THE RETROSPECTIVE APPROACH

Two bills sponsored by Senator Elizabeth Warren and other senators— the Protection of Civilians in Military Operations Act (POCIMO)[296] and the Department of Defense Civilian Harm Transparency Act (CHTA)[297] would adopt some of the same practices advocated here involving periodic assessment of civilian-harm reduction policies. However, the bills fail to highlight the role of technology in reducing civilian casualties, thereby failing to provide needed guidance. Moreover, each of the proposed bills would entail an unduly intrusive retrospective examination of the criteria for all U.S. strikes. While some review of past incidents is necessary, as this Article has urged, the intrusive review that this proposed legislation contemplates would reveal sources and methods and divert resources and bureaucratic attention from present-day innovations.

Proposed legislation such as POCIMO has much merit in setting

296. *See* To enhance protections of civilians during United States military operations, and for other purposes, S. 4108, 117th Cong. (2022); *see also* Shiel & Yager, *supra* note 199 (analyzing POCIMO).

297. *See* Dickinson et al., *supra* note 199 (discussing CHTA).

up mechanisms for auditing.[298] These measures dovetail with the benchmarking approach. However, other aspects of this proposed legislation entail crucial omissions. Neither POCIMO nor the CHTA highlights the importance of developing technology that will reduce civilian harm. For example, POCIMO does not mention the importance of benchmarks in disseminating available technology to comply with the "constant care" obligation. The legislation, to provide clearer guidance to the Department of Defense, should flag widely available technology, such as using enhanced video capabilities and AI agents to detect additional civilians at the scene of a proposed strike.

At the same time, POCIMO's focus in Article 8 on targeting criteria would require unwise and inappropriate disclosure of information. Article 8 includes a requirement that the Defense Department fund an independent inquiry into how service members conducting targeting have "differentiated between combatants and civilians in both ground and air operations since 2001" in all regions in which the United States has operated, including Afghanistan, Iraq, Syria, Somali, Libya, and Yemen.[299] The proposed legislation includes similar requirements for reviewing past civilian casualty investigations.[300] The breadth of this required review could reveal intelligence sources and methods, as well as details of targeting criteria.

As the European Court of Human Rights recently noted in *Big Brother Watch v. United Kingdom* (*Big Brother Watch II*),[301] undue specificity in the disclosure of targeting criteria can hamper ongoing operations.[302] Overly detailed descriptions of factors triggering government responses can allow a determined adversary to adapt its conduct to frustrate detection. That deficit undermines one of the two pillars of IHL, military necessity.

The POCIMO's focus in Section 8, along with the CHTA's similar provisions, provides too much information. Consider the POCIMO's requirement that independent review determine "whether military-

298. *See* To enhance protections of civilians during United States military operations, and for other purposes, S. 4108, 117th Cong. §§ 5–7 (2022) (providing for auditing strikes and establishing Center of Excellence on civilian-harm reduction).

299. *Id.* § 8(b)(1).

300. *Id.* § 8(b)(2).

301. Big Brother Watch v. United Kingdom, App. Nos. 58170/13, 62322/14 & 24960/15, ¶ 420 (May 25, 2021), https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-210077%22]}.

302. *Id.* ¶ 353; *see also* Rubinstein & Margulies, *supra* note 264, at 416–17 (discussing *Big Brother Watch*); *see generally* Ashley Deeks, *An International Legal Framework for Surveillance*, 55 VA. J. INT'L L. 291 (2015) (outlining principles to accommodate foreign surveillance within human rights law).

aged males were presumptively targetable."[303] Suppose that the independent review found and disclosed that military-age males who carried mortars in addition to small arms met the criteria for targeting. Groups of insurgents would take care to conceal mortars, impeding detection of entirely lawful targets. In this sense, both the POCIMO and the CHTA pair welcome requirements for deliberation with excess transparency mandates that would impair the effectiveness of lawful and necessary military action. In addition, POCIMO would require a massive investigation of *all* past U.S. strikes, which would divert resources, personnel, and bandwidth from the task of reform.

CONCLUSION

Civilian harm is inevitable in armed conflicts, but its current level reflects active and tacit choices by states. The law of armed conflict (LOAC), with its balancing of humanity and military necessity, has reduced civilian harm. However, LOAC's focus on individual attack planners tolerates civilian harm that is avoidable.

While the U.S. experience since 9/11 has included ebbs and flows in civilian harm, analysis of that experience points to several causal factors. Access to available technology, such as high-resolution video, is hindered by supply-chain snarls. Cognitive flaws, including confirmation bias and base-rate neglect, can distort analysis of possible targets. Overzealous targeting cells exaggerate the case for unit self-defense. Training is inadequate and institutional priorities are elsewhere, although the U.S. Defense Department's new Center of Excellence and forthcoming guidance may herald a heightened commitment.

Citing LOAC's mandate that states use "constant care" in sparing civilians, this Article proposes a systemic approach. The requirement of methodical, systemic efforts is *lex lata*—binding on states. It springs from the language and logic of the constant-care duty, and from states' obligation to comply with LOAC in good faith. Additional support comes from the International Court of Justice's *Nuclear Weapons* decision and the analysis of the U.N. Human Rights Committee, which view LOAC through the prism of human rights law's prohibition on arbitrary deprivations of life. Arbitrariness would undermine LOAC if human errors in supply-chain management, cognition, procedure, and training ratcheted up civilian harm. A systemic effort to address the

303. To enhance protections of civilians during United States military operations, and for other purposes, S. 4108, 117th Cong. § 8(b)(1)(C) (2022).

problem reinforces states' good faith and keeps arbitrary deprivations at bay.

Under this Article's account, states have flexibility in implementing this systemic duty. As a best practice that is *lex ferenda*, this Article suggests a benchmarking approach that echoes administrative law's three-pronged focus on impact assessments, reasoned explanation, and notice and comment. These principles provide a basis for evaluating recent U.S. policies, including the respective drone frameworks of presidents Obama and Trump, the 2022 Political Declaration on explosive weapons, the Responsible AI policies of the Defense Department, and the Defense Department's Civilian Harm Mitigation Response Action Plan. Applying the benchmarking model to these policies suggests a positive trend, although questions remain.

No single model will eradicate avoidable civilian harm in armed conflict. The exigency and confusion that surround this setting will always entail some risk. However, under LOAC, states have a systemic obligation to do better. The benchmarking approach encourages states to adopt habits of deliberation and accountability that will maximize those systemic efforts.