
INTERNATIONAL LAW STUDIES

PUBLISHED SINCE 1895

U.S. NAVAL WAR COLLEGE



Mind the Gap: Can Developers of Autonomous Weapons Systems be Liable for War Crimes?

Tim McFarland and Tim McCormack

90 INT'L L. STUD. 361 (2014)

Volume 90

2014

Mind the Gap: Can Developers of Autonomous Weapons Systems be Liable for War Crimes?

*Tim McFarland**

*Tim McCormack***

I. ACCOUNTABILITY CONCERNS IN RELATION TO AUTONOMOUS WEAPONS SYSTEMS

Few aspects of the emergence of autonomous weapons systems engender divergence of opinion as dramatically as questions of

* Tim McFarland BE (Hons), BEc, JD, is a PhD candidate at the Melbourne Law School and a member of the Research Team on Emerging Technologies of Warfare and Challenges to the Law of Armed Conflict at the Asia Pacific Centre for Military Law. His PhD is focused on the legal implications of developing and deploying autonomous weapons systems.

** Tim McCormack is a Professor of Law at the Melbourne Law School and an Adjunct Professor of Law at the University of Tasmania Law School. He is also Special Adviser on International Humanitarian Law to the Prosecutor of the International Criminal Court, although this article is written in his personal capacity and should not be read as reflective of the position of the Prosecutor. Professor McCormack leads the Research Team on the Australian Research Council-funded project Emerging Technologies of Warfare and Challenges to the Law of Armed Conflict.

The authors wish to thank Professor Marco Sassóli for drawing our attention to the specific issue of the actions of developers occurring before the commencement of an armed conflict. We also thank Group Captain Ian Henderson, Associate Professor Robert Mathews, Dr. Rain Liivoja, Ms. Kobi Leins and Ms. Natalia Jevglevskaja for insightful comments on earlier drafts and Ms. Madeleine Summers for her helpful research assistance.

accountability for violations of the law of armed conflict. One extreme is simplistically summarized as “the sky is falling.” Technological advances represent an epochal challenge to the existing legal framework because human decision making, and the concomitant responsibility which ensues, has never previously been removed or distanced from control of the weapons system in quite the same way as is now contemplated with autonomous weapons systems. The decision to deploy an autonomous weapons system is now not predicated on the ongoing exercise of human oversight over the system; control over the behavior of the weapons system in military operations will have been exercised much earlier in the programming of the system.¹

The other extreme is that “there is nothing new under the sun.” Machines of widely varying levels of technological sophistication, many of them far too complex for a non-specialist to understand in detail, have been wielded in combat for millennia and rarely has there been much difficulty in applying existing principles of law to their use. The law is adaptable and capable of regulation of the current and foreseeable future generations of autonomous weapons systems. There may well be new aspects here that will require particular amendment, revision and adaptation, but there is no wholesale existential threat to the existing law of armed conflict.

Intriguingly, even these two extremes (and presumably variations of them along a spectrum) share substantial common ground in acknowledging the fundamental importance of accountability to the efficacy of the law of armed conflict. That commonality is important and is the starting point for our analysis. In the development of new weapons technology, lawyers and policymakers are obliged to ensure that compliance with the law will be maintained² and that those who seriously violate the law are held accountable.³

1. *See, e.g.*, HUMAN RIGHTS WATCH, LOSING HUMANITY: THE CASE AGAINST KILLER ROBOTS (2012), available at <http://www.hrw.org/reports/2012/11/19/losing-humanity>.

2. Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts art. 36(2), June 8, 1977, 1125 U.N.T.S. 3 [hereinafter Additional Protocol I].

3. Convention (I) for the Amelioration of the Condition of the Wounded and Sick in the Armed Forces in the Field art. 49, Aug. 12, 1949, 75 U.N.T.S. 31 [hereinafter GC I]; Convention (II) for the Amelioration of the Condition of the Wounded, Sick, and Shipwrecked Members of Armed Forces at Sea art. 50, Aug. 12, 1949, 75 U.N.T.S. 85 [hereinafter GC II]; Convention (III) Relative to the Treatment of Prisoners of War art. 129, Aug. 12, 1949, 75 U.N.T.S. 135 [hereinafter GC III]; Convention (IV) Relative to the

Our intention in this article is not to attempt an exhaustive and definitive analysis of accountability issues arising from war crimes perpetrated via autonomous weapons systems. Instead, our hope here is to provoke more detailed analysis and discussion. Our aim is to explore some of the complexities in the application of the existing accountability framework likely to result from the deployment of increasingly advanced autonomous weapons systems.

Here we focus on the particular issue of accountability for developers of autonomous weapon systems. Development of any modern weapon is a complex undertaking; large teams of people and organizations are involved at all stages of the process, from creation of the concept, through the various stages of development (with extensive interaction between military decision makers providing specifications, producers responding to those specifications, and those responsible for testing and approval for production) to production and ultimately deployment. “Developers,” for the purposes of this article, refer broadly to people who play some significant role in defining the behavior of an autonomous weapons system, as opposed to “operators,” which refer to those responsible for utilizing the system in some situation during armed conflict. We are not suggesting here that operators cannot be accountable for deploying or activating an autonomous weapons system, including for actions such as inputting certain data (for example, those targets currently approved for attack) or for employing the system outside of its designed operating parameters. Our focus here, though, is on the specific issue of the accountability of the developers of autonomous weapons systems because accountability concerns so often focus on this particular category of involved personnel.

Effective implementation of individual criminal liability for violations of the law of armed conflict is fundamental to ensuring respect for the law. In the absence of effective enforcement, impunity flourishes and the normative status of the law is undermined. The key concept was first explicitly articulated in the grave breaches regime common to all four Geneva Conventions of 1949.⁴ States parties to the Conventions are obliged to criminalize the specific breaches identified as “grave” and to investigate and either prosecute or extradite for trial those allegedly responsible for perpetration of any of the grave breaches.

Protection of Civilian Persons in Time of War art. 146, Aug. 12, 1949, 75 U.N.T.S. 287 [hereinafter GC IV]; Additional Protocol I, *supra* note 2, art 85(1).

4. GC I, art. 50; GC II, art. 51; GC III, art. 130; GC IV, art. 147, all *supra* note 3.

Explicit obligations for criminalization of serious violations of the law of armed conflict were novel in 1949, although the concept of individual criminal responsibility for serious violations was not. In the aftermath of the First World War, the Allied Commission on Responsibility of the Authors of the War and on Enforcement of Penalties enumerated a list of offenses for further investigation and subsequent prosecution despite the lack of explicit criminalization of serious violations of the law of armed conflict in any of the relevant treaties.⁵ The proposed Allied High Tribunals for the prosecution of defendants from the defeated Central Powers failed to eventuate, but not because of any principled objection to the notion of individual criminal responsibility.⁶ The Geneva Conventions of 1949, in contrast, were negotiated in the immediate aftermath of the conclusion of the Nuremberg (1945–46) and Tokyo (1946–48) trials and it is hardly surprising that the breakthrough grave breaches regime reflects the post-Second World War commitment to individual criminal responsibility—at least for defeated Axis Power accused.

Additional Protocol I takes the notion of individual responsibility for violations of the law a step further by explicitly identifying the pre-eminence of military commanders in ensuring compliance with the law. Commanders who order their subordinates to perpetrate atrocities in violation of the law are, of course, liable for the ensuing crimes.⁷ But military commanders who fail to exercise their authority to stop crimes occurring, or who fail to punish their responsible subordinates if crimes have already occurred, are also liable for such crimes. Articles 86 and 87 of Additional Protocol I explicitly impose an international legal obligation on parties to an armed conflict to require commanders to take all necessary measures to prevent crimes being committed and to initiate disciplinary or penal sanctions when crimes have been perpetrated.⁸ These provisions are

5. Commission on Responsibility of the Authors of the War and on Enforcement of Penalties, *Report Presented to the Preliminary Peace Conference* (Mar. 29, 1919), reprinted in 14 AMERICAN JOURNAL OF INTERNATIONAL LAW 95, 114–15 (1920).

6. For more detail on the proposed Allied High Tribunals, see GARY J. BASS, *STAY THE HAND OF VENGEANCE: THE POLITICS OF WAR CRIMES TRIBUNALS* (2000); Tim McCormack, *Their Atrocities and Our Misdemeanours: The Reluctance to Try a State's 'Own Nationals' for International Crimes*, in JUSTICE FOR CRIMES AGAINST HUMANITY 107 (Mark Lattimer & Philippe Sands eds., 2003).

7. Rome Statute of the International Criminal Court art. 25(3)(b), July 17, 1998, 2187 U.N.T.S. 90, recognizes this mode of criminal liability.

8. Additional Protocol I, *supra* note 2, arts. 86, 87.

recognized as customary international law⁹ and are reflected, for example, in Article 28(a) of the Rome Statute.

The ICRC *Commentary* to Additional Protocol I articulates the rationale for the imposition of an explicit international legal duty on military commanders:

In fact the role of commanders is decisive. Whether they are concerned with the theatre of military operations, occupied territories or places of internment, the necessary measures for the proper application of the Conventions and the Protocol must be taken at the level of the troops, *so that a fatal gap between the undertakings entered into by Parties to the conflict and the conduct of individuals is avoided*. At this level, everything depends on commanders, and without their conscientious supervision, general legal requirements are unlikely to be effective.¹⁰

The drafters of the Protocol recognized that commanders, through the control they exercise over subordinates, have the means to ensure adherence to the law by maintaining discipline, being informed about the conduct of their forces and taking action to prevent or punish breaches.¹¹ So important is the role of the commander in the law of armed conflict, that the commentaries describe the ever-present potential for a “fatal gap” between international legal obligations on one hand and the perpetration of egregious violations of the law on the other. Military commanders have the ability to stop atrocities occurring, or at least to sanction those subordinates who commit them in order to deter others from violations of the law in the future. Those commanders who choose not to act, or choose not to utilize the authority they have to intervene and to prevent crimes from being committed, are to be held criminally responsible for the crimes they could have prevented and/or should have punished.

But all of this—individual criminal responsibility for those who violate the law and for those commanders who fail to exercise the authority they possess to prevent or to punish their subordinates responsible for violations of the law—is predicated on the notion, immortalized in the words of the Nuremberg judges, that “crimes . . . are committed by men,

9. 1 CUSTOMARY INTERNATIONAL HUMANITARIAN LAW r. 153, 558–63 (Jean-Marie Henckaerts & Louise Doswald-Beck eds., 2005).

10. COMMENTARY ON THE ADDITIONAL PROTOCOLS OF 8 JUNE 1977 TO THE GENEVA CONVENTIONS OF 12 AUGUST 1949, ¶ 3550 (Yves Sandoz, Christophe Swinarski & Bruno Zimmermann eds., 1987) (emphasis added) (footnotes omitted).

11. *Id.*, ¶ 3560.

not by abstract entities.”¹² In 1946 when the Nuremberg judgment was handed down, the abstract entity in question was the State—specifically the Third Reich—and the judges were making the point that collective policy decisions and their implementation are reducible to the acts and intent/knowledge of the individual human beings that make up the abstract collective entity. Enforcement of the law of armed conflict is dependent upon identification of those responsible for the serious violations that have occurred.

We are not suggesting that there is anything abstract about a robot, but the judges’ words apply equally well: a weapons system, regardless of its level of sophistication, is an inanimate object; its existence and its capacity to cause harm result from the efforts of weapon developers, just as with a rifle or a missile. Unlike developers of rifles and missiles though, developers of highly autonomous weapons systems will play a role arguably not entirely consistent with the current legal framework. As explained in more detail in Section II, a highly autonomous weapons system has the potential to partly or fully displace combat personnel from the roles they have traditionally occupied, so that accountability for proscribed acts committed via such systems may not be easily ascribed to those personnel or to their commanders. Rather, developers will exert greater control over not only the range of actions the weapons system is capable of performing, but over the specific actions that it in fact performs after being deployed. Accordingly, the question of developer accountability becomes pertinent.

Before considering the legal significance of machine autonomy, it is perhaps worth revisiting the early observation about divergence of opinion in relation to accountability for serious violations of the law perpetrated via autonomous weapons systems. That divergence may be at least partially explicable in terms of interpretation versus application of existing law. Proponents of the “nothing new under the sun” perspective would presumably claim that since technological advances have not resulted in a level of autonomy that replaces human control, the existing legal framework is just as adequate as it is for any existing weapons system. In contrast, proponents of “the sky is falling” perspective might argue that even with existing relatively rudimentary levels of automation (including remote-controlled unmanned combat aerial vehicles (UCAVs), for example) the current legal framework is not effectively applied, and if

12. 1 TRIAL OF THE MAJOR WAR CRIMINALS BEFORE THE INTERNATIONAL MILITARY TRIBUNAL 223 (1947).

accountability is lacking for such systems, how much less effective will it be for significantly more sophisticated levels of autonomy.

Some authors have approached this topic by suggesting that where a machine operates with a degree of independence from human oversight sufficient to raise serious questions about accountability, the machine could, perhaps, be granted legal personality and could itself be held accountable for violations of the law.¹³ This article takes the view that the level of technological and social development needed to justify talking about the legal personality of machines is far in the future and is not yet—and may never be—a practical consideration. That does not mean that identifying the potential future issue and discussing the legal ramifications is not a worthwhile endeavor, but this article is not an appropriate forum for engaging with that particular question.

II. THE ROLE OF DEVELOPERS IN RELATION TO AUTONOMOUS WEAPONS SYSTEMS

We understand autonomous weapons systems to be those with some significant capacity to manage their own operation. Specific configurations of such systems vary widely, and are likely to vary much more widely in the future, but for legal analytical purposes all such systems can be viewed as being comprised of a few high-level subsystems. The weapon itself, such as a gun or a missile, is only one of those subsystems and is not necessarily of central interest when addressing questions of accountability as the novel issues relate more to how the weapon is controlled. Similarly, one or more sensor subsystems would normally be employed, such as cameras or radar, to obtain environmental information. Of primary interest though is the autonomous weapon's control system, essentially a special-purpose computer, which manages the weapon, sensors and other subsystems. The control system's task is to receive commands or goals from an external entity (human, or perhaps another autonomous system), as well as environmental information via sensor systems, and then operate the weapon accordingly. At the heart of the control system, and of novel questions about legal accountability, is software written by weapon developers.

13. See, e.g., John Sullins, *When is a Robot a Moral Agent?*, 6 INTERNATIONAL JOURNAL OF INFORMATION ETHICS 23 (2006).

The capacity of such software to manage the operation of a weapons system without waiting for an explicit input from a human operator at every step is the basis of many of the anticipated operational advantages driving development of these advanced systems.¹⁴ The speed and accuracy of computation in modern computer systems will enable autonomous weapons to function in situations where a human operator is incapable of making timely decisions, such as where the tempo of battle is too fast¹⁵ or the amount of data to be processed is too great. Enabling vehicles and weapon installations to function without crews greatly reduces the potential for casualties among the armed force operating them. Significant cost savings can be realized by reducing the number of human soldiers to be paid, transported and for which care must be provided.

The capacity for self-managed operation is also the basis of the novel legal effects of deploying autonomous weapons systems; in particular, those effects relating to accountability for violations of the law of armed conflict. Essentially, autonomous capability in a weapons system alters the relationship between the soldier, the traditional bearer of legal responsibilities, and their weapon, the means by which those responsibilities are fulfilled. If, for example, a soldier aims a “dumb” weapon such as a rifle at a civilian and pulls the trigger, it is not contentious to say that the soldier performs that proscribed act. A separate, specific decision and action by the soldier is required for each separate activation of the rifle and there is no intervening event which may interfere with identification of that soldier as perpetrator of that action. When the rifle is replaced with an autonomous weapons system though, the weapon’s control system “steps into the shoes” of an operator to some significant extent, reducing the need—and opportunity—for human intervention and placing the weapon developer somewhat “between” the soldier and activation of the weapon. This is the essential legal distinction between systems with a significant capacity for autonomous operation and other complex weapons systems: autonomy extends to instigation of an action, not just performance. To some extent, an autonomous machine replicates

14. PETER W. SINGER, WIRED FOR WAR 126–34, 229–36 (2009); Robert Sparrow, *Robotic Weapons and the Future of War*, in NEW WARS AND NEW SOLDIERS: MILITARY ETHICS IN THE CONTEMPORARY WORLD 117 (Jessica Wolfendale & Paolo Tripodi eds., 2011); Gary E. Marchant et al., *International Governance of Autonomous Military Robots*, 12 COLUMBIA SCIENCE AND TECHNOLOGY LAW REVIEW 272, 275 (2011).

15. Robert Sparrow, *Killer Robots*, 24 JOURNAL OF APPLIED PHILOSOPHY 62, 68 (2007).

the mental (decision-making) processes leading to an act rather than just the physical processes involved in its commission.

That is not to say that autonomous weapons should be treated as systems that will operate without any real-time human interaction or oversight. The degree to which human actors may be removed from positions of direct control is likely to vary in complex ways. Machine autonomy is a capability that exists within a continuum rather than at discrete levels, ranging from complete human control over some operations to complete computer control. Military and civilian research organizations have proposed many different taxonomies of autonomous capability.¹⁶ Relevantly, these taxonomies invariably anticipate some degree of shared control, where either a human or the computer may initiate a task and the human operator may or may not be in a position to approve or override an operation initiated by the computer. Further, the capacity for autonomous operation will generally be specific to particular subsystems of a weapons system, such as navigation (in the case of a vehicle), target identification and weapon release. There is no reason to suppose that all operations of an autonomous system will be subject to the same degree of human oversight, and a system may be operating at several different levels of autonomy simultaneously with respect to different tasks. For example, a weapons system may autonomously identify and track potential targets but require a human operator to manually operate the weapon. This consideration also applies to a “system of systems” scenario wherein nominally separate systems with different levels of autonomy, such as an intelligence, surveillance and reconnaissance (ISR) system and a weapons system, communicate directly. Finally, the degree of direct human control over a weapons system may vary with mission phases and according to unexpected circumstances that arise during a mission, with either the system or the operator instigating the revision.¹⁷

Despite the complexities of determining whether and to what extent a human operator may be in a position to exercise control over a particular action of an autonomous system, it is envisaged that generally the higher the level of autonomous operation exhibited by a weapons system, the less

16. Peter A. Hancock & Stephen F. Scallen, *Allocating Functions in Human–Machine Systems*, in *VIEWING PSYCHOLOGY AS A WHOLE: THE INTEGRATIVE SCIENCE OF WILLIAM N. DEMBER* 509, 521 (Robert R. Hoffman, Michael F. Sherrick & Joel S. Warm eds., 1998).

17. DEFENSE SCIENCE BOARD, U.S. DEPARTMENT OF DEFENSE, TASK FORCE REPORT: THE ROLE OF AUTONOMY IN DoD SYSTEMS 27 (July 2012).

control will be exercised by a human operator in the field and the more control will be exercised by developers of the control software (that has reduced the level of control over the weapons system exercised by that human operator).

If the level of autonomous operation can vary from complete human control through various degrees and types of shared control to complete computer control, it is necessary to ask: in what circumstances, if any, does a system's inbuilt capacity for autonomous operation start to affect the allocation of criminal responsibility for serious violations of the law of armed conflict? It is suggested that there are two indications that such a point has been reached. The first is that the system's autonomous operation relates to a decision and action that is the subject of legal regulation. The canonical example of this would be discharging a weapon, but other acts such as crossing an international border may also be significant. The second indication is that the degree of control over the regulated action exercised by the system itself—and by extension by its developers—has increased to the point where it is not reasonable to say that a human operator alone exercises effective control; that is, the individual who would normally decide to “pull the trigger” is no longer solely responsible for that decision. At that point, some part of the effective human control over the behavior of the weapons system is built-in to the system during development rather than after deployment or at the time of the actual use of force. It would generally be the case that such development would occur prior to commencement of an armed conflict in which the weapons system is deployed.

The next question then is whether these circumstances have yet arisen, or are within the scope of development plans that have so far been made public. Weapons systems in use today generally exhibit capabilities for autonomous operation far below what would be required to raise serious questions about accountability, although many of the component technologies which may one day comprise a highly autonomous weapons system either exist already or are under development. UCAVs¹⁸ employed in strikes in the Middle East and Africa may have some capacity to autonomously perform low-level piloting tasks such as flight path stabilization,¹⁹ but legally significant actions, including firing a weapon,

18. See, e.g., *Predator UAS*, GENERAL ATOMICS AERONAUTICAL, <http://www.ga-asi.com/products/aircraft/predator.php> (last visited May 14, 2014).

19. See, e.g., *Autonomous Flight for Flight Platforms*, U.S. Patent No. 7809480 B2 (filed Sept. 26, 2005), available at <http://www.google.com/patents/US7809480>.

require an instruction to be issued by an operator who is actively monitoring the UCAV. Defensive systems such as the Patriot missile system²⁰ and the family of close-in weapons systems (CIWS), including the well-known Phalanx CIWS²¹ mounted on naval vessels deployed by various States, arguably go further in automating part or all of the targeting and firing process, but that process operates only in very specific and unambiguous circumstances.

Development plans that have been made public by various research and military organizations still focus on fielding “unmanned” rather than “autonomous” systems for the near future, but they indicate that the range of tasks assigned to both such systems is likely to expand greatly over the coming years. Basic research being conducted by universities, commercial organizations and organizations such as the Defense Advanced Research Projects Agency (DARPA) focuses more strongly on machine autonomy though, as do the long-term plans of various armed forces. Plans made public by the U.S. Air Force, for example, mention autonomous target engagement as a long-term goal to be achieved sometime after 2025,²² and other organizations have announced similar plans.²³ Perhaps of more immediate interest are plans to increase the level of autonomy in ISR systems, such as DARPA’s Military Imaging and Surveillance Technology program,²⁴ which aims to develop ISR systems that will be able to perform automated identification and recognition of potential targets. Given that one of the major motivations for developing systems such as these is the need to process the “unprecedented increase in intelligence, surveillance, and reconnaissance . . . data”²⁵ which can overwhelm analysts working in

20. *Patriot*, RAYTHEON, <http://www.raytheon.com/capabilities/products/patriot/> (last visited May 13, 2014).

21. *Phalanx Close-In Weapon System (CIWS)*, RAYTHEON, <http://www.raytheon.com/capabilities/products/phalanx/> (last visited May 13, 2014).

22. Headquarter, U.S. Air Force, Unmanned Aircraft Systems (UAS) Flight Plan 2009–2047 (2009), available at http://www.fas.org/irp/program/collect/uas_2009.pdf [hereinafter UAS Flight Plan 2009–2047].

23. See, e.g., Army Capabilities Integration Center, Robotics Strategy White Paper (2009), available at <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA496734>.

24. Strategic Technology Office, *Military Imaging and Surveillance Technology—Long Range (MIST-LR) Phase 2*, at 6, FEDBIZOPPS.GOV (Mar. 12, 2013), https://www.fbo.gov/index?s=opportunity&mode=form&id=78b0ddb382678fa9ace985380108f89&tab=core&_cvie w=0 (then DARPA-BAA-13-27_03.12.2013.docx hyperlink).

25. U.S. Department of Defense, *Autonomy Research Pilot Initiative (ARPI): Invitation for Proposals 4* (Nov. 2012), http://auvac.org/uploads/publication_pdf/Autonomy%20Research%20Pilot%20Initiative.pdf.

modern battlefields, the influence of autonomous systems on target selection may be felt well in advance of the advent of systems which remove humans from the observe-orient-decide-act (OODA) loop²⁶ altogether.

Although the timelines of these and other proposed developments are uncertain, the direction appears to support concerns about how to address possible violations of the law of armed conflict involving a highly autonomous system. Machine autonomy is a matter of degree, and the degree of autonomy needed to raise questions about the influence and accountability of weapon developers may not be particularly high if one considers the possibility of accessorial modes of liability (which are discussed below). Provided that other threshold requirements for criminal responsibility are met, a significant possibility exists that weapon developers may exercise a sufficient degree of control over proscribed acts by the systems they create such that they could be held liable for serious violations of the law. The proviso that other threshold requirements are met is important here. We turn now to consider some of the requisite elements for criminal responsibility in the context of autonomous weapons systems.

III. CHALLENGES TO DEVELOPER ACCOUNTABILITY

Given the uncertainty inherent in attempting to forecast the development paths of the relevant weapons technologies, the ways the weapons may be used and the views which courts may adopt, it is difficult to say precisely how the activities of developers may constitute acts proscribed by the law of armed conflict. We can, however, identify some of the questions that will arise and outline the most likely of the possible answers.

A. *Threshold Requirement of the Existence of an Armed Conflict*

It is axiomatic that a war crime can only be perpetrated in the context of an armed conflict. The application of the law of armed conflict is only triggered when an armed conflict exists, so it follows *ab initio* that there cannot be a serious violation of that law in the absence of the specific context in which the law applies. In many war crimes trials, it is obvious that the charges against the accused involved the perpetration of alleged

26. UAS Flight Plan 2009–2047, *supra* note 22, at 16.

acts in the context of an armed conflict. In other trials, the prosecution has a more challenging burden of establishing the existence of an armed conflict. The canonical formulation of the threshold requirement, repeatedly cited by international criminal courts and tribunals as well as in national criminal trial proceedings, was articulated by the Appeals Chamber of the International Criminal Tribunal for the former Yugoslavia (ICTY) in that Tribunal's very first trial against Duško Tadić when it held that:

[A]n armed conflict exists whenever there is a resort to armed force between States or protracted armed violence between governmental authorities and organized armed groups or between such groups within a State. International humanitarian law applies from the initiation of such armed conflicts and extends beyond the cessation of hostilities until a general conclusion of peace is reached; or, in the case of internal conflicts, a peaceful settlement is achieved. Until that moment, international humanitarian law continues to apply in the whole territory of the warring States or, in the case of internal conflicts, the whole territory under the control of a party, whether or not actual combat takes place there.²⁷

The ICTY Appeals Chamber decision, reflecting both treaty and customary international law, distinguished between international and non-international armed conflicts and articulated a different threshold test for these two types of armed conflict, with the latter having a significantly higher threshold test. Considerable effort is often expended in international criminal trials to determine whether or not an armed conflict exists and, if so, whether that armed conflict is international or non-international. At least in the case of the Rome Statute, the Prosecutor, if s/he chooses to pursue charges of war crimes, must prove the existence either of an international or of a non-international armed conflict as a requisite threshold element.

For the purposes of our analysis, let us assume the following: (1) that an armed conflict exists; (2) that it is immaterial whether that armed conflict is international or non-international; and (3) that the weapons developers completed their work before the armed conflict commenced.

27. Prosecutor v. Tadić; 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).

The critical issue arises from the third assumption and relates to the satisfaction of the threshold contextual element of a war crime—the existence of an armed conflict. It is possible that, in a protracted armed conflict lasting many years, a weapons system may be developed and deployed during the conduct of the armed conflict. Given the long lead time for development of autonomous weapons systems, that scenario, while possible (U.S. forces have already been deployed in Afghanistan for nearly 13 years, for example), will be the exception rather than the norm. In exceptional circumstances then, where all other threshold requirements are met, it may well be that developers can be prosecuted for war crimes because the threshold contextual element will be satisfied.

But what of the more likely scenario: where the development of the system has concluded well before the armed conflict in which the weapons system is deployed has commenced, such that the temporal distance of the developers' conduct from the context of armed conflict results in a serious obstacle to individual criminal liability? For the purposes of our analysis we will illustrate the issue by reference to the Rome Statute.

The Elements of Crimes document, negotiated to enumerate the requisite elements for each substantive crime within the Rome Statute, includes for every war crime a requirement that the alleged conduct “took place in the context of and was associated with an [international or non-international depending upon the precise provision of the Statute] armed conflict.”²⁸ It is difficult to imagine how the prosecution would prove that element in relation to a weapons developer whose conduct preceded the commencement of an armed conflict. Even if it were possible to argue that the development of weapons systems is “associated with an armed conflict” in the sense that the development has occurred in preparation for deployment in an armed conflict when and if the need arises, the development clearly will not have taken place “in the context of an armed conflict.”

B. Physical Perpetration and Accessorial Modes of Liability

In addition to proving all the constituent elements of a particular charge, the prosecution must also prove the criminal liability of the accused by reference to one or other of the enumerated grounds of individual criminal

28. International Criminal Court, Elements of Crimes art. 8 intro., U.N. Doc. PCNICC/2000/1/Add.2 (June 30, 2000).

liability in Article 25 of the Rome Statute. In this respect, the nature of development activities is significant and several observations can be made.

First, development means defining the behavior of a weapons system, but stops short of activating the system on the battlefield (activation or deployment is undertaken by the operator). Under this definition, even a hypothetical developer with the requisite intent and a high degree of control over a proscribed act committed via an autonomous system must act through another person or organization. In itself, weapon development can be no more than preparation for a proscribed physical act, not physical performance of the act. It is therefore unlikely that developers of weapons systems would be charged as physical perpetrators of war crimes pursuant to Article 25(3)(a), unless the allegation was that the developer exercised sufficient control to commit the crime “through another person, regardless of whether that other person is criminally responsible”; the other person being the operator that deployed the already developed weapons system years after the system had been developed.

Another hypothetical scenario, as applicable to conventional weapons and to weapons of mass destruction as it will be to autonomous weapon systems, should also be mentioned. Where evidence exists of an explicit common criminal purpose in which weapons developers intentionally develop a weapon system for a known illegal purpose, the fact that certain actions were undertaken prior to the commencement of an armed conflict in which the weapon was subsequently deployed and the intended crimes perpetrated will not preclude prosecution of the developers’ conduct. Such scenarios will, hopefully, be rare—exceptions rather than the norm—but the international community has witnessed ruthless political regimes willing to utilize all the resources available to them, including the co-opting of scientific and technical expertise, to produce a capability which they have not hesitated to use, including against their own civilian populations. The doctrine of common criminal purpose as a mode of individual criminal liability clearly envisages acts of preparation to contribute to or to further an agreed criminal activity or criminal purpose.²⁹

Second, to the extent that weapon developers may be said to instigate or control some action of an autonomous weapons system, they do so through the control software and other subsystems that they design, and the degree of control exercised by a developer depends on the degree of autonomy exhibited by the weapons system in respect of an action. Insofar

29. Rome Statute, *supra* note 7, art. 25(3)(d).

as it shapes the relationship between operators and weapons systems and excludes operators from intervening to exercise direct physical control over proscribed acts, the degree of system autonomy would also determine the potential for developers to be held accountable for actions instigated by those systems and is the likely mode for finding liability.

As explained above, as the degree of autonomy exhibited by a system increases, operators and developers move through various forms of shared control of—and responsibility for—the behavior of the system. As long as an operator remains “in the loop” or “on the loop,” and thus able to intervene in the functioning of the weapons system, control is shared between developer and operator. At some point yet to be defined it may be that system developers could occupy a position of effective control over the actions taken by a system in some situation encountered during a conflict, such that soldiers and their commanders may be effectively excluded and cannot be reasonably said to instigate those actions or be in a position to intervene. It therefore seems necessary to examine two modes of liability: where a developer via an autonomous system’s control software plays some contributory, but not decisive, role in instigating and carrying out a course of action, and where the system exercises such control that soldiers in the field cannot be said to have meaningfully contributed. Perhaps the most likely and appropriate alleged ground of liability where any degree of shared control is present would be pursuant to Article 25(3)(c) as an accessory who “aids, abets or otherwise assists in [the] commission [of the crime] . . . including providing the means for its commission.”

C. *International Jurisprudence on Ex Ante Aiding and Abetting*

Both alternative grounds of criminal liability would require the conduct constituting the crime to have been completed in the context of and be associated with an armed conflict. In relation to the deployment of an autonomous weapon system that results in a serious violation of the law in an armed conflict context, the “completion” of the requisite conduct is readily imaginable. The critical question is whether acts of preparation, which are to be completed either subsequently through the agency of another person or are to aid, abet or otherwise assist in the commission of the crime (including providing the means for its commission), can occur prior to the existence of an armed conflict.

We are unable to identify any extant jurisprudence from international war crimes trials to support the notion of individual criminal liability for war crimes where an accused's acts have occurred prior to the commencement of an armed conflict. At best, there are broad general statements as to hypothetical possibilities of accessorial liability for acts committed prior to the principal offense. By way of *obiter dicta*, the Appeals Chamber of the ICTY in the *Blaškić* judgment affirmed the possibility that: “[T]he *actus reus* of aiding and abetting a crime may occur before, during, or after the principal crime has been perpetrated, and that the location at which the *actus reus* takes place may be removed from the location of the principal crime.”³⁰

The Appeals Chamber countenanced the possibility that acts constituting aiding and abetting may occur before the principal crime has occurred, but that is a broad statement of principle and one not tested by the facts in the *Blaškić* trial. Both at the trial stage and on appeal, it was decided that aiding and abetting was not the appropriate ground of criminal liability for the charges against Blaškić and, for our purposes, the nature of the allegations against the accused certainly did not involve acts of preparation prior to the outbreak of hostilities in the Lašva Valley region of Bosnia-Herzegovina.

The Trial Chamber of the Special Court for Sierra Leone in the judgment against Charles Taylor affirmed the general statement of principle on aiding and abetting articulated by the ICTY Appeals Chamber in *Blaškić*:

The Accused may aid and abet at one or more of the “planning, preparation or execution” stages of the crime or underlying offence. The lending of practical assistance, encouragement or moral support may occur before, during or after the crime or underlying offence occurs. The *actus reus* of aiding and abetting does not require “specific direction.” No evidence of a plan or agreement between the aider and abettor and the perpetrator is required, except in cases of *ex post facto* aiding and abetting where “at the time of planning or execution of the crime, a prior agreement exists between the principal and the person who subsequently aids and abets the principal.”³¹

30. Prosecutor v. Blaškić, Case No. IT-95-14-A, Appeal Judgment, ¶ 48 (Int'l Crim. Trib. for the former Yugoslavia July 29, 2004) [hereinafter *Blaškić Appeal*].

31. See Prosecutor v. Taylor, Case No. SCSL-03-01-T, Judgment, ¶ 484 (May 18, 2012) [hereinafter *Taylor*].

Here again, the judges countenanced the possibility of *ex ante* acts constituting aiding and abetting but, in *Taylor*, unlike in *Blaškić*, the accused was convicted on the basis of this ground of individual criminal liability. However, in contrast to the scenario of a weapons developer completing all acts prior to the commencement of an armed conflict, Charles Taylor was found to have aided and abetted offenses throughout the indictment period and while the armed conflict raged within Sierra Leone. In particular, the Trial Chamber found that Taylor's steady supply of arms, ammunition, soldiers, operational support, encouragement and moral support to the RUF/AFRC forces throughout the indictment period all constituted substantial contributions to the commission of the alleged crimes.³²

In the absence of jurisprudence considering actual *ex ante* conduct, it is speculative to anticipate the hypothetical determinations of a court. But given: (1) the formulation in the Rome Statute's Elements of Crimes of the requisite contextual existence of an armed conflict for all war crimes; and (2) Article 22(2) of the Statute which requires that "the definition of a crime shall be strictly construed and shall not be extended by analogy" and that "in case of ambiguity, the definition shall be interpreted in favour of the person being investigated, prosecuted or convicted," any chamber of the International Criminal Court asked to decide the question in the future is likely to read the requisite elements narrowly.

D. Mens Rea Requirements

This section focuses on the *mens rea* requirements of the most plausible ground of developer liability discussed above—aiding and abetting. As was seen in relation to the physical element of the offense, the *ex ante* nature of developers' activities is also likely to present a significant barrier to establishing the mental element required for criminal responsibility.

Article 30 of the Rome Statute represents the first general attempt to codify the rules relating to the mental element of crimes in international criminal law. Relevantly, the first paragraph of that Article states that "[u]nless otherwise provided, a person shall be criminally responsible and liable for punishment for a crime within the jurisdiction of the Court only if the material elements are committed with intent and knowledge," on the basis that intent exists when a person means to engage in conduct or cause

32. *Id.*, ¶¶ 6907–58 (findings on aiding and abetting).

a consequence, and knowledge refers to awareness that a circumstance exists or a consequence will occur in the ordinary course of events.

Some further guidance as to the limits of the knowledge required by the aider and abettor may be gleaned from the judgment of the Trial Chamber of the ICTY in the trial of Anto Furundžija, which noted that:

[I]t is not necessary that the aider and abettor should know the precise crime that was intended and which in the event was committed. If he is aware that one of a number of crimes will probably be committed, and one of those crimes is in fact committed, he has intended to facilitate the commission of that crime, and is guilty as an aider and abettor.³³

The Appeals Chamber of the ICTY in the *Blaškić* judgment later concurred with that statement,³⁴ which seems to indicate that a developer need not be a party to detailed and specific plans to employ an autonomous system in a criminal act in order to be found guilty of aiding and abetting. That condition would likely be readily satisfied by a competent developer who imparts behavior to a weapons system which a reasonable person would be expected to see as criminal in nature. However, there are two significant challenges for prosecutors to overcome.

First, existing case law relates only to situations in which the accused is found to be aware of a criminal intent which is held by the physical perpetrator *at the time that the accused engages in conduct which assists the perpetrator*. The members of Anto Furundžija's anti-terrorist unit possessed criminal intent and were committing criminal acts while Furundžija was present and rendering assistance. Similarly Charles Taylor knew of a criminal intent which had been formed (and acted on) by his soldiers at the time he was providing supplies. Earlier, in the post-Second World War *Zyklon B* case,³⁵ in which the accused German industrialists were found guilty of supplying poison gas used to exterminate allied nationals interned in concentration camps in the knowledge that the gas was to be so used, the intent of the SS officers who exterminated the internees had been formed at the time the gas was supplied to them, and the industrialists who supplied it were aware of that intent. There is no suggestion in the

33. Prosecutor v. Furundžija, Case No. IT-95-17/1, Judgment, ¶ 246 (Int'l Crim. Trib. for the former Yugoslavia Dec. 10, 1998).

34. *Blaškić Appeal*, *supra* note 30, ¶ 50.

35. The *Zyklon B* Case, Trial of Bruno Tesch and Two Others, 1 LAW REPORTS OF TRIALS OF WAR CRIMINALS 93 (1947) (British Military Court Hamburg, Germany).

jurisprudence that the requisite mental state may extend to a developer's expectation that their work would be of assistance to a future perpetrator who may later form a criminal intent which did not exist at the time the developer completed their work.

Second, Article 25(3)(c) of the Rome Statute also specifies a special mental element which applies in relation to aiding and abetting. This is in addition to the general requirements of Article 30 that a developer's conduct would need to be done "for the purpose of facilitating the commission" of a crime within the jurisdiction of the Court. Prosecutors would be required to demonstrate that a weapon developer acts not only with awareness of the eventual physical perpetrator's intention to commit the crime and with the knowledge that his or her conduct would assist in perpetration of the offense,³⁶ but also that s/he acts for the purpose of facilitating the crime, rather than for the sole purpose of, for example, selling weapons systems for profit³⁷ or achieving another legal military objective.

Equally formidable challenges would also face a prosecutor seeking to show that a developer of a highly autonomous weapons system committed a crime "through another person, regardless of whether that other person is criminally responsible," though the picture is less clear. Despite recent signs of judicial opposition to the concept,³⁸ the International Criminal Court has generally favored the "control of the crime" approach to adjudicating on commission through another person. If that approach is adhered to, it would seem that a developer of a highly autonomous weapons system could, with criminal intent, produce a weapons system which performs a proscribed act without an instruction to that effect from its operator. However, similar problems exist here as in relation to aiding and abetting: the crime which the developer purportedly intends to commit, or knows will be committed in the ordinary course of events, can only occur later during an armed conflict. At a minimum, prosecutors would need to demonstrate that the developer understood that the weapons system would behave in an illegal manner, for example, to be

36. *Taylor*, *supra* note 31, ¶ 487.

37. ROBERT CRYER, HAKAN FRIMAN, DARRYL ROBINSON & ELIZABETH WILMSHURST, AN INTRODUCTION TO INTERNATIONAL CRIMINAL LAW AND PROCEDURE 312 (2d ed., 2010).

38. *See, e.g.*, Prosecutor v. Mathieu Ngudjolo Chui, Case No. ICC-01/04-02/12, Judgment Pursuant to Article 74 of the Statute (Van den Wyngaert, J., concurring ¶¶ 4-30) (Dec. 18, 2012).

incapable of distinguishing between military personnel and civilians by targeting and killing as yet unidentified civilian victims. Evidence of a developer intentionally building in to the weapons system a mechanism for targeting a specific individual known by the developer to be a civilian immune from attack would, of course, satisfy the *mens rea* requirements, but that type of evidence of specific intent will likely be rare.

E. Attribution of Responsibility

This article has consistently referred to weapons developers as individuals who may be held accountable for criminal acts committed via weapon systems which they have created. In practice, of course, highly sophisticated autonomous weapons systems would be—and are being—developed by organizations and by groups of organizations. This is not new. Complex weapons systems in use today are commonly comprised of hundreds or thousands of subsystems constructed by large networks of organizations: military and governmental bodies; domestic, foreign and multinational corporations; academic institutions and so on. Attribution of individual responsibility for a flaw, whether deliberate or accidental, would likely be exceedingly difficult.

IV. IMPLICATIONS OF AN ACCOUNTABILITY GAP FOR DEVELOPERS

At the outset, we identified two extreme views of an alleged accountability gap between the law of armed conflict and those responsible for serious violations of that law resulting from the deployment of autonomous weapons systems. On one hand, some say there is not just a gap but a yawning chasm “so unbridgeable, so deep and precarious” that the only way to preserve the extant normative framework is to ban autonomous weapons systems before they become any more sophisticated and the international community loses any capacity for effective accountability. On the other hand, some say that there is nothing to fear—that autonomous weapons systems are not yet sufficiently sophisticated to pose a threat to the efficacy of the normative framework and, in any case, that framework has shown itself eminently capable of adaptation to new circumstances in the past. That adaptability will manifest itself yet again when, and if, it becomes necessary to revisit the application of current norms to new weapons technologies generally and to autonomous weapons systems specifically.

As with many instances of polarized extremes, perhaps the truth lies somewhere in the middle. We stated earlier that existing weapons systems generally exhibit capabilities for autonomous operation far below what would be required to raise serious questions about accountability. Those of the “nothing new under the sun” perspective are likely to applaud this observation as precisely the reason that the adaptability of the existing normative framework is more than capable of regulating the increasing deployment of autonomous weapons systems. Two responses are important here.

First, limitations to the efficacy of the existing normative framework are exposed by the increasing incidence of attacks from existing weapons systems with relatively low levels of autonomy. For example, Philip Alston, former United Nations Special Rapporteur on Extrajudicial, Summary or Arbitrary Killings, is scathing in his assessment of the lack of accountability in the dramatically increased CIA and U.S. Department of Defense targeted killing programs. Alston writes of a “combination of high levels of secrecy, combined with poor accountability” rendering it “impossible to verify the extent to which applicable international standards are respected in practice.”³⁹ He argues that “none of the many existing oversight mechanisms have been even minimally effective in relation to targeted killings” and that the increasing use of drones and of special forces to execute the targeted killings programs have “grave consequences for the twin regimes of international human rights law . . . and international humanitarian law . . . which aim to uphold the value of human life and minimize the brutalities of warfare.”⁴⁰

We have indicated earlier that legally significant actions in relation to drones, such as the firing of a weapon, require an instruction to be issued by an operator who is actively monitoring the drone. If accountability is so ineffectual in relation to a level of automation which is legally insignificant, what hope is there for robust accountability in respect of sophisticated levels of automation? Of course, one response to Alston’s critique is that the problem is not with the normative framework itself but with ineffectual or improper implementation of it. The fact that deployment of weapons systems with legally significant levels of autonomy may not be subject to effective scrutiny and that serious violations of the law resulting from such

39. Philip Alston, *The CIA and Targeted Killings Beyond Borders*, 2 HARVARD NATIONAL SECURITY JOURNAL 283, 285 (2011), http://harvardnsj.org/wp-content/uploads/2011/02/Vol-2_Alston1.pdf.

40. *Id.* at 285, 287.

deployment may not result in substantive accountability for those responsible, does not render the weapons themselves illegal. However, the critique remains and it is important to reflect on the implications of it for the claim that the existing normative framework is more than capable of adapting to the new challenges posed by autonomous weapons systems.

Second, while existing weapons systems are generally operating at the lower end of the spectrum of autonomy, we have also observed that many of the component technologies which will one day comprise a highly autonomous weapons system either exist already or are currently under development. Given the potential for legally significant levels of autonomy in existing and/or developing technology, it seems naïve to simply dismiss accountability dilemmas on the basis of “wait and see what happens.”⁴¹ We have attempted to identify and articulate what some of those dilemmas are likely to be, at least in relation to the developers of weapons systems with legally significant levels of autonomy.

The likely scenario that developers will complete their involvement in a weapons system prior to the commencement of an armed conflict raises real concerns about the satisfaction of various requisite elements of any alleged war crime. We identified potential obstacles to developer liability for war crimes. The first of these is the threshold requirement that the alleged acts take place “in the context of an armed conflict.” That particular threshold requirement may amount to giving de facto immunity to weapons developers for subsequent war crimes. A prosecutor may seek to argue that the culmination of acts perpetrated prior to the commencement of an armed conflict “in the [subsequent] context of an armed conflict” is sufficient to satisfy the threshold element, but that is a novel argument and there is no guarantee that such an argument would be endorsed by judicial decision makers. It is true that international jurisprudence on the scope of aiding and abetting as a mode of liability envisages the sufficiency of *ex ante* acts of preparation. However, we are unable to identify any extant international jurisprudence to substantiate the proposition that acts of preparation to aid and abet a subsequent war crime can occur prior to the commencement of an armed conflict.

There are other additional obstacles to weapons developer liability, including the requisite *mens rea* requirements for aiding and abetting. A weapon developer accused of aiding and abetting would typically render

41. We agree with the rejection of this approach by Anderson and Waxman. Kenneth Anderson & Matthew Waxman, *Law and Ethics for Robot Soldiers*, 176 POLICY REVIEW (Dec. 1, 2012), <http://www.hoover.org/publications/policy-review/article/135336>.

their assistance prior to an armed conflict, whereas a criminal intent on the part of the physical perpetrator would typically emerge during the conflict. Bridging that temporal gap would require significant development of the law beyond the current jurisprudence.

Even if each of these requisite elements could be established in respect of a weapons developer, there is the further practical, perhaps practically insurmountable, obstacle that to speak of “the developer” is to misrepresent the realities of the development process. A sophisticated autonomous weapons system will not be developed by a single individual, but by many teams of developers in many organizations working on a multitude of subsystems with complex interdependencies. Attempting to identify an individual most responsible for subsequent behavior of a deployed weapons system that constitutes a war crime may simply be too difficult for the purposes of initiating trial proceedings. Each of these obstacles compound the general concerns expressed by Alston as to the inefficacy of existing accountability frameworks.

Those who favor a comprehensive multilateral treaty ban on autonomous weapons systems may well break into applause at this point. The identification of an accountability gap confirms the rationale for a prohibition on the category of weapons before they are produced and deployed in any larger numbers or, more importantly, in increasingly sophisticated (and so increasingly legally significant) forms. Several observations are relevant here.

The identification of obstacles to developer accountability in the existing normative framework does not mean that updating, amendment, revision and reform of the law is out of the question. The threshold requirement that relevant acts occur “in the context of an armed conflict” could be amended to explicitly (or could be interpreted to implicitly) include acts of preparation prior to the commencement of the armed conflict provided that the completion of the crime occurred in the relevant context. *Mens rea* requirements for developers to aid and abet subsequent offenses could also be clarified to cover the scenario we have been considering.

There is much more to accountability than the individual criminal liability of weapons developers. In situations of legally significant levels of autonomy, the operator who deploys the weapons system may well be excused of individual liability for lack of effective control over the behavior of the weapons system. However, the commander who has called for the deployment of the weapons system will be on notice once the system

behaves in such a manner as to have resulted in a serious violation of the law of armed conflict. At that point of the commander's awareness, the doctrine of command responsibility requires s/he take "all necessary and reasonable measures" to prevent further crimes.⁴² Such measures may well extend to ordering the standing down of the system pending re-programming/adjustment/tweaking to prevent recurrent offenses. More detailed analysis of the question of command responsibility in relation to the deployment of autonomous weapons systems that result in the perpetration of serious violations of the law of armed conflict is surely warranted.

There is also much more to accountability than individual criminal liability. Too little analysis has been undertaken on questions of State responsibility for the deployment of autonomous weapons systems that result in serious violations of the law of armed conflict. In particular, national reviews of any proposed acquisition or modification of autonomous weapons systems to ensure that the deployment of such systems will only occur in compliance with the law of armed conflict specifically, and international law obligations more generally, seem so much more important for weapons systems over which operators exercise relatively little, if any, effective control.⁴³ States acquiring new—or modifying existing—weapons systems may need to consider the imposition of specific product liability requirements on developers, manufacturers and vendors of such weapons systems to help cover reparation costs in the event of violations of international legal obligations.

42. Rome Statute, *supra* note 7, art. 28(a)(ii).

43. As required by Additional Protocol I, *supra* note 2, art. 36.