Volume 94 Number 886 Summer 2012

INTERNATIONAL REVIEW of the Red Cross

New capabilities in warfare: an overview of contemporary technological developments and the associated legal and engineering issues in Article 36 weapons reviews

Alan Backstrom and Ian Henderson*

Alan Backstrom, BEng, MEngSci, is an automotive engineering quality manager. He has extensive experience working with original equipment manufactures, system suppliers, subsystem suppliers, and component suppliers, with a particular focus on major design validation techniques, warranty analysis, and accident investigation.

Group Captain Ian Henderson, AM, BSc, LLB, LLM, PhD, is a legal officer with the Royal Australian Air Force.

Abstract

The increasing complexity of weapon systems requires an interdisciplinary approach to the conduct of weapon reviews. Developers need to be aware of international humanitarian law principles that apply to the employment of weapons. Lawyers need to be aware of how a weapon will be operationally employed and use this knowledge

* This paper was written in a personal capacity and does not necessarily represent the views of the Australian Department of Defence or the Australian Defence Force. Thank you to many friends and colleagues who generously provided comments on the draft.

doi:10.1017/S1816383112000707

to help formulate meaningful operational guidelines in light of any technological issues identified in relation to international humanitarian law. As the details of a weapon's capability are often highly classified and compartmentalized, lawyers, engineers, and operators need to work cooperatively and imaginatively to overcome security classification and compartmental access limitations.

Keywords: weapon, international humanitarian law, law of armed conflict, warfare, IHL, LOAC, Geneva, additional protocol, weapons review, autonomous, target recognition, reliability.

Article 36 of Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts provides:

> In the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party.[1]

As weapons become more technologically complex, the challenges of complying with this apparently simple requirement of international law become more daunting. If a lawyer were to conduct a legal review of a sword, there would be little need for the lawyer to be concerned with the design characteristics beyond those that can be observed by the naked eye. The intricacies of the production and testing methods would equally be legally uninteresting, and even a lawyer could grasp the method of employment in combat. The same cannot be said about some modern weapons, let alone those under development. The use of a guided weapon with an autonomous firing option requires an understanding of the legal parameters; the engineering design, production, and testing (or validation) methods; and the way in which the weapon might be employed on the battlefield.[2] While somewhat tongue-in-cheek, there is some truth to the view that a person becomes a lawyer due to not understanding maths, another becomes an engineer due to not understanding English, and the third a soldier due to not understanding either!

1 Opened for signature 12 December 1977, 1125 UNTS 3, entered into force 7 December 1978 (API). See generally Justin McClelland, 'The review of weapons in accordance with Article 36 of Additional Protocol I', in *International Review of the Red Cross*, Vol. 85, No. 850, June 2003, pp. 397–415; Kathleen Lawand, 'Reviewing the legality of new weapons, means and methods of warfare', in *International Review of the Red Cross*, Vol. 88, No. 864, December 2006, pp. 925–930; International Committee of the Red Cross (ICRC), *A Guide to the Legal Review of New, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977*, 2006. For a thorough discussion of what is and is not a 'weapon' for the purposes of legal review, see Duncan Blake and Joseph Imburgia, '"Bloodless weapons"? The need to conduct legal reviews of certain capabilities and the implications of defining them as "weapons"', in *The Air Force Law Review*, Vol. 66, 2010, p. 157.

2 See Michael Schmitt, 'War, technology and the law of armed conflict', in Anthony Helm (ed.), *The Law of War in the 21st Century: Weaponry and the Use of Force,* Vol. 82*, International Law Studies*, 2006, p. 142.

Our purpose in writing this article is to breakdown those barriers through a multidisciplinary approach that identifies the key legal issues associated with employing weapons, setting out important features of emerging weapons, and then analysing how engineering tests and evaluations can be used to inform the weapon review process. Through the combination of the above methods, we hope to provide a general framework by which the legal and engineering issues associated with weapon development and employment can be understood, regardless of the simplicity or complexity of the weapon.

We commence with a brief review of the key legal factors for employing and reviewing weapons, followed by three substantive parts. The first part deals with the target authorization process, regardless of the choice of weapon to be employed. The second part looks at some emerging weapons and the legal issues associated with those weapons. The final part considers the engineering issues associated with weapon reviews and, in particular, how an understanding of engineering processes can assist when reviewing highly complex weapons.

Key legal factors

The key legal steps under international humanitarian law[3] when conducting an attack can be summarized as:

1. collecting information about the target;
2. analysing that information to determine whether the target is a lawful target for attack at the time of the attack;
3. appreciating the potential incidental effects of the weapon and taking feasible precautions to minimize those effects;
4. assessing the 'proportionality' of any expected incidental effects against the anticipated military advantage of the overall attack (not just the particular attack of the individual weapon);[4]
5. firing, releasing, or otherwise using the weapon such that its effects are directed against the desired target;
6. monitoring the situation and cancelling or suspending the attack if the incidental effects are disproportionate.[5]

In addition, consideration must also be given to the type of weapon to be employed, and particularly relevant to this article is that there are also ways of employing (using) an otherwise lawful weapon that might result in a banned effect (e.g., indiscriminately firing a rifle). The key legal factors when conducting the review

3 Also known as the law of armed conflict.

4 See, for example, Australia's declaration of understanding to the effect that military advantage in Articles 51 and 57 of API, above note 1, means 'the advantage anticipated from the attack considered as a whole and not from isolated or particular parts of the attack' – reprinted in Adam Roberts and Richard Guelff, *Documents on the Laws of War*, 3rd edn, Oxford University Press, Oxford, 2000, p. 500.

5 See above note 1, Article 57(2)(b) of API.

of new weapons (including means and methods of combat) are whether the weapon itself is banned or restricted by international law;[6] and if not, whether the effects of the weapon are banned or restricted by international law.[7] Finally, the 'principles of humanity and the dictates of the public conscience' must also be kept in mind.[8]

From an operational point of view, the key points can be expressed as: achieving correct target-recognition, determining how to exercise weapon release authorization, and controlling (or limiting) the weapon effect.

With weapons of relatively simple design, the associated legal issues are simple. With the sword example above, the only real issues are whether it is a 'banned weapon';[9] and if not, whether the person who wields it does so with discrimination. Any design flaws (e.g., poorly weighted) or manufacturing defects (e.g., metal is too brittle) are unlikely to affect the legal analysis and are primarily the worry of the person using the sword. With more complex weapons like crossbows, the complexity of the weapon design introduces the potential for discrimination to be affected by:

- design errors (e.g., the weapon does not fire straight or consistent with any sighting mechanism as the design is flawed); or
- manufacturing errors (e.g., the weapon does not fire straight or consistent with any sighting mechanism as the weapon was not built, within tolerance, to the design).

These types of errors have the potential to be magnified with long-range weapons (such as artillery) and batch variation now also becomes a significant factor as any variations are magnified over the longer range of the weapon. Further, modern

6 Weapons can be banned outright, banned based on designed purpose or expected normal use, or the means of employment can be regulated (i.e., banned uses). A weapon may be totally banned through specific law (e.g., biological weapons are prohibited under the *Convention on the Prohibition of the Development, Production, and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction*, opened for signature 10 April 1972, 1015 UNTS 163, entered into force 26 March 1975), or may be banned generally if in all circumstances it is a weapon that is 'of a nature to cause superfluous injury or unnecessary suffering', see above note 1, Article 35(2) of API, and associated customary international law. Contrast this with, for example, laser weapons, which are generally lawful but are prohibited when they are specifically designed, solely or as one of their combat functions, to cause permanent blindness to unenhanced vision (*Protocol (IV) on Blinding Laser Weapons to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects*, opened for signature 13 October 1995, 35 ILM 1218, entered into force 30 July 1998). Finally, incendiary weapons are per se lawful, but, for example, may not be employed by air delivery against military objectives located within a concentration of civilians, see Article 2(2) of *Protocol III on Prohibitions or Restrictions on the Use of Incendiary Weapons to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects*, opened for signature 10 April 1981, 1342 UNTS 137, entered into force 2 December 1983.

7 ICRC, *A Guide to the Legal Review of New, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977*, above note 1, p. 11.

8 *Ibid.*

9 As there is no specific ban on swords, the issue would be a review under the general prohibition on weapons that cause unnecessary suffering pursuant to Article 35(2) of API, above note 1.

weapons have a variety of aiming mechanisms that are not solely dependent on the operator, such as inertial guidance, global positioning system (GPS), and electro-optical guidance. Finally, as discussed below, there is even the capacity for the weapon itself to select a target.

Weapon technology is advancing in many different areas and there is limited public material available on the avenues of research and the capabilities of the weapons being developed.[10] The following emerging weapons are, therefore, purely representative. In any event, the exact capabilities are of less importance to the discussion than are the general modes of operation.

Target recognition and weapon release authorization

The following discussion deals with weapons and weapon systems that have some level of functionality to discriminate between targets and, in appropriate circumstances, might attack a target without further human input. For example, a non-command-detonated landmine is a weapon that once placed and armed, explodes when it is triggered by a pressure plate, trip wire, etcetera. Such landmines have a very basic level of target recognition (e.g., a pressure plate landmine is triggered when a plate is stepped upon with a certain minimum amount of weight – e.g., 15 kilograms – and is clearly unlikely to be triggered by a mouse) and require no human weapon-release authorization.[11] More complex weapon systems purport to distinguish between civilian trucks and military vehicles such as tanks.[12] Automated and autonomous weapon systems need to be distinguished from remotely operated weapon systems. While there has been much discussion lately of unmanned combat systems, these are just remotely operated weapon platforms and the legal issues depend far more on the manner in which they are used than on anything inherent to the technology.[13] The following discussion differentiates automated weapons from autonomous weapons, briefly reviews some key legal issues associated with each type of weapon system, and concludes by outlining some methods for the lawful employment of such weapon systems.

10 See Hitoshi Nasu and Thomas Faunce, 'Nanotechnology and the international law of weaponry: towards international regulation of nano-weapons', in *Journal of Law, Information and Science*, Vol. 20, 2010, pp. 23–24.

11 Of course, this can be the very problem with landmines. Non-command-detonated landmines placed in areas frequented by civilians cannot distinguish between a civilian and a combatant activating the trigger mechanism.

12 'Anti-vehicle mines, victim-activation and automated weapons', 2012, available at: http://www.article36.org/weapons/landmines/anti-vehicle-mines-victim-activation-and-automated-weapons/ (last visited 1 June 2012).

13 For discussions of how such remotely operated systems are, legally, just like any other weapon system and are not deserving of separate categorization or treatment under international humanitarian law, see generally *Denver Journal of International Law and Policy*, Vol. 39 , No. 4, 2011; Michael Schmitt, Louise Arimatsu and Tim McCormack (eds), *Yearbook of International Humanitarian Law 2010*, Springer, Vol. 13, 2011.

Automated weapons

Automated weapon systems:[14]

> are not remotely controlled but function in a self-contained and independent manner once deployed. Examples of such systems include automated sentry guns, sensor-fused munitions and certain anti-vehicle landmines. Although deployed by humans, such systems will independently verify or detect a particular type of target object and then fire or detonate. An automated sentry gun, for instance, may fire, or not, following voice verification of a potential intruder based on a password.[15]

In short, automated weapons are designed to fire automatically at a target when predetermined parameters are detected. Automated weapons serve three different purposes. Weapons such as mines allow a military to provide area denial without having forces physically present. Automated sentry guns free up combat capability and can perform what would be tedious work for long hours and without the risk of falling asleep.[16] Sensor-fused weapons enable a 'shot and scoot' option and can be thought of as an extension of beyond-visual-range weapons.[17]

The principal legal issue with automated weapons is their ability to discriminate between lawful targets and civilians and civilian objects.[18] The second main concern is how to deal with expected incidental injury to civilians and damage to civilian objects.[19]

Starting with the issue of discrimination, it is worth noting that automated weapons are not new. Mines, booby traps, and even something as simple as a stake at the bottom of a pit are all examples of weapons that, once in place, do not require further control or 'firing' by a person. Some of these weapons also have an element of discrimination in the way they are designed. Anti-vehicle mines, for example, are

14 Not to be confused with automatic weapons, which are weapons that fire multiple times upon activation of the trigger mechanism – e.g., a machine gun that continues firing for as long as the trigger remains activated by the person firing the weapon.

15 Jakob Kellenberger, ICRC President, 'International humanitarian law and new weapon technologies', 34th Round Table on Current Issues of International Humanitarian Law, San Remo, 8–10 September 2011, Keynote address, p. 5, available at: http://iihl.org/iihl/Documents/JKBSan%20Remo%20Speech.pdf (last visited 8 May 2012). Various types of existing automated and autonomous weapons are briefly discussed, with further useful citations, in Chris Taylor, 'Future Air Force unmanned combat aerial vehicle capabilities and law of armed conflict restrictions on their potential use', Australian Command and Staff College, 2011, p. 6 (copy on file with authors).

16 South Korea is developing robots with heat and motion detectors to sense possible threats. Upon detection, an alert is sent to a command centre where the robots audio or video communications system can be used to determine if the target is a threat. If so, the operator can order the robot to fire its gun or 40 mm automatic grenade launcher. 'S. Korea deploys sentry robot along N. Korea border', in *Agence France-Presse*, 13 July 2010, available at: http://www.defensenews.com/article/20100713/DEFSECT02/7130302/S-Korea-Deploys-Sentry-Robot-Along-N-Korea-Border (last visited 6 May 2012).

17 A sensor-fused weapon is a weapon where the arming mechanism (the fuse) is integrated with a target detection system (the sensor).

18 Issues such as fratricide are not, strictly speaking, a concern of international humanitarian law. In any event, other means and methods are adopted to reduce fratricide, such as 'blue-force trackers', safe corridors, and restricted fire zones.

19 See above note 1, Article 51(5)(b) and Article 57(2)(a)(iii) of API.

designed to explode only when triggered by a certain weight. Naval mines were initially contact mines, and then advanced to include magnetic mines and acoustic mines. Of course, the problem with such mines is that there is no further discrimination between military objectives or civilian objects that otherwise meet the criteria for the mine to explode.[20] One way to overcome this is to combine various trigger mechanisms (sensors) and tailor the combination towards ships that are more likely to be warships or other legitimate targets than to be civilian shipping.

As weapons have become more capable and can be fired over a longer range, the ability to undertake combat identification of the enemy at greater distances has become more important. Non-cooperative target recognition (also called automatic target recognition) is the ability to use technology to identify distinguishing features of enemy equipment without having to visually observe that equipment.[21] A combination of technology like radar, lasers, communication developments, and beyond-visual-range weapon technology allows an ever-increasing ability to identify whether a detected object is friendly, unknown, or enemy and to engage that target. With each advance though, there is not 'a single problem but rather . . . a continuum of problems of increasing complexity ranging from recognition of a single target type against benign clutter to classification of multiple target types within complex clutter scenes such as ground targets in the urban environment'.[22] Significant work is underway to produce integrated systems where cross-cueing of intelligence, surveillance, and reconnaissance sensors allows for improved detection rates, increased resolution, and ultimately better discrimination.[23] Multi-sensor integration can achieve up to 10 times better identification and up to 100 times better geolocation accuracy compared with single sensors.[24]

With something as simple as a traditional pressure-detonated landmine, the initiating mechanism is purely mechanical. If a weight equal to or greater than the set weight is applied, the triggering mechanism will be activated and the mine will explode. This type of detonation mechanism cannot, by itself, discriminate between civilians and combatants (or other lawful targets). The potential for incidental injury at the moment of detonation is also not part of the 'detonate/do-not-detonate' equation. While this equation can be considered with

20 Except where the mine is command-detonated.

21 One example is using laser beams (an alternative is millimetre wave radar) to scan an object and then use processing algorithms to compare the image to pre-loaded 3D target patterns. Target identification can be based on specific features with up to 15cm resolution at a distance of 1000 metres. See 'Lased radar (LADAR) guidance system', Defense Update, 2006, available at: http://defense-update.com/products/l/ladar.htm (last visited 8 May 2012).

22 'RADAR Automatic Target Recognition (ATR) and Non-Cooperative Target Recognition (NCTR)', NATO, 2010, available at: http://www.rto.nato.int/ACTIVITY_META.asp?ACT=SET-172 (last visited 8 May 2012).

23 See Andy Myers, 'The legal and moral challenges facing the 21st century air commander', in *Air Power Review*, Vol. 10, No. 1, 2007, p. 81, available at: http://www.raf.mod.uk/rafcms/mediafiles/51981818_1143_EC82_2E416EDD90694246.pdf (last visited 8 May 2012).

24 Covering memorandum, *Report of the Joint Defense Science Board Intelligence Science Board Task Force on Integrating Sensor-Collected Intelligence*, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, US Department of Defense, November 2008, p. 1.

command-detonated landmines, that is clearly a qualitatively different detonation mechanism. With pressure-detonated landmines, the two main ways of limiting incidental damage are either by minimizing the blast and shrapnel, or by placing the mines in areas where civilians are not present or are warned of the presence of mines.[25]

However, the triggering mechanisms for mines have progressively become more complex. For example, anti-vehicle mines exist that are designed to distinguish between friendly vehicles and enemy vehicles based on a 'signature' catalogue. Mines that are designed to initiate against only military targets, and are deployed consistent with any design limitations, address the issue of discrimination. Nevertheless, that still leaves the potential for incidental injury and damage to civilians and civilian objects. The authors are not aware of any weapon that has sensors and/or algorithms designed to detect the presence of civilians or civilian objects in the vicinity of 'targets'. So, while some weapons claim to be able to distinguish a civilian object from a military objective and only 'fire' at military objectives, the weapon does not also look for the presence of civilian objects in the vicinity of the military objective before firing. Take the hypothetical example of a military vehicle travelling in close proximity to a civilian vehicle. While certain landmines might be able to distinguish between the two types of vehicles and only detonate when triggered by the military vehicle, the potential for incidental damage to the civilian vehicle is not a piece of data that is factored into the detonate/do-not-detonate algorithm. This is not legally fatal to the use of such automated weapons, but does restrict the manner in which they should be employed on the battlefield.

Along with discrimination there is the second issue of the potential for incidental injury to civilians and damage to civilian objects. The two main ways of managing this issue for automated weapons are controlling how they are used (e.g., in areas with a low likelihood of civilians or civilian objects) and/or retaining human overwatch. Both points are discussed further below under the heading 'Methods for the lawful employment of automated and autonomous weapons'. A third option is to increase the 'decision-making capability' of the weapon system, which leads us to autonomous weapons.

Autonomous weapons

Autonomous weapons are a sophisticated combination of sensors and software that 'can learn or adapt their functioning in response to changing circumstances'.[26] An autonomous weapon can loiter in an area of interest, search for targets, identify suitable targets, prosecute a target (i.e., attack the target), and report the point of

25 Of course, history has shown that many anti-personnel landmines were either emplaced without adequate consideration of, or worse intentional disregard for, the risk to civilians. As a result, a majority of states have agreed to a complete ban on the use of non-command-detonated anti-personnel landmines. See ICRC, 'Anti-personnel landmines', 2012, available at: http://www.icrc.org/eng/war-and-law/weapons/anti-personnel-landmines/ (last visited 8 May 2012).

26 J. Kellenberger, above note 15, p. 5.

weapon impact.[27] This type of weapon can also act as an intelligence, surveillance, and reconnaissance asset. An example of a potential autonomous weapon is the Wide Area Search Autonomous Attack Miniature Munition (WASAAMM). The WASAAMM:

> would be a miniature smart cruise missile with the ability to loiter over and search for a specific target, significantly enhancing time-critical targeting of moving or fleeting targets. When the target is acquired, WASAAMM can either attack or relay a signal to obtain permission to attack.[28]

There are a number of technical and legal issues with weapons such as the WASAAMM.[29] While most of the engineering aspects of such a weapon are likely to be achievable in the next twenty-five years, the 'autonomous' part of the weapon still poses significant engineering issues. In addition, there are issues with achieving compliance with international humanitarian law, and resulting rules of engagement, that are yet to be resolved.[30] Of course, if the WASAAMM operated in the mode where it relayed a signal to obtain permission to attack,[31] that would significantly reduce the engineering and international humanitarian law (and rules of engagement) compliance issues – but it also would not be a true autonomous weapon if operating in that mode.

An area that is related to autonomous weapons is the development of artificial intelligence assistants to help humans shorten the observe, orient, decide, act (OODA) loop. The purpose of such decision-support systems is to address the fact that while 'speed-ups in information gathering and distribution can be attained by well-implemented networking, information analysis, understanding and decision making can prove to be severe bottlenecks to the operational tempo'.[32] There is very

27 Chris Anzalone, 'Readying air forces for network centric weapons', 2003, slide 9, available at: http://www.dtic.mil/ndia/2003targets/anz.ppt (last visited 8 May 2012).

28 US Air Force, 'Transformation flight plan', 2003, Appendix D, p. 11, available at: http://www.au.af.mil/au/awc/awcgate/af/af_trans_flightplan_nov03.pdf (last visited 8 May 2012).

29 Myers also discusses some of the moral aspects, e.g., is it 'morally correct for a machine to be able to take a life'? See A. Myers, above note 23, pp. 87–88. See also ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, Report of the 31st International Conference of the Red Cross and Red Crescent, 2011, p. 40. Moral issues are also discussed in Kenneth Anderson and Matthew Waxman, 'Law and ethics for robot soldiers', in *Policy Review* (forthcoming 2012), available at: http://ssrn.com/abstract=2046375 (last visited 8 May 2012). See generally Peter Singer, 'The ethics of killer applications: why is it so hard to talk about morality when it comes to new military technology?', in *Journal of Military Ethics*, Vol. 9, No. 4, 2010, pp. 299–312.

30 *Ibid.*

31 For example, the UK 'Fire Shadow' will feature: 'Man In The Loop (MITL) operation, enabling a human operator to overrule the weapon's guidance and divert the weapon's flight path or abort the attack and return to loiter mode in conditions where friendly forces are at risk, prevailing conditions do not comply with rules of engagement, or where an attack could cause excessive collateral damage', see 'Fire Shadow: a persistent killer', Defense Update, 2008, available at: http://defense-update.com/20080804_fire-shadow-a-persistent-killer.html (last visited 8 May 2012).

32 Shyni Thomas, Nitin Dhiman, Pankaj Tikkas, Ajay Sharma and Dipti Deodhare, 'Towards faster execution of the OODA loop using dynamic decision support', in Leigh Armistead (ed.), *The 3rd International Conference on Information Warfare and Security*, 2008, p. 42, available at: http://academic-conferences.org/pdfs/iciw08-booklet-A.pdf (last visited 8 May 2012).

limited publicly available information on how such decision-support systems might operate in the area of targeting.

The key issue is how to use 'computer processing to attempt to automate what people have traditionally had to do'.[33] Using sensors and computer power to periodically scan an airfield for changes, and thereby cue a human analyst, has been more successful than using sensors such as synthetic aperture radar to provide automatic target recognition.[34] A clear difficulty is that the law relating to targeting is generally expressed in broad terms with a range of infinitely varying facts, rather than as precise formulas with constrained variables, which is why a commander's judgement is often needed when determining whether an object or person is subject to lawful attack.[35] As Taylor points out, it is this 'highly contextual nature' of targeting that results in there not being a simple checklist of lawful targets.[36] However, if a commander was prepared to forgo some theoretical capability, it is possible in a particular armed conflict to produce a subset of objects that are at any given time targetable. As long as the list is maintained and reviewed, at any particular moment in an armed conflict it is certainly possible to decide that military vehicles, radar sites, etcetera are targetable. In other words, a commander could choose to confine the list of targets that are subject to automatic target recognition to a narrow list of objects that are clearly military objectives by their nature – albeit thereby forgoing automatic target recognition of other objects that require more nuanced judgement to determine status as military objectives through their location, purpose, or use.[37]

The next step is to move beyond a system that is programmed to be a system that, like a commander, learns the nature of military operations and how to apply the law to targeting activities. As communication systems become more complex, not 'only do they pass information, they have the capacity to collate, analyse, disseminate . . . and display information in preparation for and in the prosecution of military operations'.[38] Where a system is 'used to analyse target data and then provide a target solution or profile'[39] then the 'system would reasonably

33 See above note 24, p. 47.

34 *Ibid.*, pp. 47–48. Automatic target recognition systems have worked in the laboratory but have not proved reliable when deployed and presented with real data rather than 'unrealistic controlled data for assessing the performance of algorithms', *ibid.*, pp. 47 and 53. While now somewhat dated, an article that explains how such target recognition works is Paul Kolodzy, 'Multidimensional automatic target recognition system evaluation', in *The Lincoln Laboratory Journal*, Vol. 6, No. 1, 1993, p. 117.

35 See C. Taylor, above note 15, p. 9. See generally Ian Henderson, *The Contemporary Law of Targeting: Military Objectives, Proportionality and Precautions in Attack under Additional Protocol I*, Martinus Nijhoff, Leiden, 2009, pp. 45–51.

36 See C. Taylor, *ibid.*, p. 9; see also I. Henderson, *ibid.*, pp. 49–50.

37 See above note 1, Art. 52(2) of API.

38 See J. McClelland, above note 1, p. 405. The technical issues (from as simple as meta-data standards for the sensor-collected data and available bandwidth for transmission of data, through to the far more complex) should not be downplayed, particularly with multi-sensor data. See generally, *Report of the Joint Defense Science Board Intelligence Science Board Task Force on Integrating Sensor-Collected Intelligence*, above note 24, pp. 1–9.

39 See J. McClelland, above note 1, p. 405.

fall within the meaning of "means and methods of warfare" as it would be providing an integral part of the targeting decision process'.[40]

What might a system look like that does not require detailed programming but rather learns? Suppose an artificial intelligence system scans the battlespace and looks for potential targets (let's call it the 'artificial intelligence target recognition system' (AITRS)). Rather than needing to be preprogrammed, the AITRS learns the characteristics of targets that have previously been approved for attack.[41] With time, the AITRS gets better at excluding low-probability targets and better at cueing different sensors and applying algorithms to defeat the enemy's attempt at camouflage, countermeasures, etcetera. In one example, the outcome of the process is that the AITRS presents a human operator with a simplified view of the battlespace where only likely targets and their characteristics are presented for human analysis and decision whether to attack. Importantly though, all of the 'raw information' (e.g., imagery, multispectral imagery, voice recordings of intercepted conversations, etcetera) is available for human review. In example two, while the AITRS still presents a human operator with a simplified view of the battlespace with likely targets identified for approval to attack, the human decision-maker is not presented with 'raw information' but rather analysed data.[42] For example, the human might be presented with a symbol on a screen that represents a motor vehicle along with the following:

- probability of one human rider: 99 per cent
- probability of body-match to Colonel John Smith:[43] 75 per cent
- probability of voice-match to Colonel John Smith: 90 per cent.[44]

And finally, in example three it is the AITRS itself that decides whether to prosecute an attack. Assuming the AITRS is also linked to a weapon system then the combination is an autonomous weapon system.

It would seem beyond current technology to be able to program a machine to make the complicated assessments required to determine whether or not a particular attack would be lawful if there is an expectation of collateral

40 *Ibid.*, p. 406.

41 See K. Anderson and M. Waxman, above note 29, p. 10.

42 'Automatically processing the sensor data to reduce critical information to a smaller data packet or to provide a go/no-go response could improve reaction time', in *Report of the Joint Defence Science Board Intelligence Science Board Task Force on Integrating Sensor-Collected Intelligence*, above note 24, p. 43.

43 Assume Colonel Smith is a person on the high-value target list and issues such as *hors de combat* (e.g., wounded, sick, surrendering, or otherwise out of combat) and collateral damage aside, is otherwise subject to lawful attack. This type of attack is based on identifying a target as being Colonel Smith. Contrast this with attacks based on characteristics of the target that are associated with 'enemy forces' (such as unloading explosives, gathering at certain locations, and other patterns of behaviour) without knowing the actual identity of the target. The latter are becoming known as 'signature' strikes, while the former are 'personality' strikes. See Greg Miller, 'CIA seeks new authority to expand Yemen drone campaign', in *The Washington Post*, 19 April 2012, available at: http://www.washingtonpost.com/world/national-security/cia-seeks-new-authority-to-expand-yemen-drone-campaign/2012/04/18/gIQAsaumRT_story.html (last visited 6 May 2012).

44 See also the example used by Myers, and his discussion of multi-sensor cueing. A. Myers, above note 23, p. 84.

damage.[45] Indeed, one would wonder even where to start as assessing anticipated military advantage against expected collateral damage is like comparing apples and oranges.[46] For now, that would mean any such weapon system should be employed in such a manner as to reduce the risk of collateral damage being expected.[47] However, a true AITRS that was initially operated with human oversight could presumably 'learn' from the decisions made by its human operators on acceptable and unacceptable collateral damage.[48]

As pointed out at footnote 46 above, collateral damage assessments are not just about calculating and comparing numbers – a function well suited to current computers. But instead, there is a clear qualitative assessment, albeit one where the things being compared are not even alike. How could a machine ever make such judgements? Perhaps not through direct programming but rather by pursuing the artificial intelligence route. So, along with learning what are lawful targets, our hypothetical AITRS would also learn how to make a proportionality assessment in the same way humans do – through observation, experience, correction in the training environment (e.g., war games), and so on. An AITRS that failed to make reasonable judgements (in the view of the instructing staff) might be treated the same as a junior officer who never quite makes the grade (perhaps kept on staff but not given decision-making authority), whereas an AITRS that proved itself on course and in field exercises could be promoted, entrusted with increasing degrees of autonomy, etcetera.

Another technical problem is that the required identification standard for determining whether a person or object is a lawful target is not clear-cut. The standard expressed by the International Criminal Tribunal for the Former Yugoslavia is that of 'reasonable belief'.[49] In their rules of engagement, at least two states have adopted the standard of 'reasonable certainty'.[50] A third approach,

45 ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, above note 29, pp. 39–40; William Boothby, *Weapons and the Law of Armed Conflict*, Oxford University Press, Oxford, 2009, p. 233.

46 See I. Henderson, above note 35, pp. 228–229. Many facets of military operations require commanders to exercise judgement, and this includes certain legal issues. Having determined what is the military advantage expected from an attack (not an exact quantity in itself) on a command and control node, and estimated the expected incidental civilian injury, death, and damage, somehow these two factors must be compared. The evaluation is clearly somewhat subjective and likely to differ from person to person, rather than objective and mathematical. In this respect, one can think of interpreting and complying with certain aspects of international humanitarian law as part art and not just pure science.

47 W. Boothby, above note 45, p. 233.

48 For a contrary view, see Markus Wagner, 'Taking humans out of the loop: implications for international humanitarian law', in *Journal of Law Information and Science*, Vol. 21, 2011, p. 11, available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1874039 (last visited 8 May 2012), who concludes that autonomous systems will never be able to comply with the principle of proportionality.

49 'The Trial Chamber understands that such an object [normally dedicated to civilian purposes] shall not be attacked when it is not reasonable to believe, in the circumstances of the person contemplating the attack, including the information available to the latter, that the object is being used to make an effective contribution to military action', ICTY, *The Prosecutor v Galic*, Case No IT-98-29-T, Judgement (Trial Chamber), 5 December 2003, para. 51.

50 International and Operational Law Department: The Judge Advocate General's Legal Centre & School (US Army), *Operational Law Handbook 2012,* 'CFLCC ROE Card', p. 103, available at: http://www.loc.gov/rr/frd/Military_Law/operational-law-handbooks.html (last visited 8 May 2012); ICRC, *Customary IHL*,

reflected in the San Remo *Rules of Engagement Handbook* is to require identification by visual and/or certain technical means.[51] The commander authorizing deployment of an autonomous weapon, and any operator providing overwatch of it, will need to know what standard was adopted to ensure that both international law and any operation-specific rules of engagement are complied with. It is also possible to combine the requirement for a particular level of certainty (e.g., reasonable belief or reasonable certainty) with a complementary requirement for identification to be by visual and/or certain technical means.

Presumably, for any identification standard to be able to be coded[52] into a computer program that standard would need to be turned into a quantifiable confirmation expressed as a statistical probability. For example, 'reasonable belief' would need to be transformed from a subjective concept into an objective and measurable quantity – for example, '95 per cent degree of confidence'. This would then be used as the benchmark against which field experience (including historical data) could produce an empirical equation to profile a potential target. Then new battlespace data can be compared to quantify (assess) the strength of correlation to the required degree of confidence (in the current example, 95 per cent or greater correlation). However, the uncertainty of measurement associated with the battlespace feedback sensors would also need to be quantified as a distinctly separate acceptance criterion. For example, assume in certain operational circumstances that an uncertainty of measurement results in an uncertainty of plus or minus 1 per cent, whereas in other operational circumstances the uncertainty is plus or minus 10 per cent. In the first circumstance, to be confident of 95 per cent certainty, the correlation would need to be not less than 96 per cent. In the second case, though, the required degree of confidence would never be achievable as the required degree of confidence of 95 per cent cannot be achieved due to the measurement uncertainty.[53]

Methods for the lawful employment of automated and autonomous weapons

Most weapons are not unlawful as such – it is how a weapon is used and the surrounding circumstances that affect legality.[54] This applies equally to automated and autonomous weapons, unless such weapons were to be banned by treaty

'Philippines: Practice Relating to Rule 16. Target Verification', 2012, available at: http://www.icrc.org/customary-ihl/eng/docs/v2_cou_ph_rule16 (last visited 8 May 2012).

51 See the sample rules at Series 31 'Identification of Targets', in International Institute of Humanitarian Law, *Rules Of Engagement Handbook*, San Remo, 2009, p. 38.

52 Again, a non-coding method would be through artificial intelligence.

53 In this second case, the targeting system could provide cueing for other sensors or a human operator; it just would be programmed to not permit autonomous weapon release.

54 Philip Spoerri, 'Round table on new weapon technologies and IHL – conclusions', in *34th Round Table on Current Issues of International Humanitarian Law*, San Remo, 8–10 September 2011, available at: http://www.icrc.org/eng/resources/documents/statement/new-weapon-technologies-statement-2011-09-13.htm (last visited 8 May 2012).

(e.g., like non-command-detonated anti-personnel landmines). There are various ways to ensure the lawful employment of such weapons.

> [The] absence of what is called a 'man in the loop' does not necessarily mean that the weapon is incapable of being used in a manner consistent with the principle of distinction. The target detection, identification and recognition phases may rely on sensors that have the ability to distinguish between military and non-military targets. By combining several sensors the discriminatory ability of the weapon is greatly enhanced.[55]

One method of reducing the target recognition and programming problem is to not try to achieve the full range of targeting options provided for by the law. For example, a target recognition system might be programmed to only look for high-priority targets such as mobile air defence systems and surface-to-surface rocket launchers – objects that are military objectives by nature and, therefore, somewhat easier to program as lawful targets compared to objects that become military objectives by location, purpose, or use.[56] As these targets can represent a high priority, the targeting software might be programmed to only attack these targets and not prosecute an attack against an otherwise lawful target that was detected first but is of lower priority.[57] If no high-priority target is detected, the attack could be aborted or might be prosecuted against other targets that are military objectives by nature. Adopting this type of approach would alleviate the need to resolve such difficult issues as how to program an autonomous system to not attack an ambulance except where that ambulance has lost protection from attack due to location, purpose, or use.[58]

A further safeguard includes having the weapon '"overwatched" and controlled remotely, thereby allowing for it to be switched off if considered potentially dangerous to non-military objects'.[59] Such overwatch is only legally (and operationally) useful if the operators provide a genuine review and do not simply trust the system's output.[60] In other words, the operator has to value add. For example, if an operator is presented with an icon indicating that a hostile target has been identified, then the operator would be adding to the process if that person separately considered the data, observed the target area for the presence of civilians, or in some other way did more than simply authorize or prosecute an attack based on the analysis produced by the targeting software. In other words, the operator

55 J. McClelland, above note 1, pp. 408–409.

56 See Lockheed Martin, 'Low cost autonomous attack system', in *Defense Update,* 2006, available at: http://defense-update.com/products/l/locaas.htm (last visited 8 May 2012).

57 An example would be detecting a T-72 tank but ignoring it as a low-priority target and continuing in search mode until detecting and engaging an SA-8 mobile surface-to-air missile launcher, *ibid.*

58 The presumption being that the high-priority targets are all clearly military in nature and, therefore, it would be easier to program target recognition software to identify such targets. If the high-priority targets happened to be ambulances being misused as mobile command and control vehicles, programming issues would still remain. See above note 37 and the accompanying text.

59 J. McClelland, above note 1, pp. 408–409.

60 See *Report of Defense Science Board Task Force on Patriot System Performance: Report Summary*, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, 2005, p. 2.

is either double-checking whether the target itself may be lawfully attacked, or is ensuring that the other precautions in attack (minimizing collateral damage, assessing any remaining collateral damage as proportional, issuing a warning to civilians where required, etcetera) are being undertaken. A problem arises where the operator is provided with large volumes of data,[61] as his or her ability to provide meaningful oversight could be compromised by information overload.[62] A way to manage this would be for the targeting software to be programmed in such a way that the release of a weapon is recommended only when the target area is clear of non-military objects.[63] In other circumstances, the targeting software might simply identify the presence of a target and of non-military objects and not provide a weapon release recommendation, but only a weapon release solution. In other words, the targeting software is identifying how a particular target could be hit, but is neutral on whether or not the attack should be prosecuted, thereby making it clear to the operator that there are further considerations that still need to be taken into account prior to weapon release.

Two further legal aspects of automated and autonomous weapons (and remotely operated weapons) that require further consideration are the rules relating to self-defence[64] and how the risk to own forces is considered when assessing the military advantage from an attack and the expected collateral damage.

The issue of self-defence has two aspects: national self-defence (which is principally about what a state can do in response to an attack) and individual self-defence (which is principally about what an individual can do in response to an attack).[65] Prior to an armed conflict commencing, the first unlawful use of force against a state's warships and military aircraft may be considered as amounting to an armed attack on that state, thereby allowing it to invoke the right of national self-defence. Would the same conclusion be reached if the warship or military aircraft were unmanned? Imagine an attack on a warship that for whatever reason had none of the ship's company on board at the time of the attack. What is it about attacks on warships that is of legal significance: the mere fact that it is a military vessel that is flagged to the state, the likelihood that any attack on the warship also imperils the ship's company, or a combination of the two?

Second, consider the different legal authorities for using lethal force. In broad terms, individual self-defence allows Person A to use lethal force against Person B when Person B is threatening the life of Person A.[66] Whether Persons A and B are opposing enemy soldiers or not is an irrelevant factor. Compare this to international humanitarian law, which allows Soldier A to use lethal force against

61 This could be a single system that processes and displays large volumes of data or a single operator who is given multiple systems to oversee.
62 ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, above note 29, p. 39.
63 J. McClelland, above note 1, pp. 408–409.
64 Conversations between Patrick Keane and Ian Henderson, 2011–2012.
65 In this context, individual self-defence also encompasses the issue of defending another party against an unlawful attack.
66 Domestic criminal law varies from jurisdiction to jurisdiction and the issue is more nuanced than this simple explanation.

Soldier B purely because Soldier B is the enemy.[67] Soldier B need not be posing any direct threat to Soldier A at all. Indeed, Soldier B may be asleep and Soldier A might be operating a remotely piloted armed aircraft. However, Soldier A must be satisfied, to the requisite legal standard, that the target is in fact an enemy soldier. Identification, not threat, is the key issue. However, during rules of engagement briefings military members are taught that during an armed conflict not only can they fire upon identified enemy, but also that nothing in international humanitarian law (or other law for that matter) prevents them from returning fire against an unidentified[68] contact in individual self-defence.[69] This well-known mantra will require reconsideration when briefing operators of unmanned assets. In all but the most unusual of circumstances, the remote operator of an unmanned asset will not be personally endangered if that unmanned asset is fired upon. This issue will need to be carefully considered by drafters of rules of engagement and military commanders, as generally returning fire to protect only equipment (and not lives) would be illegal under the paradigm of individual self-defence.[70] Compare this to the international humanitarian law paradigm that arguably would allow use of lethal force to protect certain types of property and equipment from attack, based on an argument that whoever is attacking the property and equipment must be either (1) an enemy soldier, or (2) a civilian taking a direct part in hostilities.[71]

Similarly, how to treat an unmanned asset under international humanitarian law when considering the 'military advantage' to be gained from an attack is not straightforward. While risk to attacking forces is a factor that can be legitimately considered as part of the military advantage assessment,[72] traditionally that has been thought of as applying to the combatants and not the military equipment. While it is logical that risk of loss of military equipment is also a factor, it will clearly be a lesser factor compared with risk to civilian life.

In conclusion, it is the commander who has legal responsibility 'for ensuring that appropriate precautions in attack are taken'.[73] Regardless of how remote in time or space from the moment of an attack, individual and state responsibility attaches to those who authorize the use of an autonomous weapon system.[74] It should be noted that this does not mean a commander is automatically

67 Subject to Soldier B being *hors de combat*. It would also be lawful under international humanitarian law for Soldier A to fire upon Person B for such time as Person B was a civilian taking a direct part in hostilities, but space does not allow a further exploration of that point.

68 Unidentified in the sense of unaware whether the person firing is an enemy soldier, a civilian, etcetera. There is still a requirement to identify the source (i.e., the location) of the threat.

69 The concept of 'unit self-defence' adds little to the present discussion, being a blend of both national and individual self-defence.

70 The legal paradigm of individual self-defence can be invoked to protect equipment where loss of that equipment would directly endanger life.

71 As long as I am satisfied that I have at least one legal basis for using lethal force against a *person (e.g.,* enemy combatant of civilian taking a direct part in hostilities), I do not have to determine which one is actually the case. Space does not allow a full discussion of this point, or the other interesting issue of using force to protect equipment as part of a national security interest under national self-defence outside of an armed conflict.

72 I. Henderson, above note 35, p. 199.

73 C. Taylor, above note 15, p. 12.

74 P. Spoerri, above note 54.

liable if something goes wrong. In war, accidents happen. The point under discussion is who could be found liable, not who is guilty.

The above discussion has focused on the intended target of a weapon. The following discussion deals with emerging weapons that highlight the legal issue of weapon effect even where the target is an otherwise lawful target.

Weapon effect

Directed energy weapons

Directed energy weapons use the electromagnetic spectrum (particularly ultraviolet through to infrared and radio-frequency (including microwave)) or sound waves to conduct attacks.[75] As a means of affecting enemy combat capability, directed energy weapons can be employed directly against enemy personnel and equipment, or indirectly as anti-sensor weapons. For example, laser systems could be employed as 'dazzlers' against aided and unaided human eyesight, infrared sensors, and space-based or airborne sensors,[76] and as anti-equipment weapons.[77] High-powered microwaves can be employed against electronic components and communications equipment. Lasers and radars are also used for target detection, target tracking, and finally for providing target guidance for other conventional weapons.

When directed energy weapons are employed against enemy communication systems, the legal issues are not significantly different from those that would arise if kinetic means were used. Is the target (e.g., a communication system) a lawful military objective and have incidental effects on the civilian population been assessed? As directed energy weapons have the clear potential to reduce the immediate collateral effects commonly associated with high-explosive weapons (e.g., blast and fragmentation),[78] the main incidental effect to consider is the second-order consequences of shutting down a communication system such as air traffic control or emergency services. While it is common to state that second-order effects must be considered when assessing the lawfulness of an attack, a proper understanding of what is 'counted' as collateral damage for the purpose of proportionality assessments is required. It is a mistake to think that any inconvenience caused to the civilian population must be assessed. That is wrong.

75 Particle weapons are also being studied but currently appear to remain in the area of theory, see Federation of American Scientists, 'Neutral particle beam', 2012, available at: http://www.fas.org/spp/starwars/program/npb.htm (last visited 8 June 2012); Carlo Popp, 'High energy laser directed energy weapons', 2012, available at: http://www.ausairpower.net/APA-DEW-HEL-Analysis.html (last visited 8 June 2012). For a good review of 'non-lethal' directed energy weapons (including acoustic weapons), see Neil Davison, *'Non-Lethal' Weapons*, Palgrave MacMillan, Basingstoke, 2009, pp. 143–219.

76 Laser systems could be employed as 'dazzlers' against space-based or airborne sensors while high-powered microwaves can be employed against electronic components, see *Defense Science Board Task Force on Directed Energy Weapons*, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, US Department of Defense, December 2007, pp. 2, 11 and 13.

77 Particularly for use against missiles, mine-clearing and as anti-satellite weapons, *ibid.*, p. 19.

78 As do other kinetic weapons such as inert concrete bombs.

Along with death and injury, it is only 'damage' to civilian objects that must be considered.[79] Therefore, a directed energy weapon attack on an air traffic control system that affected both military and civilian air traffic[80] need only consider the extent to which civilian aircraft would be damaged, along with associated risk of injury or death to civilians, and need not consider mere inconvenience, disruption to business, etcetera.[81]

Directed energy weapons are also being developed as non-lethal (also known as less-lethal) weapons to provide a broader response continuum for a controlled escalation of force.[82] For a variety of operational and legal reasons, it is preferable to have an option to preserve life while still achieving a temporary or extended incapacitation of the targeted individual. However, the very terms used to describe these weapons can cause problems beyond any particular legal or policy constraints.[83] The unintended consequences of the weapons (particularly due to the unknown health characteristics of the target) can lead to permanent injury or death. Such consequences are then used to stigmatize the concept of a non-lethal/less-than-lethal weapon. The important point to remember is that as for any other combat capability (including kinetic weapons), use of directed energy weapons during an armed conflict is governed by international humanitarian law and by any applicable rules of engagement and directions from the combat commander.[84]

Non-lethal directed energy weapons can be used in combination with traditional, lethal weapons. For example, it is reported that:

> Another weapon . . . can broadcast deafening and highly irritating tones over great distances. The long-range device precisely emits a high-energy acoustic beam as far as five football fields away. To a reporter standing across the airstrip from where it was set up in a hangar here, it sounded as if someone was shouting directly into his ear.
>
> The device 'has proven useful for clearing streets and rooftops during cordon and search . . . and for drawing out enemy snipers who are subsequently destroyed by our own snipers', the 361st Psychological Operations Company, which has tested the system in Iraq, told engineers in a report.[85]

79 See above note 1, Art. 51(5)(b) and Art. 57(2)(a)(iii) of API.

80 See ICRC, 'Cyber warfare and IHL: some thoughts and questions', 2011, available at: http://www.icrc.org/eng/resources/documents/feature/2011/weapons-feature-2011-08-16.htm (last visited 8 May 2012).

81 Space does not permit a full discussion of this point, but other factors warranting discussion are effects on neutrals and any third-order affects (e.g., the effect on emergency health-care flights), although query whether the 'ICRC might have a role in helping to generate international consensus on whether civilians have fundamental rights to information, electrical power, etc., in the same way as they have rights to life and property', *ibid.*

82 See generally, US Department of Defense, 'Non-lethal weapons program', available at: http://jnlwp.defense.gov/index.html (last visited 8 May 2012); James Duncan, 'A primer on the employment of non-lethal weapons', in *Naval Law Review*, Vol. XLV, 1998. See also Jürgen Altmann, 'Millimetre waves, lasers, acoustics for non-lethal weapons? Physics analyses and inferences', in DSF-Forschung, 2008, available at: http://www.bundesstiftung-friedensforschung.de/pdf-docs/berichtaltmann2.pdf (last visited 8 May 2012).

83 See *Defense Science Board Task Force on Directed Energy Weapons*, above note 76, p. xii.

84 *Ibid.*, p. xiii.

85 Bryan Bender, 'US testing nonlethal weapons arsenal for use in Iraq', in *Boston Globe*, 5 August 2005, available at: http://www.boston.com/news/nation/articles/2005/08/05/us_testing_nonlethal_weapons_

This form of directed energy weapon demonstrates two key issues associated with non-lethal weapon technology. First, such weapons are likely to be used against a civilian population – in this case, to clear streets and rooftops.[86] Second, the non-lethal weapon may be employed in conjunction with existing weapons to achieve a lethal effect.

Other directed energy weapons include active denial systems.[87]

> One of the weapons that has been successfully tested is a heat beam . . . that can 'bake' a person by heating the moisture in the first one-64th of an inch of the epidural layer of the skin. It was originally developed for the Department of Energy to keep trespassers away from nuclear facilities.[88]

The 'irresistible heating sensation on the adversary's skin [causes] an immediate deterrence effect';[89] because the heating sensation causes 'intolerable pain [the body's] natural defense mechanisms take over'.[90] The 'intense heating sensation stops only if the individual moves out of the beam's path or if the beam is turned off'.[91] Because flamethrowers and other incendiary weapons are only regulated and not specifically banned by international humanitarian law, there is no legal reason to deny the use of the active denial system in combat.[92]

Where active denial systems are being used as an invisible 'fence', then clearly it is a matter for the individual as to whether to approach the fence, and if so, whether to try to breach the perimeter.[93] However, if active denial systems are being aimed at a person or group to clear an area,[94] an issue that needs consideration with this type of weapon is how would a person who is being subjected to this type of attack either surrender or consciously choose to leave an area when they can neither see the beam,[95] may be unaware of even this type of technology, and are reacting to intolerable pain like the 'feeling . . . [of] touching a hot frying pan'?[96] Reacting

arsenal_for_use_in_iraq/?page=full (last visited 8 June 2012). The Long Range Acoustic Device is described in detail in Altmann, above note 82, pp. 44–53. As Altmann notes, while described as a hailing or warning device, it can potentially be used as a weapon, *ibid*., p. 52. For a discussion on attempts to avoid the legal requirement to review new 'weapons' by describing these types of acoustic devices by other names, see N. Davison, above note 75, pp. 102 and 205.

86 Concerns about using non-lethal weapons against the civilian population, or against 'individuals before it is ascertained whether or not they are combatants' are raised in Davison, above note 75, pp. 216–217.

87 *Defense Science Board Task Force on Directed Energy Weapons*, note 76, pp. 33 and 38. For more details see 'Active denials system demonstrates capabilities at CENTCOM', United State Central Command, available at: http://www.centcom.mil/press-releases/active-denial-system-demonstrates-capabilities-at-centcom (last visited 8 May 2012).

88 B. Bender, above note 85. The Active denial system is described in detail in J. Altmann, above note 82, pp. 14–28.

89 *Defense Science Board Task Force on Directed Energy Weapons*, above note 76, p. 38.

90 *Ibid.*, p. 42.

91 *Ibid.*

92 J. Altmann, above note 82, p. 27.

93 Conversation between Patrick Keane and Ian Henderson, 14 April 2012.

94 As opposed to traditional kinetic weapons where the desired effect is to disable (through either wounding or killing).

95 See J. Altmann, above note 82, p. 28.

96 *Defense Science Board Task Force on Directed Energy Weapons*, above note 76, p. 42.

instinctively to intolerable pain seems likely to make a person incapable of rational thought.[97] Employment of such weapons will need to be well regulated through a combination of the tactics, techniques and procedures, and rules of engagement to ensure that unnecessary suffering is not caused through continued use of the weapon because a person has not cleared the target area.[98] In this respect, and noting that the active denial system has 'successfully undergone legal, treaty and US Central Command rules of engagement reviews',[99] it is worth recalling that as states' legal obligations vary, and as states may employ weapons differently, the legal review by one state is not determinative of the issue for other states.[100] This may prove interesting in the sale of highly technical equipment, as the details of a weapon's capability are often highly classified and compartmentalized. The state conducting the review may not control access to the necessary data. As discussed below, this may require lawyers, engineers, and operators to work together cooperatively and imaginatively to overcome security classification and compartmental access limitations.

A similar directed energy weapon using different technology is 'a high-powered white light so intense as to send any but the most determined attackers running in the opposite direction'.[101] Concepts for employment of the weapon appear to include using it as a means to identify hostile forces, as evidenced by the statement: 'If anyone appears willing to withstand the discomfort, "I know your intent", [Colonel Wade] Hall [a top project official] said. "I will kill you."'[102] While initially such statements appear quite concerning, it is instructive to consider whether this is in reality any different from the 'traditional' warnings and escalation of force scenarios such as 'stop or I will shoot' or employment of flares and dazzlers to warn vehicles not to approach too close to military convoys.

Where directed energy weapons are used to counter (often improvised) explosive devices,[103] the issue is primarily about consequences. If the directed energy weapon is causing a detonation at a safe range from friendly forces, there is a requirement to consider whether any civilians or other non-combatants are in the vicinity of the detonation and, therefore, at risk of injury or death.[104]

97 Email April-Leigh Rose/Ian Henderson, 24 April 2012.
98 Altmann also recommends investigating risk to eyesight due to potential damage to the cornea; see J. Altmann, above note 82, p. 28.
99 *Ibid.*, p. 38.
100 See J. McClelland, above note 1, p. 411, who makes this point with respect to manufacturer's claims of legality.
101 B. Bender, above note 85.
102 *Ibid.*
103 See *Defense Science Board Task Force on Directed Energy Weapons*, above note 76, p. 40.
104 Space does not permit a full exploration of this point, but note that the issues are different if instead of causing a detonation the countermeasure prevents the explosive device from detonating.

Cyber operations

Cyber operations are:

> operations against or via a computer or a computer system through a data stream.[105] Such operations can aim to do different things, for instance to infiltrate a system and collect, export, destroy, change, or encrypt data or to trigger, alter or otherwise manipulate processes controlled by the infiltrated computer system. By these means, a variety of 'targets' in the real world can be destroyed, altered or disrupted, such as industries, infrastructures, telecommunications, or financial systems.[106]

Cyber operations are conducted via software, hardware, or via a combination of software and personnel. A recent example of a cyber operation that was essentially conducted purely by software is the Stuxnet virus. Once in place, the Stuxnet virus appears to have operated independently of any further human input.[107] Compare this to a software program that is designed to allow a remote operator to exercise control over a computer – allowing, among other things, the upload of data or modification of data on the target computer. Finally, a non-military example of a cyber operation that requires both hardware and software is credit card skimming.

The application of specific international humanitarian law rules to cyber warfare remains a topic of debate.[108] However, for the purposes of this article, it is assumed that the key international humanitarian law principles of distinction, proportionality, and precaution, apply, as a minimum, to those cyber attacks that have physical consequences (e.g., the Stuxnet virus altered the operating conditions for the Iranian uranium enrichment centrifuges, which ultimately resulted in physical damage to those centrifuges).[109] Four particular legal aspects of cyber weapons are worth mentioning.

First, cyber weapons have the distinct possibility of being operated by civilians.[110] The 'weapon' is likely to be remote from the battlefield, is technologically sophisticated, and does not have an immediate association with death and injury. The operation of the cyber weapon exposes a civilian operator to

105 Based on this definition, a kinetic attack to shut down a computer system (for example, by dropping a bomb on the building housing the computer) would not be a cyber operation.

106 ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, above note 29, p. 36.

107 See Angus Batey, 'The spies behind your screen', in *The Telegraph*, 24 November 2011; Jack Goldsmith, 'Richard Clarke says Stuxnet was a US operation', in *LawFare: Hard National Security Choices*, 29 March 2012, available at: http://www.lawfareblog.com/2012/03/richard-clarke-says-stuxnet-was-a-u-s-operation/ (last visited 18 April 2012).

108 See 'Tallinn Manual on the International Law Applicable to Cyber Warfare', 2012, pp. 17–22, available at: http://issuu.com/nato_ccd_coe/docs/tallinn_manual_draft/23 (last visited 8 June 2012).

109 ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, above note 29, pp. 36–37.

110 See Adam Segal, 'China's cyber stealth on new frontline', in the *Australian Financial Review*, 30 March 2012, available at: http://afr.com/p/lifestyle/review/china_cyber_stealth_on_ new_frontline_z6YvFR0mo3uC87zJvCEq6H (last visited 1 June 2012), referring to 'cyber-militias' at technology companies recruited by the People's Liberation Army.

lethal targeting (as a civilian taking a direct part in hostilities),[111] as well as potential criminal prosecution for engaging in acts not protected by the combatant immunity enjoyed by members of the armed forces.[112] These issues are discussed in detail in a recent article by Watts who raises, among other things, the possibility of the need for a complete rethink of how the law on direct participation in hostilities applies in the area of cyber warfare.[113] It could also be queried what training such civilian operators might have in the relevant rules of international humanitarian law.[114]

Second, cyber attacks can have consequences in the real world and not just the virtual world.[115] Where those consequences affect the civilian population by causing loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, those consequences must be considered under international humanitarian law.[116] The discussion of this point for directed energy weapon attacks applies equally to cyber attacks. A further related consideration is that where it could reasonably be expected that a virus introduced into a military system might find its way into civilian systems and cause infrastructure damage, that collateral damage must also be considered.[117] A common example of a possible cyber attack that would directly affect civilians is disabling a power station – either just by shutting it down, or by overloading or shutting down a fail-safe, thereby damaging hardware. This can potentially happen to any infrastructure maintained by software.

Third, cyber weapons need to be considered not only in relation to international humanitarian law, but also very importantly under *jus ad bellum*.[118] As Blake and Imburgia point out, even if a cyber attack has no kinetic effects, the attack might still be contrary to the UN Charter specifically or international law generally[119] and may, if amounting to an 'armed attack', legitimize the use of force by the affected state in self-defence.

111 See above note 1, Article 51(3) of API.

112 On both these points, see D. Blake and J. Imburgia, above note 1, pp. 195–196.

113 See Sean Watts, 'Combatant status and computer network Attack', in *Virginia Journal of International Law*, Vol. 50, No. 2, 2010, p. 391.

114 See J. Kellenberger, above note 15, where this point was made with respect to remotely operated weapon systems.

115 ICRC, 'Cyber warfare and IHL: some thoughts and questions', above note 80.

116 See above note 1, Art. 51(5)(b) and Art. 57(2)(a)(iii) of API. It is a matter of policy whether to consider other consequences for the civilian population such as disruption, loss of amenities, etcetera.

117 See ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, above note 29, p. 38.

118 Put simply, *jus ad bellum* is the law regulating the overall resort to the use of force, compared to international humanitarian law (*jus in bello*) that regulates the individual instances of the application of force during an armed conflict. See Matthew Waxman, 'Cyber attacks as "force" under UN Charter Article 2(4)', in Raul Pedrozo and Daria Wollschlaeger (eds), *International Law and the Changing Character of War, International Law Studies*, Vol. 87, 2011, p. 43; Sean Watts, 'Low-intensity computer network attack and self-defense', in *ibid*., p. 59; Michael Schmitt, 'Cyber operations and the *jus ad bellum* revisited', in *Villanova Law Review*, Vol. 56, No. 3, 2011, pp. 569–605.

119 D. Blake and J. Imburgia, above note 1, pp. 184–189. Discussed in more detail in M. Schmitt, *ibid.*, who also discusses the current 'fault lines in the law governing the use of force [that] have appeared because it is a body of law that predates the advent of cyber operations'.

Finally, the very nature of cyber warfare can make it hard to determine who initiated an attack, and issues of attribution go to the very heart of both state responsibility and individual accountability.[120]

Nanotechnology and weaponization of neurobiology

> Nano-weapons are hard to define, but encompass not only objects and devices using nanotechnology that are designed or used for harming humans, but also those causing harmful effects in nano-scale if those effects characterise the lethality of the weapon.[121]

An example of the latter is the Dense Inert Metal Explosive (DIME):

> DIME involves an explosive spray of superheated micro shrapnel made from milled and powdered Heavy Metal Tungsten Alloy (HMTA), which is highly lethal within a relatively small area. The HMTA powder turns to dust (involving even more minute particles) on impact. It loses inertia very quickly due to air resistance, burning and destroying through a very precise angulation everything within a four-meter range – and it is claimed to be highly carcinogenic and an environmental toxin. This new weapon was developed originally by the US Air Force and is designed to reduce collateral damage in urban warfare by limiting the range of explosive force.[122]

The 'capacity [of DIME] to cause untreatable and unnecessary suffering (particularly because no shrapnel is large enough to be readily detected or removed by medical personnel) has alarmed medical experts'.[123] The other concern with nanotechnology is that elements and chemicals that on a macro scale are not directly harmful to humans can be highly chemically reactive on the nanoscale. This may require a review of what international humanitarian law considers as chemical weapons.

Similarly, with the current advances in the understanding of the human genome and in neuroscience, there exists the very real possibility of militarization of this knowledge.[124] One of the legal consequences is a need to reappraise maintaining

120 J. Kellenberger, above note 15; ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, above note 29, p. 37.

121 H. Nasu and T. Faunce, above note 10, p. 23.

122 Whether such a weapon has been used in actual combat appears to remain a matter of speculation – see generally *Dense Inert Metal Explosive (DIME)*, Global Security, available at: http://www.globalsecurity.org/military/systems/munitions/dime.htm (last visited 8 May 2012).

123 H. Nasu and T. Faunce, above note 10, p. 22. Along with Art. 35(2) of API, above note 1, on unnecessary suffering, there is also *Protocol I of the Convention on Certain Conventional Weapons on Non-Detectable Fragments*, (10 October 1980). Amnesty International is of the view that 'further studies are required before it can be determined whether the use of DIME munitions is lawful under international law'. Amnesty International, 'Dense Inert Metal Explosives (DIME)', in *Fuelling conflict: foreign arms supplies to Israel/Gaza*, 2009, available at: http://www.amnesty.org/en/library/asset/MDE15/012/2009/en/5be86fc2-994e-4eeb-a6e8-3ddf68c28b31/mde150122009en.html#0.12. (last visited 8 May 2012). For a discussion generally of the *Protocol I of the Convention on Certain Conventional Weapons on Non-Detectable Fragments*, see W. Boothby, above note 45, pp. 196–199.

124 See generally Mark Wheelis and Malcolm Dando, 'Neurobiology: a case study for the imminent militarization of biology', in *International Review of the Red Cross*, Vol. 87, No. 859, 2005, p. 553. See also

a legal distinction between chemical and biological weapons. It may be that based on the manner in which they can be used we should legally view these weapons as part of a 'continuous biochemical threat spectrum, with the Chemical Weapons Convention and Biological and Toxin Weapons Convention (CWC and BTWC) overlapping in their coverage of mid-spectrum agents such as toxins and bioregulators'.[125]

There are competing tensions in this area. Quite understandably, chemical and biological weapons have a 'bad name'. At the same time, research is underway into non-lethal weapons such as incapacitating biochemical weapons.

> Although there is currently no universally agreed definition, incapacitating biochemical agents can be described as substances whose chemical action on specific biochemical processes and physiological systems, especially those affecting the higher regulatory activity of the central nervous system, produce a disabling condition (e.g., can cause incapacitation or disorientation, incoherence, hallucination, sedation, loss of consciousness). They are also called chemical incapacitating agents, biotechnical agents, calmatives, and immobilizing agents.[126]

A key point to note is that while traditional biological and chemical agents were used against enemy soldiers or non-cooperative civilians, and clearly would be classified as weapons, modern agents may be used to 'enhance' the capability of a state's own military forces. In such cases, it is much less likely that the agents would amount to weapons.[127] For example:

> within a few decades we will have performance enhancement of troops which will almost certainly be produced by the use of diverse pharmaceutical compounds, and will extend to a range of physiological systems well beyond the sleep cycle. Reduction of fear and pain, and increase of aggression, hostility, physical capabilities and alertness could significantly enhance soldier performance, but might markedly increase the frequency of violations of humanitarian law. For example, increasing a person's aggressiveness and hostility in conflict situations is hardly likely to enhance restraint and respect for legal prohibitions on violence.[128]

Similar concerns have already been expressed about remotely operated weapons. And in a manner similar to using directed energy weapons to disperse civilian

'Brain waves 3: neuroscience, conflict and security', in *The Royal Society*, available at: http://royalsociety.org/policy/projects/brain-waves/conflict-security (last visited 6 May 2012) for a discussion of, among other things, potential military applications of neuroscience and neurotechnology and current legal issues.

125 M. Wheelis and M. Dando, *ibid.*, p. 560.

126 Michael Crowley and Malcolm Dando, 'Submission by Bradford Nonlethal Weapons Research Project to Foreign Affairs Select Committee Inquiry on Global Security: Non-Proliferation', 2008, pp. 1–2, available at: http://www.brad.ac.uk/acad/nlw/publications/BNLWRP_FAC071108MC.pdf (last visited 8 May 2012).

127 Body armour, for example, is not classified as a weapon.

128 M. Wheelis and M. Dando, above note 124, pp. 562–563.

crowds, there is also the potential to pacify civilians in occupied territories through chemicals included in food distributions.[129] Perhaps of even more concern, as it goes directly to the ability to enforce international humanitarian law, particularly command responsibility, is the possibility of 'memories of atrocities committed [being] chemically erased in after-action briefings'.[130]

The need to understand the role of engineering in the weapon review process

The above overview of emerging weapons highlights that as weapons become more complex the ability for non-experts to understand the complex manner in which the weapon operates becomes increasingly difficult. This part of the article focuses on engineering issues and how an understanding of those issues can be factored into the legal review of weapons.

Why a weapon may not perform as intended

A weapon may not perform as intended or in accordance with the 'product design specification'[131] for a variety of reasons. Those reasons include: inadequate technical specification, design flaws, or poor manufacturing quality control (batch variation). Other factors include 'age of the munition, storage conditions, environmental conditions during employment, and terrain conditions'.[132]

A simple example of specification failure, or at least a specification that will not be 100 per cent reliable, is an anti-vehicle mine that is not intended to explode when stepped on by a human. For example, if it is a load activated mine, the load might be set to 150 kg. However, biomechanical research:

> shows very strong evidence that a human being can very easily exert an equivalent force close to and above such pressures. For example, an 8-year-old boy weighing 30 kg, running downhill in his shoes, exerts a ground force of 146 kg. A 9-year-old girl weighing 40 kg running downhill in her bare feet exerts 167 kg of force. An adult male running will exert 213 kg.[133]

Alternatively, the specification might be correct but the design, manufacturing process, or integration of systems does not consistently lead to the intended result. This may be an engineering quality issue where the implemented engineering

129 *Ibid.*, p. 565.
130 *Ibid.*, p. 565
131 The product design specification is a step before the actual technical specifications for a product. The former is about what a product should do, while the latter is concerned with how the product will do it.
132 *Defense Science Board Task Force, Munitions System Reliability*, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, US Department of Defense, Washington, DC, September 2005, p. 15, available at: http://purl.access.gpo.gov/GPO/LPS72288 (last visited 8 May 2012).
133 'Anti-vehicle mines: discussion Paper', Actiongroup Landmine.de, 2004, p. 5. (footnote omitted), available at: http://www.landmine.de/fileadmin/user_upload/pdf/Publi/AV-mines-discussion-paper.pdf (last visited 8 May 2012).

processes were inadequately robust leading to product flaws, and as such presents a reliability issue.

Where a weapon does not perform as intended, two prime consequences are:

- The desired combat effect is not achieved. If the weapon fails to perform, own forces are put at risk. If the weapon does not perform to specification, civilians and civilian property are put at risk.[134]
- Where civilians are injured or killed or civilian property damaged, liability may be incurred.[135] State liability may be incurred for an internationally wrongful act (i.e., a breach of the international humanitarian law) and criminal liability potentially attaches to the commander who authorized the use, or to the person who employed the weapon, or both.

As weapons systems become more complex, an understanding of reliability analysis will need to become part of the legal review process.

Reliability: test and evaluation

The purpose of test and evaluation is to provide an objective measurement of whether a system (or a component thereof) performs reliably to a specification. Reliability is the probability of correct functioning to a specified life (measured in time, cycles of operation, etcetera) at a given confidence level. Understanding that reliability is a key factor in weapon performance is intuitively simple but in fact has a level of complexity not always immediately grasped by those unfamiliar with reliability engineering.[136] Quantifying reliability is not a 'yes' or 'no' proposition,[137] nor can it be achieved by a single pass/fail test, but rather 'is subject to statistical confidence bounds'.[138] For example, to obtain an appropriate level of statistical confidence that the failure rate for a given weapon population is acceptable there are a minimum number of tests required. But as resources are always finite the question for responsible engineering practice is how to optimize resources and understand the minimum required resources to assure acceptable reliability? Suppose that undertaking the required number of tests will be too time-consuming or beyond budget allocation. A naïve approach would simply reduce the number of tests to meet budget requirements and presume that the test will still give some useful information. But that may not be the case. Arguably, the compromised test can only provide misleading conclusions if the result does not achieve the required level of confidence. For certification purposes, either a certain level of confidence is required or not. While the statistical confidence level may be set appropriately low

134 This has direct military effectiveness consequences, as well as effecting morale, domestic public support, international support, etcetera.

135 Liability may also arise where the means or method of warfare against combatants is unlawful, which may be the case in a defective weapon scenario, for example, firing on a combatant who is *hors de combat*.

136 See generally, *Defense Science Board Task Force on Munitions System Reliability*, above note 132.

137 'Just tell me whether it is reliable or not?' asks the hypothetical boss.

138 *Defense Science Board Task Force on Munitions System Reliability*, above note 132, p. 15.

for non-lethal weapon components where a failure has a low-operational impact and minor to no safety implications (e.g., failure of a tracer bullet), the target recognition system on an autonomous weapon may require a very high statistical confidence to minimize lethal weapon deployment on civilians while still ensuring engagement of enemy targets. If a high statistical assurance is deemed necessary for civilian safety while budgetary constraints preclude the corresponding necessary development testing, then appropriate limits should be implemented regarding the approved applications for that weapon until field experience provides appropriate reliability confidence.

How should this be applied in practice? The main steps of weapon acquisition are usefully outlined by McClelland, including the various testing stages during 'demonstration', 'manufacture', and 'in-service'.[139] As McClelland notes, this is not a legal process but rather part of the acquisition process; but nonetheless these steps provide decision points that are 'important stages for the input of formal legal advice'.[140] For testing to be meaningful, critical issues of performance must be translated into testable elements that can be objectively measured. While many smaller nations might be little more than purchasers of off-the-shelf weapons,[141] other governments are involved in envisaging, developing, and testing emerging weapons technology. While the degree of that involvement will vary, that is a choice for governments.[142] So, rather than being passive recipients of test results and other weapons data, one pro-active step that could be taken as part of the legal review process is for lawyers to input into the test and evaluation phases by identifying areas of legal concern that could then be translated into testable elements. This may be one way to at least partly address the security and compartmented access difficulties associated with high-technology weapons that were raised above. For example, it is appropriate to assign increased confidence in reliability for military applications involving higher risks factors for civilians. This could be cross-referenced against existing weapons system reliability data as an input to the decision-making process when determining whether a new targeting procedure may be considered lawful.

To be effective, the legal requirements need to be expressed in terms that are 'testable, quantifiable, measurable, and reasonable'.[143] Part of the challenge will

139 J. McClelland, above note 1, p. 401. Or during design, during initial acceptance, and as part of operational evaluation.

140 *Ibid.*, p. 402.

141 Of course, purchasers of off-the-shelf weapon systems must still satisfy themselves of the legality of a weapon. Even with a fully developed and tested weapon, this can still prove difficult for purchasers of high-technology weapons. For example, a manufacturer may refuse to disclose sufficient information about a high-technology weapon that uses encrypted proprietary software for the end-user to make an informed judgement about the algorithms used to be confident of the weapon's ultimate reliability.

142 See *Report on the Defense Science Board Task Force on Developmental Test & Evaluation*, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, US Department of Defense, May 2008, pp. 6–7, available at: www.acq.osd.mil/dsb/reports/ADA482504.pdf; wherein the recent decrease in US government involvement in design testing was highlighted, and perhaps more worryingly, government access to the contractor's test data was limited.

143 *Ibid.*, p. 38. Noting that this might initially be challenging. For example, *ibid.*, p. 39, for a discussion of where this has not occurred for the operational requirements.

be bridging the disconnect that often exists between the definitions of technical requirements and the desired operational performance. This disconnect can often be 'traced to the terminology used to define the level of performance required, under what conditions and how it is [to be] measured'.[144] This is where lawyers working with systems engineers can influence the process so that the use of tests, demonstrations, and analysis can be adopted as valid methods to predict actual performance.

Once a system is in-service, further testing may also be conducted to gain additional insights into the capability and to ensure that the system is actually meeting the requirements of the user. This phase of test and evaluation is particularly critical as it is the only phase that truly relates to the 'real world' use of a system.[145] By having lawyers provide meaningful legal criteria against which a class of weapons could be judged, the ongoing legal compliance of that weapon could be factored into an already existing process. Another area for useful input is evaluation and analysis of system and subsystem integration and interaction. When it comes to a system-of-systems, US military experience is that there is no 'single program manager who "owns" the performance or the verification responsibility across the multiple constituent systems, and there is no widely used adjudication process to readily assign responsibility for [system-of-systems] capabilities, with the exception of command and control systems'.[146] Compare this to other industries such as leading automotive companies that have highly sophisticated design, production, testing, and quality-approval processes for every component that goes into a vehicle and a resulting detailed assignment of responsibility by component, system, and whole product (comprising multiple systems). Working with systems engineers, layers of quality control process could identify the critical legal issues that require both testing and assignment of responsibility (for example, in case of non-compliance with international humanitarian law) among the weapon manufacturer and the various military stakeholders.

Reliability and automatic target recognition

Weapons that are designed to explode but fail to when used operationally, and if left on the field after the cessation of hostilities, are known as explosive remnants of war.[147] Indeed, munition reliability is even defined as 'a measure of the probability of successful detonation'.[148] Due to the effects on the civilian population of unexploded ordnance, legal regulation already exists in this area.[149] Less well

144 *Ibid.*, p. 41.

145 For example, there is anecdotal evidence that some weapon failures arise due to 'operational factors that are not assessed as part of the developmental, acceptance and surveillance testing', *Defense Science Board Task Force on Munitions System Reliability*, above note 132, p. 17.

146 *Report on the Defense Science Board Task Force on Developmental Test & Evaluation*, above note 142, p. 43.

147 See *Defense Science Board Task Force on Munitions System Reliability*, above note 132, p. 10.

148 *Ibid.*, p. 14.

149 For example, see the chapter on 'Unexploded and abandoned weapons', in W. Boothby, above note 45, pp. 297–317.

understood is that weapons reliability associated with automatic target recognition has another important aspect. It is not just about a weapon that does not explode, but also about one that selects the wrong target.

Here we are trying to determine whether it is reasonable to conclude from the analysis of reconnaissance data that the target possesses certain enemy properties or characteristics, and when is it reasonable to reach such a conclusion. Suppose the difference between the hypothesized enemy characteristic and the reconnaissance measurements is neither so large that we automatically reject the target, nor so small that we readily accept it. In such a case, a more sophisticated statistical analysis, such as hypotheses testing, may be required. Suppose that experience indicates that a 90 per cent match in reconnaissance data with existing information regarding an enemy target type has proven to be a reliable criterion for confirming an enemy target. If the data was a 100 per cent match or a 30 per cent match we could possibly come to an acceptable conclusion using common sense. Now suppose that the data match was 81 per cent, which may be considered relatively close to 90 per cent, but is it close enough to accept as a lawful target? Whether we accept or reject the data as a lawful target, we cannot be absolutely certain of our decision and we have to deal with uncertainty. The higher we set our data-match acceptance criterion the less likely an automatic target recognition system will identify non-targets as lawful targets, but the more probable that the recognition system will fail to identify lawful targets as being lawful targets.[150]

The desired level for whether or not a weapon explodes might be a 'reliable functioning rate of 95 per cent'.[151] This corresponds to an autonomous weapon system that fires at an unlawful target, due to misclassification as 'lawful', one out of every twenty times. Would this be considered acceptable performance for discriminating between lawful and protected targets? So, when a weapon system is looked at in this way, the better definition for reliability is whether the weapon system 'performs its intended function'[152] and as the 'fuzing and guidance capabilities become more integrated, the reliability of target acquisition must be measured and assessed'.[153] It has been suggested that what is required is a 'very high probability of correct target identification . . . and a very low probability of friendly or civilian targets being incorrectly identified as valid (i.e., enemy) targets'.[154] As there is an inherent trade-off between sensitivity and specificity, consideration also needs to be given to how a weapon will be employed. If a human provides go/no-go authorization based on an independent review, therefore providing additional safeguard against false recognition, then a greater number of false positives generated by the automatic recognition system may be acceptable. However, if the weapon system is autonomous, combat effect (correct employment against

150 See *Defense Science Board Task Force on Munitions System Reliability*, above note 132, p. 28.
151 *Ibid.*, p. 11. Even this level of reliability is based on controlled conditions and a lower level is allowed in operational conditions to account for 'environmental factors such as terrain and weather', *ibid.*, Appendix III, *DoD Policy Memo on Submunition Reliability*, p. 1.
152 *Ibid.*, p. 14.
153 *Ibid.*, p. 16.
154 *Ibid.*, p. 23.

identified enemy targets) must be more carefully balanced against risk to civilians. Noting that one of the purposes of automated and autonomous systems is to undertake high-volume observations that would overwhelm a human operator, where 'observations [are] in the millions . . . even very-low-probability failures could result in regrettable fratricide incidents'.[155] Confidence in the ability of an autonomous system to work in the real world might be developed by deploying such systems in a semi-autonomous mode where a human operator has to give the final approval for weapons release.[156] Rigorous post-mission analysis of data would allow, with time, a statistically significant assessment of the reliability of the system to correctly identify lawful targets.

A final point on testing:

> Achieving these gains [capability increases, manpower efficiencies, and cost reductions available through far greater use of autonomous systems] will depend on development of entirely new methods for enabling 'trust in autonomy' through verification and validation (V&V) of the near-infinite state systems that result from high levels of adaptability and autonomy. In effect, the number of possible input states that such systems can be presented with is so large that not only is it impossible to test all of them directly, it is not even possible to test more than an insignificantly small fraction of them. Development of such systems is thus inherently unverifiable by today's methods, and as a result their operation in all but comparatively trivial applications is uncertifiable.
>
> It is possible to develop systems having high levels of autonomy, but it is the lack of suitable V&V methods that prevents all but relatively low levels of autonomy from being certified for use. Potential adversaries, however, may be willing to field systems with far higher levels of autonomy without any need for certifiable V&V, and could gain significant capability advantages over the Air Force by doing so. Countering this asymmetric advantage will require as-yet undeveloped methods for achieving certifiably reliable V&V.[157]

A distinctly separate consideration from weapons testing is weapons research. Should weapons research (as opposed to development) be limited or constrained by legal issues? Generally, there is no legal reason (budgets aside) why research cannot take potential weapons as far as the bounds of science and engineering will allow, not the least of which is because laws change.[158] The time for imposing limits based on law is in the production and employment of weapons. Of course, some may, and

155 See *Report of Defense Science Board Task Force on Patriot System Performance: Report Summary*, above note 60, p. 2.

156 See A. Myers, above note 23, pp. 91–92.

157 US Air Force, 'Technology horizons', available at: http://www.af.mil/information/technologyhorizons.asp (last visited 6 May 2012).

158 See the examples of submarines and airplanes referred to in Anderson and Waxman, above note 29, pp. 6–7. While some aspects of international humanitarian law may change, this presumably does not extend to the cardinal principles of distinction, proportionality, and unnecessary suffering.

do, argue differently on moral and ethical lines.[159] That is where such arguments are best made and debated.

Conclusion

With the ever-increasing technological complexity of weapons and weapon systems, it is important that, among others, computer scientists, engineers, and lawyers engage with one another whenever a state conducts a review of weapons pursuant to Article 36 of the Protocol Additional to the Geneva Conventions of 12 August 1949 and relating to the Protection of Victims of International Armed Conflicts (API).[160] The reviews cannot be compartmentalized, with each discipline looking in isolation at their own technical area. Rather, those conducting legal reviews will require 'a technical understanding of the reliability and accuracy of the weapon',[161] as well as how it will be operationally employed.[162] While that does not mean lawyers, engineers, computer science experts, and operators need to each be multidisciplined, it does mean that each must have enough understanding of the other fields to appreciate potential interactions, facilitate meaningful discussion, and understand their own decisions in the context of impacts on other areas of development.

Those who develop weapons need to be aware of the key international humanitarian law principles that apply to the employment of weapons. Lawyers providing the legal input into the review of weapons need to be particularly aware of how a weapon will be operationally employed and use this knowledge to help formulate meaningful operational guidelines in light of any technological issues identified with the weapon in terms of international humanitarian law. Furthermore, all parties require an understanding of how test and validation methods, including measures of reliability, need to be developed and interpreted, not just in the context of operational outcomes, but also in compliance with international humanitarian law.

As the details of a weapon's capability are often highly classified and compartmentalized, lawyers, engineers, and operators may need to work co-operatively and imaginatively to overcome security classification and compartmental access limitations. One approach might be to develop clearly expressed legal

159 See Matthew Bolton, Thomas Nash and Richard Moyes, 'Ban autonomous armed robots', Article 36, 5 March 2012, available at: http://www.article36.org/statements/ban-autonomous-armed-robots/ (last visited 6 May 2012): 'Whilst an expanded role for robots in conflict looks unstoppable, we need to draw a red line at fully autonomous targeting. A first step in this may be to recognize that such a red line needs to be drawn effectively across the board – from the simple technologies of anti-vehicle landmines (still not prohibited) across to the most complex systems under development. This is not to ignore challenges to such a position – for example, consideration might need to be given to how automation functions in missile defence and similar contexts – but certain fundamentals seem strong. Decisions to kill and injure should not be made by machines and, even if at times it will be imperfect, the distinction between military and civilian is a determination for human beings to make'.

160 See P. Spoerri, above note 54.

161 K. Lawand, above note 1, pp. 929.

162 ICRC, *A Guide to the Legal Review of New, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977*, above note 1, pp. 17–18.

parameters that can be the subject of meaningful systems testing. Another approach may be to devise multi-parameter acceptance criterion equation sets. Such equation sets would allow for hypothesis testing while factoring in reliability data, confidence levels, and risk factors using input data such as anticipated military advantage, weapon reliability data, reconnaissance measurement uncertainty, and civilian risk factors.