

**Autonomous Weapon Systems In International Humanitarian Law – Simply Just
Another Weapon System.**

LLM RESEARCH ESSAY

LAWS 533: LAW OF ARMED CONFLICT

Philippus Jacobus Jooste

Abstract

This paper reviews the arguments made by the opponents of the utilization at all of so called autonomous weapons (killer robots), on the basis of alleged inherent non compliance with certain cardinal principles of IHL, against the argument that said weapons or systems, and their use, can be satisfactorily accommodated under existing IHL. It gives an overview of the issues identified by the opponents and the main arguments made in favour of banning autonomous weapons. The author looks at the characteristics of two weapon systems and concludes there are already fully autonomous weapon systems in non contentious use. The author submits that the concerns of the banning authors are not rational, the purported distinctions made between certain weapon systems cannot factually and consistently be maintained and are based on the averred existence of categories of weapons that do not exist. Finally, the author argues that the opponents' concerns and arguments are in the main and based on the normal uncertainties, inherent in all factually to be determined situations and also on slippery slope reasoning. The author emphasises the established rules of IHL and opines that the same are adequate to regulate the so called autonomous weapons and weapon systems and any liability arising from their use.

Word length

The text of this paper (excluding abstract, table of contents and bibliography) comprises 7819 words.

Subjects and Topics

International humanitarian law
Distinction
Proportionality
Military Necessity
Fully autonomous weapons
Accountability
Robots

CONTENTS (REFERENCES – TABLE OF CONTENTS)

<i>Abstract</i>		ii
CONTENTS		iii
<i>I Introduction</i>		1
<i>II Fully Autonomous and Automatic Weapon Systems: a Legal Category?</i>		2
<i>A Definitions, Distinctions or Features of Autonomous Weapon Systems</i>		3
<i>B The Difference between Automatic and Autonomous Weapon Systems in Fact</i>		3
<i>C Definition of a Fully Autonomous Weapon System</i>		4
<i>III Fully Autonomous Weapon Systems in Current Use</i>		5
<i>IV The Objections Raised by the Opponents</i>		7
<i>A The Legal Objections</i>		7
<i>B The Principles of Distinction, Proportionality and Military Necessity</i>		7
<i>C Arguments against the Objections Raised</i>		8
<i>V Accountability</i>		9
<i>A Military Commander Responsibility</i>		10
<i>B Programmer's Liability</i>		12
<i>C Manufacturer's Liability</i>		13
<i>VI Conclusion</i>		14
Bibliography		15

I Introduction

The unarguable and some say inexorable development of technologically more advanced weapons and weapon systems, in the main due to the general growth in the computing power of machines, has left some concerned with one of the foreseen and submitted inevitable outcomes of this development: fully autonomous weapons or weapon systems – or killer robots.¹

As a broad statement autonomous weapons systems are seen as robots, meaning they have to varying degrees independence or autonomy from human control or supervision. The foremost concern is that these autonomous systems and specifically fully autonomous weapon systems will kill humans, generally civilians, in breach of International Humanitarian Law (IHL).² Closely linked with this concern is the apparent issue of accountability in foreseen cases where these weapon systems and specifically fully autonomous ones breach IHL.³

The opponents of autonomous weapon systems say there are not yet fully autonomous systems, but before there are they should be banned.⁴ (This paper refers to this group as the opponents.)

Against the opponents stand those who broadly hold the view that these autonomous weapon systems, including fully autonomous ones, can satisfactorily be accommodated within the current framework of IHL.⁵

They firstly agree with the opponents, that there are in fact not yet any fully autonomous weapons systems in use,⁶ but then part ways in arguing that these weapons systems will in any event be able to comply in certain cases with IHL and that such compliance will be sufficient.⁷ Nor do they see any issue with accountability; they say current IHL and international criminal law principles suffice.⁸

This paper looks at the different definitions or features of autonomous weapon systems, then analyses whether certain weapon systems do not already in fact fall under said definitions, concludes that indeed they do, and consequently submits that the distinction made

¹ I will as far as possible only refer to weapon systems herein and not continuously and cumbersomely refer to ‘weapons and weapon systems’, as usually autonomous weapons are indeed more than just a weapon, but are indeed a system, with one of its components an actual weapon.

² Human Rights Watch Report “Losing Humanity: The Case against Killer Robots” (November 2012) Human Rights Watch <www.hrw.org> at 2 [HRW report]; Hun –Yan Liu “Categorization and legality of autonomous and remote weapon systems” (2012) 12 ICRCR No 886 at 628; Peter Asaro “On banning autonomous weapons systems: human rights, automation and the dehumanization of lethal decision making” (2012) 12 ICRCR No 886 at 690 ; and Noel E Sharkey “The evitability of autonomous robot warfare” (2012) 12 ICRCR No 886 at 788.

³ Asaro, above n 2, at 689; and HRW report, above n 2, at 42–43.

⁴ Asaro, above n 2, at 689; Sharkey, above n 2, at 791; and HRW report, above n 2, at 42–43.

⁵ Michael N Schmitt “*Autonomous weapon systems: A reply to critics*” (2012) Harv Nat’l Sec J at 2 [Schmitt]; Michael N Schmitt and Jeffrey S Thurner “*Out of the Loop: Autonomous Weapon Systems and the Law of Armed Conflict*” (2012) 4 Harv Nat’l Sec J 213; Kenneth Anderson and Matthew Waxman “*Law and Ethics for Robot Soldiers*” [2012] Policy Rev 35 at 46 ; and William Boothy “Some legal challenges posed by remote attack” (2012) 12 ICRCR No 886 at 582 and 588.

⁶ Schmitt, above n 5, at 5; and Anderson and Waxman, above n 5, at 45.

⁷ Schmitt, above n 5, at 4 and 14; Boothy, above n 5, at 588; and Anderson and Waxman, above n 5, at 46.

⁸ Schmitt, above n 5, at 4; and Anderson and Waxman, above n 5, at 48.

by both groups between autonomous weapon systems and fully autonomous weapon systems is one of expediency not fact. The paper thus takes issue with the current consensus view that there are not as yet fully autonomous weapons systems by concluding there are in fact, in all but name, already fully autonomous weapon systems.

Also it submits that the opponents are wrong and that in certain instances (and that is all they legally have to do) fully autonomous weapon systems do comply with IHL.

The other serious concern raised by the opponents is accountability.⁹ The paper looks at two case situations where the possible accountability of humans in using advanced weapon systems was involved and argues that these two concrete situations show that the accountability elements and principles are the same as in any other situation under IHL or international criminal law.

Finally, the paper submits that existing IHL principles can satisfactorily deal with these weapon systems.

II Fully Autonomous and Automatic Weapon Systems: a Legal Category?

At the outset it must be stated that there does not exist in IHL any category of weapon or weapon system called an automatic or autonomous weapon or system, nor is there a prohibited weapon or weapon system so called. Associated with this absence, IHL also does not have a general prohibition on the development and use of ‘new’ weapons or weapons systems such as ‘autonomous weapon systems’.¹⁰

Article 36 of Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Armed Conflicts of International Armed Conflicts 1125 UNTS 3 (opened for signature on 12 December 1977, entered into force 7 December 1978), (AP1) deals with new weapons and weapon systems and reads: ‘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’. I do not address this article herein, and accept the article is part of customary international law¹¹ – as to which there does seem to be some debate¹² – and hence binds all countries not just the contracting parties to AP 1. But even though the article is applicable, it does not present an answer to the issues herein. It imposes an obligation on a state to review weapon systems for their legality as such within IHL, but it does not answer the question posed by weapon systems that are not de facto prohibited or illegal per se, or which have passed art 36 review, including of course autonomous or fully autonomous weapon systems, but which could still in some cases be used unlawfully.¹³

⁹ Asaro, above n 2, at 689; Sharkey, above n 2, at 790; and HRW report above n 2, at 42–43.

¹⁰ Antonio Cassese, Guido Acquaviva, Mary D Fan and Alex Whiting *International Criminal Law, Cases and Commentary* (Oxford University Press, 2011) at 134.

¹¹ Schmitt, above n 5, at 14 and 28; and see HRW report, above n 2, at 21.

¹² Schmitt, above n 5 at 28; and HRW report, above n 2, at 21.

¹³ Schmitt, above n 5, at 8, 14 and 32–33.

Nevertheless, it is universally accepted, new, including fully autonomous, weapons and weapon systems must be able to be used in compliance with the ‘cardinal principles of humanitarian law’ including the principles of distinction and proportionality.¹⁴

A Definitions, Distinctions or Features of Autonomous Weapon Systems

A distinction is often made in the literature between automatic weapon systems (including remote controlled ones, the notorious example being unmanned aerial vehicles (UAVS) or drones) and autonomous weapon systems.¹⁵ However, remote control weapons systems are also sometimes included among autonomous weapons systems.¹⁶ The ICRC distinguishes between different stages of autonomy within weapon systems by naming three separate categories: remote controlled weapon systems, automated weapons systems, and autonomous weapons systems.¹⁷

Autonomous weapon systems again are divided by some into fully autonomous or not.¹⁸ Human Rights Watch calls all these weapon systems robots and divides them into: ‘Human-in-the-Loop Weapons: Robots that can select targets and deliver force only with a human command; Human-on-the-Loop Weapons: Robots that can select targets and deliver force under the oversight of a human operator who can override the robots’ actions; Human-out-of-the-Loop Weapons: Robots that are capable of selecting targets and delivering force without any human input or interaction.’¹⁹

In conclusion: There is thus no agreed categorisation or division of autonomous weapon systems, nor any scientific demarcation of these weapon systems. The making of these distinctions by the different parties is really for convenience, made by the parties themselves and based on specific individual opinions as to categorisation. All the same, it is submitted by opponents of autonomous weapons to be a real factual one (this is dealt with below). These distinctions are not necessitated by any legal requirement. The paper refers to autonomous and fully autonomous weapon systems as such only for the sake of convenience.

B The Difference between Automatic and Autonomous Weapon Systems in Fact

Automatic and remote control weapon systems are not really contentious in IHL, autonomous and specifically fully autonomous weapon systems are. They are seen by the opponents as having a ‘seperateness’²⁰, from the other two, being a ‘quantitative leap up’²¹, a ‘new and novel means and method of armed conflict’²². The author disagrees: these are still

¹⁴ Marina Castellaneta “New weapons, old crimes?” Fausto Pocar, Marco Pedrazzi and Michaela Frulli (eds) *War crimes and the conduct of Hostilities* (Edward Elgar, Cheltenham, 2013) 194 at 200–201; *Advisory Opinion on the legality of the Threat or Use of Nuclear Weapons* [1986] ICJ Rep at 78–79; Schmitt, above n 5, at 14; and Jean-Marie Henckaerts and Louis Doswald Beck *Customary International Humanitarian Law*, Volume 1: Rules (Oxford University Press, 2005) at 237 [ICRC Rules].

¹⁵ See HRW report, above n 2, at 2; and Asaro, above n 2, at 690.

¹⁶ See HRW report, above n 2, at 2; and Asaro, above n 2, at 690.

¹⁷ ICRC ‘International Humanitarian Law and the Challenges of Contemporary Armed Conflicts’ 31st International Conference of the Red Cross and Red Crescent, Geneva: Switzerland, 28 November–1 December 2011, 31IC/11/5.1.2, at 38–40.

¹⁸ See generally PW Singer *Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century* (Penguin Press, New York, 2009) at 67 and 124.

¹⁹ HRW report, above n 2, at 2.

²⁰ Hun–Yan Liu, above n 2, at 628.

²¹ Asaro, above n 2, at 693.

²² Hun–Yan Liu, above n 2, at 628.

only weapon systems, which are currently seen – only because they are new – as qualitatively different. Guided missiles, submarines, aircraft, machine guns were all new at some stage, yet were and are dealt with under the existing or developed IHL.²³

So this paper is in agreement with those who see autonomous weapon systems as just another weapon, technologically more advanced, but simply further along the range of technologically advanced weapons.²⁴

Automatic or automated weapons, it is stated, are ‘unsupported systems or processes that involve repetitive, structured, routine operations’²⁵, whereas the features of autonomous weapons that underpin the arguments for banning autonomous weapons are that these weapons ‘operate in dynamic, unstructured, open environment, and act on a feedback information from several sensors’.²⁶ Peter Asaro writes ‘We can thus define an ‘autonomous weapon system’ as any system that is capable of targeting and initiating the use of potentially lethal force without direct human supervision and direct human involvement in lethal decision-making’.²⁷

C Definition of a Fully Autonomous Weapon System

For purposes of this article, and then inferring from the above, it can be accepted that to be a fully autonomous weapon system, a weapons system cumulatively must be:

- 1) a weapon or weapon system;
- 2) sensually capable of selecting a target;
- 3) capable of initiating the use of force;
- 4) able to do so without direct human supervision ; and
- 5) do so without direct human involvement in the lethal decision making.

As an observation; it is conceded by all opponents of autonomous weapon systems that all of the categories of weapons however demarcated ‘follow algorithmic’²⁸ instructions that are fixed and determinate’.²⁹ This acceptance of the factual basis for the working of these weapon systems might be seen to be determinative of the question whether autonomous weapons are truly completely different, as this common and all encompassing factual characteristic with automatic weapon systems shows they are not.

The opponents’ and consensus view is that there are at present no fully autonomous weapon systems in existence.³⁰ And this is so, as there is perceived non fulfilment of conditions (4) and (5) above; in that, with current autonomous weapon systems, there is still direct human supervision and direct human involvement in the lethal decision making.³¹

²³ Anderson and Waxman, above n 5, at 45. And see James J Busittil *Naval Weapon Systems and the Contemporary Law of War* (Caledon Press, Oxford, 1998) at 191–192.

²⁴ Boothy, above n 5, at 582 and 588; and Marina Castellaneta, above n 14, at 209 - 210. And see Anderson and Waxman, above n 5, at 37–38.

²⁵ Asaro, above n 2, at 690 (n 5); and Hun –Yan Liu, above n 2, at 632.

²⁶ Asaro, above n 2, at 690 (n 5); and see Singer, above n 18, at 67.

²⁷ Asaro, above n 2, at 690. And see Schmitt, above n 5, at 4.

²⁸ Mathematically and logically predestined (my definition).

²⁹ Hun –Yan Liu, above n 2, at 628 (n 2); and Asaro, above n 2, at 692 and 694.

³⁰ Asaro, above n 2, at 689; Sharkey, above n 2, at 788; HRW report, above n 2, at 42– 43; Schmitt, above n 5, at 2 and Anderson and Waxman, above n 5, at 3–5.

³¹ Asaro, above n 2, at 689; Sharkey, above n 2, at 788.; HRW report, above n 2, at 42–43; Schmitt, above n 5,

But, the opponents say that since development of automatic and autonomous weapon systems are progressing towards the absence of human supervision and involvement, this issue needs to be addressed before there are.³²

It is the author's view that there are indeed fully autonomous weapon systems already in use, two of which are discussed below, but because these systems are currently only being used defensively, and with (albeit generally token) human override supervision, against what are clearly armed attacks by military objects, with no real issue as to non compliance with the 'cardinal principles of proportionality or distinction', these weapons are flying below the radar so to speak and not expressly identified as fully autonomous or robotic.

As it is the presence of the human as 'supervisor and over rider' that allegedly make an otherwise fully autonomous weapon system not autonomous, this is too addressed below.³³

III Fully Autonomous Weapon Systems in Current Use

Without being exhaustive, two weapons systems may be mentioned: the Israeli Iron Dome Air Defence missile system³⁴ and the MBDA Sea Wolf seaborne air defence missile system.³⁵ Both these systems are used against inter alia hostile missiles, guided and unguided, ranging from relatively slow and early identified rockets to guided missiles approaching the protected area at three times the speed of sound.³⁶ If hostile supersonic missiles are fired from close range, the reaction time of a missile defence system such as Iron Dome or Sea Wolf, from first identifying the hostile projectile to launching itself, is measured in a few seconds.³⁷ No human is able to, and indeed attempts to 'use' Iron Dome or Sea Wolf to shoot down the incoming missile in these circumstances: Iron Dome and Sea Wolf react and act automatically, from indentifying the hostile missile, tracking and targeting it, and launching and guiding its own missile against the incoming missile.³⁸

The human on the loop has in many situations with the use of these two systems only a 'split second' to veto or stop the launching of the reacting missile.³⁹ The human operator plays and has no other role to play in this process and many times indeed does not do anything at all.⁴⁰

From the aforementioned, it is clear that in instances where the human does not, in that split second, or at all, override the weapon system it acted wholly autonomously. To then

at 2; and Anderson and Waxman, above n 5, at 5 and 6.

³² Asaro, above n 2, at 693 – 694; HRW report, above n 2, at 8; and see generally Sharkey, above n 2, at 788.

³³ HRW report, above n 2, at 2.

³⁴ Rafael Advanced Defense Systems Ltd "Iron Dome: Defense against Short Range Artillery Rockets," 2010, <www.rafael.co.il> (accessed 18 September 2014).

³⁵ MBDA <www.mbda-systems.com/products> (accessed 18 September 2014).

³⁶ See ns 34 and 35 above.

³⁷ In such circumstances these systems are most likely set in automatic or autonomous or casualty mode; meaning they execute all actions themselves. See generally Singer at 124 – 125; Schmitt, above n 5, at 4, n11; and Anderson and Waxman, above n 5, at 37 and at 39, n 15.

³⁸ HRW report, above n 2, at 9. In effect, a summary of the involvement of a human operator in automatic defence systems like Iron Dome. See Schmitt, above n 5, at 5, n 11 Also see Defense Update "Iron Dome Battle Management Demonstrated" (2009), <www.defenseupdate.com> accessed 20 September 2014).

³⁹ HRW report, above n 2, at 11.

⁴⁰ HRW report, above n 2, at 11; and see Schmitt, above n 5, at 5, n 11.

argue it is not a fully autonomous system, although it is in fact agreed to be one, but it is sometimes not used that way, is logically and, even more importantly, factually wrong. This view is implicitly supported by Human Rights Watch who agrees that current and human-on-the-loop weapon systems, including Iron Dome, are *de facto* fully autonomous – “because the supervision is so limited.”⁴¹

It is noteworthy that the opponents, and even Schmitt, do not attempt to clearly articulate the reason why they say that ‘fully automatic’ or human supervised autonomous systems, like Iron Dome, are in fact not fully autonomous. All these authors seem prepared to simply accept that the mere presence of the over rider (in some cases) constitutes the feature or characteristic which removes systems like Iron Dome from the realm of fully autonomous and back into that of autonomous without addressing the fact that these systems can and sometimes do act fully autonomously.

It may be mentioned here that Iron Dome was very successfully used in August 2014 by Israel in defence to attacks by Hamas with non guided rockets, as well as mortars, during the latest Israeli incursion into Gaza . It is reported that Iron Dome shot down between eighty five to ninety percent of incoming rockets against which it was launched.⁴²

Some opponents call these two systems simply automated and not autonomous, despite the whole sequence from target acquisition to engagement complying with the requirements above and they being regulated by algorithmic, robotic processes.⁴³ They argue that there is not truly autonomous decision making by the weapon system, since there is a human supervisor.

The author submits that the opponents call these systems automated, rather than fully autonomous, to avoid accepting the necessary legal and factual conclusion following from the above, that if indeed Iron Dome and Sea Wolf are capable of being used fully autonomously, as robots in other words, and are being used with sufficient compliance with IHL in the circumstances in which they are used, then necessarily following from that at least some fully autonomous weapon systems (robots) do comply with IHL in certain conditions. This conclusion will seriously undermine, if not be fatal, to both strands of their point of view: that it is not too late to ban fully autonomous weapon systems – as there are not yet any in existence (implicitly, if there had been it might have been too late or would have been more difficult to ban them), and these weapons cannot be used in compliance with IHL – this issue is examined in detail below.

Concluding then, there are already in existence and in use fully (in all but name) autonomous weapon systems (robots) functioning in compliance with IHL in certain circumstances.

To be clear, this paper does do not argue as a result that all fully autonomous weapon

⁴¹ HRW report, above n 2, at 2.

⁴² Alon Ben David “*Iron Dome blunts 90% of enemy rockets*”. Aviation Week and Space Technology (1 September 2014) <www.aviationweek.com.>

⁴³ The HRW report, above n 2, at 13 calls them autonomous but not fully so because there is still a human on the loop. But see n 41 above regarding the HRW report’s acceptance of the *de facto* fully autonomous status of these weapons. But see Asaro, above n 2, at 693 – 694; and Sharkey, above n 2, at 788.

systems can in all circumstances be used in compliance with IHL.⁴⁴ It only contents that the alleged factual basis for distinguishing between autonomous and fully autonomous weapon systems in IHL is wrong and thus bad. Both are robots.

IV The Objections Raised by the Opponents.

Several objections are made against specifically fully autonomous weapon systems. These can be divided into legal (alleged inability to comply with IHL) and ethical (moral) ones.⁴⁵ I only deal with the legal objections herein.

A The Legal Objections

- Autonomous weapons systems are in effect the first non human combatants by virtue of the alleged fact that they themselves will function as actors in armed conflict; deciding to kill by themselves and then killing humans.⁴⁶ (As this objection posits the recognition in IHL of a completely new factual and legal category of combatants – robots, (all combatants under IHL so far have been only human) and not hitherto recognized under IHL, it is an interesting but simply speculative objection. Hence it is not dealt with herein.)
- Autonomous systems as a means of warfare will never be able to discriminate between combatants and non combatants.⁴⁷
- Autonomous systems as method of warfare will probably not ever be able to correctly apply the proportionality requirement of decision making on an armed attack.⁴⁸
- These systems are not able and will probably never be able to judge correctly the military necessity of an attack.⁴⁹ That is, whether in a military sense it is necessary to attack a specific target, weighing up the requirement of a definite military advantage in attacking a specific military object, also killing combatants, against the judgement of something being a military object in the specific circumstances.⁵⁰
- The problem of accountability: who is to be held legally accountable in the case of a breach of IHL by an autonomous weapon system?⁵¹

B The Principles of Distinction, Proportionality and Military Necessity

The above objections are then the perceived issues, as the opponents argue that autonomous weapon systems will not be able to be used in compliance with the very important, even called ‘cardinal’ rules of IHL. It is useful to briefly say what the above objections then signify in IHL before addressing the foreseen non compliance with these rules of IHL.

⁴⁴ Asaro, above n 2, at 690. And see Schmitt, above n 5, at 8 and 18 -19.

⁴⁵ Asaro, above n 2, at 691.

⁴⁶ Hun –Yan Liu, above n 2, at 629; and Asaro, above n 2, at 692. See also HRW report, above n 2, at 3 and 42 – 43; and see generally Nils Melzer *Targeted Killing in International Law* (Oxford University Press, 2008) at 299 onwards.

⁴⁷ Asaro, above n 2, at 692; Sharkey, above n 2, at 789; and HRW report, above n 2, at 3, 24–25 and 30–1.

⁴⁸ Asaro, above n 2, at 691; HRW report, above n 2, at 32–34; and Sharkey, above n 2, at 790.

⁴⁹ Asaro, above n 2, at 690; HRW report, above n 2, at 34–35; and Sharkey, above n 2, at 789.

⁵⁰ Asaro, above n 2, at 692; HRW report, above n 2, at 34–35; and Sharkey, above n 2, at 790.

⁵¹ Asaro, above n 2, at 689; and HRW report, above n 2, at 42–43.

The basic IHL rule of distinction is expressed as follows: ⁵²

In order to ensure respect for and protection of the civilian population and civilian objects, the Parties to the conflict shall at all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations against military objectives.

In order to comply with the requirement of proportionality under IHL, an attack is prohibited where the attack: ⁵³

May be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete military advantage anticipated.

Military necessity in the context of using autonomous weapons systems will generally be linked with the principle of distinction,⁵⁴ as military necessity means in this context whether there is a IHL sanctioned need to attack a specific target, be it a military objective or a combatant. AP1 article 52(2) reads: ⁵⁵

Insofar as objects are concerned, military objectives are limited to those objects which by their nature, location, purpose or use make an effective contribution to military action and whose total or partial destruction, capture or neutralization in the circumstances ruling at the time, offers a definite military advantage.

C Arguments against the Objections Raised

This paper agrees that generally fully autonomous weapon systems inherently cannot discriminate between combatants and non-combatants, persons hors de combat, wounded etc. But this is a straw man argument. Most weapons (if not all by their nature of being per definition a tool) cannot do so by themselves; ⁵⁶ the humans who use any weapon must discriminate as to when and under what circumstances that weapon may be used have to comply with this requirement.⁵⁷

One has to agree with Schmitt who writes “Self evidently, it would be unlawful to use an autonomous weapon system to directly attack civilians or civilian objects.” ⁵⁸

And also, assuming fairly that a weapon system itself in order to be IHL complaint must sometimes as a means of warfare be able to discriminate, it need not always be able to do so. In some circumstances, it only needs to be able to distinguish, or even just be able to identify

⁵² Art 48 AP 1; Melzer, above n 45, at 301 and ICRC Rules, Rules 1,6,7,10,11,12 and 15. And see Leslie C Green *The Contemporary Law of Armed Conflict* (3 ed, Manchester University Press, 2008) at 185; Ian Henderson *The Contemporary Law of Targeting* (Martinus Nijhoff, Leiden, 2009) at 44 – 45 and Héctor Olásolo *Unlawful Attacks in Combat Situations* (Martinus Nijhoff, Leiden, 2008) at 17.

⁵³ Art 57(2)(a)(iii) AP 1 and ICRC Rules, Rule 14. And see Henderson, above n 51, at 45; and Olásolo, above n 51, at 19.

⁵⁴ Above n 31 and n 32; arts 51(4) and 51(5) (b) AP1; ICRC Rules, Rules 1 and 6 and Schmitt, above n 2, at 14.

⁵⁵ See ICRC Rules, Rule 14. And see Henderson, above n 51, at 45 and Olásolo, above n 51, at 103–153.

⁵⁶ Schmitt, above n 2, at 2 and 11; and Yoram Dinstein *The Conduct of Hostilities Under the Law of International Armed Conflict* (2nd ed, Cambridge University Press, 2010) at 62.

⁵⁷ Schmitt, above n 2, at 2, 11 and 18; and Dinstein, above n 55, at 62 – 63.

⁵⁸ Schmitt, above n 2, at 16.

an enemy weapon or weapon system, not combatants, nor civilians.⁵⁹ Schmitt refers to the targeting of a hostile tank or tanks on the battlefield.⁶⁰ I prefer to look at Iron Dome again; all it needs to do to be compliant with IHL, in the factual circumstances in which it is used, is to recognise incoming rockets, artillery shells, mortar shells and target the same. No arguments have been made that Iron Dome as used as a defensive weapon system recently was in breach of this principle.

The same answer and arguments in principle and in fact apply as to the lack of ability on the part of autonomous weapon systems themselves to apply the principle of proportionality and in judging military necessity. Humans must apply the principle; humans can work with and judge when and how, with what degree of autonomy to apply the robots.⁶¹

Also again, in specific circumstances where there is no human involvement, as with Iron Dome for example being used against enemy weapons, or other autonomous weapon systems on the conventional battlefield, outside of an urban area, one with tanks, armoured vehicles, missiles and conventional forces, it is generally not necessary for an autonomous weapon system (robot) to apply proportionality.⁶²

Military necessity as a requirement to be complied with is answered in the same way. Fully autonomous weapon systems would not be unlawful per se, as is stated by Schmitt “because it is clear that autonomous weapon systems may be used in situations in which they are valuable militarily – that is, militarily necessary”.⁶³ Iron Dome, for example, is militarily necessary in fully autonomous mode to defend against short range saturation rocket attacks.

The last issue is accountability. This concern is addressed separately below, as this is the crux of the matter and at the heart of the concerns raised by the opponents. If one accepts, that autonomous weapon systems are simply another weapon system that has to be used in compliance with IHL, but that these systems are very complex and advanced as to their level of operation, and indeed act quite autonomously, which human actor if any, of all those possibly involved in their use is liable when possible breaches of IHL occur? Restated, if someone can legally be held liable for breaches of IHL flowing from the use of fully autonomous weapon systems, is there then still any problem at all (including an accountability problem) with their use? Likely not, as then the situation is exactly the same as in any other situation of possible IHL liability for the use of any other weapon or weapon system.

V Accountability

The objection as to the possible lack of an accountable actor is succinctly stated by Sharkey at 789–790:⁶⁴

⁵⁹ Schmitt, above n 2, at 11, 18 and 36; and Dinstein, above n 55, at 63 – 64.

⁶⁰ Schmitt, above n 2, at 16.

⁶¹ Schmitt, above n 2, at 8, 11 and 14.

⁶² The autonomous weapon system for example does not have to balance up definite military advantage against civilian lives foreseen to be lost; there are no civilians. Schmitt, above n 5, at 21. But see Anderson and Waxman, above n 5, at 17.

⁶³ Schmitt, above n 5, at 22.

⁶⁴ Sharkey, above n 2. See also HRW report, above n 2, at 4 and 42 – 43; and Anderson and Waxman, above n 5, at 11.

A robot does not have agency, moral or otherwise, and consequently cannot be held accountable for its actions. Moreover, if autonomous robots were used in limited circumstances in the belief that they could operate with discrimination, it would be difficult to decide exactly who was accountable for mishaps. Some would say that the commander who gave the order to send the robot on a mission would be responsible (last point of contact). But that would not be fair since it could be the fault of the person who programmed the mission, the manufacturer who made the robot, or the senior staff or policy makers who decided to deploy it. Or it could be claimed that the device was tampered with or damaged.

HRW identifies the following possibly liable human actors in the use of autonomous weapon systems:⁶⁵

- Military commanders
- Programmers
- Manufacturers

The starting point is as Human Rights Watch observes that there is a duty under IHL to prosecute cases of ‘grave breaches of IHL or war crimes’.⁶⁶ For example, the use of a weapon system that causes the indiscriminate killing of civilians can be a war crime or grave breach of IHL.⁶⁷

A *Military Commander Responsibility*

Liability under the legal doctrine of “command responsibility” allows for a military commander to be held accountable for war crimes committed by a subordinate. This liability arises where the commander knew or should have known that the subordinate planned or most likely will commit a war crime but does not take any action to prevent it.⁶⁸

Human rights Watch correctly declare that a commander would likely be liable if he “was aware in advance of the potential for unlawful actions against civilians and still recklessly deployed a fully autonomous weapon.”⁶⁹ Human Rights Watch then asks but what if he was not aware in advance of the weapon system’s unlawful potential? The simple answer of course is then he is not, and should not be, liable: the same as the commander not being liable in any other armed conflict situation where he was not aware in advance of any subordinate’s intention or planning to commit a war crime. From this it does not follow that no one is liable, although that too can happen of course.

A case that is of moment here in illustrating a commander’s liability, under command responsibility, for the unlawful use of a specific weapon system in circumstances where the specific weapon system could not be used with sufficient discrimination to avoid imposition of criminal responsibility for the wilful targeting of civilians, or the indiscriminate use of a weapon system is the decision of the International Criminal Tribunal for the former Yugoslavia (ICTY) in *Prosecutor v Martić*.⁷⁰

⁶⁵ HRW report, above n 2, at 42 – 44.

⁶⁶ HRW report, above n 2, at 42 (n 168). See list of grave breaches in Geneva Convention Relative to the Protection of Civilian Persons in Time of War, opened for signature on 12 August 1949, 75 UNTS (entered into force on 21 October 1950) art 147; and AP 1, art 85.

⁶⁷ *Prosecutor v Martić (Judgement)* ICTY Appeals Chamber IT - 95 - 11- T, 8 October 2008 (*Martić* appeal judgement).

⁶⁸ Arts 86 (2) and 87 AP 1; see Green, above n 51, at 333 – 335; and Olásolo above n 51, at 189 –193.

⁶⁹ HRW report, above n 2, at 43.

⁷⁰ *Prosecutor v Martić (Judgement)* ICTY Trial Chamber IT - 95 - 11 - T, 12 June 2007, (*Martić* trial

Milan Martić, the President and overall commander of Serbia Krajina was charged with the war crime of attacks on civilians in violation of the laws of war, and this count was based on article 51(2) of API.⁷¹ These charges arose out of attacks on the Croatian city of Zagreb.

The Court reiterated, relying on previous decisions of the Appeal Chamber of the ICTY that there is an absolute prohibition under IHL against attacking civilians,⁷² that this prohibition may not be derogated from because of military necessity and only military objects may be lawfully attacked.⁷³ This also, of course applies to the use of fully automatic and autonomous weapon systems.⁷⁴

The Court also confirmed the proportionality principle and stated that “indiscriminate attacks, that is attacks which affect civilian objects and military objects without distinction, may also qualify as direct attacks on civilians”.⁷⁵ Such an indiscriminate attack may too be inferred from the indiscriminate character of the weapon used.⁷⁶ Once more, all of this is highly relevant in cases where fully autonomous weapon systems are used.

The weapon system used by the Serbs in *Martić*’s case was a rocket system styled the M-87 Orkan, this is a non-guided rocket cluster munitions weapon system.⁷⁷ Of substantial importance to this article is the fact that the prosecution led expert evidence on the features and characteristics of the M-87 Orkan, and that this evidence, which the court accepted, showed that the M-87 Orkan, by virtue of “its characteristics and the firing range in this specific instance [...] is an indiscriminate weapon, the use of which in densely populated civilian area ... in the infliction of severe civilian casualties”.⁷⁸ To play an old melody again, these considerations would be relevant too in finding liability where autonomous weapon systems are involved.

The Court further found mens rea, as wilfulness, in the form of indirect intent due to the awareness on the part of Martić of the effect of the weapon system’s use the day before.⁷⁹ Similarly evidence of a fully autonomous weapon system’s previous use and failures, or that the weapon system had never been subjected to art 36 of API review, or indeed other appropriate testing, and again expert evidence as to its working and features for example will be highly relevant.

It may be mentioned here that in two other judgements of the ICTY, namely, the decision in *Prosecutor v Gallić*, the weapons and weapons systems used being a 120 mm mortar system, and in *Prosecutor v Blaškić*, the weapons called baby bombs or home

judgement); confirmed on appeal in the *Martić* appeal judgement.

⁷¹ *Martić* trial judgment, above n 68, at [67].

⁷² *Martić* trial judgment, above n 68, at [68]; *Prosecutor v Blaškić (Judgement)* ICTY Appeals Chamber IT-95-14-A, 29 July 2004, at [62] (*Blaškić* appeal judgement); and *Prosecutor v Gallić (Judgement)* ICTY Appeals Chamber IT-98-29-A, 30 November 2006, at [109].

⁷³ *Martić* trial judgment, above n 68, at [68].

⁷⁴ Schmitt, above n, at above 8–11. And see Castellaneta, above n 14, at 205 – 210.

⁷⁵ *Martić* trial judgment, above n 68, at [68].

⁷⁶ *Martić* trial judgment, above n 68, at [69].

⁷⁷ The use of cluster weapons systems was not prohibited at the time.

⁷⁸ *Martić* trial judgment, above n 68, at [69].

⁷⁹ At [71] – [72].

made mortars, expert evidence was led on the nature and characteristics of these weapon systems and such expert evidence was determinative on the findings in both cases as to the alleged indiscriminate attacks on civilians.⁸⁰

Of course in these cases the weapons systems were not fully autonomous weapon systems, but I submit, and really there can be no doubt, what will be determinative indeed in future possible liability cases on the use of fully autonomous weapon systems will be expert evidence.

B Programmer's Liability

Following on, Human Rights Watch mentions the possibility of a design or programming flaw, raising the possibility of programmer liability.⁸¹ They argue that the weapon will be fully autonomous and so it would be unfair to hold the programmer liable as the programmer “could not predict with complete certainty the decisions a fully autonomous robot might eventually make in a complex battlefield situation.”⁸² Another straw man is set up. No programmer in any event can know with “complete certainty” what their autonomous weapon system will do, nor can any such programmer know with complete certainty what their man-on-the-loop weapons system or other technologically advanced weapon system will always do. If factually it cannot be proved that in any case the programmer acted wilfully, which incorporates recklessness, the applicable standard for liability under IHL for war crimes,⁸³ then he or she will not be liable.

This is moreover no foreign or extraordinary situation under current IHL, specifically, and as a notorious example, where anti aircraft missiles, under continual or ultimate human control, are fired at apparently hostile warplanes but due to poor or defective technology as to the identifying of friendly or hostile aircraft, a friendly or civilian aircraft is shot down. This leads to our next case study.

The United States has an air defence missile system called the Patriot Air defence System. It is similar to Iron Dome in many ways (but is apparently not so effective) and it can operate, but is not currently so used, in fully autonomous mode.⁸⁴ It has a human on-the-loop except in saturation hostile missile attacks when it can act fully autonomously. The Patriot was used extensively by the United States in both the Gulf war (1991) and the Iraq invasion (2003).

The Patriot weapon system is credited by the United States in shooting down eight out of nine Scud ballistic missiles fired by Iraq during 2003. It was however also involved in what was called by the United States Department of Defence Science Board in its report on the efficiency of the Patriot autonomous weapon system “three regrettable

⁸⁰ *Prosecutor v Gallic (Judgement)* ICTY Trial chamber IT-95-14-T, 5 December 2003, at [441]-[453] (*Gallic* trial judgment); and *Prosecutor v Blaskic (Judgement)* ICTY Trial Chamber IT-98-28-T Trial Chamber, 3 March 2000, at [512].

⁸¹ HRW report, above n 2, at 43 – 44.

⁸² HRW report, above n 2, at 43– 44.

⁸³ Article 85(3) of AP1; *Gallic* trial judgment at [54] ; Cassese, above n 10 , at 168 ; Schmitt, above n 5, at 33; and see Castellana, above n14, at 208; see also Francesco Moneta “Direct attacks on civilians and indiscriminate attacks as war crimes” in Fausto Pocar, Marco Pedrazzi and Micaela Frulli (eds) *War crimes and the conduct of Hostilities* (Edward Elgar, Cheltenham, 2013) 59 at 65 – 67.

⁸⁴ <www.fas.org/spp/starwars/program/patriot> And see Schmitt, above n 5, at 4 and 5.

fratricide incidents.”⁸⁵

In these incidents Patriot shot down two coalition (friendly) warplanes. The Patriot weapon system in one instance mistook the coalition aircraft for an incoming ballistic missile, and an attacking anti-radiation missile in another.⁸⁶ The missile system thus showed in these instances, in IHL parlance, an inability to discriminate between enemy military objects and its own. Albeit then that these were not cases of a failure to discriminate between civilian and military objects it could possibly have been. For example, in 1988, a United States warship, the USS Vincennes shot down an Iranian civilian airliner by mistake, when it failed to identify the plane as civilian. At fault was not the ship’s own Aegis autonomous weapon system,⁸⁷ but nevertheless this was one of the most technologically advanced warships at the time and she was unable to correctly discriminate between a civilian airliner and an Iranian fighter aircraft in the short time period that the decision to shoot had to be made.⁸⁸

The Defence Science Board held that the shooting down of the coalition aircraft by Patriot occurred each time during a “complex chain of events” and it was difficult to determine the exact causes of failure. Nevertheless, it found that there had been problems with the *technology* in respect of identifying friendly aircraft, albeit that this happened in only a few out of possibly millions friendly aircraft observations by the Patriot system.⁸⁹

The final shortfall identified by the board was the “Patriot system operating philosophy, protocols, displays and software, seemed to be a poor match for the conditions”. The operating protocol was “largely automatic, and the operators were trained to trust the system’s software; a design which would be needed for heavy missile attacks.”⁹⁰

The above consequently raises the possibility of programmer’s liability for the software’s failure to correctly identify or discriminate. No doubt too in the case of fully autonomous weapon systems it could or would be difficult to hold a programmer liable for an intentional or wilful act of designing or programming defective software.⁹¹ Again, however, this is not a new problem; already as to most modern weapon systems, such as guided missile systems like Patriot or Iron Dome, which incorporate computer technology and hence software, there is needed expert evidence to prove defective software. This, however, in principle is no different from having to prove factually that the software on a civilian aircraft or motor vehicle is defective. As in such cases it is difficult to establish fault, especially wilfulness, but that has never been a reason to prohibit anything.

C Manufacturer’s Liability

This brings us finally then to possible manufacturer’s liability. Similarly as with the

⁸⁵ *Report of the Defense Science Board on Patriot System Performance*, Department of Defense <www.acq.osd.mil> January 2005 (Patriot Report Summary), at 2 (accessed on 12 October 2014).

⁸⁶ Patriot Report Summary, above n 83, at 2.

⁸⁷ Schmitt, above n 5, at 13; and Singer, above n 18, at 124.

⁸⁸ See the report and generally Singer, above n 5, at 124.

⁸⁹ Patriot Report Summary, above n 83, at 2.

⁹⁰ Patriot Report Summary, above n 83, at 2.

⁹¹ HRW report, above n 2, at 44.

possible liability of programmers it will be difficult to establish wilfulness on the part of a manufacturer. Individual or corporate international criminal law liability will be difficult to establish. But again this is not a new problem under IHL.

If an autonomous weapon system is wilfully deployed where it should not have been because of the system's known inability to discriminate between civilians and combatants, then those who were aware of this inability, from commander to user, may be held liable under IHL.⁹²

In summary as to accountability: sometimes it is very simple to factually establish a chain of causation, at times it is not; sometimes it is relatively easy to determine as fact wilfulness or intent, at other times it is very difficult. There are sometimes accidents and there are sometimes malfunctioning robots or cars or computers. Underlying it all, as any trial lawyer will say, evidence is needed to establish the facts. In future, to prove the facts may require extensive expert evidence on autonomous systems, but this is already so as to other automatic, autonomous or electronic systems.

VI Conclusion

The analysis above of the definitions of autonomous weapon systems, their factual make up, and their de facto use show that, despite the non recognition thereof, there are already fully autonomous weapon systems in use. And, these are being used in restricted and predetermined circumstances.

Furthermore, these weapons systems can be and are being used in compliance with IHL in the circumstances in which they are used. Fully autonomous weapon systems need not be inherently complaint with IHL in all circumstances, only in their intended use and in specific circumstances. Also, compliance by autonomous weapon systems with IHL principles lays probably more on a continuum in specific circumstances, rather than a dichotomy of or absolutely can or cannot. Iron Dome is a currently in use (in all but name) fully autonomous weapon system which does in certain circumstances clearly comply with IHL. Fully autonomous weapon systems can be used in compliance with IHL.

The same principles as currently applied in IHL to find liability or accountability can be applied in the case of the use of fully autonomous weapon systems. Liability will lie where a fully autonomous system is wilfully used where it should not have been. This will each time be a factual question to be determined on a case by case basis.

⁹² See Schmitt, above n 5, at 32; and *Martić* trial judgment, above n 68, at [69]–[72].

Bibliography:

Primary Sources

A Cases

Advisory Opinion on the legality of the Threat or Use of Nuclear Weapons [1996] ICJ Rep 226.

Legal Consequences of the construction of a Wall in Occupied Palestinian Territory (Advisory Opinion) [2005] ICJ Rep 131.

Prosecutor v Blaškić (Judgement) ICTY Trial Chamber IT-98-28-T, 3 March 2000.

Prosecutor v Blaškić (Judgement) ICTY Appeals Chamber IT - 95-14-A, 29 July 2004.

Prosecutor v Gallić (Judgement) ICTY Trial Chamber IT-95-14-T, 5 December 2003.

Prosecutor v Gallić (Judgement) ICTY Appeals Chamber IT-98-29- A, 30 November 2006.

Prosecutor v Martić (Judgement) ICTY Trial Chamber IT 95-11-T, 12 June 2007.

Prosecutor v Martić (Judgement) ICTY Appeals Chamber IT - 95- 11- T, 8 October 2008.

B Books and Chapters in Books

James J Busittil *Naval Weapon Systems and the Contemporary Law of War* (Caledon Press, Oxford, 1998).

Antonio Cassese, Guido Acquaviva, Mary D Fan and Alex Whitin *International Criminal Law, Cases and Commentary* (Oxford University Press, 2011).

Andrea Canone “The use of prohibited weapons and war crimes” in Fausto Pocar, Marco Pedrazzi and Micaela Frulli (eds) *War Crimes and the Conduct of Hostilities* (Edward Elgar, Cheltenham, 2013) 173.

Marina Castellaneta “New weapons, old crimes?” Fausto Pocar, Marco Pedrazzi and Michaela Frulli (eds) *War Crimes and the Conduct of Hostilities* (Edward Elgar, Cheltenham, 2013) 194.

Yoram Dinstein *The Conduct of Hostilities under the law of International Armed Conflict* (2nd ed, Cambridge University Press, 2010).

Leslie C Green *The Contemporary Law of Armed Conflict* (3 ed, Manchester University Press, 2008).

Jean-Marie Henckaerts and Louis Doswald Beck *Customary International Humanitarian Law, Volume 1: Rules* Oxford University Press, 2005).

Ian Henderson *The Contemporary Law of Targeting* (Martinus Nijhoff, Leiden, 2009).

Antonio Leandro “Arms transfer and complicity in war crimes” Fausto Pocar, Marco Pedrazzi and Micaela Frulli (eds) *War Crimes and the Conduct of Hostilities* (Edward Elgar, Cheltenham, 2013) 225.

Francesco Moneta “Direct attacks on civilians and indiscriminate attacks as war crimes” in Fausto Pocar, Marco Pedrazzi and Micaela Frulli (eds) *War Crimes and the Conduct of Hostilities* (Edward Elgar, Cheltenham, 2013) 59.

Nils Melzer *Targeted Killing in International Law* (Oxford University Press, 2008).

Héctor Olásolo *Unlawful Attacks in Combat Situations* (Martinus Nijhoff, Leiden, 2008).

PW Singer *Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century* (New York: The Penguin Press, 2009).

C Journal Articles

Peter Asaro “On banning autonomous weapons systems: human rights, automation and the dehumanization of lethal decision making” ICRCR No 886 (Volume 12, 2012) 682.

William Boothy “Some legal challenges posed by remote attack” ICRCR No 886 (Volume 12, 2012) 579.

Hun –Yan Liu “Categorization and legality of autonomous and remote weapon systems” ICRCR No 886 (Volume 12, 2012) 628.

Noel E Sharkey “The inevitability of autonomous robot warfare” ICRCR No 886 (Volume, 2012) 788.

Michael N Schmitt and Jeffrey S Thurner “Out of the Loop: Autonomous weapon Systems and the Law of Armed Conflict” (2012) 4 Harv Nat’l Sec J 213.

Michael N Schmitt “Autonomous weapon systems: A reply to critics” (2013) Harv Nat’l Sec J 2.

D Treaties

Geneva Convention Relative to the Protection of Civilian Persons in Time of War, opened for signature on 12 August 1949, 75 UNTS (entered into force on 21 October 1950).

Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Armed Conflicts of International Armed Conflicts 1125 UNTS 3 (opened for signature on 12 December 1977, entered into force 7 December 1978)

E Reports

Human Rights Watch Report “Losing Humanity: the case against killer robots” (Human Rights Watch, November 2012 <www.hrw.org>)

Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia (2000) ,< www.un.org/icty > At 16 September 2014.

Report of the Defence Science Board on Patriot System Performance, Department of Defence <www.acq.osd.mil> January 2005 (Patriot Report Summary).

Defence Update “Iron Dome Battle Management Demonstrated” (2009), <www.defenseupdate.com> accessed 20 September 2014).

F Internet Sources

Alon Ben David “*Iron Dome blunts 90% of enemy rockets*” Aviation Week & Space Technology (1 September 2014) <www.aviationweek.com>

Rafael Advanced Defence Systems Ltd “Iron Dome: Defence against Short Range Artillery Rockets,” 2010, <www.rafael.co.il> (accessed 18 September 2014).

MBDA< www.mbda-systems.com/products > (accessed 18 September 2014).