THE STATUS OF AUTONOMOUS WEAPON SYSTEMS UNDER INTERNATIONAL HUMANITARIAN LAW

TIMOTHY MCFARLAND

ORCID Identifier: 0000-0002-8740-6980

Submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy

April 2017

Melbourne Law School

The University of Melbourne

Abstract

Autonomous weapon systems (AWS) are an emerging class of advanced weapons which utilise technologies associated with robotics and artificial intelligence to select and attack targets without human intervention. Various States are developing AWS for possible use in armed conflict over the coming years. This thesis analyses the legal implications of that path of weapon development, primarily in relation to international humanitarian law.

The main forum for debate on legal issues relating to AWS is within the framework of the Convention on Certain Conventional Weapons (CCW). States parties to the CCW are engaged in an ongoing series of meetings about AWS with a view to regulating their use. This thesis discusses three unresolved questions arising from those meetings:

- How should machine autonomy be regarded, and how should AWS be defined for the purpose of legal regulation? One of the main challenges facing participants in the legal debate about AWS has been to understand the nature of the technical developments to which the law must respond. This part of the thesis presents an understanding of how the emerging technical capabilities of AWS relate to the applicability and efficacy of relevant legal rules.
- How can States meet their obligations under international humanitarian law when using weapon systems which operate autonomously? In particular, the thesis examines the basic principles of distinction, proportionality and precaution in attack, a proposed requirement for 'meaningful human control' over AWS, and the novel challenges of reviewing AWS for legal compliance.
- How can States and non-State entities be held responsible for proscribed acts done in armed conflict via AWS? This question is primarily about whether the required accountability standards could be upheld under the current law of State responsibility and international criminal law. Additionally, some alternative accountability regimes are surveyed.

The thesis is based primarily on legal doctrinal research and cross-disciplinary technical research in relevant subject areas. In addition, fieldwork was conducted in the form of discussions with a range of operational military personnel, policy analysts and academic researchers in relevant legal and technical fields, in Australia and the United States.

Declaration

This is to certify that:

1. the thesis comprises only my original work towards the PhD except where indicated in the Preface;
2. due acknowledgement has been made in the text to all other material used; and
3. the thesis is less than 100,000 words in length, excluding the front matter and bibliography.

TIM McFARLAND

Preface

Chapter 3, entitled 'Understanding Weapon Autonomy', is largely based on a paper I wrote in 2013 soon after commencing my PhD candidature. The paper was published in the *International Review of the Red Cross* in 2016 under the title 'Factors Shaping the Legal Implications of Increasingly Autonomous Military Systems'.

Section 7.3.3 discusses criminal accountability of weapon developers. Parts of that section are drawn from a paper written jointly with my supervisor Professor Tim McCormack on that topic. The paper was published in the US Naval War College's *International Law Studies* journal in 2014. I would like to acknowledge Professor McCormack's assistance in identifying and refining discussion of the issues covered in that section.

Acknowledgements

This research was supported by the Australian Research Council's Discovery Projects funding scheme (project DP130100432).

To my supervisors, Professor Tim McCormack and Dr Rain Liivoja, many thanks for being constantly supportive and accommodating. Thanks also to all the members of the Program on the Regulation of Emerging Military Technologies for their camaraderie during the project.

This thesis has been greatly enhanced by the many valuable insights I have gained from interviews and conversations with officers of the Australian Defence Force and staff of the Defence Science and Technology Group, as well as with military officers and academic researchers of the United States Naval War College in Newport, Rhode Island and of the National Defense University in Washington, D.C.

Many other people generously made time to read and give feedback on elements of the thesis. Special thanks in this respect goes to Group Captain Ian Henderson and Associate Professor Hitoshi Nasu.

Finally, I would like to thank my wife Sandra for her continual love and support.

Contents

Chapter 1 Introduction 1

1.1 Purpose and Scope 2

1.2 Methodology 4

1.3 Chapter Outline 7

Chapter 2 Background 9

2.1 Legal Background 9

2.2 History of the Debate about AWS 22

2.3 Broader Concerns about AWS 28

2.3.1 AWS Replacing Human Combatants 29

2.3.2 Technical Capabilities of AWS 32

2.3.3 Accountability 34

2.3.4 Accessibility to Non-State Actors 35

2.3.5 Effects on Civilian Populations 36

2.4 Questions Discussed in this Thesis 36

2.4.1 Defining and Mapping Autonomy 37

2.4.2 Challenges to IHL 41

2.4.3 Accountability 45

Chapter 3 Understanding Weapon Autonomy 48

3.1 The Technology of Autonomous Machines 49

3.1.1 Autonomy and Control 51

3.1.2 Adaptive and Intelligent Control 57

3.1.3 Autonomy is a Matter of Degree 60

3.1.3.1 Degree of Computer Control 63

3.1.3.2 Methods of Human-Machine Interaction and Task Allocation 65

3.1.3.3 Variations by Function 66
3.1.3.4 Variations by Circumstance 67
3.1.3.5 Summary 67
3.1.4 Summary 70
3.2 Autonomy as a Property of a Platform, Not a Weapon 70
3.2.1 Autonomy Alters the Relationship between Machine and Operator 71
3.2.2 Autonomy will Extend beyond Weapon Systems 73
3.2.3 Autonomous Systems will Collaborate with Each Other 75
3.3 Conclusion 78
Chapter 4 Identifying Legal Issues 79
4.1 Categorisation of New Technology 79
4.2 Applications of Weapon Autonomy 94
4.2.1 Applications of Autonomy within a Weapon System 94
4.2.1.1 Functional Scope of Autonomous Capability 94
4.2.1.2 Facility for Human Oversight 95
4.2.2 Applications of AWS 96
4.2.2.1 Anti-personnel or Anti-materiel Roles 96
4.2.2.2 Physical Domain 97
4.2.2.3 Complexity of the Environment 97
4.2.2.4 Others 97
4.3 Effects on Decision-making Processes 98
4.4 Effects on Task Performance 100
4.4.1 Predictability of Performance and Failure 100
4.4.1.1 Complexity 100
4.4.1.2 Brittleness 103
4.4.2 Nature and Consequences of AWS Failure 105

4.4.2.1 Possibility of Runaway Failures 106
4.4.2.2 Common Vulnerabilities among Weapon Systems 106
4.5 Conclusion 107
Chapter 5 Weapons Law 109
5.1 Autonomous Systems Subject to Rules of Weapons Law 109
5.2 Weapons Law and Targeting Law 112
5.3 Law Relating Specifically to AWS 115
5.4 Legality *per se* 116
5.4.1 Threshold Levels of Autonomy 116
5.4.1.1 Anthropomorphic Views of Machine Autonomy 118
5.4.1.2 Means and Ends 119
5.4.1.3 Example Scenarios 119
5.4.1.4 Implicit Requirement for Human Subjectivity 121
5.4.1.5 Conflating Nature and Use of Weapon Systems 122
5.4.2 Distinction 123
5.4.2.1 Content of the Principle 123
5.4.2.2 Application to AWS 126
5.4.2.2.1 Striking Protected Persons or Objects 130
5.4.3 Superfluous Injury and Unnecessary Suffering 133
5.4.3.1 Content of the Principle 133
5.4.3.2 Application to AWS 133
5.4.4 Summary 134
Chapter 6 Targeting Law 136
6.1 Autonomous Attacks 136
6.1.1 Preparing for an Attack 139
6.1.1.1 Substantive Obligations 139

6.1.1.2 Standard of Precautions 140
6.1.1.3 Bearers of Obligations 141
6.1.2 How is an 'Attack' Defined? 141
6.1.2.1 'Acts of Violence' 143
6.1.2.1.1 Scale 145
6.1.2.1.2 Severity 147
6.1.2.2 'Against the Adversary' 147
6.1.2.3 'Whether in Offence or in Defence' 148
6.1.2.4 Application to Delayed-Action Weapons 149
6.1.2.4.1 Mines 149
6.1.2.4.2 Cyber-Attack 151
6.1.3 Applying Current Law to AWS 153
6.1.3.1 Identifying the Attack 153
6.1.3.2 Preparatory Obligations Relate to the Combined Actions of Software and Humans 159
6.1.3.3 Extended Operations 159
6.1.3.4 Other Precautions in Extended Operations 163
6.1.3.5 Summary 164
6.2 Legal Use of AWS 165
6.2.1 Summary 172
6.3 Meaningful Human Control 173
6.3.1 History 173
6.3.2 Meaning 178
6.3.3 Alternatives 185
6.3.4 Arguments against MHC 186
6.3.5 Controllability 190

6.3.6 Summary 191
Chapter 7 Accountability 192
7.1 General Statement on Accountability for AWS Behaviour 193
7.2 State Responsibility 196
7.2.1 Applying the Law of State Responsibility to AWS 199
7.2.1.1 Lack of Human Judgement 200
7.2.1.2 AWS as Agent 202
7.2.2 Sufficiency of the Law of State Responsibility 202
7.3 Individual Criminal Responsibility 203
7.3.1 Prosecuting an Offence Involving an AWS 206
7.3.1.1 Material Elements 208
7.3.1.2 Mental Elements 212
7.3.2 Altering the Mental Element 215
7.3.3 Developer Responsibility 217
7.3.3.1 Threshold Requirement of the Existence of an Armed Conflict 219
7.3.3.2 Physical Perpetration and Accessorial Modes of Liability 221
7.3.3.3 International Jurisprudence on *Ex Ante* Aiding and Abetting 223
7.3.3.4 *Mens Rea* Requirements 225
7.3.3.5 Attribution of Responsibility 227
7.3.3.6 Conclusions on Responsibility of Developers 227
7.3.4 Command or Superior Responsibility 229
7.4 Other Forms of Accountability 230
7.4.1 Retrospective Accountability 231
7.4.1.1 War Torts 232
7.4.2 Prospective Accountability 233
7.4.2.1 Weapon Reviews 234

7.5 Conclusion 239
Chapter 8 Conclusions and Recommendations 240
8.1 Defining AWS 240
8.2 Regulating AWS 242
Bibliography 246

Chapter 1 Introduction

The progressive development of international humanitarian law (IHL) represents an ongoing effort to accommodate multiple, often competing,[1] goals of States and the international community: principally, the need of States to effectively pursue and protect their national interests through military means when peace has failed, and their parallel obligation to serve broader humanitarian goals by limiting the harm inflicted on those involved in armed conflict.[2] Maintaining a suitable balance of interests[3] in the face of the constantly changing practice of warfare means revising, reinterpreting and supplementing the prevailing body of law as the actors, the battlefields and the means and methods of warfighting evolve.

In many cases, such evolution can be traced directly or indirectly to progress of a scientific, mathematical or technical nature which confers some military advantage on the armed forces which wield it.[4] Where that military advantage is related to the type or extent of harm inflicted on those against whom it is wielded, it is necessary to reassess the legal checks that are in place, to ensure that none of the various goals of IHL are unduly compromised. This thesis is motivated by one example of such progress. It responds to efforts by several militarily advanced States to develop increasingly autonomous weapon systems (AWS) for use in armed conflict: a development path which, it will be seen, has implications for both the means and the methods that armed forces will employ in conflict. While many existing weapon systems possess some capacity for operating without human intervention, current development efforts indicate that those capabilities will expand enormously in coming years. That prospect has already sparked considerable controversy, with a number of advocacy groups calling for

---

[1] See generally Michael N Schmitt, 'Military Necessity and Humanity in International Humanitarian Law: Preserving the Delicate Balance' (2010) 50 *Virginia Journal of International Law* 796.

[2] Jean de Preux, 'Protocol I – Article 35 – Basic Rules' in Yves Sandoz, Christophe Swinarski and Bruno Zimmermann (eds), *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (Martinus Nijhoff, 1987) 389, 392-3 [1389].

[3] References to 'balance' between military necessity and humanity do not imply that the thesis takes any particular position on the question of whether or not they are necessarily in conflict. For a detailed discussion of two conflicting views on that question see Nobuo Hayashi, 'Military Necessity as Normative Indifference' (2013) 44 *Georgetown Journal of International Law* 675. For current purposes it is only necessary to note that military necessity and humanity are both significant drivers of weapon development and of the corresponding regulatory responses.

[4] See generally Braden R Allenby, 'Are New Technologies Undermining the Laws of War?' (2014) 70(1) *Bulletin of the Atomic Scientists* 21.

a pre-emptive ban[5] and many States expressing concern in United Nations fora. Yet the motive to increase levels of autonomy in weapon systems is strong: increased operational effectiveness,[6] safety for one's own soldiers, decreased personnel requirements and financial savings are among the benefits that may be realised by States which possess AWS.

This thesis asks: how will use of increasingly autonomous weapon systems affect a State's ability to meet its obligations under IHL, and what, if any, legal responses are required to ensure that the law continues to progress towards achieving its aims?

1.1 Purpose and Scope

The international debate about whether and how to regulate development of AWS is very broad. Proponents and opponents of AWS alike have made arguments based on a wide range of considerations: military operational effectiveness, ethics, politics, international relations, and a number of different legal frameworks including IHL, human rights law,[7] international criminal law and arms control law. At this early stage it is possible that any of those might ultimately determine the international response to AWS development. A brief overview of the various arguments is presented in Section 2.3, but there is far too much material to cover the entire spectrum of concerns in detail in one thesis. Instead, this thesis addresses only issues arising under IHL. Four further restrictions have been applied.

First, the thesis is written in abstract terms. It is not a detailed case study of a specific weapon system or a specific conflict, although examples of both are employed where appropriate. Rather, it addresses an abstract phenomenon, being a capacity for autonomous operation of military weapon systems, however that capacity may be achieved and in whatever type of weapon system. An abstract approach is necessary, given that the motive for the thesis is to respond to plans for a broad path of future weapon development. The precise characteristics of future AWS are unknown at this

[5] In particular, the Campaign to Stop Killer Robots (http://www.stopkillerrobots.org), a coalition of non-governmental organisations which has been active since 2012, has participated extensively in multilateral discussions about regulation of AWS.

[6] See, eg, Michael W Byrnes, 'Nightfall: Machine Autonomy in Air-to-Air Combat' (2014) (May-June) *Air & Space Power Journal* 48.

[7] For a general discussion of AWS and human rights law, albeit in a law enforcement context, see Christof Heyns, 'Human Rights and the Use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement' (2016) 38 *Human Rights Quarterly* 350.

time, so the discussion must relate to weapon autonomy as an abstract capability that may be present in potentially any type of weapon system. Similarly, the characteristics of the conflicts in which future AWS may be employed are difficult to reliably anticipate,[8] so discussions about applying the rules of IHL to AWS simply assume that *jus ad bellum* requirements (law governing the legality of the resort to violence) have been satisfied and that it is legal for the State operating the AWS to employ military force against its enemy. Consistent with the independence of the *jus ad bellum* and the *jus in bello*, no assumptions are made about either the causes of, or the legal merits of, participation in armed conflicts.

Second, the rules of IHL discussed below are those that apply to an international armed conflict, unless otherwise specified. The points of difference between the rules applicable to international and non-international conflicts are not at issue in respect of AWS, and significant space would need to be devoted to discussing the less well defined legal concepts which apply to non-international armed conflict.[9] Greater utility may be realised from a deeper discussion of AWS and the law of international armed conflict, with the conclusions being readily applicable to situations of non-international armed conflict.

Third, the discussion relates primarily to autonomy in the critical functions of selecting and engaging targets. While other functions of a weapon system may also be done with a degree of autonomy, such as navigation, for the purposes of this thesis they are of interest only insofar as they contribute to the weapon system's autonomously selecting and engaging targets.

Fourth, it is assumed that the weapon part of the system (the gun, missile or other munition) is not subject to any specific prohibition or regulation. If a class of weapon is prohibited in some or all circumstances, such as a biological weapon[10] or a chemical

---

[8] There is a large and growing volume of literature available on the many aspects and implications of the changing nature of armed conflict. See, eg, Hew Strachan and Sibylle Scheipers (eds), *The Changing Character of War* (Oxford University Press, 2011).

[9] These would include the definitions of lawful objectives and combatants, among other matters. See Ian Henderson, *The Contemporary Law of Targeting* (Martinus Nijhoff, 2009) 12.

[10] *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) ('*Biological Weapons Convention*').

weapon,[11] there is no *prima facie* reason to think the addition of autonomous capabilities would affect the applicability of that ban. In cases where a weapon is subject to some special regulation, such as an anti-personnel mine, the interaction of those regulations with the changes wrought by a degree of autonomy would require separate investigation. This thesis is concerned only with the interaction between weapon autonomy and general IHL.

It is not the intention here to advocate either for or against continued development of AWS. The thesis attempts to map part of the legal landscape that would be encountered by States fielding increasingly autonomous weapons, but is deliberately neutral as to whether States should or should not take that path. As noted above, the decision to pursue a path of weapon development must be based on considerations far broader than just IHL. It is hoped that this thesis will inform decisions about development and use of AWS, but it cannot, and does not attempt to, prescribe a path of development.

1.2 Methodology

There are three main components to the research. The first establishes a technical understanding of machine autonomy from which its legal significance can be determined. Many of the findings of this part of the research are not specific to weapon systems; the technologies that are relevant to this thesis are common to many types of autonomous machines. However, the discussion here focuses primarily on their application to weapon systems. The research draws upon textbooks, journal articles and conference presentations from scientific and engineering fields relating to autonomous machines and their component technologies.

The second part surveys the characteristics of current and future AWS employing the technologies covered in the first part. The purpose of this phase is to discover how those technologies are being, and will be, employed by armed forces. It draws primarily on publications of military technical and research organisations and weapon system developers. Considerable assistance has been gained from a series of conversations with officers and academics at the US Naval War College in Newport, Rhode Island, and at

[11] *Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on their Destruction*, opened for signature 13 January 1993, 1974 UNTS 45 (entered into force 29 April 1997) ('*Chemical Weapons Convention*').

the National Defense University in Washington, DC. Journal articles and conference presentations from the broader academic community are also used.

The third part applies existing rules of IHL to weapon systems exhibiting those characteristics in order to identify constraints and incompatibilities and to make recommendations. A traditional doctrinal approach is taken. Conventions, cases and commentary have been consulted to establish which rules are at issue and how those rules are currently interpreted. Here also, conversations with colleagues at the Naval War College and at the National Defense University were very helpful.

The legal frame of reference for this research is IHL as embodied in the First Additional Protocol to the Geneva Conventions of 1949.[12] It is hoped that this study will be useful to any State which is considering the development, purchase or use of AWS in coming years. In practice, those decisions would have to be consistent with all of a State's legal obligations, but it is not possible in the time and space available to address all the combinations of international and domestic obligations which a State may bear. Instead, API was chosen as representing relevant legal rules which directly bind a large number of States and which, in some instances, codify customary rules which bind all States.

At the time of writing, API conventionally binds 174 States Parties, with three other States being signatories.[13] The relevant articles of API which are considered, with varying degrees of acceptance, to codify customary law, are as follows.

Article 35(1), which provides that the right of the parties to a conflict to choose methods or means of warfare is not unlimited, and Article 35(2), which prohibits means and methods of warfare that are of a nature to cause superfluous injury or unnecessary suffering, both echo provisions in the Regulations annexed to the Fourth Hague

[12] *Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I)*, opened for signature 8 June 1977, 1125 UNTS 3 (entered into force 7 December 1978) ('*API*').

[13] *Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the protection of victims of international armed conflicts (Protocol I)*, United Nations Treaty Collection <https://treaties.un.org/Pages/showDetails.aspx?objid=08000002800f3586>.

Convention of 1907[14] (with the possible exception of the term 'methods of warfare', being a phrase not present in the Hague Regulations).[15]

Article 48, which sets out the principle of distinction, restates a general rule of international law which is universally acknowledged as binding on all States. Articles 51 and 52 protecting civilians and civilian objects respectively from dangers arising from military operations also restate customary law. Of particular interest for this thesis are the customary prohibitions on inherently indiscriminate weapons codified in Articles 51(4)(b) ('means of combat which cannot be directed at a specific military objective') and 51(4)(c) ('means of combat the effects of which cannot be limited …'). However, it has been argued that some codifications in API employ more specificity than is generally seen as customary.[16] In particular, the definition of military objectives in Article 52(2) introduces clarifications which may be subject to different interpretations.[17]

Article 57, which sets out the precautions which an attacker must take in preparing for and conducting an attack, together with Article 58, have been described by the ICTY as representing customary law:

> [The principles of distinction and proportionality] have to some extent been spelled out in Articles 57 and 58 of the First Additional Protocol of 1977. Such provisions, it would seem, are now part of customary international law, not only because they specify and flesh out general pre-existing norms, but also because they do not appear to be contested by any State, including those which have not ratified the Protocol.[18]

---

[14] *Convention respecting the Laws and Customs of War on Land and its Annex: Regulations concerning the Laws and Customs of War on Land*, signed 18 October 1907 [1910] ATS 8 (entered into force 26 January 1910) ('*Hague Convention 1907*').
[15] Fausto Pocar, 'To What Extent Is Protocol I Customary International Law?' (2002) 78 *International Law Studies* 337, 344.
[16] Ibid 345.
[17] Ibid 348.
[18] *Prosecutor v Kupreškić (Judgement)* (International Criminal Tribunal for the Former Yugoslavia, Trial Chamber, Case No IT-95-16, 14 January 2000) [524].

The International Committee of the Red Cross (ICRC) has also argued for the customary status of Article 57.[19] On the other hand, Greenwood has argued that some of its provisions go beyond customary law.[20]

This thesis does not critique existing law or argue for a particular interpretation in areas of uncertainty, except to an extent in respect of the definition of 'attack'.[21] The intent instead is to argue for a particular understanding of how the law applies to a new set of technologies. The precise interpretation of many of the provisions of API is a matter of contention. In some cases that is because 'the specificity of Protocol I's provisions add new elements to principles that, while well established in customary law, leave margins of discretion to belligerent States.'[22] In others it is because the expression of new, non-customary rules allows a range of interpretations. In all cases an attempt has been made to utilise the most commonly accepted interpretation, in order to keep the focus on the novel ways in which the law must be applied to AWS. In respect of customary law, repeated reference is made to the ICRC's Customary IHL Study.[23] However, the extent to which that study accurately captures the current state of customary IHL is itself the subject of debate.[24]

Finally, the thesis employs many examples of particular weapon systems and incidents in which they have been, or could be, involved. No special status should be accorded to those systems or incidents due to their being mentioned here. They are used purely to demonstrate relevant points.

1.3 Chapter Outline

Chapter 1, the current introductory chapter, defines the purpose and scope of the thesis and the methodology employed.

---

[19] Jean-Marie Henckaerts and Louise Doswald-Beck, *Customary International Humanitarian Law* (Cambridge University Press, 2005) vol 1, 51.
[20] Christopher Greenwood, 'Customary Law Status of the 1977 Geneva Protocols' in Astrid J M Delissen and Gerard J Tanja (eds), *Humanitarian Law of Armed Conflict: Challenges Ahead: Essays in Honour of Frits Kalshoven* (Martinus Nijhoff, 1991) 111.
[21] See Section 6.1.
[22] Pocar, above n 15, 347.
[23] Henckaerts and Doswald-Beck, above n 19.
[24] See, eg, Jann Kleffner, 'Sources of the Law of Armed Conflict' in Rain Liivoja and Tim McCormack (eds), Routledge Handbook of the Law of Armed Conflict (Routledge, 2016) ch 4 s 2.

Chapter 2, ‘Background’, briefly outlines the development of the relevant law and the progress of the debate about regulating AWS, then introduces the topics discussed in this thesis.

Chapter 3, ‘Understanding Weapon Autonomy’, explains the technology of autonomous systems. Its purpose is to establish a clear understanding of what it means for a machine to be ‘autonomous’ and to provide a basis for identifying the legal challenges that will accompany increases in the autonomous capabilities of weapon systems.

Chapter 4, ‘Identifying Legal Issues’, links the technical discussion of Chapter 3 to the legal discussion to follow in the remaining chapters, explaining how the technologies which make a weapon system ‘autonomous’ relate to the law which governs their use.

Chapter 5, ‘Weapons Law’, discusses the legality *per se* of AWS. It covers the ways in which the rules of IHL might constrain the design and development of increasingly autonomous weapon systems.

Chapter 6, ‘Targeting Law’, discusses legal use of AWS in armed conflict. It investigates the meaning of an ‘attack’ when human involvement in operating a weapon may be reduced or absent, then steps through a typical targeting process that might be applied to an attack via AWS. Finally, the emerging notion of ‘meaningful human control’ over AWS is described and critiqued.

Chapter 7, ‘Accountability’, discusses how existing accountability regimes may apply when proscribed acts are committed by a weapon under the direct control of a computer rather than a human. Some alternative regimes are surveyed.

Chapter 8, ‘Conclusions and Recommendations’, proposes some steps to address concerns about development of AWS.

Overall, the thesis aims to assist States whose interests are affected by the possibility of increasingly autonomous weapon systems to resolve one aspect of the issue, being the question of compatibility with IHL.

Chapter 2 Background

Machine autonomy is defined and discussed in detail in Chapter 3; it is sufficient to note here that the weapon systems of interest in this thesis are those 'that, once activated, can select and engage targets without further intervention by a human operator.'[25] This chapter frames the discussion to follow by outlining the applicable law and the history and current state of the debate about AWS.

The first section provides a brief account of the development of legal principles governing the nature and design of weapons used in armed conflict. The second and third sections summarise the growth of public interest in the increasing autonomy of weapon systems and the broad range of concerns raised by States, NGOs and other researchers. The fourth section introduces the three main questions addressed in the remainder of the thesis.

2.1 Legal Background

The primary motive for States to develop or acquire new types of weapons for use in armed conflict is, of course, to achieve a military advantage over their enemies when it has become necessary to resort to violence as a means of settling disputes. A State's interest in seeking such advantage is undoubtedly legitimate, indeed part of the *raison d'être* of the State, but the measures permitted to be taken in pursuit of that interest are not unlimited. Countervailing humanitarian obligations borne by States – Dinstein refers to the 'global Zeitgeist'[26] – have led to the drafting of treaties and adoption of practices, some of which have become customary law, that limit the extent to which a desire for military advantage can legally justify the deployment of new and more terrible weapons in conflict.

The balancing point between these opposing interests, insofar as they relate to weapon development and use, can be seen in the interplay between the notion of military necessity and the cardinal principles underlying the law governing armed conflict: distinction, proportionality and the avoidance of superfluous injury or unnecessary suffering.

---

[25] US Department of Defense, 'Autonomy in Weapon Systems' (Directive No 3000.09, Office of the Under Secretary of Defense for Policy, 21 November 2012) 13 ('*Directive 3000.09*').
[26] Yoram Dinstein, *The Conduct of Hostilities under the Law of International Armed Conflict* (Cambridge University Press, 2nd ed, 2010) 4.

The idea of military necessity in IHL is an acknowledgement that a State which is fighting a war has a legitimate interest in winning that war.[27] 'Winning the war', in the restricted scope of concern of IHL, means overcoming the military forces of the enemy, without reference to whatever broader political goals may be sought. Actions that may reasonably be deemed essential to overcome the enemy's military forces are considered militarily necessary, and States and their military commanders are given considerable latitude in taking those actions.

Implicit in the concept of military necessity are three types of limitations on the measures that States may take in pursuit of a military victory. First, acts of violence directed against persons or objects other than those associated with an enemy's military forces are forbidden: this is the principle of distinction. Second, acts of violence against enemy forces that go beyond what is necessary to weaken their ability to fight are forbidden: this is the principle of avoidance of superfluous injury or unnecessary suffering. Third, acts of violence against enemy forces that may be expected to cause excessive incidental harm to civilians or civilian objects are forbidden: this is the principle of proportionality.

The content and customary nature of these principles has developed over time in response to changes in the practice of war. By way of preparing for an analysis of AWS, it is useful to review the points at which technological developments that lead to new means (weapons) or methods (tactics) of warfare have goaded States to tighten or reinforce the associated legal restrictions. A full historical account of the evolution of the law governing warfare is beyond the scope of this investigation, and has been presented most thoroughly elsewhere.[28] To properly situate this research question it is sufficient to review the major milestones in the development of the principles and rules of weapons law and targeting law as part of modern IHL, since the middle of the nineteenth century. As for the law prior to that time, it will be noted only that the law of war as it was understood then reflected customary rules stretching back for hundreds, or thousands, of years including, arguably, a customary ban on poison and poisoned

[27] Whether or not the war itself is being fought for a legitimate reason is a *jus ad bellum* question.
[28] See, eg, John Carey, William V Dunlap and R John Pritchard, *International Humanitarian Law: Origins* (Brill, 2003).

weapons.[29] The impetus for the move towards codification and extension of the law at that time has been attributed to several interrelated factors:[30]

- introduction of compulsory military service in many countries, which greatly enlarged armies;
- improvements in arms technology, which made warfare more destructive;
- a consequent increase in the number of victims of armed conflict;
- 'a growing conviction … that civilization was rapidly advancing and that it was therefore imperative "to restrain the destructive force of war"';[31]
- parallel moves towards codification of private law in Europe.

Many treaties and events of importance to the legal regulation of armed conflict have been omitted from this account. Prior to the 1970s, the law developed largely along two parallel paths: 'Geneva' law related to protection of victims of armed conflict and 'Hague' law related to limitations on the means and methods of war (the central concern of this thesis). Developments in Geneva law during that time are omitted or mentioned only briefly.

Despite the central importance of the principle of distinction it receives little mention prior to 1977. The conduct of warfare during early development of the law of weaponry was such that battlefields tended to be largely free of civilians, so attacks directed against military targets were generally unlikely to also impact civilians. As a principle of targeting law guiding the decision to direct an attack only at military targets, distinction has always been central. Once that decision had been made, it was not seen as a major consideration in the choice of weapon with which to conduct the attack until developments in the means and methods of war brought civilians more often into the line of fire.

The story of the modern law of weaponry arguably began in Russia in the mid-1860s. At the invitation of Tsar Alexander II, representatives of 17 States met in St Petersburg in November 1868 to discuss whether to ban use of a specific type of exploding

[29] Dinstein, above n 26, 68.
[30] Dietrich Schindler and Jiri Toman, *The Laws of Armed Conflicts: A Collection of Conventions, Resolutions and Other Documents* (Martinus Nijhoff, 4th rev ed, 2004) v.
[31] Ibid vi.

projectile.[32] The projectile, which had been invented by Russian military authorities some years previously for use by artillery against military equipment such as powder magazines, had recently been modified and made smaller for use against human targets. Fired from a hand-held weapon, the new projectile contained a small explosive charge which would detonate on contact with soft targets, such as a human body. While the impact of the projectile itself would cause a serious enough wound, the explosive charge would scatter fragments of the casing through the body of the victim, greatly aggravating their suffering and the difficulty of treating the wound.

Consensus was reached quite readily at the conference and the participating States agreed to prohibit among themselves use of the anti-personnel exploding rounds in conflict. That agreement itself was an important step in the regulation of armed conflict at the time, being the first formal agreement between States to ban use of a particular weapon, but it is significant to the current investigation for two other reasons. First, the conference directly addressed the specific technical properties of the weapon in question which the participants felt made a ban appropriate. In prohibiting 'any projectile of a weight below 400 grammes, which is either explosive or charged with fulminating or inflammable substances',[33] the conference had 'by common agreement fixed the technical limits at which the necessities of war ought to yield to the requirements of humanity'.[34] That process of engaging with the physical and operational details of a weapon in order to set a balance between military necessity and humanity would become increasingly important as the level of technical sophistication seen in military systems increased exponentially over the following decades.

Second, the written document produced by the conference, which became known as the Declaration of St Petersburg,[35] set out in its preamble the considerations which led to the parties' agreement, and which express the foundational principles that underlie the law of weaponry today. In particular, the Preamble states:

---

[32] Hans-Peter Gasser, 'A Look at the Declaration of St. Petersburg of 1868' (1993) 33 *International Review of the Red Cross* 511, 512.

[33] Exploding projectiles heavier than 400 grams were intended for anti-materiel applications and so were excluded from concerns about unnecessary suffering caused to their targets.

[34] On the face of it, the Declaration bans all use of exploding projectiles. It has been argued, though, that based on the purpose of the Declaration and subsequent State practice, the better view is that the ban only applied to exploding rounds used in an anti-personnel role: Henderson, above n 9, 12 n 64.

[35] *Declaration Renouncing the Use, in Time of War, of Explosive Projectiles under 400 Grams Weight*, [1901] ATS 125 (signed and entered into force 11 December 1868).

> That the only legitimate object which States should endeavour to accomplish during war is to weaken the military forces of the enemy;
>
> That for this purpose it is sufficient to disable the greatest possible number of men;
>
> That this object would be exceeded by the employment of arms which uselessly aggravate the sufferings of disabled men, or render their death inevitable;
>
> That the employment of such arms would, therefore, be contrary to the laws of humanity; …

These passages articulate the ideas of military necessity (military activity should be limited to that which is necessary to achieve the ends of the war) and unnecessary suffering (combatants should avoid causing injury or death beyond what is necessary to achieve the ends of the war). Importantly, the Preamble links those ideas to each other, expressing the need to maintain a balance. It also establishes that the choice of weapons to use in armed conflict can be limited by legal principles, an idea which was made explicit eight years later in the Brussels Declaration of 1874.[36]

When the delegates of 15 European countries met in Brussels in July of 1874, again on the invitation of Tsar Alexander II, they were attempting to secure agreement on a more wide-ranging statement of the laws and customs of war.[37] The Brussels Declaration included a section describing limitations on acceptable means of warfare which read in part:

> Art. 12. The laws of war do not recognize in belligerents an unlimited power in the adoption of means of injuring the enemy.
>
> Art. 13. According to this principle are especially 'forbidden':
>
> (a) Employment of poison or poisoned weapons;
>
> …

[36] *Project of an International Declaration concerning the Laws and Customs of War*, signed 27 August 1874 (reprinted in: *The Proceedings of the Hague Peace Conferences: The Conference of 1899* (James Brown Scott trans, William S Hein & Co, 2000) 564 [trans of: *Conférence internationale de la paix. La Haye, 18 mai-29 juillet 1899.* (first published 1907)]) ('*Brussels Declaration*').
[37] Schindler and Toman, above n 30, 21.

> (e) The employment of arms, projectiles or material calculated to cause unnecessary suffering, as well as the use of projectiles prohibited by the Declaration of St. Petersburg of 1868;

The foundational principle expressed in Article 12, a *sine qua non* of weapons law, is repeated substantially unchanged in all major instruments from that time forward, including those which bind States today (most significantly for the purposes of this thesis, API). It is from that principle that the specific rules of the law of weaponry are derived.

The Brussels Declaration reiterated the prohibition on exploding rounds that had been made at St Petersburg,[38] adding a ban on 'poison or poisoned weapons'[39] (although that was arguably already considered customary law at the time),[40] and made an additional statement that would prove to be of lasting significance: the prohibition in article 13(e) on '[t]he employment of arms, projectiles or material calculated to cause unnecessary suffering', which links the basis of legal restrictions on weapons to the causing of unnecessary suffering, is perhaps the first example of States attempting to set out a general principle that could guide development and use of new types of weapons.

Although the draft text was adopted by the Conference with only minor changes, it was not accepted as a binding convention and was not ratified. It did, however, have some effect on development of the law in that it informed the writing of the Oxford Manual of the Laws and Customs of War, released in 1880,[41] and together those two documents contributed to the most significant events in the evolution of weapons law until that time, the Hague Peace Conferences of 1899 and 1907.

The First Hague Peace Conference of 1899 was convened on the initiative of Tsar Nicholas II of Russia, largely in response to concerns about military build-ups and developments in weapons technology in Europe, 'with the object of seeking the most effective means of ensuring to all peoples the benefits of a real and lasting peace and,

---

[38] *Brussels Declaration* art 13(e).
[39] *Brussels Declaration* art 13(a).
[40] Henckaerts and Doswald-Beck, above n 19, 251.
[41] *The Laws of War on Land: Manual Published by the Institute of International Law*, adopted 9 September 1880 (reprinted in: *Resolutions of the Institute of International Law Dealing with the Law of Nations* (James Brown Scott trans, William S Hein & Co, 2003) 26.

above all, of limiting the progressive development of existing armaments'.[42] (A somewhat less idealistic view of the motive behind the 1899 Conference is presented in several publications:[43] in early 1898 word reached St Petersburg that Austria was about to acquire a particularly effective quick-firing gun that the French and German armies had begun using the previous year.[44] Realising that budgetary restrictions would inhibit its ability to keep pace, the Russian government instead proposed the Peace Conference, with the intention of fixing the military *status quo* in Europe and freeing Russia to devote resources to other challenges.)

Unlike the Conference of 1874, the participating States, 26 in all, were able to reach agreement on a new set of rules governing land warfare. The Final Act of the Conference[45] was signed by the delegates as an authoritative account of the results achieved and was later ratified by States, becoming binding law. The Conference produced three Conventions covering general matters relating to the law of war[46] and three Declarations prohibiting the use of specific weapons: projectiles launched from balloons,[47] projectiles intended to deliver asphyxiating or poisonous gases,[48] and expanding ('dum-dum') bullets.[49]

The second conference, in 1907, was attended by representatives of 44 governments, 'for the purpose of giving a fresh development to the humanitarian principles which served as a basis for the work of the First Conference of 1899.'[50] Like the first Conference, its Final Act was ratified by States, and the thirteen Conventions and one

---

[42] 'Russian Circular Note Proposing the Program of the First Conference' in James Scott Brown (ed), *The Hague Conventions and Declarations of 1899 and 1907* (Oxford University Press, 2nd ed, 1915) xvii.
[43] See, eg, Arthur Eyffinger, *The 1899 Hague Peace Conference: The Parliament of Man, the Federation of the World* (Martinus Nijhoff, 1999) 20; Michael A Hennessy and B J C McKercher (eds), *War in the Twentieth Century: Reflections at Century's End* (Praeger, 2003) 46.
[44] That gun was most likely the *Canon de 75 Mle Modèle 1897*, 'history's first truly modern artillery piece and the most advanced field gun in any army at the start of World War I': Spencer C Tucker, *Instruments of War: Weapons and Technologies That Have Changed History* (ABC-CLIO, 2015) 135.
[45] 'Final Act of the International Peace Conference' in Brown, above n 42, 1.
[46] *International Convention for the Pacific Settlement of International Disputes*, signed 29 July 1899, [1901] ATS 130 (entered into force 4 September 1900); *International Convention with respect to the Laws and Customs of War on Land*, signed 29 July 1899, [1901] ATS 131 (entered into force 4 September 1900) ('*Hague Convention 1899*'); *International Convention for Adapting to Maritime Warfare the Principles of the Geneva Convention of 22 August 1864*, signed 29 July 1899, [1901] ATS 132 (entered into force 4 September 1900).
[47] 'Declaration IV, 1 to Prohibit for the Term of Five Years the Launching of Projectiles and Explosives from Balloons, and Other New Methods of a Similar Nature' in Brown, above n 42, 220.
[48] 'Declaration IV, 2 Concerning Asphyxiating Gases' in Brown, above n 42, 225.
[49] 'Declaration IV, 3 Concerning Expanding Bullets' in Brown, above n 42, 227.
[50] 'Final Act of the Second International Peace Conference' in Brown, above n 42, 1, 1.

Declaration signed by the delegates were the most comprehensive and detailed statement of the laws of war at the time and 'undoubtedly represented an advance over existing international law'.[51] Later in the twentieth century they were recognised as being declaratory of customary law.[52]

Of particular interest for this thesis is the *Hague Convention 1907*, which covers the laws and customs of war on land. Convention IV furthered the development of weapons law along two paths, by reinforcing and expanding on the general principles by which new weapons would be judged, and by adding to the list of prohibitions and restrictions on specific types of weapons. Section II of that Convention, entitled 'On Hostilities', states in part:

> Art. 22. The right of belligerents to adopt means of injuring the enemy is not unlimited.
>
> Art. 23. In addition to the prohibitions provided by special Conventions, it is especially forbidden
>
> (a) To employ poison or poisoned weapons;
>
> …
>
> (e) To employ arms, projectiles, or material calculated to cause unnecessary suffering;

Interestingly, in the 1899 Convention II, which also covered the laws and customs of war on land, the equivalent provision to art 23(e) above prohibited 'arms, projectiles, or material *of a nature* to cause superfluous injury'[53] (emphasis added). This difference in terminology is not explained in relation to the Hague Conferences, but today the 1907 version is broadly considered to be based on a mistranslation from the original French.[54] The difference in position reappeared in later years in the form of controversy about whether an 'effects based' or 'design purpose' approach should be used to determine

[51] *Trial of the Major War Criminals before the International Military Tribunal, Nuremberg, 14 November 1945 – 1 October 1946* (1947) vol 1, 253-4.
[52] Ibid 254.
[53] *Hague Convention 1899* at 23(e).
[54] 'International Criminal Court: Making the Right Choices – Part V: Recommendations to the Diplomatic Conference' (Position Paper No IOR 40/010/1998, Amnesty International, 30 April 1998) 14.

whether a weapon violates this prohibition,[55] although the former approach is generally accepted today.

To these provisions, otherwise substantially unchanged from those in the Brussels Declaration of 1874, were added prohibitions on the 'discharge of projectiles and explosives from balloons'[56] and restrictions on the use of submarine mines. The Convention Relative to the Laying of Automatic Submarine Contact Mines[57] continued the practice, begun in the Declaration of St Petersburg, of linking the legality of weapons to particular technical characteristics. Specifically, the Convention forbids:

> 1. To lay unanchored automatic contact mines, except when they are so constructed as to become harmless one hour at most after the person who laid them ceases to control them;
>
> 2. To lay anchored automatic contact mines which do not become harmless as soon as they have broken loose from their moorings;
>
> 3. To use torpedoes which do not become harmless when they have missed their mark.[58]

These provisions represent progress in the law's ability to address technological change in weaponry in that they describe specific behaviours that mines must follow in order to comply with the law, but stop short of describing any particular technical implementation of those behaviours. That drafting approach makes it possible to apply the written law to technologies unforeseen at the time.

Relatively little progress of relevance to this thesis was made in the field of weapons law over the next half century, despite the horrors of two world wars. The next major development occurred in 1974 when the Government of Switzerland convened the Diplomatic Conference on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts.[59] The outcome of that Conference,

---

[55] See, eg, W Hays Parks, 'Conventional Weapons and Weapons Reviews' (2005) 8 *Yearbook of International Humanitarian Law* 55, 86 n 123.
[56] 'Declaration XIV Prohibiting the Discharge of Projectiles and Explosives from Balloons' in Brown, above n 42, 220.
[57] *International Convention relative to the Laying of Automatic Submarine Contact Mines*, signed 18 October 1907, [1910] ATS 11 (entered into force 26 January 1910).
[58] Ibid art 1.
[59] Federal Research Division, *Official Records of the Diplomatic Conference on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts* (16 July 2010) Library of Congress <https://www.loc.gov/rr/frd/Military_Law/RC-dipl-conference-records.html>.

the two Protocols Additional to the Geneva Conventions of 12 August 1949,[60] united Geneva law and Hague law, the two main strands of the law of war that had been developing in parallel until that time.

True to its name, the Conference both reaffirmed and developed the law in important ways. Article 35, introducing the section covering methods and means of warfare, provides in part:

> 1. In any armed conflict, the right of the Parties to the conflict to choose methods or means of warfare is not unlimited.
>
> 2. It is prohibited to employ weapons, projectiles and material and methods of warfare of a nature to cause superfluous injury or unnecessary suffering.

These rules essentially reiterate the foundational principles of weapons law.[61] Importantly, Article 35(2) also broadens the scope of the general prohibition to include methods of warfare (tactics, operating procedures and so forth) as well as means (weapons and associated equipment).

Two other developments of the law by the drafters of API are relevant to this thesis. First, article 51(4) provides:

> 4. Indiscriminate attacks are prohibited. Indiscriminate attacks are:
>
> (a) Those which are not directed at a specific military objective;
>
> (b) Those which employ a method or means of combat which cannot be directed at a specific military objective; or
>
> (c) Those which employ a method or means of combat the effects of which cannot be limited as required by this Protocol;
>
> and consequently, in each such case, are of a nature to strike military objectives and civilians or civilian objects without distinction.

---

[60] *API*; *Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II)*, opened for signature 8 June 1977, 1125 UNTS 609 (entered into force 7 December 1979) ('*APII*').

[61] de Preux, above n 2, 390-9 [1382]-[1409].

The principle of distinction expressed in this article, already generally recognised as a customary rule of armed conflict, was crystallised into a written rule which specifically regulated the design of weapons (or the choice of weapons for a particular attack) for the first time in API. Earlier documents had referred to the need to avoid attacking civilian buildings,[62] to avoid bombarding undefended towns,[63] and similar injunctions, but none had explicitly stated that the legality of an attack depends on directing it at a 'specific military objective' and using a weapon which is designed to be sufficiently discriminate in the circumstances. Taken together, paragraphs (b) and (c) above impose a requirement that the weapons used in an attack must be sufficiently capable of being directed at the military objective being attacked, and not at civilians or civilian objects, and the sufficiency of that capability is to be judged in the circumstances of that particular attack.[64]

The International Court of Justice later described the principle of distinction as a 'cardinal' principle of humanitarian law.[65] In reference to both distinction and the principle forbidding superfluous injury or the infliction of unnecessary suffering, the court stated that 'these fundamental rules are to be observed by all States whether or not they have ratified the conventions that contain them, because they constitute intransgressible principles of international customary law.'[66]

The second relevant development in API is the imposition of an express obligation to take precautions in relation to an attack. The most significant rules for the purposes of this thesis are given in Article 57:

> 2. With respect to attacks, the following precautions shall be taken:
>
> (a) Those who plan or decide upon an attack shall:
>
> (i) Do everything feasible to verify that the objectives to be attacked are neither civilians nor civilian objects and are not subject to special protection …;

[62] See, eg, *Brussels Declaration* art 17.
[63] See, eg, *Hague Convention 1899* art 25, *Hague Convention 1907* art 25.
[64] William H Boothby, *Weapons and the Law of Armed Conflict* (Oxford University Press, 2009) 79.
[65] *Legality of the Threat or Use of Nuclear Weapons (Advisory Opinion)* [1996] ICJ Rep 226, 257 [78].
[66] Ibid 257 [79].

(ii) Take all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event to minimizing, incidental loss of civilian life, injury to civilians and damage to civilian objects;

(iii) Refrain from deciding to launch any attack which 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 and direct military advantage anticipated;

To an extent, Article 57 reaffirms rules already laid down in other articles,[67] but the explicit statement that an attacker must take precautions in selecting their means and methods of attack is a new development, even if it was arguably implicit in prior rules.[68]

Article 58 sets out the complementary obligation to take precautions against the effects of an attack by separating civilians and civilian objects from military objectives, as well as 'other necessary precautions'. Although it is significant in the overall development of IHL, Article 58 is not directly relevant to questions about AWS and will not be discussed further.

In parallel with the Conference's work on IHL in general, an ad hoc committee on weaponry met to consider proposals related to regulation of particular weapons. At the same time, the ICRC convened Conferences of Governmental Experts on the Use of Conventional Weapons in 1974 and 1976,[69] building on work it had been doing since the 1950s. Those efforts did not directly bear fruit during the Diplomatic Conference, but did result in a resolution[70] to convene a separate conference to consider regulation of specific weapons. That conference resulted in an important step forward in the regulation of weapons in armed conflict: the Convention on Certain Conventional

[67] Claude Pilloud and Jean Pictet, 'Protocol I – Article 57 – Precautions in Attack' in Sandoz, Swinarski and Zimmermann, above n 2, 677, 679 [2189].
[68] Michael Bothe, Karl Josef Partsch and Waldemar A Solf, *New Rules for Victims of Armed Conflicts* (Martinus Nijhoff, 2013) 403 [2.2].
[69] International Committee of the Red Cross, *Conference of Government Experts on the Use of Certain Conventional Weapons (Lucerne, 24.9-18.10.1974): Report* (1975); International Committee of the Red Cross, *Conference of Government Experts on the Use of Certain Conventional Weapons (Second Session - Lugano, 28.1-26.2.1976): Report* (1976).
[70] *Official Records of the Diplomatic Conference on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts* (1978) vol 1, 215 Resolution 22 (IV).

Weapons of 1980.[71] The CCW process and its relevance to AWS is discussed in more detail in Section 2.2.

This brief overview has omitted mention of a range of conventions, protocols and other legal instruments which, while significant in the context of IHL in general and weapons law in particular, are not directly relevant to this thesis. Those omitted instruments include major legal landmarks such as the 1925 Geneva Protocol prohibiting the use in war of chemical and bacteriological weapons,[72] the 1972 Biological Weapons Convention, the 1993 Chemical Weapons Convention, and a range of others of greater or lesser significance to the conduct of armed conflict today.

The final legal development relevant to this thesis is the establishment, in 1998, of the International Criminal Court (ICC). The ICC is the first permanent treaty-based institution with jurisdiction, as specified in its governing document,[73] over 'the most serious crimes of concern to the international community as a whole',[74] including several war crimes related to violations of weapons law. The Rome Statute was negotiated on the basis that its provisions would reflect customary law,[75] excluding some conventional law governing armed conflict. Violation of treaty obligations through use of weapons prohibited by some of the instruments discussed above is therefore not in itself a crime within the jurisdiction of the ICC. Specifically, Article 8 of the Statute states:

> 2. For the purpose of this Statute, "war crimes" means:
>
> …
>
> (b) Other serious violations of the laws and customs applicable in international armed conflict, within the established framework of international law, namely, any of the following acts:

---

[71] *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 October 1980, 1342 UNTS 137 (entered into force 2 December 1983) ('*CCW*').
[72] *Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous or other Gases, and of Bacteriological Methods of Warfare*, opened for signature 17 June 1925, [1930] ATS 6 (entered into force 8 February 1928).
[73] *Rome Statute of the International Criminal Court*, opened for signature 17 July 1998, 2187 UNTS 90 (entered into force 1 July 2002) ('*Rome Statute*').
[74] Ibid art 5(1).
[75] Boothby, above n 64, 21.

…

(xvii) Employing poison or poisoned weapons;

(xviii) Employing asphyxiating, poisonous or other gases …

(xix) Employing bullets which expand or flatten easily in the human body …

(xx) Employing weapons, projectiles and material and methods of warfare which are of a nature to cause superfluous injury or unnecessary suffering or which are inherently indiscriminate in violation of the international law of armed conflict, provided that such weapons, projectiles and material and methods of warfare are the subject of a comprehensive prohibition and are included in an annex to this Statute …[76]

Thus, the only weapons specifically mentioned are those prohibited in older treaties that have since been generally acknowledged as being representative of customary law. There is provision in article 8(2)(b)(xx) to criminalise use of other weapons that violate the customary prohibitions on weapons that are 'of a nature to cause superfluous injury or unnecessary suffering or which are inherently indiscriminate', but any such prohibitions must be negotiated among the States Parties and annexed to the Statute. At the time of writing, no such prohibitions have been negotiated, so there are no other types of weapons the mere use of which is a crime under the Statute.

2.2 History of the Debate about AWS

As a source of military advantage, there is obvious appeal in a machine, particularly a weapon system, which can do its job without needing human assistance. Military interest in machines with a capacity to operate unattended accordingly stretches back many years. Some weapon systems in use today, which can operate without human interaction once activated, have been in use since the 1970s.[77] Isolated examples can be identified in much earlier years.[78] Due largely to limitations in the associated

[76] These provisions from Article 8(2)(b) of the Rome Statute applicable to international armed conflict have been replicated verbatim in Article 8(2)(e) applicable to non-international armed conflict as a result of amendments to the Rome Statute made at the Review Conference in Kampala in 2010.
[77] Some examples are described briefly in Section 4.1.
[78] However, few, if any, played significant roles in early conflicts. See, eg, Kenneth P Werrell, *The Evolution of the Cruise Missile* (Air University Press, 1985); Markus Wagner, 'The Dehumanization of International Humanitarian Law: Legal, Ethical, and Political Implications of Autonomous Weapon Systems' (2014) 47 *Vanderbilt Journal of Transnational Law* 1371, 1375.

technologies though, their roles in battle have been restricted to carefully controlled situations, and have been somewhat peripheral to mainstream human-operated weapons.

The recent and sudden prominence of AWS may be attributed largely to a conflation of advances in a range of technologies, driven in part by political and financial pressures on armed forces. Those advances are making possible the production of weapon systems with much higher capacities for autonomously performing useful tasks in complex environments than has been seen previously. Concerns about their destructive potential, and about their effect on the practice of warfighting, have closely followed those advances and have been prominent in the debate.[79]

Broad public interest in AWS, and the resulting political and legal debate, arguably began in the early 2000s with publication of military development plans for uninhabited vehicles that included references to increasing levels of autonomy.[80] Soon after, papers by well-known roboticists began the process of addressing some of the associated ethical and legal issues.[81] The public debate began in earnest in 2009 when a range of books, articles and reports were published, discussing the technical developments that had been taking place in armed forces and delving further into the ethical and legal implications.[82] Since then, interest in the issue has spread to law- and policy-making bodies.

---

[79] Those concerns are outlined in Section 2.3.

[80] See, eg, Office of the Secretary of Defense, 'Unmanned Aircraft Systems Roadmap, 2005 – 2030' (US Department of Defense, 4 August 2005) <https://fas.org/irp/program/collect/uav_roadmap2005.pdf>.

[81] See, eg, Ronald C Arkin, 'Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture' (Technical Report No GIT-GVU-07-11, Mobile Robot Laboratory, College of Computing, Georgia Institute of Technology, 2007); Noel Sharkey, 'Cassandra or False Prophet of Doom: AI Robots and War' [2008] (July/August) *IEEE Intelligent Systems* 14; Noel Sharkey, 'Grounds for Discrimination: Autonomous Robot Weapons' [2008] (October) *RUSI Defence Systems* 86; Jason Borenstein, 'The Ethics of Autonomous Military Robots' (2008) 2(1) *Studies in Ethics, Law and Technology* 13; Patrick Lin, George Bekey, and Keith Abney, 'Autonomous Military Robotics: Risk, Ethics, and Design' (Report No ADA534697, US Department of Navy, Office of Naval Research, 20 December 2008) <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA534697>.

[82] See, eg, Peter W Singer, *Wired for War: The Robotics Revolution and Conflict in the 21st Century* (Penguin Press, 2009); United States Air Force, 'United States Air Force Unmanned Aircraft Systems Flight Plan, 2009-2047' (Report No ADA505168, Headquarters, United States Air Force, 18 May 2009) <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA505168> ('*USAF UAS Flight Plan*'); Ronald C Arkin, *Governing Lethal Behavior in Autonomous Robots* (Chapman and Hall, 2009); Wendell Wallach and Colin Allen, *Moral Machines: Teaching Robots Right from Wrong* (Oxford University Press, 2009); Armin Krishnan, *Killer Robots: Legality and Ethicality of Autonomous Weapons* (Ashgate, 2009).

In August 2010, the then Special Rapporteur of the Human Rights Council on extrajudicial, summary or arbitrary executions, Philip Alston, submitted an interim report to the 65th session of the United Nations General Assembly.[83] The section of the report relating to robotic technologies begins by noting that '[o]ver the past decade, the number and type of unmanned or robotic systems developed for, and deployed in, armed conflict and law-enforcement contexts has grown at an astonishing pace' and '[i]n the foreseeable future, the technology will exist to create robots capable of targeting and killing with minimal human involvement or without the need for direct human control or authorization.'[84] After giving an overview of the state of autonomous technologies at the time, Alston notes the 'possible'[85] or 'hypothetical'[86] advantages flowing from the use of increasingly autonomous systems and outlines the areas of concern that, in his view, require in-depth examination: definitions of key terms, international and criminal responsibility, safeguards and standards for deployment, civilian support and *jus ad bellum* considerations.[87] While no answers are provided, Alston recommends formation of 'a group of military and civilian representatives from States, leading authorities in human rights and humanitarian law, applied philosophers and ethicists, scientists and developers to advise on measures and guidelines designed to promote [compliance with international law].'[88] Whether because of that report or not, the States Parties to the CCW have commenced a series of meetings, about which more will be said later, under the auspices of the United Nations Office at Geneva to discuss those very matters.[89]

Alston's successor as Special Rapporteur, Christof Heyns, reiterated many of those points in his first report some months later[90] including, relevantly, 'the need for greater definitional uniformity in relation to the types of technology being developed … and the

---

[83] Philip Alston, *Interim Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, UN GAOR, 65th sess, Agenda Item 69(b), UN Doc A/65/321 (23 August 2010).
[84] Ibid 10 [17].
[85] Ibid 16 [30].
[86] Ibid 16 [31].
[87] Ibid 17-20.
[88] Ibid 21 [48].
[89] Alston also surveys some of the early views expressed by international actors: Philip Alston, 'Lethal Robotic Technologies: The Implications for Human Rights and International Humanitarian Law' (2011) 21(2) *Journal of Law, Information and Science* 35, 55.
[90] Christof Heyns, *Statement by Mr Christof Heyns, Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, UN GAOR, 3rd Comm, 65th sess, Agenda Item 68(a) (22 October 2010).

fundamental question of whether lethal force should ever be permitted to be fully automated.'[91]

Following those discussions in the General Assembly, and perhaps because of them, States began to formulate national positions on AWS. The United Kingdom[92] and the United States[93] have both published formal statements of policy, setting out in broad terms their plans for integrating increasingly autonomous weapons into their armed forces. Other States have expressed their views informally and many appear to be working towards formulating policy positions, although none have yet published official positions.

In April 2013, Christof Heyns produced another report on AWS (referring to them as 'lethal autonomous robotics' (LARs)).[94] The report views AWS as a potentially serious threat, stating that they could 'create serious international division and weaken the role and rule of international law – and in the process undermine the international security system.'[95] Heyns notes the difficulty of regulating new military technologies:

> 'As with any technology that revolutionizes the use of lethal force, little may be known about the potential risks of the technology before it is developed, which makes formulating an appropriate response difficult; but afterwards the availability of its systems and the power of vested interests may preclude efforts at appropriate control.'[96]

On that basis, he recommends that:

> The Human Rights Council should call on all States to declare and implement national moratoria on at least the testing, production, assembly, transfer, acquisition, deployment

[91] Ibid 5.
[92] Ministry of Defence, 'The UK Approach to Unmanned Aircraft Systems' (Joint Doctrine Note 2/11, 30 March 2011) ('*JDN 2/11*'); Joint Doctrine Notes are "raised to either promote debate, place 'markers in the sand' or capture and disseminate best practice. They are not endorsed national doctrine" <https://www.gov.uk/government/groups/development-concepts-and-doctrine-centre>.
[93] *Directive 3000.09*; A US Department of Defense directive '[e]stablishes policy, delegates authority, and assigns responsibilities [to DoD Components]' in relation to the subject matter of the directive: DoD Instruction 5025.01, 1 August 2016.
[94] Christof Heyns, *Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, Human Rights Council, 23rd sess, Agenda Item 3, UN Doc A/HRC/23/47 (9 April 2013).
[95] Ibid 6 [30].
[96] Ibid 7 [33].

> and use of LARs until such time as an internationally agreed upon framework on the future of LARs has been established;[97]

and that a High Level Panel be convened to examine the technical, legal, ethical and policy issues related to AWS.[98]

As of the time of writing, no State actively involved in development of AWS has declared any such moratorium, but Professor Heyns' call for expert examination of the issue has been more successful. Through 2013, approximately 30 States[99] spoke publically about their views on AWS: at the Human Rights Council debate following submission of Professor Heyns' report, at a September seminar on AWS convened by France in cooperation with the United Nations Office for Disarmament Affairs,[100] and at the 68th session of the United Nations General Assembly First Committee on Disarmament and International Security in October.[101] The statements revealed a range of positions for and against AWS, but most noted that a cautious approach is appropriate and expressed support for further discussion in an appropriate forum.

At the 68th UNGA session, the report of the Advisory Board on Disarmament Matters recommended that

> The Secretary-General should … promote coordinated efforts in an existing forum, such as the Conventional Weapons Convention … to address the possible need for disarmament measures in respect of potential future fully autonomous systems. Such a consensus-oriented approach could result in a code of conduct, a road map or other tools and should involve Governments and relevant stakeholders from industry, research, academia and civil society.[102]

---

[97] Ibid 21 [113].
[98] Ibid 21 [114].
[99] Mary Wareham, "Report on Outreach on the UN Report on 'Lethal Autonomous Robotics'" (Report, Campaign to Stop Killer Robots, 31 July 2013) <http://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC_ReportHeynsUN_Jul2013.pdf>.
[100] *France convenes Seminar at the UN* (26 September 2013) Campaign to Stop Killer Robots <http://www.stopkillerrobots.org/2013/09/france-seminar/>.
[101] First Committee, *Disarmament and International Security - Documents of the 68th Session*, General Assembly of the United Nations <http://www.un.org/en/ga/first/68/documentation.shtml>.
[102] *Work of the Advisory Board on Disarmament Matters: Report of the Secretary-General*, UN GAOR, 68th sess, Provisional Agenda Item 101, UN Doc A/68/206 (26 July 2013) 10 [46](d).

That advice was heeded, and beginning with the 2013 Meeting of the High Contracting Parties to the CCW, the main international forum for discussion of AWS has been a series of Informal Meetings of Experts held in Geneva as part of the CCW process.

At the 2013 meeting it was resolved to 'convene in 2014 a four-day Informal Meeting of Experts … to discuss the questions related to emerging technologies in the area of lethal autonomous weapons systems, in the context of the objectives and purposes of the Convention.'[103] That meeting, held in May 2014, was attended by 74 States Parties as well as a range of observers, political and educational bodies, and civil society organisations.[104] The participants first exchanged general views on the use and regulation of autonomous technologies in military contexts. The bulk of the meeting consisted of expert presentations and discussion of developments in AWS technology and their legal, ethical, political and security implications. Subsequent meetings covering substantially the same aspects were held in April 2015[105] and April 2016.[106]

The CCW discussions are still at an early stage. Although some participants have expressed strong views about the prospect of increasingly autonomous weapons, there is a clear consensus that much further work is needed before recommendations on regulatory measures can be made. At the 2016 Informal Meeting of Experts it was resolved to formalise the discussion somewhat by establishing a Group of Governmental Experts:

> The Informal Meeting of Experts recommends that the [CCW States Parties] may decide to establish an open-ended Group of Governmental Experts (GGE) … The GGE should meet … to explore and agree on possible recommendations on options related to emerging technologies in the area of [AWS], in the context of the objectives and

[103] *Final Report of the Meeting of the High Contracting Parties 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*, 2013 sess, Agenda Item 13, UN Doc CCW/MSP/2013/10 (16 December 2013) 4 [32].

[104] *Report of the 2014 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS)*, 2014 sess, Provisional Agenda Item 8, UN Doc CCW/MSP/2014/3 (11 June 2014).

[105] *Report of the 2015 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS)*, 2015 sess, Provisional Agenda Item 8, UN Doc CCW/MSP/2015/3 (2 June 2015).

[106] *Report of the 2016 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS) (Advanced Version)* ('*2016 CCW Report*') <http://www.unog.ch/80256EDD006B8954/(httpAssets)/DDC13B243BA863E6C1257FDB00380A88/$file/ReportLAWS_2016_AdvancedVersion.pdf>

> purposes of the Convention, taking into account all proposals – past, present and future.[107]

The GGE will meet for the first time in August 2017. It will focus on many of the same aspects of AWS that the participants in the Informal Meetings of Experts have been examining: 'definitions, the consideration of instruments for transparency and confidence-building measures, and building on the legal principles and rules applicable to [AWS].'[108]

2.3 Broader Concerns about AWS

In addition to IHL matters, participants in the CCW process have debated interrelated questions raised by AWS development relating to human rights law, ethics and international security, with many questions spanning more than one of those areas. The main questions being asked in each of those areas are outlined below. No attempt is made to discuss those broader concerns in depth in this thesis, although it is hoped that some parts of the analysis which follows, particularly the understanding of weapon autonomy presented in Chapter 3, will be useful to researchers examining those matters.

This section, like the remainder of the thesis, focuses on issues peculiar to AWS. Some of the points raised in the debate so far are broadly applicable to weapon development in general, such as the possibility that continued AWS development may lead to an arms race. Unless an argument is specific to AWS, or applies differently to AWS than to other types of weapons, it has been omitted from this investigation, the purpose of which is solely to clarify the IHL implications of applying technologies of machine autonomy to military weapon systems. Concerns which relate only to the limited technical capabilities of today's weapon systems are also ignored, as the state of the art is progressing rapidly and arguments which claim only that a particular technical capability is not suitable for a particular application, even if valid today, are not likely to remain relevant for long.

---

[107] *Recommendations to the 2016 Review Conference* 1 [3] <http://www.unog.ch/80256EDD006B8954/(httpAssets)/6BB8A498B0A12A03C1257FDB00382863/$file/Recommendations_LAWS_2016_AdvancedVersion+(4+paras)+.pdf>.

[108] *Report of the 2016 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS) (Advanced Version)* 4 [24].

The points made in the debate so far have been quite general, and are evolving as the discussion progresses. The challenge of agreeing on a precise definition of AWS for regulatory purposes has not yet been overcome, and the task requires participants to look ahead to a broad path of development rather than at a specific machine or technology. The arguments made about regulating AWS therefore focus on the most basic, intrinsic characteristics of AWS, and relate primarily to a general notion of 'fully' autonomous weapons, or weapons that would operate without 'meaningful human control'.[109] The precise meaning of those terms and their relevance to IHL and armed conflict is discussed later.[110] It is sufficient for now to say that concerns about regulation generally arise in relation to still-hypothetical weapons that would partly or completely replace humans in combat roles and utilise advanced software which independently implements decision processes relating to the use of force. The defining characteristic of such weapons is that their control software would make the final determination about whether and how to perform an act of violence, rather than simply carrying out a direct command from a human combatant. AWS which allow a sufficient degree of manual human intervention in the use of force, on the other hand, are generally seen as less problematic and are even seen as potentially offering some advantages beyond the military operational advantages which primarily drive AWS development. For example, they may aid fulfilment of human rights and targeting obligations by allowing States to take advantage of more accurate targeting capabilities and more precise control of the force that is applied.[111]

2.3.1 AWS Replacing Human Combatants

The fact of a machine usurping the role of a human soldier in combat to any significant extent provokes a range of questions, independently of how effectively the machine may operate in that role.

One such question is about the human right to life, 'the bedrock of international human rights law.'[112] The right to life is codified in Article 6(1) of the International Covenant

[109] See, eg, Christof Heyns (Note, CCW Meeting of Experts on LAWS: Human Rights and Ethical Issues, April 2016); Amnesty International (Statement, CCW Meeting of Experts on LAWS: Human Rights and Ethical Issues, April 2016).
[110] See sections 3.1.3 and 6.3 respectively.
[111] Heyns, above n 109, 1.
[112] Human Rights Watch, *Shaking the Foundations: The Human Rights Implications of Killer Robots* (2014) 8.

on Civil and Political Rights[113] ('Every human being has the inherent right to life. This right shall be protected by law. No one shall be arbitrarily deprived of his life.') and in many other human rights instruments. It has been argued that AWS would necessarily violate the right to life of anyone who is subject to their attack because allowing a machine to 'decide' to kill a human and act on that decision without human intervention amounts to an arbitrary choice on the part of those who deploy the AWS.[114] It is argued that this is particularly the case where it may be difficult to hold any person accountable for the actions of an AWS.

The other major human rights concern is about the right to dignity.[115] 'The concept of human dignity lies at the heart of international human rights law.'[116] Although it is not a standalone right, the inherent dignity of the human person is seen as the basis of human rights[117] and as a concept which influences the interpretation of all rights. In the context of AWS, human dignity is most often linked to the right to life, and sometimes to the right against inhumane treatment. It is argued that '[a] human being in the sights of a fully autonomous machine is reduced to being an object'.[118] That is, a machine cannot comprehend the value of life or the significance of its loss, so when there is no human whose role is to deliberate over whether it is necessary and justifiable to target a person, the value of the targeted person's life is compromised and their human dignity is lost.

The dignity of combatants responsible for operating AWS, as well as those targeted by them, may also be at issue:

> If the decision whether to use deadly force is taken out of the hands in whose name it is being done, they cease to be moral agents: people who take a decision and assume responsibility for it. It is no wonder that so many military officers are reluctant about the introduction of AWS. Their dignity is on the line.[119]

---

[113] *International Covenant on Civil and Political Rights*, opened for signature 16 December 1966, 999 UNTS 171 (entered into force 23 March 1976) ('*ICCPR*').
[114] Christof Heyns (Comment, CCW Meeting of Experts on LAWS: Overarching Issues, April 2015) 5.
[115] For an overview see Robin Geiss, 'The International-Law Dimension of Autonomous Weapons Systems' (Study, Friedrich-Ebert-Stiftung, International Policy Analysis, October 2015) 17.
[116] Human Rights Watch, above n 112, 23.
[117] *ICCPR* Preamble.
[118] Heyns, above n 114, 5.
[119] Ibid 6.

A more novel argument is based on the idea that modern warfare should be seen as 'a form of governance, in which [S]tate-executive power is exercised vis-à-vis individuals',[120] and consequently States involved in conflicts bear obligations stemming from administrative law.[121] One such obligation is to exercise suitable discretion in relation to each administrative decision affecting an individual.[122] The argument is that

> AWS, in this context, reflect the use of executive power against individuals through a computerized proxy, and thus without proper administrative discretion. … As long as they are incapable of *human*-like metacognition, each of their actions is a case where executive decision is made on the basis of discretion bound in advance. Operating on the basis of bound discretion, especially where human life is concerned, is *per se* arbitrary and … contradicts basic notions of administrative law.[123]

Thus use of AWS would violate the human right to just administrative action.[124]

Much of the ethical debate about AWS has so far been 'a circular argument between result-oriented approaches and deontological, "moral" approaches.'[125] The arguments made under the deontological approach, both for and against further development of AWS, have largely coincided with those made in the human rights debate. On the other hand, three recurring result-oriented concerns relate to the likely effects of weapon autonomy on the behaviour of weapon system operators. The first effect is the well-known phenomenon of automation bias, being the tendency of operators to place excessive trust in computer-made decisions.[126] The second is that AWS may act as a moral buffer which gives rise to an 'operator's tendency to morally disengage from robotic actions or behaviour.'[127] 'As the philosopher Emmanuel Levinas has shown, the

---

[120] Eliav Lieblich (Note, CCW Meeting of Experts on LAWS: Human Rights and Ethical Issues, April 2016) 1.
[121] For a non-conflict discussion of automated decision-making and administrative law see Danielle Keats Citron, 'Technological Due Process' (2008) 85 *Washington University Law Review* 1249.
[122] For an argument against AWS on the basis of an obligation to exercise administrative discretion see Eliav Lieblich and Eyal Benvenisti, 'The Obligation to Exercise Discretion in Warfare: Why Autonomous Weapons Systems are Unlawful' in Nehal Bhuta et al (eds), *Autonomous Weapons Systems: Law, Ethics, Policy* (Cambridge University Press, 2016) 245.
[123] Ibid.
[124] Heyns, above n 109, 2.
[125] Lieblich, above n 120, 1.
[126] See, eg, Mary L Cummings, 'Automation and Accountability in Decision Support System Interface Design' (2006) XXXII *The Journal of Technology Studies* 23.
[127] Daniele Bourcier (Presentation, CCW Meeting of Experts on LAWS: Human Rights and Ethical Issues, April 2016) 29; Wagner, above n 78, 1410.

face-to-face encounter with the "other" is one of the fundamental experiences which arouses the moral conscience and responsibility.'[128] Together, these effects may reduce the likelihood that a human operator would intervene to prevent undesirable actions by an AWS, thereby possibly allowing the machine to perform proscribed acts or apply force unnecessarily.

Unnecessary use of force may also result from lower barriers to entry to armed conflict. On the one hand, the safety of one's own forces is an important and legitimate consideration for a State, and use of weapon systems which can replace soldiers in combat situations offers clear advantages in that respect.[129] However, the corollary is that 'use of [AWS] increases the probability of triggering conflicts, as they do not entail the loss of human lives on the part of the user'.[130] The lower political cost of losing a machine in combat rather than a human may lead to States being more ready to use military action as a means of settling disputes, and to armed forces personnel being more ready to use force on the battlefield.

2.3.2 Technical Capabilities of AWS

A separate set of questions relates to the standard to which an AWS might be able to perform the combat tasks assigned to it. Weapons in use today that have some ability to identify and attack targets unattended are able to do so only in narrow clearly defined circumstances. Military plans for future AWS indicate that they will need to be able to operate in much more varied and complex scenarios. They will face daunting technical challenges in identifying valid targets and calibrating the force that is to be applied. Where AWS misidentify targets, their operators risk violating the right to life of any protected persons attacked by the weapons. Those which do not correctly calibrate the amount and type of force applied risk doing disproportionate harm.[131] Other commentators point out that autonomous targeting is not just a threat to protected persons, but also an opportunity to afford greater protection.[132] Future AWS are likely

[128] Holy See (Text, CCW Meeting of Experts on LAWS: Overarching Issues, April 2015) 4.
[129] See, eg, Wolfgang Richter, 'Utility and Limitations of the Use of LAWS in Military Operations' (Text, CCW Meeting of Experts on LAWS: Operational and Military Aspects, May 2014) 5.
[130] Holy See, above n 128, 3; Wagner, above n 127, 1418.
[131] Heyns, above n 109, 1.
[132] Alan Backstrom and Ian Henderson, 'New Capabilities in Warfare: An Overview of Contemporary Technological Developments and the Associated Legal and Engineering Issues in Article 36 Weapons Reviews' (2012) 94 *International Review of the Red Cross* 483, 489.

to be equipped with advanced sensors and software that could exceed human capabilities. Given the rapidly increasing pace and complexity of combat, it is possible that AWS might be able to make superior targeting decisions,[133] and retaining direct human involvement 'during times of hostile action, is of little value and perhaps even dangerous effect in the sense that humans are unable to rapidly process some forms of information required to maintain last minute veto action.'[134]

On the other hand, a failure in an advanced AWS would be likely to have a much greater damage potential than an equivalent failure by a human wielding a manual weapon:

> If [a manually operated] weapon system begins to fail, the human controller can modify the weapon system's operation or halt the engagement before further damage is done.
>
> With an [AWS], however, the damage potential before a human controller is able to intervene could be far greater. In the most extreme case, an [AWS] could continue engaging inappropriate targets until it exhausts its magazine, potentially over a wide area. If the failure mode is replicated in other [AWS] of the same type, a military could face the disturbing prospect of large numbers of [AWS] failing simultaneously, with potentially catastrophic consequences.[135]

The possibility of 'mass fratricide' as a result of an equipment failure is a novel risk unique to AWS. Nor is the possibility of failure limited to the hardware and software of the machine itself:

---

[133] Indeed, a stark example of the potential advantages of an automated targeting process was the well-known 1988 incident involving the *USS Vincennes*. The *Vincennes*, a naval cruiser, was fitted with the Aegis combat system, one of the functions of which was to detect and assess possible incoming air attacks and, if authorised by a crew member, launch a missile to intercept the incoming missile or aircraft. On July 3, 1988 in the Strait of Hormuz, the *Vincennes'* Aegis system detected an Iranian airliner and correctly identified it as a civilian aircraft. Tragically, the Aegis was overridden by the *Vincennes* crew and the airliner was shot down because of concerns that the system had failed to identify an incoming military aircraft which posed a direct threat to the US fleet: William Fogarty, 'Formal Investigation into the Circumstances Surrounding the Downing of Iran Air Flight 655 on 3 July 1988' (Investigation Report, 19 August 1988) <http://www.dod.mil/pubs/foi/Reading_Room/International_Security_Affairs/172.pdf>; Kristen A Dotterway, *Systematic Analysis of Complex Dynamic Systems: The Case of the USS Vincennes* (Masters Thesis, Naval Postgraduate School, 1992) <http://www.dtic.mil/get-tr-doc/pdf?AD=ADA260260>.

[134] Jai Galliott (Presentation, CCW Meeting of Experts on LAWS: Security Issues, April 2016) 1; for other examples of failures of human judgement see Daniel Sukman, 'Lethal Autonomous Systems and the Future of Warfare' (2015) 16 *Canadian Military Journal* 44, 50.

[135] Paul Scharre, 'Autonomous Weapons and Operational Risk' (Center for a New American Security, Ethical Autonomy Project, February 2016) 5 <https://www.cnas.org/publications/reports/autonomous-weapons-and-operational-risk>.

> causes of failure may reflect interactions with the user or operator, as well as the overall context including the environment, the behaviour of adversaries and friendly forces, as well as the socio-technical system in which any technology emerges and is used. … any of these factors might interact with any other — or others — to compound the initial failure.[136]

Outside of malfunctions and equipment failures, the general issue of controllability is particularly pertinent to AWS, many of which will be designed to operate unattended and remain deployed for long periods of time. It has been argued that '*[o]ne of the conditions to be able to use armed robots is the fact that we can be assured that it will never produce behaviour that is a priori prohibited by its user.*'[137] Further, '*it is fundamentally immoral to utilize a weapon the behaviour of which we cannot completely control.*'[138] While one may take issue with the absolute terms in which those arguments have been put, predictability of behaviour will certainly be a contentious matter in deploying AWS. The difficulty of predicting the behaviour of complex software systems, and the impossibility of exhaustively testing them, particularly when they will operate in complex and chaotic environments, is well known and extends far beyond weaponry.

2.3.3 Accountability

Many observers believe that allowing a machine to both select a target and 'pull the trigger' without referring the decision to a human would create an accountability gap: as no human played a direct part in attacking that particular target, nobody could be held accountable for the outcome. The human right to life requires that where a person is arbitrarily deprived of life, someone must be in a position to be held accountable. 'A lack of accountability in itself constitutes a violation of the right to life.'[139]

An inability to hold a person responsible for a wrongful act also deprives wronged parties of a remedy, in violation of the widely recognised human right to a remedy, as expressed in the ICCPR's requirement that each State party 'ensure that any person whose rights or freedoms as herein recognized are violated shall have an effective

---

[136] John Borrie (Presentation, CCW Meeting of Experts on LAWS: Security Issues, April 2016) 2.
[137] Holy See, above n 128, 8 (emphasis in original).
[138] Ibid (emphasis in original).
[139] Heyns, above n 109, 2.

remedy'.[140] A lack of accountability prevents States from investigating, prosecuting or punishing rights violations and prevents victims from receiving reparations.

Some commentators also relate human dignity to 'accountability, remedy, and respect',[141] where there is a lack of meaningful human control:

> Without that control, it's unclear who or what is responsible for wrongful targeting and errors. Without being able to identify responsible parties, there's no one to demand remedy from or compensation for wrongful injury or death. And making decisions to kill another human being – one of the most profoundly serious decisions we can ever make – demands respect in the form of reflection and meaning that machines cannot give.[142]

More broadly, a lack of accountability may also make it easier for States to resort to force:

> to the extent that a system is entrenched whereby the exercise of the critical functions of force delivery are out of the hands of human beings, including their political leaders, it may become easier for such entities to take the decision to use force – and when things go wrong for those in power to relay [sic] on the defence – explicitly or implicitly - that 'the machine did it'.[143]

2.3.4 Accessibility to Non-State Actors

Since the end of the Cold War, the nature of armed conflict has shifted in favour of greater participation by non-State armed groups including insurgents and terrorist groups. If development of AWS continues, it would be prudent to assume that advanced weapons may be acquired by non-State actors, although perhaps after some delay, as the cost is driven down and the technologies made more accessible through wider adoption by States.[144] There are already indications that the precursor technologies to autonomous Unmanned Aerial Vehicles (UAVs) are favoured by non-State groups:

> … in July 2014, a drone laden with explosives, hovering over Mafraq, northeast of Amman was destroyed by the Jordanian Armed Forces. It is believed that ISIS operated

[140] *ICCPR* art 2(a).
[141] Patrick Lin (Text, CCW Meeting of Experts on LAWS: Overarching Issues, April 2015) 2.
[142] Ibid.
[143] Heyns, above n 109, 2.
[144] Collin S L Koh (Note, CCW Meeting of Experts on LAWS: Security Issues, April 2016) 3-4.

> it. [In March 2015], a coalition aircraft destroyed a drone operated by ISIS that was probably used for surveillance and reconnaissance purposes near the Iraqi city of Fallujah. One of these drones was bought 'off- the-shelf', while the other was either captured or smuggled into the country and modified so that it could carry explosives. On March 22, a Shiite militia group, Saraya Al Salam, staged a demonstration of two unmanned ground vehicles equipped with machine guns in Iraq.[145]

Given the dual use nature of many of the component technologies of AWS, it is likely to be difficult to tightly control access to the hardware and software that could be used to construct a weapon system. Possession of advanced weapons by non-State actors would potentially affect the progress of armed conflicts and other situations of violence.

2.3.5 Effects on Civilian Populations

Armed forces cannot afford to ignore the broader effects of weaponry on the peoples and regions against which they are directed. Weapons have a psychological, as well as physical, impact on the populations which suffer their effects. 'Certain weapons … are of a nature which provokes or augments the stress of civilian populations.'[146] That propensity to cause stress affects the ethical evaluation of a weapon system as well as the security situation in the region and the difficulty of restoring the normal functioning of society after the conflict. The resentment among civilians caused by the United States' UAV operations in the Middle East has been widely acknowledged:[147] 'We are absolutely … sowing the seeds for the next round of people to hate us. The next round of terrorists, as we call them within the intelligence community.'[148] More highly autonomous UAVs would be well suited to continue those activities, and other AWS are likely to be used for similar operations which require stealth and persistence in order to search for desired targets.

2.4 Questions Discussed in this Thesis

The debate about AWS is too broad to be properly covered in a single thesis. This research focuses on only one aspect, being the question of whether and in what

---

[145] Jean-Marc Rickli (Text, CCW Meeting of Experts on LAWS: Overarching Issues, April 2015) 4.
[146] Holy See, above n 128, 4-5.
[147] See, eg, Peter Bergen and Katherine Tiedemann, 'Washington's Phantom War: The Effects of the US Drone Programs in Pakistan' (2011) 90(4) *Foreign Affairs* 12; Sarah Kreps, 'Flying Under the Radar: A Study of Public Attitudes Towards Unmanned Aerial Vehicles' (2014) 1(1) *Research & Politics*.
[148] Vadim Kozyulin (Presentation, CCW Meeting of Experts on LAWS: Security Issues, April 2016) 5.

circumstances States can continue to observe their obligations under IHL while wielding increasingly autonomous weapon systems in armed conflict. Specifically, three unresolved questions currently facing the participants in the CCW discussions are addressed.

2.4.1 Defining and Mapping Autonomy

The first question is as much technical as legal. The task is to understand machine autonomy as it will apply to weapon systems, and the ways in which it might give rise to novel legal questions. The CCW participants face three significant challenges in this endeavour.

First, the subject matter is highly technical, with legal issues potentially arising from the nature of a host of technologies, most notably those relating to robotics and artificial intelligence. In addition to being specialised subjects well outside the typical field of expertise of lawyers, the state of the art in these areas is advancing rapidly and any detailed analysis is at risk of becoming obsolete before it is complete.

The second, as mentioned above, is that the CCW mandate is restricted to 'emerging technologies', rather than including existing weapon systems, which forces the investigation to be somewhat hypothetical. Precursors to those future technologies exist today in deployed weapon systems and laboratory prototypes. However, with few exceptions, systems currently in use appear to have a very low capacity for autonomous operation, generally below that needed to raise significant legal questions. Published plans give only general descriptions of the future forms of more advanced systems, the technologies driving them and the manner in which they will be used.

The third is that, unlike other weapons regulated under the CCW, AWS are not a discrete, well-defined set of devices or technologies. Machine autonomy is rather amorphous, a capability that may be present in any type of weapon system, affecting any function of the weapon system, and to any degree. Taken together, these challenges make the definitional phase of regulating AWS both more important and more difficult than it has been in relation to other weapons.

From the outset of the discussions there was an understanding among delegations that an AWS is a weapon system that can manage its own operation in some way. The Swiss

delegation has described AWS as 'weapon systems that are capable of carrying out tasks governed by IHL in partial or full replacement of a human in the use of force, notably in the targeting cycle.'[149] Other delegations likewise described AWS as weapon systems that might operate 'without human intervention',[150] 'without meaningful human control'[151] or 'without any human control or oversight',[152] but reaching agreement on a more nuanced understanding has proved difficult. Chris Jenks noted the complexity of autonomy in his 2016 presentation:

> Autonomy is not a discrete property. Autonomy reflects task allocation between human and machine across several spectrums. And depending on the system, functions are performed concurrently as well as sequentially and with varied allocations to human and machine.[153]

Experts at the Informal Meetings of Experts have presented a range of views on what autonomy means in terms of weapon systems, and how it may be mapped or quantified.

Leon Kester noted that 'the basis for autonomy is the automation of situation assessment and situation management.'[154] Situation assessment in this context refers to a machine's ability to sense its environment, identify objects and their natures and relationships, and evaluate risks and benefits. Situation management uses the results of the assessment process to formulate plans of action and carry them out. That is, Kester sees an autonomous system as one that can perceive and assess its environment and then formulate and execute appropriate actions without relying on a human operator.

Similarly, Raja Chatila described autonomy as the '[c]apacity of a system to decide and act without the assistance of another agent'.[155] He noted that a robotic system is a machine 'endowed with four basic capacities' (data acquisition, data interpretation, decision-making and action execution),[156] and that '[a]utonomy can be defined in

---

[149] Switzerland (Statement, CCW Meeting of Experts on LAWS: Towards a Working Definition of LAWS, April 2016) 1.
[150] Germany (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, May 2014) 2-3.
[151] Norway (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, May 2014) 1.
[152] Sweden (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, May 2014) 1.
[153] Chris Jenks (Notes, CCW Meeting of Experts on LAWS: Towards a Working Definition of LAWS, April 2016) 2.
[154] Leon Kester (Notes, CCW Meeting of Experts on LAWS: Mapping Autonomy, April 2016) 1.
[155] Raja Chatila (Text, CCW Meeting of Experts on LAWS: Technical Issues, May 2014) 5.
[156] Ibid 3.

relation to the aforementioned capacities and their integration.'[157] Professor Chatila also distinguished between 'operational autonomy', being autonomy in executing actions, and 'decisional autonomy', being autonomy is making non-trivial decisions about which actions to perform.[158]

One of the fundamental questions being addressed is how to distinguish weapon systems that should be considered 'autonomous', and therefore possibly subject to regulatory measures, from those that are merely 'automated' or 'automatic'. Some experts and delegations perceive discrete levels of autonomy. Paul Scharre sees the degree of human control exercised over an AWS as spanning three levels: a 'human in the loop' (semi-autonomous) system waits for human input before taking action; a 'human on the loop' (supervised autonomous) system allows, but does not require, human intervention in real time; a 'human out of the loop' (fully autonomous) system does not allow for real-time human intervention.[159] Various speakers presented considerably more complex taxonomies of autonomy, some of which will be outlined in Section 3.1.3 below.

For those who hold the view that autonomy exists in discrete levels defined by some important variable, such as the facility for human intervention, the question is which levels of autonomy should be subject to regulation? For example, France, in a working paper submitted to the 2016 Meeting, claimed that '[r]emotely operated weapons systems and supervised weapons systems should not be regarded as [AWS] since a human operator remains involved, in particular during the targeting and firing phases' and '[AWS] should be understood as implying a total absence of human supervision, meaning there is absolutely no link (communication or control) with the military chain of command.'[160] Other delegations expressed the view that semi-autonomous weapon systems might also be subject to appropriate regulations.

Some delegations and experts eschewed the idea of discrete levels of autonomy, arguing instead that '[m]achine automation/autonomy exists on a continuum'[161] and '[t]here is a

[157] Ibid 5.
[158] Ibid.
[159] Paul Scharre (Text, CCW Meeting of Experts on LAWS: Technical Issues, May 2014) 4.
[160] France, 'Characterization of a LAWS' (Working Paper, CCW Meeting of Experts on LAWS, April 2016) 1.
[161] Sweden, above n 152.

continuum between the deterministic and predictable automaton and the autonomous system.'[162] In this case, the question is where to place the threshold degree of autonomy which qualifies an AWS as being of legal interest.

Still others note that the attempt to assign a static measure of autonomy to a machine, whether as a discrete class or a point on a continuum, is itself unproductive: 'because of the dynamic nature of functions within a system, attempting to quantify levels of autonomy oversimplifies and though superficially attractive is, in the end, unhelpful.'[163]

Another relevant aspect of the discussions has been about the dimensions along which autonomy may be measured. Dahlmann noted that autonomy may be seen in the human-machine command and control relationship, the complexity of the machine and the type of decision being placed under machine control.[164] Jenks argued that while many functions of a weapon system could be made autonomous to some degree (navigation, diagnostics, refuelling, and so on), autonomy in the critical functions of selecting and engaging targets is most important for regulatory purposes.[165] Chatila noted that a machine's level of autonomy depends on the complexity of the environment in which the machine is operating and the complexity of the task it is performing as much as on intrinsic capabilities of the machine itself.[166]

Finally, the point was made that any working definition would need to be broad and flexible enough to allow for future technological developments and that, while the question of a definition has political consequences, it should not be used as a tool to prejudge the outcome of the discussions.[167] In particular, Switzerland argued that

> the CCW should aim for a purpose-oriented working definition that corresponds to where we are in this debate. … It should not seek to draw a line between desirable, acceptable or unacceptable systems, and should not prejudge the question of appropriate regulatory response for such systems down the line.[168]

---

[162] Chatila, above n 155, 6.
[163] Jenks, above n 153, 2.
[164] Anja Dahlmann (Presentation, CCW Meeting of Experts on LAWS: Towards a Working Definition of LAWS, April 2016) 3.
[165] Jenks, above n 153, 2.
[166] Chatila, above n 155, 6.
[167] *2016 CCW Report*, above n 106, 6 [37].
[168] Switzerland, above n 149, 2.

Chapter 3 argues for a view of weapon autonomy which accurately captures its legal significance. A 'bottom up' approach is used, beginning with an outline of the technological ingredients of an autonomous machine, and then discussing some aspects of planned military applications. The view of autonomy resulting from this process is technically accurate, legally useful and politically neutral. It is not based directly on the CCW discussions, but attempts to provide the participants in those discussions with tools for moving the debate forward.

2.4.2 Challenges to IHL

The second question is about the ability of States to meet their obligations under IHL while utilising highly autonomous weapon systems. In the area of public international law governing resort to lethal force, a distinction is maintained between the *jus ad bellum*, governing the legality of the resort to violence, and the *jus in bello*, governing the legality of conduct in armed conflict. The principles and rules by which the legality of a weapon system and its use in armed conflict are to be judged are derived mainly from the *jus in bello*, in which they form the sub-disciplines commonly known as weapons law[169] (on the nature and design of weapons) and targeting law[170] (on the use of weapons). This thesis is primarily concerned with these sub-disciplines of the *jus in bello*. While there is common agreement that IHL is fully applicable to AWS, there is considerable disagreement about whether the capacity of a weapon to select and attack targets without human intervention would, in some or all circumstances, violate the rules of either weapons law or targeting law. At present there is no conventional law specifically addressed to AWS, nor do any generally recognised customary rules expressly mention them. The rules of most interest in this thesis are therefore the generally recognised customary and conventional rules of weapons law and targeting law. Many of the treaties mentioned in this thesis also contain elements of arms control law, being 'that part of public international law which deals both with the restrictions internationally placed upon the freedom of behaviour of States with regard to their

[169] See generally Boothby, above n 64.
[170] See generally Henderson, above n 9.

national armaments, and with the applicable supervisory mechanisms.'[171] Those arms control elements will be discussed where relevant.

There has been relatively little progress made in the CCW Informal Meetings of Experts on assessing the legal challenges presented by increasingly autonomous weapon systems. Doubtless this is due largely to the difficulty of defining and mapping weapon autonomy, the results of which will underlie the eventual legal analysis. Many States and other organisations have maintained the positions they took at the first meeting in 2014, although the discussions and expert presentations have led to generally better informed arguments and a degree of consensus on some issues, as described below. Three broad areas of concern have emerged in the discussions so far.

The main concern in relation to IHL is about the technical possibility for AWS to comply with basic IHL requirements relating to distinction, proportionality and precautions in attack. On this point a wide range of views have been expressed. At one extreme are those who claim that AWS will never be able to comply with those legal principles. Pakistan's statement at the 2016 meeting captured the essence of this view:

> [AWS] cannot be programmed to comply with [IHL], in particular its cardinal rules of distinction, proportionality, and precaution. These rules can be complex and entail subjective decision making requiring human judgment. The introduction of fully autonomous weapons in the battlefield would be a major leap backward on account of their profound implications on norms and behaviour that the world has painstakingly arrived at after centuries of warfare. We firmly believe that developments in future military technologies should follow the established law and not vice versa.[172]

Many other participants were less forthright, taking cautious views while not rejecting the possibility of legally compliant AWS outright. Germany's statement reflects the general reluctance of many States to cede control over use of lethal force to a machine:

---

[171] Guido den Dekker, *The Law of Arms Control: International Supervision and Enforcement* (Martinus Nijhoff, 2001) 37.

[172] Pakistan (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016) 1-2.

> There was a widely shared view [in previous CCW discussions] that the decision over the use of lethal force shall never be delegated to [AWS] without the possibility of human intervention.[173]

However, there was general acknowledgement that development of AWS is at an early stage. The delegation from the Center for a New American Security (CNAS) noted that future technical developments could affect the legal arguments:

> Technology changes. In the field of machine intelligence, the past few years have seen rapid progress. Arguments in favor of banning weapons based on the state of technology today make little sense. What if technology improves? Could some of the same sensors and autonomy that can allow a self-driving car to avoid pedestrians be used to avoid civilian casualties in war? Perhaps. We should not preemptively close ourselves off to opportunities to decrease human suffering in war.[174]

Israel also expressed some cautious optimism about development of AWS, citing possible legal and ethical advantages:

> We should also be aware of the military and humanitarian advantages that may be associated with [AWS], both from operational as well as legal and ethical aspects. These may include better precision of targeting which would minimize collateral damage and reduce risk to combatants and non-combatants.[175]

These differences of opinion are likely to be attributable in part to the unsettled question of what constitutes an AWS, a problem that will diminish as that phase of the CCW discussions continues.

An associated question is about the importance of ensuring that States thoroughly review AWS for compatibility with IHL, and the difficulty of establishing procedures and standards for conducting those reviews. The software that will be used to control future AWS will be highly complex, as will the tasks they may be given and the environments in which they are likely to operate; development of testing procedures for such complex scenarios is an unsolved problem, and is an area of active research in

[173] Germany (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016) 2.
[174] CNAS (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016) 1.
[175] Israel (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016) 3.

civilian as well as military settings.[176] The question of how to conduct AWS weapon reviews was raised at the outset of the CCW discussions and has become consistently more prominent as discussions have progressed.

At the 2016 Informal Meeting of Experts several States outlined their current weapon review procedures,[177] although not with specific reference to AWS. The details of those review procedures vary from State to State but they showed a considerable degree of commonality among those States which chose to discuss their procedures, and general agreement that some work must be done to ensure that existing procedures are, or can be made, appropriate for reviewing future AWS:

> We see a number of challenges related to [AWS] distinct from traditional weapons reviews. … Beyond the purely technical challenge of assessing IHL compliance of an [AWS], there is also a conceptual challenge related to the fact that an autonomous system will assume an increasing number of determinations in the targeting cycle which traditionally are being taken care of by a human operator. … New evaluation and testing procedures may need to be conceptualized and developed to meet this particular challenge.[178]

In addition to incorporating various targeting decisions into an AWS control system, the prospect that States may eventually field AWS that 'learn' after deployment presents novel challenges for a weapon reviewer:

> Self-learning [AWS] that updated their functioning would be problematic from a legal review perspective because the weapon would change its characteristics over time, effectively requiring a continued legal review.[179]

Other States and organisations were more pessimistic about the efficacy of weapon reviews generally as a means of ensuring that AWS would comply with the law:

---

[176] For general information on the challenges of testing complex software systems, see: James A Whittaker, 'What is software testing? And why is it so hard?' (2000) 17(1) *IEEE Software* 70; Walt Truszkowski et al, 'NASA's Swarm Missions: The Challenge of Building Autonomous Software' (2004) 6(5) *IT Professional Magazine* 47; Boris Beizer, *Software Testing Techniques* (Dreamtech, 2003).

[177] Belgium, Germany, Israel, United Kingdom, Netherlands, Switzerland, Sweden (Presentations/Statements, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016).

[178] Switzerland, above n 177, 3.

[179] Gilles Giacca (Notes, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016) 4.

> However, better weapons reviews will not solve the problems associated with [AWS] for a number of reasons. … A successful international process to increase the effectiveness of weapons reviews will require a significant amount of time … weapons reviews were designed for a very different type of weapon than [AWS] … [AWS] blur the line between weapon and soldier to a level that may be beyond the ability of a weapons review process.[180]

States generally acknowledge that much work is yet to be done before conclusions can be drawn:

> Determining what kinds of actions should or should never be ceded to machines; establishing guidelines around operational safeguards; clarifying contextual implications that may impact upon the employment of particular systems; working to promote and implement existing mechanisms for ensuring compliance with international law; and better fleshing out conceptual notions of 'meaningful human control' or appropriate human judgement are but some examples of concrete, pragmatic and useful ways in which we as an international community can continue to grapple with the challenges and possibilities posed by [AWS]. Indeed, over time, these approaches - individually and collectively – may help us develop norms of responsible behaviour regarding the development, testing, deployment and use of weapons systems with significant levels of autonomy.[181]

2.4.3 Accountability

The third question is about accountability for proscribed acts committed via AWS. The fear of an 'accountability gap' is at the heart of many objections to further development of AWS. The essential concern is that, if a legally significant decision such as selecting a target to attack is made via a robot rather than directly by a human, how can someone be held accountable if the target selected is a person or object which should be protected from attack, such as a civilian or a surrendering combatant? If an AWS's target selection decision is made by the weapon system's software some time after the last human involvement in the attack, who is responsible if that decision results in a violation of the law? Various commentators have suggested that blame might fall on one or more of: the military commander who authorised the attack, the programmer who

[180] Mines Action Canada (Statement, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016) 1.
[181] Canada (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016) 2-3.

encoded the machine's decision process, the weapon manufacturer, or that State responsibility might be the only viable accountability mechanism.

The submission from Human Rights Watch at the 2016 CCW meeting outlined this concern:

> [AWS] would have the potential to create an accountability gap because under most circumstances a human could not be found legally responsible for unlawful harm caused by the weapons. Operators, programmers, and manufacturers would all escape liability.
>
> A human could presumably be held criminally responsible if he or she intentionally created or used a fully autonomous weapon in order to commit a war crime. But a more likely and troublesome scenario would be one in which the robot unexpectedly committed an unlawful act.
>
> The robot could not itself be held liable because machines cannot be punished. There would also be no human to hold directly responsible for the decision to attack because the robot, not a human, made the decision.[182]

Concerns about accountability are widespread among States and NGOs, but not universal. The UK offered a different view:

> If in the future [AWS] that could comply with an Article 36 Review were ever to exist, we do not believe that accountability would or should be any different from [the current regime]. The person who decides to deploy the weapon would ultimately be responsible for the consequences of its use. Accountability might even be improved if we assume that the automated record systems that an autonomous system would need in order to operate may provide better evidence to support subsequent investigation.[183]

Israel also expressed confidence that weapon autonomy would not inhibit the process of establishing accountability for violations of the law:

> In Israel's view, it is safe to assume that human judgment will be an integral part of any process to introduce [AWS], and will be applied throughout the various phases of the research, development, programming, testing, review, approval, and decision to employ them. [AWS] will operate as designed and programmed by humans. In cases where

[182] Human Rights Watch (Statement, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016) 1.
[183] United Kingdom, above n 177, 3-4.

> employment of [AWS] would involve a violation of the law, individual accountability would be sought in accordance with the law.[184]

Chapters 3 through 7 address these three questions. The findings of each of those chapters are utilised in Chapter 8 to make a set of recommendations for moving the CCW discussions on AWS forward.

[184] Israel, above n 177, 4.

Chapter 3 Understanding Weapon Autonomy

The first step toward formulating an appropriate regulatory regime for AWS is to develop a clear understanding of the nature of the underlying technological changes. From that understanding one can discern the behaviour of devices employing those technologies, the capabilities they will bestow on those who wield them and the uses to which they are likely to be put. Only then can their legal significance be understood with any clarity. The purpose of this chapter is to address these matters in relation to the path of AWS development that has occurred in recent years and is projected to continue for the foreseeable future. The technical discussion here provides the basis for the legal analysis to follow in the remaining chapters.

The task of understanding the technologies underlying AWS does not require a lawyer to work at the level of detail that an engineer or a scientist would engage in; in the context of a legal analysis it is unnecessary to focus in great depth on the technical means by which autonomous behaviour is achieved, only on the resulting nature of the behaviour. Indeed, it may be undesirable to venture too deeply into technical matters. Legal rules based on fine technical detail or a specific implementation of a technology are likely to lack durability and wide applicability. This is particularly so early in a technology's life cycle, when implementation details tend to change rapidly and may bear little resemblance to their eventual mature form. It is, however, important to attain a degree of understanding about the fundamentals of AWS sufficient to support a discussion of their general capabilities, limitations and uses, insofar as the operation of applicable bodies of law may be impacted.

This chapter contains two main sections. The first section addresses machine autonomy as a technical concept: what it means for a machine to have a capacity for autonomous operation, a brief overview of how autonomous behaviour is achieved, and discussion of some aspects of autonomous machines which are important for understanding the legal discussion to follow. This section is not about weapon systems in particular, but about electromechanical systems in general; the technological means employed do not differ in any significant way.

The second section discusses some aspects of how autonomous systems are and will continue to be used for military applications: autonomy will extend to systems ancillary

to combat operations, and autonomous systems will be employed in roles wherein they effectively 'collaborate' with human soldiers and with each other. The purpose is to show that assessing the legal consequences of using autonomous systems in military operations is not simply a matter of studying the properties of a new type of weapon; increasing use of machine autonomy will have wider effects on the conduct of military operations.

3.1 The Technology of Autonomous Machines

The concept of autonomy carries very different meanings in different fields of study, and even within the debate about use of AWS, authors have adopted a wide range of definitions. That is unfortunate, as disagreement about the nature of the defining characteristic of AWS is an impediment to rational discussion of their legal significance. This thesis takes the position that, as the challenge is to discern how a path of technical development will affect the operation of the law, one must begin with a clear understanding of what 'autonomous' means in the relevant technical fields, most notably robotics. Unfortunately, even when used in a strictly technical sense, autonomy is not susceptible to thorough and concise definition. It is therefore helpful to examine several definitions which highlight different aspects of autonomy which will be relevant to the legal discussion, and which will be investigated in more detail below.

The fundamental notion is of a machine that is able to manage its own operation in some significant way. For example:

> Autonomy (in machines). The capacity to operate in the real-world environment without any form of external control, once the machine is activated and at least in some areas of operation, for extended periods of time.[185]

Similarly:

> A system with a high level of autonomy is one that can be neglected for a long period of time without interaction.[186]

[185] Lin, Bekey and Abney, above n 81, 4.
[186] Michael A Goodrich and Alan C Schultz, 'Human-Robot Interaction: A Survey' (2007) 1 *Foundations and Trends in Human-Computer Interaction* 203, 217.

These definitions capture the self-management aspect of machine autonomy; they make it clear that a capacity for autonomous operation relates to how a machine's operator interacts with the machine. The first definition above also touches on another important point: that autonomy is not an all-or-nothing quality. In general, a machine would only perform specific functions autonomously, and not necessarily to the complete or permanent exclusion of human interaction. Thus, autonomy can also be described as

> the capacity of a system to select and decide within limits its own behavior. … This concept of autonomy is essentially based on the relative possibility of the system to control its relation with its environment.[187]

More briefly:

> Autonomy is a capability (or a set of capabilities) that enables a particular action of a system to be automatic or, within programmed boundaries, 'self-governing.'[188]

That last proviso, 'within programmed boundaries', is important. That a machine is able to operate without human interaction does not mean there are no human-imposed restrictions on the machine's behaviour. Indeed, machine autonomy essentially describes an approach to exercising human control over machines in circumstances where direct or real-time human interaction is infeasible or undesirable:

> *Autonomous* means having the power for self government. *Autonomous controllers* have the power and ability for self governance in the performance of control functions. They are composed of a collection of hardware and software, which can perform the necessary control functions, without external intervention, over extended time periods.[189]

The following sections discuss in more detail the aspects of machine autonomy mentioned in the above definitions.[190]

---

[187] Charles François (ed), *International Encyclopedia of Systems and Cybernetics* (K G Saur, 2nd ed, 2004) vol 1, 51.
[188] *Directive 3000.09*, 1.
[189] Panos J Antsaklis, Kevin M Passino and S J Wang, 'An Introduction to Autonomous Control Systems' (1991) 11(4) *IEEE Control Systems* 5 (emphasis in original).
[190] For another approach to defining AWS, based on the functions they are intended to perform, see Michael C Horowitz, 'Why Words Matter: The Real World Consequences of Defining Autonomous Weapons Systems' (2016) 30 *Temple International and Comparative Law Journal* 85, 94.

3.1.1 Autonomy and Control

A weapon system, whether in a simple form like a spear or a gun, or a complex form like a modern robotic system, is a tool. It is an implement used by a person or group to accomplish some goal. Its operation must, therefore, be directed by the operator toward that goal. In discussions about autonomous systems they are often described as 'self-governing' or operating 'without external control' as in the definitions given above, but those phrases can be misleading. A lack of interaction with an operator while a machine is operating does not mean that the behaviour of the machine has not been defined by a person. Rather, it means that the intended behaviour was defined in advance of the machine's activation and is then enforced by some part of the machine itself, typically by way of a computer-based control system attached to or in communication with the machine. That control system monitors the operation of the machine (such as a weapon) and issues commands to it as needed, in order to achieve the behaviour desired by the machine's human operator.

The discipline that deals with regulating the behaviour of a machine, such as a robot or a weapon, over time is *control theory*,[191] and its applications extend far beyond weaponry. Mechanisation of industrial processes has occurred in almost all industries in modern economies over the last century or more. The replacement of human labour with mechanical and electrical devices of ever increasing complexity, and the drive to automate the operation of those devices, has given rise to the need for a formal, structured means of controlling complex machines. The mathematical discipline of control theory and the related engineering discipline of control systems engineering are the bodies of knowledge that have grown to fill that need.[192] Their aims are to ensure that the machine being controlled behaves in a desired manner in the absence of manual intervention by a human operator, such that the advantages of automation can be realised. Those advantages, noted as far back as the early 20th century, are among those being cited today in the debate about AWS:

> In this age characterized by huge resources of mechanical and electrical power, these agencies have in many fields almost completely replaced human muscular power. In a

[191] See generally James R Leigh, *Control Theory: A Guided Tour* (Institution of Engineering and Technology, 2012).
[192] For a general introduction to autonomous control systems, see Antsaklis, Passino and Wang, above n 189.

similar way the functions of human operators are being taken over by mechanisms that automatically control the performance of machines and processes. Automatic control is often more reliable and accurate, as well as cheaper, than human control. Consequently the study of the performance of automatic control devices is of particular interest at the present time.[193]

Control system designers generally represent the systems they design as consisting of two main components: the 'plant' is the machine or process to be controlled and the 'controller' or 'control system' is the device which directly governs the behaviour of the plant.[194] The term 'plant' is carried over from chemical engineering; in a military context it would refer to the equipment which, if not capable of autonomous operation, would be directly operated by a human, such as a vehicle or a gun turret or similar. The controller of an autonomous system consists of the hardware and software that manages the vehicle, weapon or other device according to a program provided by a developer. Figures 1 and 2 give a conceptual outline of how these components work together in hypothetical manually operated and autonomous weapon systems. The solid arrows show the typical interactions between the various components.

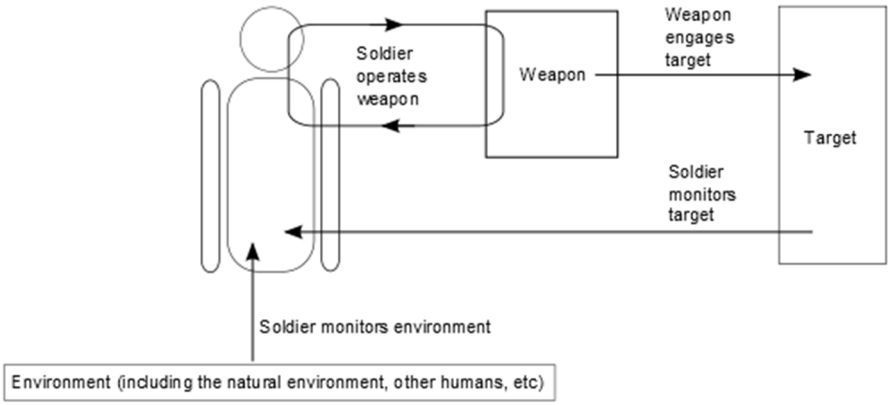

Figure 1: Manual weapon system

[193] H L Hazen, 'Theory of Servo-Mechanisms' (1934) 218(3) *Journal of the Franklin Institute* 279, 279.
[194] See generally Zdzislaw Bubnicki, *Modern Control Theory* (Springer, 2005) 3–4.

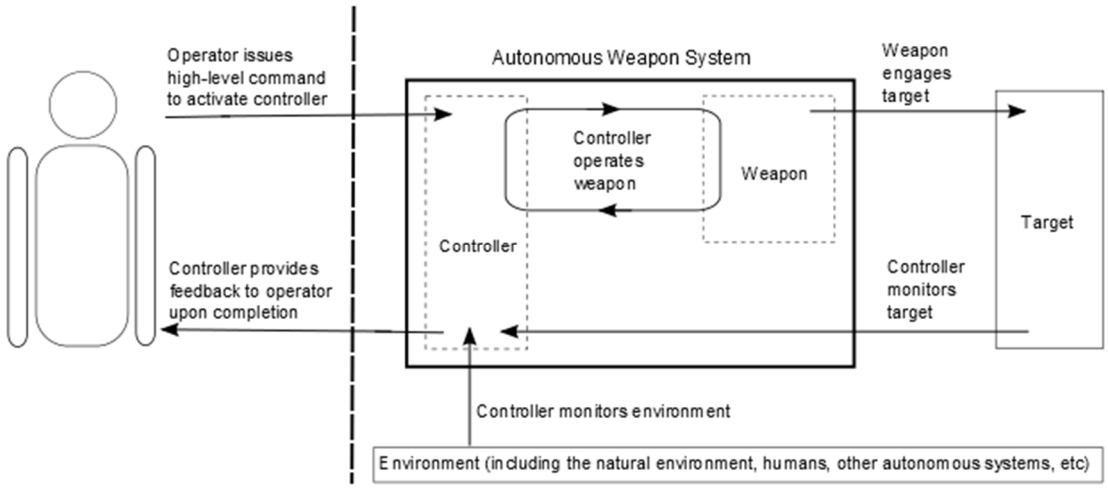

Figure 2: Autonomous weapon system

The main point to note in those diagrams is that when a manual system or process is replaced with a system capable of some degree of autonomous operation, the control system 'steps into the shoes' of the human operator of the manual system to some extent. The operator's (or a system developer's) understanding of how to control the machine is expressed in software (typically) and programmed into the control system. The physical means by which the operator manipulates the machine is converted to a set of actuators which the control system can activate (actuators are devices through which a control system controls a machine; one example would be an electric motor which pivots a gun turret based on a signal from the turret's software-based control system). Some form of sensor or feedback mechanism is provided by which the control system can monitor the machine. Other sensors may also be provided which allow the control system to monitor relevant environmental factors. The control system processes information from all these sensors in accordance with its programming and generates output signals which are sent to the actuators to regulate the machine. This process is encapsulated in the well-known 'sense-think-act' paradigm which is often used as the operational definition of a robot.[195]

The software-based control systems used to achieve autonomous operation in AWS and other machines are essentially special-purpose computers running programs which

[195] See, eg, George A Bekey, *Autonomous Robots: From Biological Inspiration to Implementation and Control* (MIT Press, 2005) 2.

control the machine in place of a human operator. These control systems, although they may be highly specialised in design and purpose, are nevertheless forms of ordinary stored-program computers, a class of devices which also includes common items such as personal computers. The defining characteristic of stored-program computers is that instructions entered by a human programmer are stored in the machine's memory and drawn upon to govern its operation.[196] Barring a major technological shift, tomorrow's autonomous systems will employ essentially the same technology; the systems will still be controlled by software written by human developers.

The fact that even very complex programs are just sets of pre-defined instructions is often obscured in discussions about sophisticated weapon systems and indeed it is not always apparent to an observer that a complex machine is merely executing instructions rather than operating independently. Even systems with only simple and limited capabilities are often driven by programs with instructions of the form

```
if <X happens> then <do action A> else <do action B>
```

Such instructions can make it appear that the system itself is 'choosing' between two alternative courses of action, when in fact the choice was made in advance by the person or organisation which wrote the program; the expression of that party's will was merely waiting within the system's memory for the previously determined trigger to be detected. For example, if a hypothetical UAV has cameras and an image recognition program which matches people within its field of view against a database of pictures of known insurgents, an instruction like

```
if <camera image matches image in database> then <aim and
fire> else <keep searching>
```

would make it appear that the UAV itself is selecting targets, when actually the targets and the conditions under which they would be attacked were selected in advance by the system developers.

[196] William Aspray, 'The Stored Program Concept' (1990) 27(9) *IEEE Spectrum* 51.

This reference to computing technology is included here because repeated references to weapon systems having the capacity for 'choice' or 'truly autonomous'[197] operation in the debate so far are potentially misleading. No computer is able to choose for itself whether or not to run a program stored in its memory, or to exercise discretion about whether or not to execute a particular instruction within a program; any such appearance of choice can only be the result of other instructions embedded in software. Fundamentally, the only function of a computer is to run whatever software is installed on it.

From this it can be seen that autonomy, in a technical sense, is simply the ability of a system to behave in a desired manner or achieve the goals previously imparted to it by its operator, without needing to receive the necessary instructions from outside itself on an ongoing basis.

Military development proposals often discuss autonomy in terms of the 'observe, orient, decide, and act (OODA) loop',[198] the model of a combatant's recurring decision-making cycle developed by USAF Colonel John Boyd. The OODA model describes the ongoing mental and physical processes involved in observing one's environment and responding to changes therein in pursuit of some goal (where a high level goal may involve several sub-tasks each with its own OODA loops nested within the higher level loop). With a manual system all steps in a loop are completed by a human: observing the environment to extract raw information; orienting oneself in relation to the environment by processing that information to form a useful model; making a decision based on that model and acting on the decision. In terms of the OODA loop the purpose of developing autonomous systems is to assign part or all of the loop to a machine in order to realize some operational advantage such as greater speed or endurance, lower cost or less risk to an operator's life. A highly autonomous system is one that can execute most or all of the OODA loops required to achieve some goal, using only the operator's high level instruction as guidance in the decision stage of a loop, such that the 'nearest' human functions similarly to a commander. A system with a lower level of autonomy is only able to execute the lower level loops, or only certain parts of a loop, and must work

[197] See, eg, Chantal Grut, 'The Challenge of Autonomous Lethal Robotics to International Humanitarian Law' (2013) 18(1) *Journal of Conflict & Security Law* 5.
[198] See, eg, *USAF UAS Flight Plan*, above n 82, 16.

together with a human operator to achieve a high level goal, casting that person as more of a collaborator. One example of such a system might be a radar warning receiver in an aircraft: a simpler system might only be able to detect a possible threat and issue a warning to a human operator, while a more highly autonomous system might be able to implement countermeasures such as dispensing chaff and flares without human intervention. The OODA loop may be seen as another expression of the 'sense-think-act' loop, and some literature on human-machine interaction also refers to an OODA-like loop to describe four types of functions that may be performed by an autonomous system: information acquisition, information analysis, decision selection and action implementation.[199]

One existing weapon system which exhibits a level of autonomous capability, and to which several references will be made later in this thesis, is the Israel Aerospace Industries 'Harpy' anti-radiation loitering weapon system:[200]

> Combining capabilities of a UAV and a missile, HARPY NG searches, identifies, acquires, attacks and destroys enemy radar targets. Not dependent on real-time intelligence, HARPY NG is highly effective against a wide spectrum of modern air defense systems.[201]

The Harpy's purpose is to seek out and destroy radar-emitting devices such as surface-to-air missiles in a defined search area. It is able to operate without human interaction after being launched, identifying, verifying and attacking targets autonomously:

> After the rocket-assisted launch, the drone flies autonomously enroute to its patrol area, predefined by a set of navigational waypoints. … Its radar seeker head constantly search [sic] for hostile radars, both along and across the flight track. Once suspicious radar is [sic] acquired, Harpy compares the signal to the library of hostile emitters, and prioritizes the threat. If the target is verified, the drone enters an attack mode … If the radar is turned off before Harpy strikes, the drone can abort the attack and continue loitering. If no radar was spotted during the mission, the drone is programmed to self

---

[199] See, eg, Raja Parasuraman, Thomas B Sheridan and Christopher D Wickens, 'A Model for Types and Levels of Human Interaction with Automation' (2000) 30(3) *IEEE Transactions on Systems, Man and Cybernetics* 288.
[200] *Harpy NG*, Israel Aerospace Industries <http://www.iai.co.il/2013/36694-16153-en/IAI.aspx>.
[201] *Harpy NG Anti-Radiation Loitering Weapon System*, Israel Aerospace Industries, MBT Missiles Division <http://www.iai.co.il/Sip_Storage//FILES/5/41655.pdf>.

> destruct over a designated area. Follow-on systems which are already proposed … are calling for a combination of seeker and killer drones that will enable visual identification and attack of targets even after they turn off their emitters.[202]

The Harpy is certainly lower on the autonomy spectrum than the types of future weapons planned by developers and debated in the CCW and other forums, but it autonomously performs at least two operations which are the subject of IHL obligations: it assesses the military status of possible targets,[203] and can abort an attack if conditions change.[204]

3.1.2 Adaptive and Intelligent Control

Understanding machine autonomy as a form of control rather than as the absence of control is a necessary step toward identifying the legal effects of AWS. It is also useful to go a step further and examine the workings of the types of machines which are of central concern in the debate about AWS: those which, beyond simply operating without human intervention, may appear to exhibit some behaviour which has not been explicitly programmed by a person. Objections to highly autonomous weapons based on fears that they may select the wrong target or otherwise act in undesirable ways generally refer either explicitly or implicitly to this type of machine.

AWS, like many other machines for which automation is seen as an advantage, must be able to operate in complex and chaotic environments. Indeed, the exigencies of combat operations arguably make this a more critical requirement for military machines than for those in civilian applications, as well as a greater challenge. A control system for an AWS must be able to maintain acceptable performance of the weapon being controlled when there is a very high degree of uncertainty in the behaviour of the weapon system and the environment in which it operates. Actions of adversaries, damage inflicted on the weapon system, unexpected environmental changes and other events may all interfere with the operation of a weapon system such that some corrective action is needed outside of whatever is normal for the task being undertaken. In the case of an autonomous system, that corrective action must be initiated by the control system rather

---

[202] *Harpy Air Defense Suppression System* (4 March 2006) Defense Update: International Online Defense Magazine <http://defense-update.com/directory/harpy.htm>.
[203] *API* art 57(2)(a)(i).
[204] *API* art 57(2)(b) (specifically, if a radar which the Harpy has targeted is deactivated before being struck).

than by a human operator. That is, when either the plant changes (perhaps due to damage) or the environment changes (perhaps due to unexpected enemy interference) such that the 'rules' the control system is using no longer work, the control system must adjust those rules to restore or maintain acceptable performance. For example, the control system of an autonomous car may need to handle the vehicle safely and effectively in the event of a punctured tyre or interference from another vehicle, either of which may happen unexpectedly. That might, depending on the circumstances, require the control system to alter the rules by which it operates the vehicle's engine or steering, for example, without an explicit instruction from a human operator. The discipline which addresses the ability of a control system to alter its behaviour in response to changing circumstances is *adaptive control*,[205] and it is one source of the behaviours which define the weapon systems that are of most interest. Essentially, adaptive control is a systematic approach for enabling a control system to automatically tune the way in which it manages the operation of the plant, in response to changes in either the environment or the plant itself, such that it continues to behave in the way desired by its human designers.

Under adaptive control and related approaches which enable a system to tune its own behaviour, the system may be operating as it was designed to even if the precise rules by which it is operating at a given time were not explicitly provided by a human operator (and may not even be precisely known to a human operator). Essentially, adaptive control relies on higher level logic in the control system to generate the required lower level operative rules as circumstances change. That higher level logic represents the operator's instructions to the weapon system, and by altering its behaviour according to those higher level rules, the system is behaving as intended.

Although adaptive control techniques enable a machine to alter its behaviour to an extent, their usefulness is limited by the complexity of the machine and its environment. They rely on the control system designer having a high degree of *a priori* knowledge about the machine, its task and the environmental changes and disruptions that the machine might encounter, such that they can be mathematically modelled and represented in the control system software. In highly complex or uncertain

[205] See, eg, I D Landau, R Lozano and M M'Saad, *Adaptive Control* (Springer, 1998) 1.

environments, or where the task to be completed is complicated, or even where the machine itself is very complicated, it is often not feasible to construct such a model in sufficient detail. In that case, another class of control techniques is likely to be employed, which is generally referred to as *intelligent control*.

Intelligent control is a diverse field which draws on a range of techniques that enable a machine to operate in environments that are too complex or unpredictable, or about which too little is known, to be susceptible to the mathematical modelling required by traditional control techniques. Generally, intelligent control software works by emulating various aspects of biological cognitive processes, based on the premise that biological entities are often able to operate effectively with incomplete knowledge, in complex and ambiguous environments.[206] The specific techniques employed are many, and the details are beyond the scope of this thesis; the most well-known techniques, which may be employed separately or in combination, are perhaps neural networks, fuzzy logic, expert systems and genetic algorithms. The relevant advantage which all such techniques afford to the control system designer is that they do not require detailed foreknowledge of all combinations of circumstances which the machine may encounter. They allow the control system designer to employ heuristics, approximation techniques and optimisation techniques to adapt the machine's behaviour to circumstances which could not be precisely foreseen.

A related issue is that of 'learning' control systems.[207] Learning, in this context, refers to the process of finding a generalised model which accounts for a set of observations, so that model can be applied when similar observations are made in the future.[208] Rather than just responding to unexpected changes in a machine or its environment, a learning controller is one that can improve its abilities over time by adjusting its 'rules' according to accumulated experiential knowledge; that is, allow information such as the performance of the machine at previous tasks to be retained and used to tune behaviour in future tasks. Online learning (being learning that happens after a machine is put into

---

[206] See, eg, Katalin M Hangos, Rozália Lakner and Miklós Gerzson, *Intelligent Control Systems: An Introduction with Examples* (Kluwer Academic Publishers, 2004) 1.
[207] D A Linkens and H O Nyongesa, 'Learning Systems in Intelligent Control: An Appraisal of Fuzzy, Neural and Genetic Algorithm Control Applications' (2002) 143(4) *IEE Proceedings - Control Theory and Applications* 367.
[208] William Bialek, Ilya Nemenman and Naftali Tishby, 'Predictability, Complexity, and Learning' (2001) 13 *Neural Computation* 2409.

operation, as opposed to offline learning which happens during a development phase) is a considerably more ambitious control technique that is useful when the complexity or uncertainty of a situation prevents *a priori* specification of an optimal control algorithm. It is another mechanism by which the rules a control system is using at a given time may not be rules explicitly provided by a human operator. As with adaptive control though, generation of those rules according to the higher-level learning process is the behaviour intended by the system's operator.

Intelligent control is one application of techniques from the broader field of artificial intelligence (AI).[209] AI aims to understand the factors that make intelligence possible, and to employ that knowledge in creation of machines that can operate in ways which, if observed in living beings, would be considered intelligent. That is, machines that can respond appropriately to changing circumstances and goals, take appropriate actions when provided with incomplete information and, if needed, 'learn' from experience.

Despite the complexity of existing and proposed AWS and their behaviour, they are not fundamentally distinct from simpler automated or manually operated weapon systems. Inherent in the notion of control is the aim of achieving some human-defined goal. Regardless of its complexity, control system software amounts to a set of instructions guiding the machine towards such a goal. Those instructions may give a machine a capacity for complex actions and responses, including the ability to formulate actions and responses according to new information encountered during operations which may not be precisely foreseeable to a human operator. Autonomous control systems employing the types of techniques described above can enable a weapon system to respond in more complex ways and operate for extended periods without having to interact with a human, but that does not constitute independence from human control in any sense. Rather, it is best seen as control applied in a different way, in advance rather than in real time.

3.1.3 Autonomy is a Matter of Degree

Legal authors commonly refer to AWS as a discrete category of devices which are easily distinguishable from non-autonomous systems and refer to the 'level' of

[209] See generally Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Prentice Hall, 3rd ed, 2009).

autonomy that a particular system exhibits as though such levels are intrinsic properties of particular systems.[210] Some common taxonomies, referred to by participants in the CCW discussions, were outlined in Section 2.4.1 above.

One particularly popular classification scheme, often cited in the CCW discussions and elsewhere, distinguishes between 'automatic', 'automated', and 'autonomous' weapons. These terms have been used somewhat differently by different authors, but the essential distinctions are as follows:[211]

- 'Automatic' refers to very simple devices which perform well-defined tasks and may have pre-set responses to external disturbances; a non-weapon example would be a thermostat set to maintain a defined temperature by turning a heater on or off in response to changes in temperature.
- 'Automated' systems overlap with automatic systems but are likely to follow more complex sets of rules in normal operation and in responding to disturbances, such that they can perform more complex tasks or operate in more complex environments. Examples include automated telephone support lines that can respond in limited ways to various queries, or some existing weapon systems.
- The varying uses of the term 'autonomous' among authors reflects the uncertainty that surrounds the nature of these new technologies. The general view is that autonomous systems go beyond automated systems in some way, but the precise criteria vary. Some authors describe systems as 'autonomous' when they exhibit some ability to adapt their own behaviour in response to changing circumstances.[212] Others use the term to indicate that some threshold level of complexity in the system, its task or its operating environment has been reached.[213] Still others say autonomous systems are those with some degree of 'independence' from their human operators.[214] A further subdivision within this category is between 'semi-autonomous' and 'fully autonomous' systems. The

[210] See, eg, Wagner, above n 78, 1379.
[211] Scharre, above n 135, 12.
[212] Kenneth Anderson and Matthew C Waxman, "Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can" (Hoover Institution, 9 April 2013) 6.
[213] Rebecca Crootof, 'The Killer Robots Are Here: Legal and Policy Implications' (2015) 36 *Cardozo Law Review* 1837, 1854.
[214] Grut, above n 197.

> claimed difference is that fully autonomous systems are those which are designed to operate entirely without human involvement once activated while semi-autonomous systems would require some form of active human involvement in relation to some or all functions.[215] A range of views have been expressed over whether fully autonomous systems would entirely disallow human involvement, or simply not require it, and about the extent of human involvement required in semi-autonomous systems.

It is possible to define finer distinctions within each of those levels depending on the degree of necessity of the human operator's contribution, the realistic possibility for successful human intervention in battlefield conditions, and so forth.[216]

Finally, the terms 'intelligent' and 'learning' are sometimes used in relation to autonomous systems, either as descriptions of some degree of requisite human-like behaviour or in reference to capabilities of the system control software. Heather Roff expresses concern about AWS with some learning capability:

> The types of weapons under discussion here are fully autonomous *learning* machines. While contemporary militaries currently possess lethal automated weapons, that is weapons that are capable of targeting and firing on their own without human supervision or interaction, they respond to a preprogrammed set of constraints. … [AWS] that learn, and are able to target individuals and other military objectives that may not emit signals or travel at particular speeds or trajectories, would require a meaningful level of artificial intelligence. I reserve the word 'autonomous' for these weapons. [AWS], if they are to be capable of achieving the stated goals of being a force multiplier, cost-effective in both monetary and human terms, and robust enough to withstand cyberattacks, must be autonomous in this sense.[217]

Attempts to define taxonomies of autonomy are further complicated by differing perceptions about whether categories should be based on the degree and type of human interaction with the system, or the complexity of the system and its behaviour. The two

---

[215] See, eg, *Directive 3000.09*, 13-14.

[216] Paul Scharre, *Between a Roomba and a Terminator: What is Autonomy?* (18 February 2015) War on the Rocks <https://warontherocks.com/2015/02/between-a-roomba-and-a-terminator-what-is-autonomy/?singlepage=1>.

[217] Heather Roff, 'The Strategic Robot Problem: Lethal Autonomous Weapons in War' (2014) 13(3) *Journal of Military Ethics* 211, 212-13 [citation omitted].

variables are both plausible bases for categorisation, but they are independent; a system can be highly autonomous in that it requires no human interaction, but be very simple in itself or the task it performs.[218] The well-known Roomba robotic vacuum, for example, can perform its duties with virtually no human interaction once it has been set up, but is hardly the image of the type of complex, powerful machine that is often seen as 'autonomous'.

Regardless of how those challenges are met, simplistic discrete taxonomies such as these do not reflect the range of capabilities of systems that exist today and do not correspond with the development roadmaps[219] that have been made public by various armed forces. In fact, the levels of autonomy exhibited by proposed military systems may be expected to vary in complex ways. Possible degrees of autonomy vary widely, as do the ways in which tasks are allocated between an operator and an autonomous system, and the behaviour of a system may be expected to change according to both the specific task being performed and the circumstances in which the system is operating. Establishing the relative degrees of control exercised by a human operator and a computer control system in respect of a particular action for legal or other purposes may be a complex process. The aim of the brief descriptions in this section is not to provide the reader with all the technical knowledge necessary for such an undertaking, but to demonstrate the fluid nature of control over tasks in an environment where humans interact with advanced autonomous systems.

3.1.3.1 Degree of Computer Control

Machine autonomy is not an all-or-nothing capability; there is a continuum ranging from complete human control over some function or another to complete machine control. Many ways of describing this continuum have been proposed.[220] One of the best-known categorisations, which recognizes ten levels of automation, is reproduced in

[218] Scharre, above n 135, 12.
[219] See, eg, Tank-Automotive Research, Development, and Engineering Center, 'Robotics Strategy White Paper' (Army Capabilities Integration Center, US Army, 19 March 2009) (http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA496734>; *USAF UAS Flight Plan, above n 82*.
[220] Peter A Hancock and Stephen F Scallen, 'Allocating Functions in Human–Machine Systems' in Robert R Hoffman, Michael F Sherrick and Joel S Warm (eds), *Viewing Psychology as a Whole: The Integrative Science of William N Dember* (American Psychological Association, 1998) 521.

Table 1.[221] Once again though, these 'levels' are merely ways of describing points on a continuum; there are no discrete levels of machine autonomy in reality.[222]

Table 1: Sheridan & Verplank's ten levels of automation

100% human control	1. Human does the whole job up to the point of turning it over to the computer to implement;
	2. Computer helps by determining the options;
	3. Computer helps determine options and suggests one, which human need not follow;
	4. Computer selects action and human may or may not do it;
	5. Computer selects action and implements it if human approves;
	6. Computer selects action, informs human in plenty of time to stop it;
	7. Computer does whole job and necessarily tells human what it did;
	8. Computer does whole job and tells human what it did only if human explicitly asks;
	9. Computer does whole job and tells human what it did and it, the computer, decides he should be told;
100% computer control	10. Computer does whole job if it decides it should be done, and if so tells human, if it decides he should be told.

The US National Institute of Standards and Technology developed a somewhat more complex taxonomy, the Autonomy Levels for Unmanned Systems (ALFUS) framework.[223] The ALFUS framework characterizes autonomy levels in terms of human

[221] Thomas B Sheridan and William L Verplank, 'Human and Computer Control of Undersea Teleoperators' (Technical Report, Massachusetts Institute of Technology, 14 July 1978) 8–17.
[222] For a note on how to interpret Sheridan & Verplank's taxonomy see: Defense Science Board, 'The Role of Autonomy in DoD Systems' (Task Force Report, US Department of Defense, July 2012) 23-4.
[223] NIST Engineering Laboratory, 'Autonomy Levels for Unmanned Systems' (National Institute of Standards and Technology, 6 June 2010) <http://www.nist.gov/el/isd/ks/autonomy_levels.cfm>.

independence, mission complexity and environmental complexity, with each of those factors broken down further into a series of lower level considerations.

Other organisations have proposed alternative taxonomies of autonomy levels intended for various purposes, such as NASA's eight-level model for its Spacecraft Mission Assessment and Re-planning Tool (SMART),[224] and the US Army Science Board's eleven-level model from its study on human-robot interface issues.[225]

3.1.3.2 Methods of Human-Machine Interaction and Task Allocation

Even highly autonomous systems that may need no significant human input during completion of their allocated tasks will be operating alongside other entities, both human and electronic, and will need to exchange information with those entities to receive instructions, coordinate efforts, report results and so on. The study of human-computer interaction, and more recently human-robot interaction, are multi-disciplinary fields which have received considerable and increasing attention over the last two decades as computing and robotic capabilities have expanded.[226] Several alternative paradigms have emerged to describe how this interaction may be structured, and how tasks may be allocated to either entity, with implications for the roles of humans. Very broadly, humans may occupy either supervisory or collaborative roles when working with autonomous systems, and may move between those roles during execution of a series of tasks. For example, one taxonomy lists five roles that humans may play when working with robots: supervisor, operator, mechanic, peer and bystander.[227] These roles may not be static; one design paradigm called mixed-initiative interaction 'refers to a flexible interaction strategy where each agent can contribute to the task what it does best.'[228] In the mixed-initiative paradigm, the respective roles of human and computer (or robot) are often not determined in advance, but are 'negotiated' in response to evolving circumstances. At different times either the human operator or the machine

---

[224] Ryan W Proud, Jeremy J Hart and Richard B Mrozinski, 'Methods for Determining the Level of Autonomy to Design into a Human Spaceflight Vehicle: A Function Specific Approach' (Report, NASA Johnson Space Center, September 2003) 4 <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA515467>.
[225] US Army Science Board, 'Ad-hoc Study on Human Robot Interface Issues' (September 2002) 16 <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA411834>.
[226] For an overview see Goodrich and Schultz, above n 186.
[227] Jean Scholtz, 'Theory and Evaluation of Human Robot Interactions' (Paper presented at 36th Annual Hawaii International Conference on System Sciences, Hawaii, 6–9 January 2003).
[228] Marti A Hearst, 'Trends & Controversies: Mixed-Initiative Interaction' (1999) 14(5) *IEEE Intelligent Systems* 14, 14.

might have direct control over a task with the other assisting, or they might be working independently.

Human-robot interaction is an active field of research with new principles emerging continually. While it is safe to assume that humans collectively will continue to exercise a high level of control over robotic weapons employed in armed conflict, some caution is required when assessing the roles of, and the degree of control exercised by, particular individuals.

3.1.3.3 Variations by Function

Autonomy is commonly seen as a property of a system, but in practice technologies giving rise to a capability for autonomous operation will be, and are being, applied to individual subsystems which form parts of military hardware and software; in respect of robotic systems, development is occurring separately in the areas of autonomous navigation, autonomous targeting, and other functions required of an advanced weapon system. As systems with greater capabilities for autonomous operation are developed over the coming years, there is no reason to suppose that all their functions will be subject to the same degree of direct human supervision;[229] it is probable that some functions, presumably those for which the cost/benefit analysis of automation is more favourable in some way, will be made 'more autonomous' than others. A system may therefore be operating at more than one 'level' of autonomy simultaneously, with respect to different tasks. It is necessary to assess the behaviour of the system, and the levels of human and machine involvement, in relation to particular functions of interest such as course planning, navigation or weapon release.

A military targeting process, such as the 6-step process employed by the Australian Defence Force,[230] can be used to illustrate. A weapon system may have the ability to autonomously locate and observe potential targets, and perhaps take some precautions to minimise collateral damage without needing human input. It requires considerably more trust to allow that weapon system to release a weapon autonomously. Some input from a human operator might be needed to assess whether a potential target is a valid military objective and whether any expected collateral damage would be

[229] This phenomenon is emerging in some current weapon systems: Rain Liivoja, Kobi Leins and Tim McCormack, 'Emerging Technologies of Warfare' in Liivoja and McCormack, above n 24, ch 35 s 2.1.
[230] Henderson, above n 9, 237.

disproportionate to the anticipated military advantage of an attack. If such a system were to be involved in an incident which led to an inquiry about whether a civilian might have been targeted, it would be important to establish at which stage an error may have occurred, and what levels of human and machine control were exercised at that stage. Similar considerations would also apply to a 'system of systems' situation involving communication between different systems with varying levels of autonomy.

3.1.3.4 Variations by Circumstance

As noted above, even where a designer has determined that a particular system function will be subject to a level of computer control, that level may vary according to circumstances, where the circumstances in question may be defined in terms of, for example:

- the phase of a mission, such as planning, initiation, implementation or termination;
- a disruptive event, such as where a UAV that is ordinarily remotely operated may be programmed to autonomously return to base, fly in circles or even defend itself if it loses contact with the operator;
- an unexpected opportunity arising during a mission.

The US Department of Defense summarizes this variability: 'The key point is that humans and computer agents will interchange initiative and roles across mission phases and echelons to adapt to new events, disruptions and opportunities as situations evolve.'[231]

3.1.3.5 Summary

Machine autonomy is a much more nuanced phenomenon than popular images of super-advanced robots would lead one to believe. As a rapidly advancing set of technologies, some still in their infancy, many issues which will one day prove critical are no doubt yet to emerge, but two in particular stand out for legal investigations today. First, a machine's autonomous capabilities directly affect primarily its supervisors and operators, not necessarily (in the case of a weapon) those against whom it is directed. In particular, the role of the operator is altered in important ways but not eliminated; there

[231] Defense Science Board, above n 222, 27.

is no such thing as 'complete' autonomy in the sense of a machine operating entirely independently of any human. Second, the precise form of the relationship between operator and machine is likely to change across different systems, at different times, in relation to different tasks. It is difficult to draw reliable inferences purely on the basis of a weapon being described as 'autonomous'. One may gain a sense of the possible complexity of the relationships between an autonomous system and its operators and supervisors by examining the US DoD Defense Science Board's recently proposed 'Autonomous Systems Reference Framework,'[232] a system for allocating functions and responsibilities to either the computer or one of several human operators or supervisors during the system design phase. This framework considers how autonomy supports each of the various human users of a particular system in a military hierarchy, how communication between those users may be facilitated, how allocation of tasks to human or machine may vary over the course of a mission and several other factors. The Defense Science Board also offers this view of the variability of degrees of autonomy from a cognitive science perspective:

> Cognitively, system autonomy is a continuum from complete human control of all decisions to situations where many functions are delegated to the computer with only high-level supervision and/or oversight from its operator. Multiple concurrent functions may be needed to evince a desired capability, and subsets of functions may require a human in the loop, while other functions can be delegated at the same time. Thus, at any stage of a mission, it is possible for a system to be in more than one discrete level simultaneously.[233]

Situating autonomy on a continuum leads naturally to one of the questions underlying this thesis: is there a threshold level at which a weapon system's autonomous capabilities become legally significant? *Prima facie*, it seems likely that such a threshold would exist. Weapon system autonomy is not a concept that was contemplated in any significant way by the drafters of any of today's written IHL, and it certainly did not play a role in establishing any State practice that would today be considered customary. The idea of a weapon stepping into the shoes of its operator and acting in some deliberate way without direct human control seems, at first glance, to undermine

[232] Ibid 24.
[233] Ibid 4.

the law's attempt to constrain the effects of armed conflict by regulating the conduct of States and their (human) agents. It seems plausible that there may be some limit of autonomous capability beyond which the law would cease to operate effectively, and it is not obvious where such a limit might be. If some weapon systems that are in use today have some small capacity for autonomous operation, should their legal status be evaluated in light of that capacity?

In practice, there is little concern among States, NGOs or other parties about the capacity for autonomous operation in today's weapon systems, and no serious contention that their use transgresses the law because of that capacity. Concerns about AWS generally take the form of trepidation about future developments. Indeed, many parties who view autonomy in terms of discrete levels do not acknowledge any existing weapon systems as being autonomous at all, instead seeing machine autonomy as a futuristic quality requiring intelligence and a high degree of independence of action. The focus on emerging technologies is reflected in the mandate of the CCW Informal Meetings of Experts and the Group of Governmental Experts.[234]

However, a decision by CCW States Parties to examine only emerging technologies in the area of AWS does not amount to a finding that development has reached a threshold, such that the technologies emerging in the field today are necessarily those which will place AWS beyond the scope of what is legally permissible. Indeed, some participants in the CCW discussions have expressed the view that IHL is and will continue to be sufficient to regulate increasing levels of autonomy in weapons. To establish whether, and where, a threshold of legal acceptability may exist, it is necessary to look at the spectrum of autonomous capabilities with reference only to the rules of IHL to find where, if at all, a capacity for autonomous operation raises legal concerns. One of the primary challenges facing CCW discussants at the present time is to establish a definition of AWS which can then be used to support further investigations of the legal, ethical and operational implications of further development. The results of the investigation in this thesis will help to establish a definition that is useful for legal purposes.

---

[234] *Final Report of the Meeting of the High Contracting Parties 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*, 2013 sess, Agenda Item 13, UN Doc CCW/MSP/2013/10 (16 December 2013) 4 [32].

3.1.4 Summary

A capacity for autonomous operation is imparted to a machine via its control system which is essentially an encoded representation of a set of instructions that would otherwise have been issued by a human operator. That is, the control system embodies parts of the behaviour of the machine being controlled within the machine itself.

Of course, all machines are inherently bound to follow some pattern of behaviour; it is the reason that some weapons are customarily prohibited for being 'by nature indiscriminate' [235] or 'of a nature to cause superfluous injury or unnecessary suffering'.[236] In a sense, a capacity for autonomous operation that may raise novel legal questions about a weapon system merely refers to a more complex pattern of built-in behaviour, one that may include proactive, reasoned behaviours rather than only passive responses to external stimuli. More relevantly though, autonomy is a complex form of built-in behaviour that emulates specified decisions that a human operator would otherwise have been required to make as the need arises during the weapon system's operation.

Relieving a human operator of the need to intervene in, or even supervise, a weapon system's operation in that way is variously described as giving the weapon system the ability to 'make decisions', making it 'independent', placing it 'beyond human control', and so forth. None of those images is completely accurate and some are misleading in a legal discussion. In many cases, the complexity of operations on a modern battlefield will often be beyond current feasible methods of constructing autonomous systems, so that in practice it will be necessary for humans and AWS to work as a team, with AWS being responsible only for the parts of an operation that can be brought within their limited capabilities. Even where, hypothetically, an AWS is able to perform the entirety of some operation unassisted, the behaviour of that AWS is still the product of human decisions.

3.2 Autonomy as a Property of a Platform, Not a Weapon

Development of the technologies of weapon autonomy has often been regarded as simply development of a new type of weapon, a class of devices with identifiable

[235] Henckaerts and Doswald-Beck, above n 19, 244.
[236] Ibid 237.

behaviours and effects, not unlike development of gunpowder-based weapons, blinding lasers, or a host of others. Accordingly, debate about whether and how to regulate AWS has often been framed as a question about the acceptable uses of a specific type of device.

The discussion in section 3.1 above shows that the question is properly understood as being also about a new means of controlling weapons, at least as much as it is about a new class of weapons.[237] The hardware and software comprising an AWS control system may be regarded as part of a weapon platform rather than as a component of the weapon proper. Recommendations for regulatory responses should reflect that understanding. This section expands on the idea that weapon autonomy is about controlling weapons; its manifestations and effects are likely to be more pervasive than they would be if it was just about constructing a new type of weapon.

3.2.1 Autonomy Alters the Relationship between Machine and Operator

In the context of robotics in general, and the proposals for AWS in particular, autonomy refers to a capability, not to a specific technology nor to a particular device nor to a certain behaviour. It is, fundamentally, the ability of a system to perform its function, whatever that may be, with less interaction with a human operator than a manual system would require. Autonomy is thus concerned with the relationship between the system and its human operator, not the nature of the system's task nor the manner in which it performs that task. Two points in particular are relevant for the identification of legal issues.

First, autonomous systems will perform some or all mission tasks in place of humans, but not necessarily differently to humans. There is nothing in the concept of machine autonomy itself that supports an inference that an autonomous system must necessarily perform a task in a different manner than would a human or team of humans performing the same task manually. Of course, one of the motivations for developing increasingly autonomous weapon systems is to realise performance benefits: the persistence of systems that do not require constant human interaction; the ability to quickly integrate data from many sources; and the ability to take greater risk than could a manual system;

[237] Hin-Yan Liu, 'Categorization and Legality of Autonomous and Remote Weapons Systems' (2012) 94 *International Review of the Red Cross* 627.

among other things, will certainly aid military operations. However, such differences, while very important operationally, are somewhat peripheral to the legal aspects of autonomy. UAVs, for example, already allow for a high level of persistence without necessarily exhibiting any capabilities associated with a high level of autonomy and without raising the same legal issues. In assessing the legal implications of a particular development path, or a particular set of technologies, the focus must be kept on the capability of interest rather than on other capabilities that may be present in the same weapon system. In the case of machine autonomy, it is not useful to attempt to attribute specific behaviours to a system merely on the basis of it being described as having some capacity for autonomous operation; all that one can reliably say on that basis is that the human operator's direct involvement in part or all of the system's OODA loop(s) will be reduced or removed. The mere fact of reassigning a task from a human to a computer does not necessarily alter the performance of that task.

Second, it is incorrect to describe autonomous systems as being 'independent'[238] machines that operate 'without human control';[239] the relationship between human and machine is not severed, it is only modified. Choices made by hardware and software developers in the design stage will shape the behaviour of the systems from then on. A weapon system is only one component in a State's armed forces and its use must be in accordance with higher level plans and established procedures as well as with the capabilities and practices of other units and support structures. Mission planners and others will impose constraints on each mission. For example, Boothby describes the level of human involvement that would be required in conducting an attack with an autonomous UAV:[240]

> A flight plan will have been prepared and filed by a person who will decide on the geographical area that is to be searched, the time period within which the search may take place, the areas where the aircraft may loiter and for how long, and that person will programme these important requirements into the flight control software. The platform will be fuelled by a person thus defining the maximum endurance of the mission.

[238] Markus Wagner, 'Taking Humans Out of the Loop: Implications for International Humanitarian Law' (2011) 21(2) *Journal of Law, Information and Science* 155, 159.
[239] Gary E Marchant et al, 'International Governance of Autonomous Military Robots' (2011) XII *Columbia Science and Technology Law Review* 272, 273.
[240] William Boothby, 'How Far Will the Law Allow Unmanned Targeting to Go?' in Dan Saxon (ed), *International Humanitarian Law and the Changing Technology of War* (Martinus Nijhoff, 2013) 56.

> Operational planners will decide what weapons will be carried and how they are to be fused, and stores will be loaded by people before takeoff. The sensors on which the autonomous aspect of the mission depends will have been designed and built by people and will be controlled by similarly designed software. Those designers and/or the mission planners will have prescribed the level of mechanical recognition that is to be achieved before an object is recognised as a target and, thus, before an attack will be undertaken.

It may be assumed that analogous tasks would be performed by humans in relation to other operations involving AWS. In these ways a human hand always provides some degree of guidance despite a possible lack of direct supervision. The Defense Science Board of the United States Department of Defense expresses the dependence of autonomous systems on humans more generally:

> It should be made clear that all autonomous systems are supervised by human operators at some level, and autonomous systems' software embodies the designed limits on the actions and decisions delegated to the computer. Instead of viewing autonomy as an intrinsic property of an unmanned vehicle in isolation, the design and operation of autonomous systems needs to be considered in terms of *human-system collaboration*.[241]

3.2.2 Autonomy will Extend beyond Weapon Systems

Even within an analysis which is focussed specifically on IHL it is necessary to account for the effects of autonomous capabilities in systems which may not themselves perform any hostile act but may still have some impact on a decision or action which bears legal consequences.

The prime example of this type of system would be an autonomous ISR system which locates, identifies and tracks potential targets. For example, the US Department of Defense recently announced its Autonomy Research Pilot Initiative (ARPI) which 'seeks to promote the development of innovative, cross-cutting science and technology for autonomous systems able to meet future DOD system and mission requirements.'[242] The ARPI invitation for proposals identifies ISR as one of its technical challenge areas: '[b]y increasing the level of machine perception, reasoning and intelligence on ISR

[241] Defense Science Board, above n 222, 1-2.

[242] Department of Defense, 'Autonomy Research Pilot Initiative (ARPI)' (Invitation for Proposals, November 2012) 1 <http://auvac.org/uploads/publication_pdf/Autonomy%20Research%20Pilot%20Initiative.pdf>.

platforms themselves, a more efficient workload balance can be achieved. This includes the management and closed loop control of ISR assets to adapt to their environments and mission circumstances to collect appropriate and relevant data.'[243] Preliminary moves in this direction are already occurring; DARPA's Military Imaging and Surveillance Technology (MIST) program aims to 'develop a fundamentally new optical ISR capability that can provide high-resolution 3-D images to locate and identify a target at much longer ranges than is possible with existing optical systems.' Systems developed under this program would be able to perform automated identification and recognition of potential targets.[244]

The motivation for pursuing this line of research as part of ARPI is that

> Today's battlespace is creating an unprecedented increase in intelligence, surveillance, and reconnaissance (ISR) data. The [processing, exploitation and dissemination] analyst can become overwhelmed in trying to integrate and analyze these various data inputs (imagery, video, communication and human ISR data) while also trying to track targets, infer sources and provide analysis feedback (in real-time or post-analysis). [245]

It may be inferred that the increased 'intelligence' of ISR platforms will be used to perform some processing of raw data before communicating the result to a human, thus placing the ISR system in a position of influence over a human operator's perception of the battlespace. Such processing of sensor data has obvious implications for the impact which an advanced ISR system may have on a decision to fire a weapon or perform some other action that carries significant legal consequences.

Somewhat further removed causally from combat operations, but still with some peripheral influence on outcomes, are autonomous vehicles which will carry cargo and even people. There are many examples of efforts being made to automate resupply operations, such as the United States Office of Naval Research's Autonomous Aerial

---

[243] Ibid 4.

[244] Strategic Technology Office, Defense Advanced Research Projects Agency, 'Military Imaging and Surveillance Technology — Long Range (MIST-LR) Phase 2' (Broad Agency Announcement No DARPA-BAA-13-27, 12 March 2013) 6 <https://www.fbo.gov/index?s=opportunity&mode=form&id=78b0ddbf382678fa9ace985380108f89&tab=core&_cview=0>.

[245] Department of Defense, above n 242, 4.

Cargo/Utility System (AACUS) Program.[246] Following on from the success of the K-MAX unmanned cargo resupply helicopter[247] in Afghanistan, the aim of the AACUS program is 'the development of advanced autonomous capabilities to enable rapid cargo delivery by unmanned and potentially optionally-manned Vertical Take Off and Landing (VTOL) systems.'[248] The program will produce a system which can be installed in suitable aircraft to respond to calls from deployed units, autonomously (but under some supervision) planning its route, avoiding obstacles and bad weather and choosing a suitable landing site to deliver supplies and, eventually, evacuate casualties.[249] While vehicles such as those controlled by AACUS-like systems will not carry weapons, they will still be large, fast-moving objects that must operate in proximity to people and other vehicles and may carry hazardous materials or present other dangers.[250] They therefore introduce some legal risk, for example, if used to carry wounded soldiers.

The presence of autonomous capabilities in systems ancillary to combat operations will therefore necessitate a broader view of autonomy than simply treating it as a property of a new type of weapon. In addition, as discussed in the next section, the ability of the AACUS and similar systems to negotiate directly with other unmanned systems[251] may complicate the task of even identifying which systems are relevant to a particular investigation.

3.2.3 Autonomous Systems will Collaborate with Each Other

As discussed above, autonomous systems alter the relationship between weapon and operator. Of similar importance is the relationship between nominally separate autonomous systems. Most development roadmaps published by military and

---

[246] Office of Naval Research, A*utonomous Aerial Cargo/Utility System Program* (27 September 2013) <http://www.onr.navy.mil/en/Science-Technology/Departments/Code-35/All-Programs/aerospace-research-351/Autonomous-Aerial-Cargo-Utility-AACUS.aspx>.
[247] Lockheed Martin Corporation, *K-MAX* <http://www.lockheedmartin.com/us/products/kmax.html>.
[248] Mary Cummings and Angelo Collins, 'Autonomous Aerial Cargo/Utility System (AACUS)' (Concept of Operations, Office of Naval Research) 2 <http://www.onr.navy.mil/~/media/Files/Funding-Announcements/BAA/2012/12-004-CONOPS.ashx>.
[249] Ibid.
[250] Anthony Finn and Steve Scheding, *Developments and Challenges for Autonomous Unmanned Vehicles* (Springer, 2010) 156.
[251] Cummings and Collins, above n 248, 2.

government organisations make it clear that autonomous systems must collaborate with each other,[252] and this interoperability may take a range of forms.

Many current research projects, both military and civilian, focus on the behaviour of groups of decentralised cooperating robots, colloquially known as 'swarms', which work together as a single system in pursuit of some goal. One example of this is the US Army's Micro Autonomous Systems and Technology (MAST) program[253] which aims to create 'systems of diverse autonomous mobility platforms equipped with miniature sensors to quickly, quietly, and reliably explore and find targets';[254] that is, teams of small air and ground vehicles which will be used by soldiers to explore and map complex environments such as urban settings, caves and so on. The possibility of equipping swarms with weapons has also been raised.[255]

A somewhat different form of collaboration between autonomous systems is employed in development programs such as DARPA's Hydra program.[256] The Hydra program 'aims to develop a distributed undersea network of unmanned payloads and platforms',[257] essentially a system of unmanned submersible platforms which would be used to launch a variety of unmanned aerial and undersea vehicles (UAVs and UUVs) close to enemy operations. Although still in its very early stages, this program envisions a system in which the submersible 'mothership' interacts directly with its UAV and UUV payloads: for example, UUVs would dock with and recharge from the mothership, collect intelligence information for use on their mission and afterwards transfer information acquired on the mission back to the mothership to be sent on to authorities.[258]

---

[252] See, eg, US Department of Defense, 'Unmanned Systems Integrated Roadmap FY2011-2036' (No 11-S-3613) 49–50.

[253] United States Army Research Laboratory, *Micro Autonomous Systems and Technology (MAST)* (25 February 2011) <http://www.arl.army.mil/www/default.cfm?page=332>.

[254] Micro Autonomous Systems and Technology (MAST), *Research Thrusts* <https://alliance.seas.upenn.edu/~mastwiki/wiki/index.php?n=Main.Objectives>.

[255] Singer, above n 82, 229.

[256] Federal Business Opportunities, *Hydra* (22 August 2013) <https://www.fbo.gov/index?s=opportunity&mode=form&id=4cc32f06144bd6f3eba18655135d6155&tab=core&_cview=1>.

[257] DARPA Tactical Technology Office, Defense Advanced Research Projects Agency, *Hydra* <http://www.darpa.mil/Our_Work/TTO/Programs/Hydra.aspx>.

[258] John Keller, *DARPA Considers Unmanned Submersible Mothership Designed to Deploy UAVs and UUVs* (23 July 2013) Military & Aerospace Electronics <http://www.militaryaerospace.com/articles/2013/07/darpa-uuv-mothership.html>.

These and similar projects and the tactics and strategies that will be associated with them are part of a broad general trend toward more closely interconnecting formerly disparate components of a military force. Perhaps the most widely known expression of this trend is the doctrine of Network Centric Warfare (NCW),[259] an approach to conducting hostilities which emphasizes the importance of sharing information between force elements in order to best utilize that information and the force's capabilities. Originally developed in the United States, similar approaches to interconnecting military systems are being pursued, or at least have been proposed, by the North Atlantic Treaty Organization (NATO)[260] as well as in Australia,[261] the United Kingdom[262] and other countries.[263] One of the expected benefits of such information-sharing is the possibility of a degree of decentralized operation, or 'self-synchronisation',[264] wherein two or more entities can interact directly and coordinate their efforts without employing a traditional hierarchical command and control structure. Relevantly, where some of those integrated systems have significant levels of autonomous operation, the sharing of information will not necessarily be conducted or directly supervised by a human so that, for example, if intelligence gathered by one autonomous system is utilized in an aggressive action by another autonomous system, all the systems involved may become part of a legal investigation.

While concepts such as NCW are not intrinsically linked to development of autonomous capabilities in military systems, they offer visions of the type of environment in which such advanced systems are likely to operate, and of the degree of integration between

---

[259] See generally David S Alberts, John J Garstka and Frederick P Stein, *Network Centric Warfare* (Department of Defense Command and Control Research Program, 2nd revised ed, 1999).

[260] *NATO Network Enabled Capability* (27 October 2010) North Atlantic Treaty Organization <http://www.nato.int/cps/de/SID-815535E4-57782C82/natolive/topics_54644.htm>.

[261] Defence Science and Technology Organisation, *Network-Centric Warfare* (3 September 2007) Australian Government Department of Defence <http://www.dsto.defence.gov.au/research/4051/>.

[262] 'Network Enabled Capability' (JSP 777 ed 1, United Kingdom Ministry of Defence, 2005) <http://webarchive.nationalarchives.gov.uk/20121026065214/http://www.mod.uk/NR/rdonlyres/E1403E7F-96FA-4550-AE14-4C7FF610FE3E/0/nec_jsp777.pdf>.

[263] Tim McKenna et al., 'Science & Technology for Australian Network-Centric Warfare: Function, Form and Fit' (Australian Government Department of Defence, 2005) 2 <http://docsfiles.com/view.php?view=http://www.dsto.defence.gov.au/attachments/ST%20for%20Australian%20NCW.pdf&keyword=the%20australian%20approach%20to%20warfare&count=>.

[264] Alberts, Garstka and Stein, above n 259, 175; for a fuller discussion of the status of self-synchronisation see B J A van Bezooijen, P J M D Essens and A L W Vogelaar, 'Military Self-Synchronization: An Exploration of the Concept' (Paper presented at 11th International Command and Control Research and Technology Symposium, Cambridge, UK, 26–28 September 2006) <http://www.dodccrp.org/events/11th_ICCRTS/html/papers/065.pdf>.

military systems which may become the norm. If this trend continues as appears to be planned, it will become increasingly difficult to separate one autonomous system from another in respect of involvement in some incident. Indeed, attempts to draw distinctions between separate systems, such as between lethal and non-lethal systems, may become increasingly artificial. Each device which forms part of a 'system of systems' will exhibit both individual behaviour as a system in itself and group behaviour as a component in a larger network.

3.3 Conclusion

The purpose of this chapter has been to explain what machine autonomy is and how autonomous capabilities are likely to be integrated into weapon systems over the coming years. The most important point is that machine autonomy is to be seen as a form of control rather than a weakening of control, or the absence of control, and is typically done by encoding the machine's behaviour in software installed in the machine's control system.[265] It primarily affects the relationship between human operators and their weapons and it is likely to be present in military systems other than weapons. The next chapter links these conclusions to IHL to identify the ways in which development and use of AWS will require examination of the ability of the law to regulate their use.

[265] For a general discussion of the pros and cons of encoding legal rules in software see James Grimmelmann, 'Regulation by Software' (2005) 114 *Yale Law Journal* 1719.

Chapter 4 Identifying Legal Issues

The foregoing discussion of machine autonomy and its manifestations in military applications outlined the various aspects of autonomy as it relates to weapon systems. The next step is to determine the manner and circumstances in which AWS might raise novel legal questions.

The first section addresses the threshold question of how to classify, for legal purposes, military systems which appear at first glance to combine aspects of both weapon and combatant. The next three sections outline the various ways in which autonomous behaviour of weapons might raise legal concerns: through new applications, through its effects on decision-making processes, and through its effects on the manner in which tasks are performed.

4.1 Categorisation of New Technology

The technical change to which law must respond is a set of new self-governance capabilities being granted to weapon systems, with a consequent change in the role of the combatants who operate those weapons. The legal concepts corresponding to weapons and combatants are therefore the logical starting points for assessing the appropriate legal response. If AWS conform to the currently accepted legal concept of a 'weapon' and those persons who deploy and operate them fit the description of 'combatants', the normative effect of rules governing weapons and the ways in which they can be used is unchanged: that is a question of definition. If the definitions are consistent, those rules can then be applied to proposed AWS developments to decide whether and in what circumstances their use would comply with IHL: that is a question of legality:

> It is one thing to seek to determine whether a concept is known to a system of law, in this case international law, whether it falls within the categories proper to that system and whether, within that system, a particular meaning attaches to it: the question of the existence and content of the concept within the system is a matter of definition. It is quite another matter to seek to determine whether a specific act falling within the scope

> of the concept known to a system of law violates the normative rules of that system: the question of the conformity of the act with the system is a question of legality.[266]

If there are ways in which AWS do not conform to the law's concept of a 'weapon', or their operators might play a role inconsistent with the concept of a 'combatant', then more difficult questions must be asked about how the law should respond. Should development of AWS be restricted to stay within the current legal framework? Should the legal concepts of 'combatant' and 'weapon' be extended? Should they be combined in some way such that an AWS might be seen as both weapon and combatant? Should new concepts and/or rules be developed?

The first task is therefore to decide whether AWS are properly classified as 'weapon systems' for legal purposes as that term is understood in IHL. Does the capability to perform some tasks traditionally assigned to human combatants mark an AWS as being in a different legal category from ordinary, manual weapons? There is no precise definition of 'weapon' or 'weapon system' contained in API, but given the Protocol's intent to 'reaffirm and develop the provisions protecting the victims of armed conflicts',[267] a plain reading of the provisions relating to weapons and their use indicates that the terms must be broadly construed in order to ensure such protection. The ICRC Commentary to Article 35 of API, a provision which reaffirms the customary principle that 'the right of the Parties to the conflict to choose methods or means of warfare is not unlimited', states that means of warfare include 'weapons in the widest sense'.[268] Discussions of the material scope of application of the weapons review mechanism invoked by Article 36 also indicate that the concept is to be understood broadly, with various States providing expansive definitions of 'weapon'. Australia's Department of Defence has defined 'weapon', for the purpose of its regular process of weapon reviews, as

[266] *Fisheries Jurisdiction (Spain v Canada) (Judgment)* [1998] ICJ Rep 432, 460 [68].
[267] *API* Preamble.
[268] de Preux, above n 2, 398 [1402].

> an offensive or defensive instrument of combat used to destroy, injure, defeat or threaten. It includes weapon systems, munitions, sub-munitions, ammunition, targeting devices, and other damaging or injuring mechanisms.[269]

Similarly, the US Department of Defense Law of War Working Group has defined 'weapons' as

> all arms, munitions, materiel, instruments, mechanisms, or devices that have an intended effect of injuring, damaging, destroying or disabling personnel or property[270]

The same Group defines 'weapon system' as

> the weapon itself and those components required for its operation, including new, advanced or emerging technologies which may lead to development of weapons or weapon systems and which have significant legal and policy implications. Weapons systems are limited to those components or technologies having direct injury or damaging effect on people or property …'[271]

These definitions do not appear to present any barrier to considering AWS to be weapons, especially in light of the need to view the concept broadly. Nevertheless, despite the final proviso above limiting weapon systems to 'those components or technologies having direct injury or damaging effect', the outer limits of what may be considered a weapon or weapon system are poorly defined. In particular, the extent to which the legal concept of a weapon or weapon system can encompass traditionally human tasks associated with target selection (those corresponding to legal obligations borne by combatants) is almost entirely unexplored and is absent from existing definitions. It is therefore necessary to look beyond definitions and attempt to draw analogies between existing IHL concepts and the types of autonomous technologies under investigation.

---

[269] 'Legal Review of New Weapons' (Defence Instruction (General) OPS 44-1, Australian Department of Defence, 2 June 2005) sub-s 3(a).
[270] International Committee of the Red Cross, 'A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977' (2006) 88 *International Review of the Red Cross* 931, 937 n 17.
[271] Ibid.

'Particularly in the context of changing technologies, the law almost always considers new technology as merely a new form of something else.'[272] Analogical reasoning is a common approach to dealing with new legal situations, including regulation of new technologies,[273] and it can be applied to regulation of AWS. The key point is to identify the appropriate analogy to use, as the choice of analogy determines the regulatory regime which will apply.[274]

Richards and Smart illustrated the importance of the choice of analogy in their recent article on robotics and law,[275] by reference to the 1928 United States Supreme Court case of *Olmstead v United States*.[276] *Olmstead* came on appeal by several persons convicted of violating the *National Prohibition Act*[277] through their involvement in a large operation focussed on transporting and selling alcohol. The convictions had been secured by relying heavily on evidence obtained by wiretapping private telephone conversations between the convicted persons and other conspirators. Wiretapping had been practised at least since the 1890s[278] but its constitutionality had not previously been tested. The appellants in *Olmstead* argued that the use of wiretapped evidence violated, relevantly, the Fourth Amendment of the United States Constitution prohibiting unreasonable searches and seizures. The majority opinion, written by Chief Justice Taft, drew an analogy between tapping a telephone call and a 'physical conception of a search'[279] in rejecting the appeal:

> The Amendment does not forbid what was done here. There was no searching. There was no seizure. The evidence was secured by the use of the sense of hearing, and that only. There was no entry of the houses or offices of the defendants.[280]

---

[272] Neil M Richards and William D Smart, 'How Should the Law Think about Robots?' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar, 2016) 17.
[273] Dan Hunter, 'Reason is Too Large: Analogy and Precedent in Law' (2001) 50 *Emory Law Journal* 1197, 1231.
[274] For an outline of a similar problem in a different context see Craig H Allen, 'The Seabots are ~~Coming~~ Here: Should they be Treated as "Vessels"?' (2012) 65 *The Journal of Navigation* 749.
[275] Richards and Smart, above n 272.
[276] 277 US 438, 466 (1928).
[277] 41 Stat 305, ch 85.
[278] William Lee Adams, *Brief History: Wiretapping* (11 October 2010) Time Magazine <http://content.time.com/time/magazine/article/0,9171,2022653,00.html>.
[279] Richards and Smart, above n 272, 13.
[280] *Olmstead v United States,* 277 US 438, 464.

(The wiretaps had been performed 'without trespass upon any property of the defendants',[281] by connecting listening devices to the telephone wires in publically accessible places.)

Justice Brandeis took a different view in his dissenting opinion, relying in part on a broader view of the Fourth Amendment as protecting privacy rather than just physical property, and in part on an analogy between telephone calls and postal mail:

> In Ex parte Jackson, 96 U. S. 727, it was held that a sealed letter entrusted to the mail is protected by the Amendments. The mail is a public service furnished by the Government. The telephone is a public service furnished by its authority. There is, in essence, no difference between the sealed letter and the private telephone message.[282]

The two viewpoints thus relied on different readings of the Fourth Amendment, but supported by a focus on different aspects of telephone tapping (Taft relying on the physical process of tapping a telephone line and Brandeis relying on its effect on the privacy of the monitored party). The ruling in this case stood until 1967 when it was overturned in *Katz v United States*.[283] The decision in *Katz* again relied on the functional rather than physical aspects of wiretapping, finding that a focus on the physical aspects 'is, in the present day, bad physics as well as bad law, for reasonable expectations of privacy may be defeated by electronic as well as physical invasion.'[284]

A similar choice, albeit in relation to very different subject matter, arises in the legal characterisation of AWS. By embedding a subset of an operator's actions in a weapon's control system, AWS arguably invite analogies with both weapons and combatants. This is a novel issue in IHL: the legal concepts corresponding to weapons and combatants are not 'on the same scale' such that one would normally ask which of the two is most applicable to some matter, as one might ask whether a conflict is properly described as international or non-international.

Participants in the legal debate about AWS have generally adopted one of these two views, either explicitly or implicitly, and many have shifted between views in making different arguments. The fact that States generally have consented to discussing the

[281] Ibid 458.
[282] Ibid 475.
[283] 389 US 347.
[284] Ibid 362.

legal aspects of AWS as part of the CCW process indicates at least a tacit consensus that AWS are to be treated as weapons. Indeed, many participants in the legal debate refer to AWS strictly as weapons used by human combatants, in the same terms as other weapons:

> While it is undisputed that [AWS] must be capable of *being used* in accordance with IHL, difficulties in interpreting and applying these rules to *new weapons* may arise in view of their unique characteristics, the intended and expected circumstances of their use, and their foreseeable consequences in humanitarian terms.[285]

Nevertheless, many of the participants in those discussions have expressed views that, intentionally or otherwise, appear to be inconsistent with the premise that they see AWS as simply a type of weapon. Opponents of AWS development frequently refer to AWS in terms that suggest the machine would bear obligations under IHL:

> robots would appear to be incapable of abiding by the key principles of [IHL]. They would be unable to follow the rules of distinction, proportionality, and military necessity …[286]

Christof Heyns referred explicitly to robots as combatants in his 2013 report to the Human Rights Council:

> The robotics revolution has been described as the next major revolution in military affairs … But in an important respect [AWS] are different from these earlier revolutions: their deployment would entail not merely an upgrade of the kinds of weapons used, but also a *change in the identity of those who use them. With the contemplation of [AWS], the distinction between weapons and warriors risks becoming blurred, as the former would take autonomous decisions about their own use*.[287]

---

[285] International Committee of the Red Cross, 'Autonomous Weapon Systems: Implications Of Increasing Autonomy In The Critical Functions Of Weapons' (Report on Expert Meeting, Versoix, Switzerland, 15-16 March 2016) 81 (emphasis added).

[286] Human Rights Watch, *Losing Humanity: The Case Against Killer Robots* (November 2012) 30 ('*Losing Humanity*').

[287] Christof Heyns, *Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, Human Rights Council, 23rd sess, Agenda Item 3, UN Doc A/HRC/23/47 (9 April 2013) 5-6 [28] (emphasis added).

Likewise, Heather Roff has stated that with AWS, 'the weapon is also the combatant, and the decision to target and fire lies with the machine'[288] and '[t]hese machines would amount to individual commanders, as well as [Judge Advocate General] officers, weaponeering officers and intelligence officers, on a battlefield.'[289] Human Rights Watch has also taken an anthropomorphic view of AWS: 'While traditional weapons are tools in the hand of a human being, fully autonomous weapons would make their own determinations about the use of lethal force.'[290] These and similar arguments present AWS somewhat as replacements for humans in a legal sense, rather than as tools used by humans.

Commentators also refer to a robot's lack of human emotions, implicitly positioning AWS as substitutes for a human decision-maker in situations where the human's emotional state may influence their actions. Emotion-based arguments have been made both for and against use of AWS. Crootof summarises those which would support AWS development:

> [AWS] may be preferable to human soldiers in certain situations or for certain activities. … They don't get hungry, tired, bored, or sick. … They need not be motivated by self-preservation, and so may "be used in a self-sacrificing manner if needed and appropriate." They are immune from psychological "scenario fulfillment," which causes human beings to process new information in line with their preexisting beliefs … They don't act out of fear or anger, for vengeance or vainglory.[291]

Human Rights Watch also refers to emotion in presenting the opposing view:

> robots would not be restrained by human emotions and the capacity for compassion, which can provide an important check on the killing of civilians … While proponents argue robots would be less apt to harm civilians as a result of fear or anger, emotions do not always lead to irrational killing. In fact, a person who identifies and empathizes with another human being, something a robot cannot do, will be more reluctant to harm that individual.[292]

---

[288] Roff, above n 217, 211.
[289] Ibid 222.
[290] Human Rights Watch, above n 112, 5.
[291] Crootof, above n 213, 1867 [citations omitted].
[292] *Losing Humanity*, 4.

Finally, concerns about a possible inability to hold any person accountable for a proscribed act committed by an AWS necessarily position the AWS as an entity which in some way removes a legal burden from the person who would ordinarily bear that burden. Allocation of responsibility for proscribed acts done with a weapon is not problematic under IHL or international criminal law; if AWS introduce novel challenges in allocating responsibility then they are acting as something other than mere weapons.

Drawing a legal analogy between an AWS and a human combatant requires a significant conceptual leap, but that analogy is at least superficially plausible. The types of target selection functions that are encoded in AWS control system software today and into the future include tasks which would, in the case of a manual weapon system, be done by human weapon system operators. Such AWS as exist today have the ability to identify valid targets based on descriptions (such as radar signatures) provided by human supervisors, as would the operator of a manual weapon system. Future AWS may be expected to take over more complex or ambiguous decisions which today must be referred to a human. In some cases, those targeting functions correspond to legal obligations explicitly assigned to combatants, such as the precautionary obligations in Article 57 of API. Equally, the fact that human operators would presumably no longer be performing those tasks (or, at least, would share the tasks with the AWS in some way) forces one to ask where the legal responsibility for them must lie. The weapon system programmed to perform them appears to be a logical candidate: if an AWS usurps some part of a human combatant's role, perhaps the machine should be considered a combatant, or quasi-combatant, depending on its level of autonomous capability?

However, a degree of plausibility is not sufficient to categorise an AWS as a combatant. As with the legal concept of 'weapon', formal definitions are unclear on whether the concept of 'combatant' could possibly extend to AWS. The extent of the concept is comparatively well explored in terms of its human limits (that is, which humans, by their status or conduct, may be considered combatants) but the question of whether and in what circumstances a non-human entity could qualify as a combatant is new.

The most appropriate analogy may be chosen by surveying the aspects of AWS which plausibly relate to their legal characterisation.[293] A comparison of physical properties clearly favours the view that AWS are to be viewed as weapons. There is no indication that either the weapons or the digital computer hardware employed in control systems of AWS differ in any legally relevant way from those used in weapons without significant capacity for autonomous operation. Nor does the physical form of the weapon system differ in any significant way, although proposed anthropomorphic designs may lead some observers to feel otherwise. The development history of AWS, from simple 'automatic' weapons to today's systems which can operate autonomously in simple environments to the more ambitious plans for tomorrow, shows that the defining characteristics of autonomy are in the software, not the hardware. Certainly that software relies on developments in sensing, mobility, communications and other technologies that require specific hardware, but those things exist in support of autonomous capabilities, they are not the essence of autonomy.

Neither does an examination of the functional aspects of AWS suggest that their path of development has made, or will make, them other than weapon systems. On their input side, their role is still to accept a command from a human source and execute it. The nature of the command may change as the level of autonomous capability in the weapon system increases; a system with low level capabilities may require more specific low level commands such as 'move to this location', 'fire this projectile at these coordinates' while a system with higher level capabilities may be able to formulate those steps for itself based on specification of a higher level objective and information from its environment. While those increased capabilities may enable significant changes in the execution of military operations, there does not appear to be any discontinuity between different levels of autonomous capability which would place them in different legal categories based on the types of commands they are able to accept from a human operator. Today's sophisticated weapon systems are undoubtedly merely machines which execute instructions encoded in software, and it is argued here that future highly autonomous systems envisioned by designers will not be anything more than that. Certainly they will be more complex and more capable in many ways; they will likely be able to utilise information from sources beyond the reach of today's systems and

[293] Hunter, above n 273, 1214.

process that information in the face of more uncertainty, and will function effectively in more chaotic and demanding environments. Such enhancements will, however, be due to improvements within the system's software and the hardware it controls; they will not fundamentally change the nature of the system such that it should be regarded as something more than a computer-controlled machine. The system's capabilities and limitations will still result, either directly or indirectly, from human decisions and actions.

Similarly, on their output side, the purpose of all weapon systems, autonomous or otherwise, is to activate a weapon and cause harm to an enemy military target.

Some commentators contend that the ability to gather environmental information, perhaps learn from it, and use it to break a high level command into a series of sub-tasks that were not directly communicated by a human operator, is what separates AWS from other weapon systems and makes them legally something other than just weapons.[294] However, that information processing by the AWS does not change the fact that, at some level, a human does issue the command which the AWS carries out. The control system code which the AWS executes in breaking down that high level command to simpler instructions is itself a mechanism of human design rather than a source of behaviour independent of humans.

If there is a plausible argument in favour of analogising AWS to combatants, it is that transferring target selection functions from a human operator to the weapon's control system appears to also transfer the legal burden associated with those functions away from the person who is no longer performing them. It is natural to reason that if an AWS is acting in place of a human combatant in respect of selecting and engaging targets to some extent then the AWS must shoulder a corresponding part of the legal burden for those actions, perhaps even effectively becoming the combatant.

This is an example of the common human tendency to anthropomorphise nonhuman objects.[295] It has been shown that anthropomorphic responses to a machine are provoked by a range of factors: unassisted motion, an illusion of intentionality, or, of course, a

[294] See, eg, Roff, above n 288, 211.
[295] See generally Susan R Fussell et al, 'How People Anthropomorphize Robots' (2008) *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction* 145.

humanoid form, among other things.[296] In particular, it is common in technical circles to discuss autonomous systems, robots and similar devices in anthropomorphic terms. Scientists and engineers may speak of computers 'thinking' and 'making decisions', AWS 'selecting targets', and so on. It is a natural metaphor to use, when automating a system essentially means taking decisions and tasks normally handled by a human operator and making them a capability of the system itself. As mentioned above, some legal authors have also regarded autonomous systems in human-like terms, speaking of AWS as something akin to a soldier and weapon combined.

The anthropomorphic notion of AWS making decisions about their own use indicates that proponents of this view regard AWS as something that would replace their human operators not only in a physical or operational sense, but in a legal sense. That is, the weapon system itself is depicted either as the combatant which is burdened with abiding by the law in targeting decisions and similar undertakings, or at least as a device which takes those decisions out of human hands.

Variations of anthropomorphic conceptions of autonomy have highlighted one or more human-like qualities or behaviours that autonomous systems may exhibit. Most common are definitions that focus on the operational independence of systems under autonomous control; for example, 'a weapon system that, based on conclusions derived from gathered information and pre-programmed constraints, is capable of independently selecting and engaging targets';[297] '[w]hat distinguishes their functioning is the ability to independently operate, identify and engage targets without being programmed to target a specific object.'[298] In a technical sense, independence here refers to the fact that the AWS does not need an explicit instruction from a human to initiate every action (because those instructions, or the means to generate them, have been previously encoded in the weapon's control system software). However, studies that rely on this view to support legal arguments have implied it means independence of action in the sense that the AWS is not under human control, or that no human is responsible for its

[296] Karolina Zawieska (Presentation & Text, CCW Meeting of Experts on LAWS: Overarching Issues, April 2015).
[297] Crootof, above n 213, 1837.
[298] Markus Wagner, 'Autonomy in the Battlespace: Independently Operating Weapon Systems and the Law of Armed Conflict' in Saxon, above n 240, 105.

behaviour.[299] As explained above, human control is not absent from autonomous systems.

Legal arguments based on an anthropomorphic conception of machine autonomy are not supportable. By casting a weapon system in the legal role of a soldier to any significant extent, an anthropomorphic view ignores the important fact that soldiers and the weapons they wield are, legally, distinct types of entities, even if autonomous capabilities in the weapon appear to blur the distinction in an operational sense.[300] The strongest argument against analogising AWS to combatants is that combatants are legal persons under IHL, while weapons, computers, software and the other components of an AWS, are not.

The entire body of IHL assumes (although nowhere is it explicitly stated) that the legal category of 'combatants' is restricted to human beings. Combatants, as legal persons, bear obligations and rights which go to the heart of IHL and which have evolved along with IHL to enable the law to achieve its ends. The assumption that those obligations and rights are borne by the persons who exercise control over acts of violence in armed conflict is embedded deep within the law, as is the assumption that those persons can be punished for transgressions. It is by regulating their behaviour that the effects of armed conflict are confined to legitimate military targets, those persons not participating are protected, and the other aims of the law are achieved. To position AWS as combatants in a legal sense is to place them in the role of exercising some original control over decisions to commit violent acts. The foregoing discussion of the technology of autonomous systems explained that any such appearance of independence from human control is an illusion. AWS, like other weapons, are legally dissimilar to combatants and a legal analogy between them cannot stand.

Of course, non-human entities can acquire or be granted legal personality if it is considered desirable, under IHL or any other body of law. Corporate legal personality is the obvious example. Although it would be an unprecedented step, there does not appear to be any insurmountable barrier to granting an artificial entity a status which

---

[299] See, eg, Human Rights Watch, *Mind the Gap: The Lack of Accountability for Killer Robots* (April 2015).

[300] Marco Sassòli, 'Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified' (2014) 90 *International Law Studies* 308, 323.

would support a transfer to it of the legal burden of whatever targeting or other tasks its control system is programmed to perform.[301] However, such a move would be highly ill advised. Humans would still retain full control over AWS through manufacturing and programming, regardless of the legal status of the AWS, meaning that AWS would effectively be a legal proxy for their creators. Assuming that legal responsibility is also assigned to AWS, one must surely question the wisdom of allowing organisations to manufacture machines which can act in accordance with a person's designs and take responsibility for the outcome.

This risk of abuse, along with the dichotomy deep within the law between the legal personhood of combatants and the lack thereof for weapons, seems also to discourage ideas of seeing AWS as some combination of weapon and combatant. Such a move would also raise further difficult questions of what degree of autonomy or other traits would qualify an AWS as falling into that new category.

To summarise, the position taken in this thesis is that one very significant barrier prevents AWS from being effectively analogised to combatants: combatants are legal persons, subjects of the law. Regardless of any similarity to soldiers that they may eventually possess in terms of abilities or operational roles, AWS are artefacts, pieces of equipment constructed and utilised by people. Neither existing law nor policy considerations support the proposition that legal personality has been or should be extended to an artefact. For legal purposes, AWS must be treated simply as weapons.

A related question is whether use of a particular AWS, or AWS as a class, could be considered a method of warfare, where 'method' refers to 'the way in which [weapons] are used'.[302] Given the amorphous nature of machine autonomy and the many ways in which it may be employed in an operation, an extremely broad notion of 'method' would be needed in order to describe use of AWS per se as a specific method of warfare. Certainly, far broader than it has been used in other contexts.

More plausible is the idea that some specific behaviour encoded into a particular AWS could amount to a method of warfare. However, even in that case it is arguable that one

[301] See, eg, Lawrence B Solum, 'Legal Personhood for Artificial Intelligences' (1992) 70 *North Carolina Law Review* 1231.
[302] de Preux, above n 2, 398 [1402].

must adopt an overly expansive concept of what it means to 'use' a weapon. Every weapon has some intrinsic behaviour, and that behaviour is treated as part of the 'means' of warfare, not the method. For example, the behaviour of the guidance system of a missile or the trigger mechanism of a mine is a trait of that means of warfare. The programmed behaviour of an AWS is not fundamentally different, although perhaps considerably more complex. A 'method' of warfare, the way in which a weapon is used, is what happens between a human and the weapon, not the behaviour of the weapon itself. That use may be further removed from directly discharging the weapon when the weapon has a degree of autonomy, but the distinction remains: decisions about which AWS to use for an operation, when and where to deploy it, how to arm it and similar choices collectively constitute the method of warfare, while the AWS's inbuilt, programmed behaviour, even if it imitates something a human might otherwise have done, are part of the means.

A question still remains about the degree of autonomous capability that is required to raise legal concerns. It is certainly not the case that all weapon systems which can manage some aspect of their own operation are legally significant because of it. Many weapon systems which have been in operation for decades exhibit some limited degree of autonomous behaviour and no serious argument has been made that their use is, or should be, illegal due to that capability. For example:

- Close-In Weapon Systems (CIWS): This class of point defence weapon system is commonly mounted on naval vessels at sea or military bases on land. Once activated by a human operator, the weapon system scans a defined area for threats that match pre-defined signatures, using radar and infra-red sensors and a computer-based control system to detect incoming missiles and other close-in air and surface threats such as helicopters or high-speed surface vessels. Selected targets are automatically fired upon with a rapid-fire gun or a missile.[303] CIWS systems are intended to be a last line of defence against incoming attacks, in situations where a human would be unable to respond quickly enough or accurately enough.

[303] *Phalanx Close-In Weapon System: Last Line of Defense for Air, Land and Sea*, Raytheon Australia <http://www.raytheon.com.au/capabilities/products/phalanx/>.

- The Aegis Combat System[304] is another missile defence system deployed primarily on naval vessels. Like the CIWS, Aegis utilises advanced radar capabilities to identify and track threats, and to guide missiles launched in response. While it has the ability to select the target to be attacked and launch the missile, it does so with the oversight of a human operator,[305] making it what is sometimes termed a 'human-supervised [AWS]'.[306]
- The Patriot[307] system is a ground-based, long-range air defence system. It shares high-level technical similarities with the Aegis system in that it relies on radar to detect and track incoming missiles and aircraft, and can respond automatically by launching a guided missile to intercept the threat.
- Israel's Iron Dome system operates on similar principles to 'detect, assess and intercept incoming rockets, artillery and mortars.'[308] It can assess incoming threats based on their trajectory and the likelihood that they will strike an occupied area, to decide whether or not to intercept them by launching a missile in response.
- Perhaps counterintuitively, the argument has sometimes been made that even ordinary land mines qualify as AWS in a strictly technical sense.[309] After being armed and emplaced by a human operator, a mine can 'select' and engage its target entirely without human interaction. In simple mines of the type which led to formulation of the *Ottawa Convention*[310] the weapon's targeting system is very crude, consisting solely of a pressure-sensitive trigger which 'selects' targets that exert sufficient pressure and ignores those that do not. Of course, that is a very ineffective method of distinguishing military from civilian targets and is the reason their use is so widely proscribed. Nevertheless, in a strictly

---

[304] *Aegis Combat System*, Lockheed Martin <http://www.lockheedmartin.com.au/us/products/aegis.html>.
[305] Michael N Schmitt, 'Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics' (2013) *Harvard National Security Journal Features* 4; Curtis E A Karnow, 'The Application of Traditional Tort Theory to Embodied Machine Intelligence' in Calo, Froomkin and Kerr, above n 272, 54.
[306] *Directive 3000.09*, 14.
[307] *Global Patriot Solutions*, Raytheon <http://www.raytheon.com/capabilities/products/patriot/>.
[308] *Iron Dome Weapon System*, Raytheon <http://www.raytheon.com/capabilities/products/irondome/>.
[309] See, eg, Marchant et al, above n 239, 276.
[310] *Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on their Destruction*, opened for signature 18 September 1997, 2056 UNTS 211 (entered into force 1 March 1999) ('*Ottawa Convention*').

technical sense, such a device does select and engage targets without human intervention.[311]

If more advanced AWS are to raise legal concerns *due to their capacity for autonomous operation*, then that capacity, or the circumstances in which it is used, must exceed or diverge in some way from the autonomous capabilities of existing systems such as those listed above. There are many dimensions along which those differences may occur, some of which are interdependent. Broadly, there are three ways in which an expanded capacity for autonomous operation in a weapon system might give rise to legal concerns: through new military applications, through their effects on decision-making processes, and through their effects on task performance. Each of these is discussed next.

4.2 Applications of Weapon Autonomy

As autonomous technologies are developed further, they are likely to be integrated into weapon systems in different ways, and those weapon systems are likely to be deployed in an expanded range of scenarios in armed conflict.

4.2.1 Applications of Autonomy within a Weapon System

Two fundamental properties of an AWS relate directly to the legal implications of using it in the course of an armed conflict: the scope and the degree of its autonomous capabilities. In other words, its legal status would depend heavily on which specific functions are made autonomous, and what degree and type of human oversight is feasible in the intended and actual circumstances of its use.

4.2.1.1 Functional Scope of Autonomous Capability

As noted in Section 3.1.3.3 above, autonomy is a property of specific functions of a weapon system, such as target selection or navigation, rather than of the system as a whole. Weapon systems of legal interest are likely to be those with autonomous capability in a function to which legal obligations relate. In the context of weapon systems, the functions of most interest are, obviously, selecting and engaging targets. This has become the dominant area of concern in the debate about AWS. It is reflected in the definition put forward by the United States Department of Defense:

[311] The technology of mines is also evolving towards greater autonomy: Matthew Bolton, 'From Minefields to Minespace: An Archeology of the Changing Architecture of Autonomous Killing in US Army Field Manuals on Landmines, Booby Traps and IEDs' (2015) 46 *Political Geography* 41.

> autonomous weapon system. A weapon system that, once activated, can select and engage targets without further intervention by a human operator. This includes human-supervised [AWS] that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation.[312]

This definition reflects the weapon system functions of most legal interest ('select and engage targets') and acknowledges that autonomy is a matter of degree ('includes human-supervised [AWS]'). It has been cited on various occasions in the CCW meetings about AWS and has been used by many commentators in publications on the topic.

It is possible that legal consequences may also flow from autonomous performance of other functions of a weapon system. For example, a mobile platform with autonomous navigation capabilities deployed by a party to a conflict may cross an international border and enter the territory of a neutral State. That would violate the prohibition on belligerents moving war materiel across neutral land territory.[313] Such possibilities are not directly addressed in this thesis, but the analysis presented in the following chapters may be of some assistance in evaluating the consequences of autonomous operations other than selecting and engaging targets.

4.2.1.2 Facility for Human Oversight

The second dimension is the degree of autonomous capability or, equivalently, the amount and type of human interaction that is possible or required. As explained in section 3.1.3 above, this thesis takes the view that machine autonomy exists on a continuum, wherein the relative degrees of control exercised by human and machine may vary during operation. In this view, the relative degrees of human and machine involvement in selecting and executing actions are not readily divisible into unambiguous, discrete levels. Rather, the extent to which a human is both nominally in control of some function of the weapon system and practically able to exercise control over that function is subject to infinitely many variations depending on the precise design and use of an AWS and the circumstances in which the AWS is operating.

---

[312] *Directive 3000.09* 13-14

[313] *Convention (V) Respecting the Rights and Duties of Neutral Powers and Persons in Case of War on Land*, signed 18 October 1907, 205 CTS 299 (entered into force 26 January 1910) art 2.

Current weapon systems which exhibit some autonomous capability require, or at least allow, a high degree of human involvement in relation to each use of force via the weapon. In some cases that involvement is in the form of supervision with the ability to override operation of the AWS control system, such as with CIWS systems. In other cases, it is in the form of initial target selection and activation of the weapon system, such as with some advanced missiles.[314]

Development plans for future AWS generally indicate that the degree of human involvement required by AWS will decrease. Indeed, that is the source of both the mooted operational advantages and the legal and humanitarian concerns about AWS. Whether those concerns are realised will depend largely on the extent to which each of the various decisions and actions associated with operation of a weapon system are assigned to the weapon's control system rather than to a human.

In respect of tasks assigned to a weapon's control system, much will depend on whether it remains feasible for a human to intervene in order to take corrective action, update information for the AWS, abort an operation or otherwise prevent an illegal act from occurring. Systems and circumstances which prevent human intervention, rather than simply reduce or eliminate the need for it, are more likely to exceed legal limitations which place obligations on those who operate weapon systems.

4.2.2 Applications of AWS

Today, weapon systems with a capacity for autonomous operation are utilised only in a limited range of scenarios. Development plans that have been made public point to a greatly expanded range of applications for future systems.

4.2.2.1 Anti-personnel or Anti-materiel Roles

With the exception of anti-personnel mines, the existing AWS listed in Section 4.1 are all used in anti-materiel roles, in which legal targets are relatively easy to identify and distinguish from civilian objects, and the chance of harm to a civilian is likely to be less than in anti-personnel roles. Although arguably more relevant to studies in other fields such as human rights or ethics, the question of whether AWS should be deployed in

[314] See, eg, *Advanced Medium Range Air-to-Air Missile (AMRAAM)*, Naval Air Systems Command <http://www.navair.navy.mil/index.cfm?fuseaction=home.display&key=D3FAC4AB-2E9F-4150-9664-6AFBC83F203E>.

anti-personnel roles is sometimes raised in debates about their acceptability under IHL.[315] All else being equal, an AWS directed against human targets is likely to pose a greater risk to civilians than the same AWS directed against a vehicle, building or other object due simply to the greater risk of a 'false positive' target identification.[316]

4.2.2.2 Physical Domain

Some of the rules of IHL as it stands today are applicable only to warfare in specific physical domains: land, sea, air, space and the information or cyber-warfare domain. Use of weapon systems must conform to all rules applicable to the domain in which the weapon is used. It is possible that circumstances may arise in which the legal status of an AWS or a particular autonomous capability depends on the domain in which the AWS is to be used.

4.2.2.3 Complexity of the Environment

Complexity of the environment refers to the collective effect of all the variables which might challenge the ability of an AWS control system to complete its assigned task. Those include, but are not limited to, the following:

- The presence or absence of civilian persons and objects. Opponents of AWS development frequently cite the difficulty of distinguishing an enemy combatant from a civilian or friendly combatant as a factor limiting the legitimate uses of AWS in armed conflict.[317]
- The strength, technical sophistication and degree of organisation of the enemy.
- The ease of navigation and effectiveness of sensors in the operating environment (such as the different challenges of urban versus remote areas, or of land, sea, underwater, or aerial operation, and so on).

4.2.2.4 Others

It is not contended that this list of considerations is exhaustive. More focussed investigation based on a specific AWS or autonomous capability may reveal other ways

[315] See, eg, Israel (Text, CCW Meeting of Experts on LAWS: Characteristics of LAWS, April 2015) 2; Frank Sauer, *Banning Lethal Autonomous Weapon Systems (LAWS): The Way Forward* (13 June 2014) International Committee for Robot Arms Control <http://icrac.net/2014/06/banning-lethal-autonomous-weapon-systems-laws-the-way-forward/>.

[316] For that reason an AWS intended for use in an anti-personnel role is more likely to excite legal concern: Chris Jenks, 'False Rubicons, Moral Panic, & Conceptual Cul-De-Sacs: Critiquing & Reframing the Call to Ban Lethal Autonomous Weapons' (2016) 44(1) *Pepperdine Law Review* (Article 2).

[317] See, eg, *Losing Humanity* 4; see also Sassòli's rebuttal: Sassòli, above n 300, 333.

in which the uses, and consequently the legal status, of AWS may vary in the future. In general, a relevant factor is one which affects the risk of a proscribed act occurring. A given AWS may not raise any significant concerns when used for a purpose and in a situation where the risk of performing a proscribed act is low but the same AWS may be deemed unacceptable in circumstances where, given the capabilities of the system, the risk of transgressing the law is higher. Possible examples may include the remoteness of the AWS operating environment, the duration of the operation, the time available to react to an enemy action, and so on.

4.3 Effects on Decision-making Processes

One key consideration in evaluating the legal status of an AWS is that autonomy operates on the decision to perform some action via a weapon system, such as launch an attack, whereas other capabilities traditionally associated with weapon development have an effect only after a decision has been made. As explained in Chapter 3, this does not mean that sufficiently autonomous systems will themselves make decisions.[318] Rather, the character of decisions made in operating autonomous machines, and the process of making those decisions, differs from those made in operating an equivalent manual system. Additionally, the responsibility for the outcome of those decisions may be shifted.

The control system that regulates the behaviour of an AWS forms part of the 'kill chain', in that its software encodes decision processes relating to attack precautions, target selection and other attack-related tasks that have been delegated to the weapon. Depending on the circumstances, use of an AWS in place of an equivalent manually operated weapon may affect how the decision to launch an attack is made and/or the conduct of individuals in carrying out an attack. It will not, however, necessarily affect the substance of any decision or the conduct of any attack.

Encoding operational decisions in a weapon's control system necessarily alters the character of those decisions in three interrelated ways:

- Generality: when human decisions made 'on the spot' in respect of specific situations are replaced with, or supplemented by, programmatic instructions that

[318] See, contra, Krishnan, above n 82, 33.

had been previously provided to a machine, decisions about individual acts in specific situations are replaced with broader policy-like choices applicable to the range of situations which match whatever parameters had been previously provided to the machine.

- Timing: decisions about whether and how to perform an action via an AWS are effectively made at the time the relevant behaviour is programmed into the machine and at the time the decision is made to employ the weapon, rather than at the time the situation arises in a conflict. This is important because, firstly, it may have implications for the temporal scope of application of IHL[319] and, secondly, it requires an assumption that situations in which the machine is deployed in the future will not differ in important respects from those envisioned at the time the machine was developed and tested.
- Basis: decisions implemented via an AWS cannot be based on direct (or even indirect) human observation of the situation to which the decision relates; rather, they must be based on whatever information is available through experience and foresight at the time the machine is programmed, and then confirmed when the decision is made to employ the weapon.

One result of these changes is that the causal link between a specific human decision and a specific (repetition of an) action or a specific outcome, such as a particular target being fired upon, may be weakened when the decision is enacted partly or fully via a highly autonomous system.

A capacity for autonomous operation becomes legally problematic if the magnitude of those effects takes a decision-making process leading to weapon activation outside that which is contemplated by IHL. That may happen in at least two ways:

- when the resulting decision process violates a legal rule;
- when the resulting decision process challenges an assumption on which the law is based.

[319] For a discussion of one such question see Section 7.3.3; also: Tim McFarland and Tim McCormack, 'Mind the Gap: Can Developers of Autonomous Weapons Systems be Liable for War Crimes?' (2014) 90 *International Law Studies* 361.

4.4 Effects on Task Performance

For the foreseeable future, autonomous systems will continue to be based on digital computers and robotic devices. Although machine autonomy is principally about how an operator interacts with a machine rather than about the primary function of that machine, higher levels of autonomy will nevertheless affect the performance of actions which are delegated to AWS. That is, a task that is done through an autonomous machine is likely to succeed or fail in different ways and in different circumstances than it would if done manually by a human. As levels of autonomy increase and humans move 'further' away from manually completing tasks, the intrinsic properties and behaviours of digital autonomous systems will play a more dominant role in the manner in which those tasks are performed by an AWS.[320] In terms of legal compliance, the most significant considerations in this area relate to the ability to predict how an AWS will behave once activated, and the circumstances and manner in which an AWS may fail to perform as required by IHL.

4.4.1 Predictability of Performance and Failure

The ability for a weapon designer or operator to predict how a weapon system will behave once activated is fundamental to meeting legal obligations which are to be met via that system's behaviour. The nature and purpose of AWS, being to remove humans from positions of directly performing or overseeing combat operations, makes predictability of behaviour a particularly pressing concern, even more likely to be a limiting factor governing the legal acceptability of increasingly autonomous weapons. It is critical for policy makers and armed forces personnel to understand whether, and in what circumstances, limitations on the ability to predict the behaviour of AWS in combat might restrict the uses to which they can legally be put.

4.4.1.1 Complexity

Ultimately, the ability to predict how an AWS will behave in a given set of circumstances is limited by complexity arising from various sources. First, the intrinsic complexity of the weapon systems themselves may be expected to keep increasing. Modern weapon systems, like other digital systems, comprise large numbers of hardware and software components which tend to be complex in themselves as well as

[320] For a discussion see Defense Science Board, above n 222.

being interconnected with, and dependent upon, other components in complex ways. That trend of increasing complexity is, to an extent, a natural consequence of increasing the functionality of weapon systems: equipping them to undertake more sophisticated tasks in more difficult environments means enabling more responses to more combinations of command inputs and environmental stimuli in order to perform more, and more complex, operations. That in turn means the control systems of modern robotic devices such as AWS rely on large and extremely complex software that is very challenging to write and to maintain. The scale of the challenge presented may be appreciated by considering the rate at which errors typically occur in software. While there is considerable variation, studies have shown the average software industry error rate to be approximately 15 – 50 errors per 1,000 lines of code written.[321] As a rough indication of the number of errors that may be encountered in practice, one modern weapon system, the F-35 Joint Strike Fighter, uses over 20 million lines of code.[322] Further, the difficulty of reliably predicting weapon system behaviour is increased exponentially by the combined effects of increasingly complex tasks, difficult and chaotic operating environments and the often unpredictable actions of sophisticated adversaries.[323]

The second source of difficulty is the fact that AWS behaviour must, in many cases, be predicted over a much longer period of time than would be the case with other weapon systems. One of the major motivating factors for increasing levels of autonomy is to gain persistent capabilities; that is, the ability for operations to be sustained over longer periods of time than would be possible or feasible using existing weapon systems. Even today, armed forces employing UAVs in combat operations have noted the benefits of persistent ISR capabilities that allow them to maintain visibility of the operational theatre, in order to monitor the activities of adversaries and so on. Hopefully further advances in autonomous technologies may increase those benefits, yielding 'potentially unlimited persistent capabilities without degradation due to fatigue or lack of attention.'[324] In order to realise those advantages, operators of AWS must be able to be

[321] Scharre, above n 135, 13.
[322] Ibid 13 n 13.
[323] For a general discussion of risk and AWS reliability see Roberto Cordeschi, 'Automatic Decision-Making and Reliability in Robotic Systems: Some Implications in the Case of Robot Weapons' (2013) 28 *AI & Societ*y 431.
[324] Defense Science Board, above n 222, 15.

confident that the machines will operate as intended over long periods, including periods in which it may not be possible for humans to intervene to correct errors, due to communications being denied, or physical access to the AWS being infeasible.

For some systems, complexity can rise to the level where the precise behaviour of the system as a whole is simply incomprehensible to operators. This point has been made in relation to the accident at Three Mile Island, in which a cooling malfunction in a nuclear power plant in Pennsylvania, USA, caused part of the core to melt and the reactor was destroyed:

> This highlights an important point about complex systems: Their sheer complexity can make their behavior seem incomprehensible to human operators. The human operators at Three Mile Island were attempting to understand the system through a dizzying array of gauges and indicators — some 2,000 alarms in the control room. … Some of the corrective actions they took to manage the accident were, in retrospect, incorrect. It would be improper to call their actions human error, however. They were operating with the best information they had at the time. They simply could not have known better.
>
> The incomprehensibility of complex systems is a key feature of normal accidents and a major obstacle in human operators' ability to intervene and regain control of a complex highly automated or autonomous system. … it is simply not feasible to map all of the possible interactions of the system with itself, its operators, and its environment. (This problem is even further exacerbated in competitive situations where adversaries will try to seek out weaknesses in the system to hack or exploit.)[325]

A natural response to the above might be that AWS simply need to be tested sufficiently to remove uncertainty about how they will behave once deployed. Unfortunately though, testing provides only a partial solution to the problem of complexity. To remove all uncertainty about how an AWS would behave in operation, it would have to be exhaustively tested: every possible combination of inputs from operators, the environment and adversaries would have to be anticipated and replicated, along with every interaction between components of the system; the AWS's responses would have to be observed and, where needed, corrected and the system re-tested. For a sufficiently complex system, that task is infeasible with reasonable development resources within

[325] Scharre, above n 135, 27-8.

any reasonable development time.[326] Coupled with the fact that it is very unlikely that complete knowledge of operating environments and adversary actions could be anticipated, the complete certainty that comes with exhaustive testing is not practically achievable, and some risk of unintended behaviour will always remain. Certainly, considerable design effort is expended in reducing and mitigating the risk of unintended behaviour, but not all risk can be designed away. It is therefore necessary to consider whether the predictability of behaviour of a given AWS is adequate to enable operators of the AWS to fulfil their legal obligations.

4.4.1.2 Brittleness

One property of automated and autonomous systems that is of particular interest is their potential for 'brittleness'.[327] Brittleness here refers to the commonly encountered problem of systems failing when faced with challenges for which they were not programmed. As explained earlier, stored-program computers must execute the programs stored in them and are unable to do anything else. They are, therefore, able to function in the situations which their programmers foresaw and programmed them to function in, and no others. When faced with situations at all outside the scope of their programming, they are very likely to fail; computers cannot exercise 'common sense' in order to deal with unfamiliar situations in the way that humans can[328] (hence the idea of 'brittleness'; there is a rigidly defined set of situations in which the program can operate successfully and it is unable to 'bend' itself to fit situations not in that set). While programmers naturally attempt to anticipate all situations to which their software will be exposed, the complexity of the problems often limits the success of those attempts. Brittleness may be encountered for various reasons:

> the inability of the designer to anticipate and design for all of the scenarios that could arise during the use of the system, a deliberate decision by the designer to use an oversimplified model of the decision task (due to cost, time or technological limitations),

[326] Ibid 25.
[327] Benjamin Kastan, 'Autonomous Weapons Systems: A Coming Legal "Singularity"?' [2013] *Journal of Law, Technology & Policy* 45, 51.
[328] Scharre, above n 135, 6.

> a failure of the designer to correctly anticipate the behavior of the system in certain situations, or a failure to correctly implement the intended design.[329]

The idea of brittleness is introduced here because it describes likely failure modes of AWS, and failure of weapon systems to operate as intended is of obvious relevance to a discussion of their legal status. AWS would generally fail differently to how human soldiers would fail; humans may fail to adequately perform some task for which they are, in theory, 'programmed' (trained) due to inattention, fatigue and a host of other human factors, but when a human encounters situations outside their circle of expertise, they are able to apply common sense, imagination and other abilities to lessen the negative impact. Computers, on the other hand, do not typically fail to follow their programming but, being bound to stay strictly within the limits of that programming, tend to 'fail' more suddenly when faced with unanticipated circumstances.

It may be tempting to say that existing automated weapon systems are also brittle. They certainly lack any real ability to adapt themselves to unexpected situations. The observation is correct in a narrow sense, but brittleness is a significant consideration only where flexibility and adaptability are expected or required. Typically, that means only where 'high level' decisions are delegated to an autonomous system; It is not necessary to consider adaptability in a simple system which is not expected to be able to adapt itself.

One significant motivation for research into artificial intelligence and further development of intelligent control systems is to alleviate the brittleness of autonomous systems.[330] While those efforts will certainly have some effect, the extent to which the benefits may be realised in real world scenarios remains to be seen. Efforts to endow AWS with abilities resembling intelligence as a remedy for brittleness also introduce another limitation to predicting their behaviour: the broader an AWS's ability to autonomously formulate responses to unexpected situations, the more difficult it will be for operators to anticipate how the AWS will behave after being activated. It will be necessary to carefully consider how to maintain the desired balance between military

---

[329] Philip J Smith, C Elaine McCoy and Charles Layton, 'Brittleness in the Design of Cooperative Problem-Solving Systems: The Effects on User Performance' (1997) 27(3) *IEEE Transactions on Systems, Man and Cybernetics* 360, 360.
[330] See, eg, Matt Ginsberg, *Essentials of Artificial Intelligence* (Newnes, 2012) 5.

necessity and humanity, given the possibilities of greater operational effectiveness and superior adherence to rules of IHL, and the novel challenge of anticipating and controlling AWS behaviour to realise those advantages without stepping outside the limits of those same rules.

4.4.2 Nature and Consequences of AWS Failure

Autonomy in weapon systems offers the possibility of both superior performance and superior adherence to legal rules, but it also introduces a risk that, if an AWS should perform in an unintended way, the magnitude and consequences of that failure could greatly exceed an equivalent failure by a human soldier.

One can gain a sense of the potential consequences of an AWS failure by reviewing any of several unfortunate incidents involving Patriot missile systems in recent conflicts. Two such incidents occurred during the 2003 invasion of Iraq.

The first occurred on March 24. A combination of three failures led to a Patriot missile battery firing upon a British Tornado fighter, destroying the aircraft and killing the crew. First, the Patriot's targeting system misidentified the Tornado, marking it as an incoming anti-radiation missile, which would have been a legitimate target. Second, the Identification Friend or Foe (IFF) system which should also have identified the Tornado as a coalition aircraft 'performed very poorly'.[331] Third, the human operator whose role was to approve the Patriot's target identification and authorise the engagement, accepted the incorrect identification in a manner which was later described as 'trusting the system without question'.[332]

The second incident occurred on April 2. A Patriot battery detected a false signal which appeared to be an incoming missile but was 'likely the result of electromagnetic interference, possibly due to employing radars in a non-standard configuration in the field.'[333] Following the previous incident, Patriot systems in Iraq had been placed in a standby mode to prevent automated engagements, but when the system alerted its operator to the apparent incoming missile, the operator placed the battery in auto-fire

[331] Defense Science Board, 'Patriot System Performance' (Report Summary, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, January 2005) 2 <http://www.dtic.mil/get-tr-doc/pdf?AD=ADA435837>.
[332] Scharre, above n 135, 31 n 42.
[333] Scharre, above n 135, 31.

mode in preparation for an engagement. For reasons that were unclear to investigators, the Patriot fired upon a US Navy F/A-18C Hornet fighter, bringing down the aircraft and killing the pilot.

It is not suggested that the Patriot system should be subject to some special regulation due to its capacity for autonomous operation. These incidents are mentioned here because, lacking deployed examples of the type of highly autonomous weapon system being proposed, it is necessary to extrapolate from the examples of autonomous technologies that are available. Two observations can be made.

4.4.2.1 Possibility of Runaway Failures

In both of the above incidents there was considerable human involvement, arguably human error, contributing to the unintended engagements. However, that same human involvement also acted to limit the extent of the unintended engagements. The fact that human approval was required prior to the Patriot system firing at the aircraft, and the fact that a human was on hand to give that approval, meant that the system's unwanted behaviour was stopped after one engagement, as soon as the error became known to the human operator. Hypothetically, what might have happened if the role of the human operator had been subsumed into the Patriot system? Take the April 2 incident, for example. As far as the Patriot system was aware, a missile was incoming; that is what its sensors told it was happening and that is the only information it had to act upon. If a human was not 'in the loop' (or 'on the loop') and able to intervene, would there have been another way to halt the Patriot's operation? Assuming that whatever source of electromagnetic interference continued to generate false signatures of incoming missiles, would the Patriot have continued firing?

A human operator, being separate from the weapon systems they use, and availed of different senses and sources of information, acts as a fail-safe device when a weapon system malfunctions. If some part of the role of the human is absorbed into the weapon system, thought must be given to whether, and how, to retain a fail-safe function in order to contain malfunctions.

4.4.2.2 Common Vulnerabilities among Weapon Systems

Weapon systems are, of course, mass produced, with effectively identical hardware and software being shared among large numbers of systems. Hardware and software flaws

would be replicated also. The Defense Science Board of the US Department of Defense wrote in relation to use of Patriot systems in Iraq:

> The number of coalition aircraft flights in [Operation Iraqi Freedom] was enormous, 41,000, and the Patriot deployment was large, 60 fire units, so the possible Patriot-friendly aircraft observations were in the millions and even very-low-probability failures could result in regrettable fratricide incidents.[334]

The risk associated with absorbing some part of a human operator's role into a weapon's control system should not be viewed in terms of individual systems, but in terms of the aggregate risk arising from the total number of such systems in use.

Humans, on the other hand, are idiosyncratic. Different people act differently in the same situation, even where they have been given the same training. This tendency is frequently the cause of mistakes, but it also acts to isolate those mistakes to individuals rather than allowing them to be characteristic of a whole population.

4.5 Conclusion

The purpose of this chapter has been to explain how the technologies which make a weapon system 'autonomous' relate to the law which governs their use. The conclusions are that AWS are to be regulated as weapons, nothing more, and there are at least three ways in which legal questions about their use in armed conflict might arise: through their application in an expanded range of scenarios, their effects on military decision-making processes, and their effects on the manner in which tasks may be performed in comparison to tasks done by human operators. The next chapters apply the foregoing material to rules and principles of IHL, presenting an assessment of the legal challenges that lie ahead as weapon systems are made increasingly autonomous.

As autonomy affects the manner in which a weapon system is operated as much as the nature of the weapon system itself, legal questions are most likely to arise in the areas of weapons law and targeting law. Chapters 5 and 6 focus on those two areas.

The fact that the decision process leading to weapon activation changes when the weapon operates autonomously raises questions about the accountability of those legal

[334] Defense Science Board, above n 331, 1.

persons who take part in that process. Chapter 7 discusses questions about State and individual accountability for acts committed via AWS.

Chapter 5 Weapons Law

This chapter contains a study of the constraints that must be imposed on the design of AWS in order for their employment in armed conflict to comply with legal requirements.[335]

The chapter is in four sections. The first section defines the set of autonomous systems that are of interest given the scope of the discussion; that is, the set of systems that must be regulated as weapon systems. The second section discusses the unique way in which the two sets of rules governing the nature of weapons and the use of weapons interact in relation to AWS. The third section discusses law relating specifically to AWS. The fourth section applies rules about weapon system legality *per se* to AWS.

5.1 Autonomous Systems Subject to Rules of Weapons Law

This thesis is concerned with autonomous machines playing a meaningful role in the application of force during an armed conflict. Systems that directly incorporate weapons, such as autonomous UAVs or gun turrets, obviously fall within that scope. Other types of machines that contribute to the decision to use force, or play some part in carrying it out, are also of interest. One example might be an autonomous ISR system which provides information about potential targets directly to a separate weapons platform. A useful guide as to the scope of autonomous machines of interest is Article 36 of API, which 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. [emphasis added]

Broadly, autonomous systems which may be described as weapons, means or methods of warfare under Article 36 would enliven legal considerations associated with use of weapons, and are of interest in this thesis. There is some disagreement as to whether

[335] For a specific proposal for using legal requirements as constraints on the design of AWS see Tony Gillespie and Robin West, 'Requirements for Autonomous Unmanned Air Systems set by Legal Issues' (2010) 4(2) *The International C2 Journal* 1.

Article 36 precisely reflects customary international law,[336] but most States, including some non-States Parties to API, accept that there is an obligation to review new means of warfare for compliance with their obligations under IHL. There is less consensus on whether new methods of warfare must be reviewed.[337]

The set of devices that may be considered a weapon or means of warfare is broad. Unfortunately, the phrase 'weapons, means or method of warfare' is not formally defined in API, so a reasonable interpretation must be applied.[338] The ICRC Commentary on Article 35 of API, which sets out the basic rules relating to means and methods of warfare, states that '[t]he words "methods and means" include weapons in the widest sense, as well as the way in which they are used.'[339]

'Weapon' is understood to be meant in its ordinary sense: 'Deciding whether an item of equipment is a weapon will be a relatively straightforward process. The term connotes an offensive capability that can be applied to a military object or enemy combatant.'[340] In relation to that offensive capability, Boothby notes: 'The means whereby this is achieved will involve a device, munition, implement, substance, object, or piece of equipment, and it is that device, etc that is generally referred to as a weapon.'[341] 'Means of warfare' is generally understood to be a somewhat broader term that includes weapons as well as 'weapons platforms, and associated equipment used directly to deliver force during hostilities.'[342] 'Methods of warfare' refers to the ways in which weapons are used; it includes the 'tactics, techniques and procedures (TTP) by which hostilities are conducted.'[343] At the 2016 CCW Informal Meeting of Experts, Dr Gilles Giacca of the ICRC gave an example of the scope of material subject to a review:

> Defensive or offensive device with the capability to kill, injure, disable people and/or destroy or damage property. Ranging from rifles, plateforms [sic], sighting equipment, laser designators, target acquisition equipment, data links and software used for

[336] *Losing Humanity*, 21 n 82.
[337] Michael N Schmitt and Jeffrey S Thurnher, '"Out of the Loop": Autonomous Weapon Systems and the Law of Armed Conflict' (2013) 4 *Harvard National Security Journal* 231, 271.
[338] Justin McClelland, 'The Review of Weapons in Accordance with Article 36 of Additional Protocol I' (2003) 85 *International Review of the Red Cross* 397, 404.
[339] de Preux, above n 2, 398 [1402].
[340] McClelland, above n 338, 404.
[341] Boothby, above n 64, 4.
[342] Ibid.
[343] Schmitt and Thurnher, above n 337, 271.

> processing target data all require legal review. It would also include new military doctrine that applies to weapons.[344]

The fact that means of warfare includes weapon platforms and components of weapon systems is important in the context of autonomous systems. It is to be expected that the control systems which define the behaviour of AWS would often exist partly or fully in a weapon platform, or in a separate device attached to or in communication with a weapon system. That is not to suggest that there is doubt that AWS should be regulated as weapons. The ICRC Commentary to API states that automation of the battlefield was an area of direct concern to the drafters:

> The use of long distance, remote control weapons, or weapons connected to sensors positioned in the field, leads to the automation of the battlefield in which the soldier plays an increasingly less important role. The counter-measures developed as a result of this evolution, in particular electronic jamming (or interference), exacerbates the indiscriminate character of combat. In short, all predictions agree that if man does not master technology, but allows it to master him, he will be destroyed by technology.[345]

If there is doubt about whether a particular autonomous system should be considered a weapon for legal purposes, it would be most likely to stem from one of two factors. The first is the nature of the system's task. A UAV with autonomy in the functions of selecting and engaging targets is unequivocally an AWS, while an unarmed cargo transport vehicle with autonomous navigation capabilities is unequivocally not a weapon; the status of other autonomous systems may be less clear. An autonomous ISR system might be treated as a component of a weapon system if it provides targeting information directly to a weapons platform, but what if it provides information only to some intermediate device, or to a human who then issues commands to the weapons platform? The strength of the causal relationship between the autonomous system and infliction of harm on an enemy is a matter of degree, and those people tasked with reviewing weapons and means of warfare must apply judgement. It should be remembered, though, that the scope of devices to be treated as weapons or means of

[344] Giacca, above n 179, 4.
[345] Jean de Preux, 'Protocol I – Article 36 – New Weapons' in Sandoz, Swinarski and Zimmermann, above n 2, 421, 427-8 [1476].

warfare is intended to be broad, and States should be cautious in ensuring they are taking appropriate steps to meet their legal obligations.

The second factor that may cause doubt is the relative degree of human and computer control over the activities of the autonomous system. Section 3.1.3 explained that a human operator may exercise varying degrees and forms of control over an autonomous system at different times, in relation to different functions, in different circumstances. If a human is expected to exercise sufficient control over, for example, an ISR system, such that the human is effectively an intervening entity between the ISR system and a weapon platform, then the autonomous ISR system may not be considered sufficiently 'close' to the weapon that it warrants being treated as part of a weapon system.

The need to review new methods of warfare is likely to be of less relevance to development or acquisition of AWS. One can theorise that the existence of an autonomous capability might lead to some new tactic being employed in using an AWS, but it is likely that would be encompassed in a review of the AWS as a means of warfare rather than as a separate matter. The set of autonomous systems that may be considered weapons or means of warfare is therefore wide, but not unlimited.

5.2 Weapons Law and Targeting Law

When discussing the legal status of a weapon system, two sets of rules belonging to IHL must be analysed. The first set, commonly known as 'weapons law', deals with whether a weapon system is legal *per se*. That is, whether there are any circumstances in which the weapon system could be used in armed conflict in accordance with a State's legal obligations. These rules deal with the intrinsic nature of the weapon system, its behaviour and how it affects people or objects against which it is directed. The second set of rules, known as 'targeting law', deals with how a weapon system is used. It relates to whether the weapon system can be legally used in a particular manner or in a particular situation. Rules of targeting law prescribe how operators of a weapon system must behave in order to use the weapon system in accordance with the law.

Some confusion may arise in applying the rules of weapons law and targeting law to AWS because the existing rules were developed without reference to autonomous technologies which enable a machine to decide upon and manage significant parts of its own operation. Endowing a machine with a capacity for autonomous operation means

taking certain operational responsibilities out of the hands of a human operator and making them part of the behaviour of the machine itself. In other words, some behaviours that would otherwise be attributed to the way in which the machine is used, must instead be attributed to the nature of the machine. For example, weapon system autonomy means that some tasks that would otherwise be done by a human operator as part of the targeting process (selecting a target, assessing collateral damage, or similar tasks) instead become part of the intrinsic behaviour of the weapon system. The rules of weapons law would otherwise relate to what, in an AWS, is only one part of the weapon system (the weapon or munition), while some targeting law considerations are encoded into another part of the weapon system (the control system). Current views on the types of factual situations which give rise to issues under weapons law, and which enliven targeting law obligations, may therefore need to be reassessed when dealing with AWS.

Another way to view the distinction between weapons law and targeting law is that the burden of abiding by the rules of weapons law generally falls primarily on those who develop and procure weapon systems (or who approve acquisition of the weapon systems on the basis of an assessment of their compatibility with relevant international legal obligations), while targeting law places obligations primarily on attack planners, commanders and weapon system operators. This is a useful guide for applying existing law to AWS. Behaviours which are intrinsic to the weapon system and which cannot feasibly be altered, or are not intended to be altered, in the field in the course of preparing for and conducting an attack, are to be seen as a matter of weapons law and must be evaluated according to those rules. Behaviours which would generally be determined by an attack planner or weapon system operator once the weapon system has been deployed remain a matter of targeting law and must be evaluated according to those rules. For instance, a hypothetical autonomous Unmanned Combat Air Vehicle (UCAV) being prepared for a sortie might require an operator only to select a target profile, such as a description of a military installation or vehicle, a search area and a mission start time and duration, with all subsequent decisions after launch being within the purview of the UCAV's control system. In that case, rules of targeting law would apply to the decision to use that particular weapon system in that time and place, in light of whatever knowledge is available about the battlefield, and about military and civilian activity in the area, and about how the UCAV is reasonably expected to behave in those

conditions. The behaviour of the UCAV and its control system when operating without a human in the loop is a matter of weapons law. If the UCAV was to refer some decision to a human operator (perhaps if an unexpected situation arises) then the human input to that decision would be a matter of targeting law.

To be clear, legal responsibility for targeting is not transferred to an AWS or to those responsible for its programming and capabilities. The fact that some target selection software is incorporated into an AWS control system does not mean that selection of a target becomes a matter of weapons law rather than targeting law. It simply means that the process of selecting a target involves different steps when done via manually operated and autonomous weapon systems. The responsibility for ensuring that targets are lawful remains with planners and commanders in the field. The steps those planners and commanders would take when prosecuting an attack with an AWS would need to include ensuring the chosen AWS incorporates targeting software that will select the desired (lawful) target (and, concomitantly, will not select an unlawful target), and that whatever procedures are involved in directing the AWS to attack that target are done correctly. Defining the operation of the targeting software incorporated into the AWS, ensuring that it is not of a nature to select unlawful targets or select targets indiscriminately or otherwise cause the AWS to, by its nature, operate in violation of IHL, is a matter of weapons law. Validation and verification of the operation of that software would fall within the scope of a weapon review.

In operational terms, weapon autonomy may appear to blur the distinction between operator and weapon. However, the legal framework which places certain obligations on soldiers in the field and other obligations on those who manufacture, procure, review and authorise the use of weapon systems does not change in response to evolving weapon system capabilities. The important question is whether, and how, States and members of their armed forces can meet their existing legal obligations while employing AWS in a conflict. The next sections discuss the substantive legal requirements imposed on autonomous systems that constitute weapons or means of warfare.

5.3 Law Relating Specifically to AWS

There are various classes of weapons to which specific conventional or customary restrictions or prohibitions apply. These include poisons, chemical and biological weapons, mines, cluster munitions,[346] blinding lasers,[347] weapons that leave non-detectable fragments in the human body,[348] and more. As at the time of writing there is no conventional or customary international law specifically directed at AWS or their use in armed conflict. Arms control law prohibiting or restricting the production, possession, transfer or use of specific weapons would apply to the weapon or weapons attached to an AWS as it applies to those weapons in other configurations, but that law is peripheral to the question of how the legal status of a weapon system is affected by having a capacity for autonomous operation. The law applicable to AWS today and for the near future is therefore limited to general conventional and customary IHL relating to weapons and their use. If, in the future, written law specifically directed at AWS is formed, it appears most likely either to be generated by the CCW process or by some other diplomatic track along the lines of the processes that led to the negotiation of the Ottawa Convention and of the Cluster Munitions Convention. However, the CCW discussions are at an early stage and are unlikely to reach a conclusion for some years.

The task of applying weapon-specific regulations to AWS will not be an easy one. It perhaps presents different challenges than existing regimes have. Regulatory regimes that prohibit or restrict particular types of weapons currently do so on the basis of some specific and well-defined property of the system; that is, some well-defined type of harm inflicted on the target. Blinding lasers are prohibited because of the intended effect of the laser beam on the human eye. Mines and cluster munitions are regulated because their regular modes of deployment necessarily preclude accurate and precise selection of military targets. Non-detectable fragments necessarily complicate medical treatment of injured persons.

---

[346] *Convention on Cluster Munitions*, opened for signature 30 May 2008, 2688 UNTS 39 (entered into force 1 August 2010) ('*Cluster Munitions Convention*').

[347] *Additional Protocol 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 (Protocol IV, entitled Protocol on Blinding Laser Weapons)*, opened for signature 13 October 1995, 2024 UNTS 163 (entered into force 30 July 1998).

[348] *CCW*, opened for signature 10 October 1980, 1342 UNTS 137 (entered into force 2 December 1983) *(Protocol I, entitled Protocol on Non-Detectable Fragments)*.

Enabling a weapon system to operate autonomously is a different form of development, in that autonomy does not give rise to any definite and specific behaviour of the weapon system, nor any specific effect on the targets of the weapon. A capacity for autonomous operation does not dictate precisely what a weapon system does, only the degree and type of operator interaction that is involved in doing it. Further, machine autonomy is a capability that may be achieved to varying degrees, through a range of different technologies, and many existing weapons already possess that capability to some extent. It seems that a significant challenge to formulating regulations specifically for AWS will be defining precisely what technology, or weapon effect, or mode of operation is to be restricted or prohibited.[349]

5.4 Legality *per se*

Given that there is, at present, no law specifically prohibiting use of AWS in armed conflict, the next question is whether and in what circumstances AWS might fall foul of one of the fundamental norms of IHL which render certain weapons illegal *per se*.

5.4.1 Threshold Levels of Autonomy

Objections to the legality of AWS are sometimes expressed as blanket statements about weapon systems with a certain degree of autonomous capability. In particular, the charge is sometimes made that use of highly autonomous or 'fully autonomous' weapon systems would necessarily be in violation of IHL, because a sufficiently highly autonomous weapon system would be inherently incapable of behaving in accordance with IHL. Human Rights Watch took this position in their 2012 report, *Losing Humanity*:

> fully autonomous weapons would not only be unable to meet legal standards but would also undermine essential non-legal safeguards for civilians. Our research and analysis strongly conclude that fully autonomous weapons should be banned and that governments should urgently pursue that end.[350]

In this context, '[t]he term "fully autonomous weapon" refers to both out-of-the-loop weapons and those that allow a human on the loop, but that are effectively out-of-the-

[349] However, see the discussion of 'meaningful human control' in Section 6.3, as a candidate for a general principle which may form the basis of specific regulatory measures.
[350] *Losing Humanity*, 1-2.

loop weapons because the supervision is so limited.'[351] The report claims that fully autonomous weapons would not comply with IHL because:

> fully autonomous weapons would lack the human qualities necessary to meet the rules of [IHL]. These rules can be complex and entail subjective decision making, and their observance often requires human judgment. For example, distinguishing between a fearful civilian and a threatening enemy combatant requires a soldier to understand the intentions behind a human's actions, something a robot could not do.[352]

Several State delegations to the CCW Informal Meetings of Experts have expressed similar views.[353] Pakistan stated in 2016:

> the introduction of [AWS] would be illegal, unethical, inhumane and unaccountable as well as destabilizing for international peace and security with grave consequences. Therefore, their further development and use must ideally be pre-emptively banned … the states currently developing such weapons should place an immediate moratorium on their production and use.[354]

Various academic authors have made a range of similar arguments. Roff's objection is that AWS with sufficient learning ability to be useful in an operational context would be unpredictable and uncontrollable: 'given the current state of machine learning, it is the very ability of a machine to overwrite its code that makes it, in a sense, unpredictable and uncontrollable.'[355] According to Roff, use of such machines in armed conflict 'would undermine the existing command and control structures put in place to ensure compliance with the laws of war and reduce miscommunications and redundancies resulting in harm.'[356]

Peter Asaro's concerns focus on the removal of human involvement from lethal decision-making:

> It is my view that [AWS] represent a qualitative shift in military technology, precisely because they eliminate human judgement in the initiation of lethal force. Therefore they

[351] *Losing Humanity*, 2.
[352] *Losing Humanity*, 4.
[353] 'Country Views on Killer Robots' (Campaign to Stop Killer Robots, 13 December 2016) <http://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC_CountryViews_13Dec2016.pdf>.
[354] Pakistan, above n 172, 1-3.
[355] Roff, above n 217, 221.
[356] Ibid 222.

> threaten to undermine human rights in the absence of human judgement and review. There are good reasons to clarify IHL and IHRL by explicitly codifying a prohibition on the use of [AWS].[357]

Asaro's central objection to AWS is that

> an implicit requirement for human judgement can be found in [IHL] governing armed conflict. Indeed, this requirement is implicit in the principles of distinction, proportionality, and military necessity that are found in international treaties, such as the 1949 Geneva Conventions, and firmly established in international customary law.[358]

Removing human involvement from decisions to apply lethal force would therefore be unacceptable:

> The very nature of IHL, which was designed to govern the conduct of humans and human organizations in armed conflict, presupposes that combatants will be human agents. It is in this sense anthropocentric. Despite the best efforts of its authors to be clear and precise, applying IHL requires multiple levels of interpretation in order to be effective in a given situation. IHL supplements its rules with heuristic guidelines for human agents to follow, explicitly requires combatants to reflexively consider the implications of their actions, and to apply compassion and judgement in an explicit appeal to their humanity. In doing this, the law does not impose a specific calculation, but rather, it imposes a duty on combatants to make a deliberate consideration as to the potential cost in human lives and property of their available courses of action.[359]

Other authors have presented arguments that differ in various ways but share many of the same elements as those summarised above. Blanket arguments against AWS that have been presented thus far fail for several reasons.

5.4.1.1 Anthropomorphic Views of Machine Autonomy

Human Rights Watch and Asaro both refer to the inability of AWS to make decisions about the application of lethal force as a reason to restrict use of autonomy in weapon systems. By so doing, they (implicitly or explicitly) say that a human decision-maker is not involved, and instead they place the weapon system in the role of decision-maker

---

[357] Peter Asaro, 'On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making' (2012) 94 *International Review of the Red Cross* 687, 693.
[358] Ibid 687.
[359] Ibid 700.

for legal purposes. The flaws of attempting to conflate the roles of soldier and weapon were explained in Section 4.1. As a general rule, arguments which cast AWS as legal decision-makers by claiming that they would be unable to make targeting or other decisions to the standard required by law, or that it would be inherently illegal to have them attempt to do so, fail by implying that legal burdens would be transferred to the machine. Natural humans and organisations with corporate legal personality remain the bearers of legal obligations. The pertinent question is whether the people who develop and operate AWS will be able to meet their obligations when acting through an autonomous machine.

5.4.1.2 Means and Ends

Arguments that autonomous operation beyond a certain level renders a weapon system inherently illegal fail not only because of what they say about machine autonomy, but because of what they implicitly assume about the rules of weapons law. Weapons law imposes restrictions on the designed behaviour of a weapon system: whether it is able to be directed at a specific target in the context of an attack (distinction), and the type and extent of damage it may inflict (superfluous injury / unnecessary suffering). The law has nothing to say about the means by which that behaviour can, or cannot, be achieved. Autonomy, on the other hand, is intrinsically only about how weapon behaviour is achieved; controlling a weapon by encoding its behaviour into a computer rather than manipulating it manually. Arguing that an AWS control system cannot control the weapon to which it is attached in accordance with the rules of IHL amounts to arguing that there is some fundamental limit on what can be achieved through robotics, artificial intelligence and related technologies, and that limit is somewhere below what human operators of manual weapon systems can achieve in armed conflict scenarios. No reasoning has been presented to support the existence of such a limit (other than references to scenarios which would be infeasible given today's technology; that is discussed next) and the claim is far too contentious to stand on its own.

5.4.1.3 Example Scenarios

AWS, like any weapon system, must of course be able to perform to an adequate standard in order to be used legally: if an AWS is intended to be used to select and engage enemy personnel or vehicles in a particular setting, its targeting systems must be capable of identifying those targets well enough to ensure that attacks are not directed

against civilian targets or friendly forces. Some arguments against AWS rely on pointing out the very serious technical challenges that face weapon developers attempting to construct AWS which could conceivably replace human soldiers in complex scenarios. From Human Rights Watch:

> fully autonomous weapons would not possess human qualities necessary to assess an individual's intentions, an assessment that is key to distinguishing targets. … For example, a frightened mother may run after her two children and yell at them to stop playing with toy guns near a soldier. A human soldier could identify with the mother's fear and the children's game and thus recognize their intentions as harmless, while a fully autonomous weapon might see only a person running toward it and two armed individuals. The former would hold fire, and the latter might launch an attack. Technological fixes could not give fully autonomous weapons the ability to relate to and understand humans that is needed to pick up on such cues.[360]

Similarly, from Noel Sharkey:

> There is also the Principle of Proportionality and again there is no sensing or computational capability that would allow a robot such a determination, and nor is there any known metric to objectively measure needless, superfluous or disproportionate suffering. They require human judgement.[361]

The observations made in those and similar examples would undoubtedly be correct in respect of today's autonomous systems technology if it were to be used in complex combat scenarios, but there are two reasons that they do not support the conclusion that use of AWS in general is necessarily illegal.

The main problem is that the strength of the inductive generalisation is weak due to the sample of scenarios being heavily biased towards the types of complex, ambiguous problems that most seriously challenge today's autonomous systems. Proposed uses of future AWS extend far beyond anti-personnel missions in environments where civilians are in immediate danger of becoming involved, and in the near future it appears likely that most development of AWS will occur in domains in which the technical challenges are more tractable.

---

[360] *Losing Humanity*, 31-2 [citation omitted].
[361] Sharkey, 'Grounds for Discrimination: Autonomous Robot Weapons', above n 81, 88 [citation omitted].

For example, one of the areas currently attracting the most research effort is autonomy of systems operating in the undersea domain, far from the difficult combatant-vs-civilian identification/distinction problems that characterise urban and counter-insurgency fighting. This is partly because the advantages of autonomy in undersea systems are particularly clear. Water interferes with radio communications, forcing undersea vehicles to rely on either very low radio frequencies or acoustic communications.[362] Both options suffer from low data transmission rates and risk betraying the vehicle's position. An inability to operate for long periods without a support vessel nearby similarly inhibits undersea vehicles, providing another motive for developing autonomous capabilities. However, it is also because undersea operation does not realistically threaten to involve civilians. Other operational domains distant from civilian activity provide similarly appealing environments for autonomous systems, such as the ocean surface, deserts and even cyberspace.

It should also be noted that the types of scenarios most commonly used in arguments against AWS development are in fact the scenarios in which AWS are least likely to be employed, according to current development plans:

> neither the United States nor any other country is contemplating the development of any systems that would simply hunt down and kill or destroy enemy personnel and objects without restrictive engagement parameters, such as area of operation or nature of the target.[363]

5.4.1.4 Implicit Requirement for Human Subjectivity

Asaro states that 'an implicit requirement for human judgement can be found in [IHL]',[364] and other authors either state or imply the same.[365] In a sense, it is a trite notion; the purpose of the law is to regulate the decisions and conduct of people and organisations, so it is difficult to say there is no requirement for human judgement in those decisions. Certainly there is an assumption that human judgement underlies the decisions regulated by IHL, as Asaro notes, and indeed there has never been a practical

[362] Milica Stojanovic, 'Underwater Acoustic Communications: Design Considerations on the Physical Layer' (Paper presented at Fifth Annual Conference on Wireless on Demand Network Systems and Services, Garmisch-Partenkirchen, Germany, 23-25 January 2008).
[363] Schmitt, above n 305, 6.
[364] Asaro, above n 357, 687.
[365] See, eg, Jack M Beard, 'Autonomous Weapons and Human Responsibilities' (2014) 45 *Georgetown Journal of International Law* 617 (writing in relation to State responsibility for acts done via AWS).

alternative to a human decision-maker. However, a requirement for human judgement in IHL is relevant to a discussion of the legality of AWS only if use of AWS would amount to removing human judgement. As demonstrated in Chapter 3, that is not the case. Machine autonomy is a form of human control over machines which differs from manual control in that it encodes human judgement in a software-based control system rather than applying that human judgement directly to the controlled machine. Enabling a weapon system to act autonomously does not remove human judgement, it merely applies that judgement differently. The judgement to which Asaro refers must be applied by the bearers of legal obligations under IHL: the parties to the conflict, those personnel who plan or decide upon attacks, and so on. Those bearers of legal obligations are still present and still applying their human judgement, albeit in different ways.

5.4.1.5 Conflating Nature and Use of Weapon Systems

To argue that a weapon system is illegal *per se* under IHL is to make a statement about the nature of that weapon system, that there are no circumstances in which it can be used in compliance with the law. Arguments made with reference to specific scenarios such as anti-personnel roles in urban environments, or specific types of computational tasks such as distinguishing a child with a toy gun from an enemy combatant, are arguments about specific uses of AWS, not their nature. Even where the arguments can be substantiated, they say nothing more than that a weapon system with specific capabilities should not be used in some specific scenario for which it is not suited. That is an acknowledgement that attack planners would bear responsibilities to select appropriate AWS as they must select other weapon systems, and it is an indication of the sort of performance metrics that should be established in a weapons review of AWS. It is not a statement about the nature of AWS in general.

Although a level of autonomous capability does not in itself render a weapon system illegal, it is certainly possible that an AWS with specific encoded behaviours may transgress the restrictions imposed by weapons law just as biological weapons or unguided long-range missiles do. The following sections discuss the factors which weapon developers must consider to ensure that AWS operate within the limitations imposed by rules of weapons law.

5.4.2 Distinction

5.4.2.1 Content of the Principle

This principle, one of the 'cardinal principles'[366] of IHL, requires parties to a conflict to 'at all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives.'[367] The principle therefore imposes requirements and restrictions on the means and methods of warfare that may be employed by combatants. This section discusses how to interpret those requirements and restrictions in relation to AWS, and whether any special requirements or restrictions are applicable to AWS.

There are two main facets to the principle of distinction: one concerns restrictions on the nature or design of weapons, the other concerns restrictions on the use of weapons. Restrictions on the use of AWS are discussed in detail in Chapter 6. The focus of this section is on the first facet, relating to weapons law: weapons that, by their nature or design, cannot be made to discriminate between military and civilian targets in the circumstances of a planned attack, cannot legally be used in that attack. There are two categories of such weapons, as described in Article 51(4) of API:

> Indiscriminate attacks are:
>
> …
>
> (b) Those which employ a method or means of combat which cannot be directed at a specific military objective; or
>
> (c) Those which employ a method or means of combat the effects of which cannot be limited as required by this Protocol;

The first type of weapons inherently illegal because they are incapable of compliance with the rule on distinction are those which cannot be aimed with sufficient accuracy in the circumstances in which they are intended to be used.[368] The prohibition in API states that indiscriminate attacks are '[t]hose which employ a method or means of combat

[366] *Legality of the Threat or Use of Nuclear Weapons (Advisory Opinion)* [1996] ICJ Rep 226, 257 [78].
[367] *API* art 48.
[368] Michael N Schmitt, 'The Principle of Discrimination in 21st Century Warfare' (1999) 2 *Yale Human Rights and Development Journal* 143, 147.

which cannot *be directed* at a specific military objective'.[369] Perhaps the clearest example of this type of weapon, close to the current topic, would be a long range missile with limited or no guidance capabilities, such that it is not possible to be sufficiently sure of where it will land. The ICRC commentary on the Additional Protocols to the Geneva Conventions lists the German V1 and V2 rockets used enduring the latter stages of the Second World War as two such missiles.[370]

Of course, 'sufficient accuracy' is an imprecise criterion. Some weapons are more prone to indiscriminate behaviour than others and there is no bright line cleanly separating inherently indiscriminate weapons from others. The test is context-dependent: given the generic applications for which the weapon was designed and the manner in which it is intended to be used, do any reasonable scenarios exist in which the weapon could be used in a discriminate manner?[371] Views on whether a particular weapon is inherently indiscriminate sometimes differ. Illustrative of the challenges are the debates about anti-personnel mines and cluster munitions. These are two of the best known examples of weapons that are sufficiently prone to strike unintended targets that they have been prohibited by many States.[372] However, both weapons can in fact be used in a legally compliant manner[373] in specific circumstances when subject to appropriate restrictions.[374] The use of Scud missiles by Iraq in 1991 has also been cited as an

---

[369] *API* art 51(4)(b) [emphasis added].

[370] Claude Pilloud and Jean Pictet, 'Protocol I – Article 51 – Protection of the Civilian Population' in Sandoz, Swinarski and Zimmermann, above n 2, 613, 621 [1958]; Stuart Casey-Maslen and Sharon Weill, 'The Use of Weapons in Armed Conflict' in Stuart Casey-Maslen (ed), *Weapons Under International Human Rights Law* (Cambridge University Press, 2014) 248.

[371] Boothby, above n 64, 1.

[372] *Ottawa Convention*, *Cluster Munitions Convention*.

[373] *Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices as amended on 3 May 1996 (Protocol II, as amended on 3 May 1996) annexed 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 3 May 1996, 2048 UNTS 93 (entered into force 3 December 1998) ('*CCW Protocol II*'); Secretary of Defense, 'DoD Policy on Cluster Munitions and Unintended Harm to Civilians' (Memorandum, United States Department of Defense, 19 June 2008) <http://www.acq.osd.mil/tc/treaties/ccwapl/DoD%20Policy%20on%20Cluster%20Munitions.pdf>.

[374] For a general discussion of *CCW Protocol II* (albeit prior to the 1996 amendment) see A P V Rogers, 'A Commentary on the Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices' (1987) 26 *Military Law and Law of War Review* 185; Burris M Carnahan, 'The Law of Land Mine Warfare: Protocol II to the United Nations Convention on Certain Conventional Weapons' (1984) 105 *Military Law Review* 73; Peter J Ekberg, 'Remotely Delivered Landmines and International Law' (1995) 33 *Columbia Journal of Transnational Law* 149.

example,[375] although perhaps more controversially. Scud missiles have been described as 'highly inaccurate theater ballistic missiles' which 'can cause extensive collateral damage well out of proportion to military results'[376] and their deployment against Israeli and Saudi cities was widely criticised as indiscriminate.[377] In other circumstances though, SCUD missiles can be used in a discriminate manner: 'For example, if employed in the vast expanses of the desert against troops, military equipment, or bases far removed from population centers, little danger of random destruction of protected persons or objects exists.'[378]

The second type of inherently illegal weapons are those with effects that cannot be controlled well enough to restrict them to military targets. API describes these weapons as 'those which employ a method or means of combat *the effects of which cannot be limited as required by this Protocol*; and consequently, in each such case, are of a nature to strike military objectives and civilians or civilian objects without distinction'.[379] Whereas the first type of inherently illegal weapon is prohibited due to an inability to accurately direct its effects at a military target, this type is prohibited due to an inability to subsequently confine those effects to the target at which they were directed. The proviso 'as required by this Protocol' indicates that the effects referred to here are not necessarily only the direct effects of the weapon, but any effects which trigger concerns under other articles of API. Those include:[380] Articles 35(3) and 55, which prohibit 'widespread, long-term and severe damage to the natural environment', indicating that more than just the immediate effects of the weapon are at issue; Article 56, which prohibits attacks on dams, dykes and other 'works and installations containing dangerous forces', indicating that second-order effects of weapons may be considered; Articles 51(5)(b) and 57(2)(a), which prohibit use of weapons which would inflict any excessive damage in violation of the principle of proportionality.

---

[375] Eitan Barak (ed), *Deadly Metal Rain: The Legality of Flechette Weapons in International Law* (Martinus Nijhoff, 2011) 210; Schmitt, above n 368, 148.
[376] United States Department of Defense, 'Report to Congress on International Policies and Procedures Regarding the Protection of Natural and Cultural Resources During Times of War' (19 January 1993) reproduced in: Patrick J Boylan, *Review of the Convention for the Protection of Cultural Property in the Event of Armed Conflict* (UNESCO, 1993) 203 <http://unesdoc.unesco.org/images/0010/001001/100159eo.pdf>.
[377] See, eg, M N Schmitt, 'Future War and the Principle of Discrimination' (1999) 28 *Israel Yearbook on Human Rights* 55.
[378] Schmitt, above n 368, 148.
[379] *API* art 51(4)(c) [emphasis added].
[380] Bothe, Partsch and Solf, above n 68, 347.

The canonical example of a type of weapon prohibited by Article 51(4)(c) is bacteriological weapons; regardless of how accurately a bacteriological substance is delivered to a target initially, there is no known means of confining the spread of disease to enemy combatants only, so the effects of this type of weapon cannot be controlled as required by IHL. Poisoning wells containing drinking water is another example.

5.4.2.2 Application to AWS

The prohibition on weapons with effects that cannot be limited as required is less intimately related to autonomy, although it may be relevant in some cases. The prohibition relates more to the physical means by which the weapon part of the system causes harm to the target, such as the type of ammunition or harmful substance employed.[381] Autonomy relates to how that weapon is controlled, so an inability to limit the effects which is attributable to the autonomous control system is best seen as an inability to direct the weapon at a specific military target.

The focus here is on the requirement that weapons must be capable of being directed at specific military objectives as per API Article 51(4)(b). This rule relates to the nature of the weapon system and is a threshold requirement specifying only a minimum standard that must be met by a weapon system in the context of an attack: 'IHL mandates no specific degree of accuracy in either weapons or tactics; it simply bars those that cannot be aimed';[382] 'it would seem clear that weapons must be capable of direction at individual military objectives and that this requirement must be understood in the context of the attack.'[383] Use of a weapon not designed to be sufficiently discriminate is a violation of the law.

Giving a weapon system some capacity to select its own targets means that part of the process of directing the weapon is taken out of the hands of a human operator and encoded into the AWS control system. The legal responsibilities associated with those parts of the process of directing a weapon at a target, which would otherwise be understood as relating to how a weapon is used, such as identifying a potential target as a military objective, instead relate to the nature of the weapon system. Novel legal

[381] Pilloud and Pictet, above n 370, 623 [1963], [1965].
[382] Michael N Schmitt, *Essays on Law and War at the Fault Lines* (T M C Asser Press, 2012) 115.
[383] Boothby, above n 64, 79.

questions arise from that because the parts of the process of directing the weapon which are encoded into an AWS control system cease to be the domain of weapon system operators, governed by the rules of targeting law, and become part of the built-in behaviour of the weapon system, governed by the rules of weapons law. They must therefore be considered in assessing whether an AWS meets the threshold requirement in API Article 51(4)(b).

Legal assessment of AWS also requires a generalised notion of accuracy, one that includes elements of precision in target selection in addition to what is normally considered accuracy of a weapon. This is admittedly a novel extension of the concept of accuracy of a weapon and may appear counterintuitive, particularly if one subscribes to an anthropomorphic notion of AWS playing the role of soldiers in selecting targets. However, it is consistent with, and required by, the view that autonomy is a form of control built into weapon systems. The main concern from a legal perspective is still whether a weapon system can be directed at a specific military objective and the degree to which the effects of the weapon can be confined to the target it strikes. That requires a measure of accuracy which includes the overall behaviour of the weapon system from the time the operator directs it at a target or targets, not just the behaviour of the final weapon component.

The process of directing a weapon essentially amounts to constraining the possible targets of the weapon so as to reach a level of certainty that only the chosen target will be struck, rather than any other potential targets. That is done through a combination of operator actions (choice of target, choice of weapon system, choice of when and where to activate it, provision of facilities for human supervision or intervention, and so on) and weapon system behaviour (performance of sensors and targeting systems, physical characteristics of the weapon and ammunition used, and so on). The principle of distinction requires that the combination of the operator's actions and the weapon system's behaviour lead to designation of a legal target and realisation of the conditions under which it can be legally attacked; the weapon system must then behave so as to direct harm at the chosen target and not at any other.

Where a simple, unguided weapon system such as a rifle or an artillery piece is used, attack planners and weapon system operators are entirely responsible for ensuring that

the weapon is directed against legal targets; that would typically be done by manipulating the weapon to deliver its ammunition to a location and time where it is known with sufficient certainty that a legal target will be found. Although the weapon system operator might have one particular target (such as an enemy combatant) in mind, from the weapon system's point of view it is not constrained to striking that individual as such; it is constrained to striking whatever set of targets might turn out to be present at the place and time at which the weapon's bullet or shell arrives. Accordingly, the role of the weapon system is not to deliver its projectile to the target that was in the operator's mind, but to deliver it to the specified location and time, with sufficient accuracy (or deliver it to the location specified by laser illumination, GPS signal, or other precision guidance system, in the case of precision-guided munitions). The accuracy of a weapon system is the tightness with which the projectile's final destination can be constrained; for a manual weapon system it is typically expressed using a simple spatial metric such as circular error probable (CEP).[384]

When a greater part of the target selection process is encoded into the weapon's control system, the allocation of tasks between operator and weapon system changes. Consider missile-defence systems such as Israel's Iron Dome or the Phalanx CIWS. The weapon system operator still applies a 'time and place' constraint to the set of targets that may be engaged, but that human-specified constraint is much looser than it would be in the case of a simple unguided weapon system. Rather than directing the weapon system's projectile to a precise location at which a specific target is expected to be found, the decision to activate the weapon system at a given time and place combined with the range of its weapons constrains only the volume of space and time in which targets may be attacked. Once that outer limit is set, control passes to the weapon's control system, which further constrains the set of potential targets by matching radar signatures of detected objects to those of incoming missiles and similar threats. The fire-control system then aims and fires the weapon. The overall accuracy of the weapon system therefore has two components: the accuracy with which it determines that a potential target is within the set of targets that it is programmed to strike, and the accuracy with which it is able to actually strike that target.

[384] William Nelson, 'Use of Circular Error Probability in Target Detection' (Technical Report, MITRE Corporation, May 1988) <http://www.dtic.mil/dtic/tr/fulltext/u2/a199190.pdf>.

The same reasoning can be applied to use of more highly autonomous, still-hypothetical weapon systems which may be based on emerging technologies. At some level an attack planner or weapon system operator will form an intention to attack some target or targets, whether a specific target, an enumerated set, all targets matching some general description within a defined area, or some other specification, and will activate an AWS. That AWS may have more freedom of action than current weapon systems (increased range or persistence, increased ability to respond to changing circumstances, ability to undertake more complex reasoning about potential targets, and so on) but will still be subject to some limitations (its maximum range and duration of operation will be finite and linked to the point at which it is activated, it can select targets only from among those present in its operational range, and so on). The attackers' obligation to attack only military targets will not change, but the weapon system's portion of the process of meeting that obligation will gradually increase.

As long as it is feasible for an operator to constrain the weapon system's choice of targets such that it is sufficiently certain to engage only targets that can be legally attacked in the circumstances, then the weapon system satisfies the threshold requirement in API Article 51(4)(b). Specifically, the operator must consider two questions:

- How well can the targeting system identify legal targets in the circumstances of the attack? This is about the performance of the targeting system. Depending on the circumstances this question may involve a number of subsidiary questions such as: how well can the targeting system distinguish combatants from civilians? Active combatants from those *hors de combat*? Facilities being used for military purposes from those being used for civilian purposes? In practice, attack planners will of course ask how well the AWS can identify the specific (type of) target at which the attack is directed, but that goes beyond the threshold legal requirement of Article 51(4)(b).
- How accurately can a chosen target be engaged? This is about the performance of the weapon and is the same consideration that applies to assessment of manual weapon systems.

Provided that a sufficient level of accuracy in target identification and weapon operation can be achieved, and the weapon used is not one with uncontrollable effects, then the weapons law aspects of the State's obligation to attack only military targets are satisfied.

It may be intuitively appealing to take the view that some obligations associated with targeting law, such as the precautionary obligation to do everything feasible to verify that a target can be legally attacked,[385] are absorbed into the weapon system along with the associated targeting functions. If that were the case, it might be further surmised that the standard applicable to those obligations becomes applicable to the AWS targeting software; that is, the software must do everything that is practically possible in the circumstances of the attack.[386] That is not so; the standard applicable to the extra functions of AWS targeting software is the same as the standard applicable to any other software integrated into a weapon system. One of the precautionary obligations on 'those who plan or decide upon an attack' is to '[t]ake all feasible precautions in the choice of means and methods of attack'[387] to minimise expected collateral damage. Whether the targeting functions of a particular AWS are adequate to discharge that duty is a matter for the attack planners to decide in the circumstances of the attack.

Development plans for AWS point to some of the specific, low-level tasks associated with those precautionary obligations being encoded directly into weapon systems, but a change in weapons technology does not in itself affect the governing legal framework. Any targeting functions encoded into an AWS must be done to the standard required by the weapons law rules relating to distinction. The steps which attack planners must take to meet their obligations will change accordingly, such that attacks conducted with AWS still comply with the State's IHL obligations.

5.4.2.2.1 Striking Protected Persons or Objects

If the concepts of accuracy and directing a weapon at a specific military objective must be expanded to encompass whatever targeting functions are encoded into an AWS control system, it is also necessary to ask how to regard incidents in which an AWS kills or harms a protected person or object, such as a civilian or a civilian object, a military medical unit, or medical or religious personnel. Opponents of AWS

[385] *API* art 57(2)(a)(i).
[386] Pilloud and Pictet, above n 67, 681 [2198].
[387] *API* art 57(2)(a)(ii).

development frequently cite (hypothetical) concerns about AWS firing upon protected persons, describing such incidents as indiscriminate attacks,[388] and as evidence that AWS are incapable of being used in compliance with IHL. But would an AWS firing at a protected person or object necessarily amount to a violation of the principle of distinction? There is a range of factors which may lead to a protected person being harmed by an AWS; many are shared with other weapon systems and some are due to characteristics of AWS. It is argued here that all can be accounted for by existing law, and none suggests that AWS cannot be used legally in armed conflict.

The legal character of harm to protected persons depends on its cause, and the possible causes can be categorised by viewing an attack with an AWS as comprising three high-level components: a human operator who, at least, activates the weapon system; a targeting system which is part of the AWS control system; and the actual weapon. Neither the operator's actions nor the weapon's behaviour would necessarily differ between an AWS and a manual weapon system. It is therefore not necessary to discuss situations in which the operator deliberately directs an attack at a protected target (which would violate API Article 51(4)(a)), or negligently activates a weapon system in circumstances for which it was not designed, or the weapon simply misses its mark due to limited accuracy as a simple manually operated weapon might. The AWS built-in targeting functions are the element of interest and those functions are not at issue where the operator or weapon is to blame so there is no reason to assume those cases would be treated differently from similar situations involving manually operated weapons.

One possibility is that an AWS might simply malfunction, experiencing an unpredictable or accidental software or hardware failure such that its targeting system functions abnormally in directing the weapon at a protected target. It is not apparent that this failure would have any legal implications beyond those which would attach to an equivalent software or hardware error in a manual weapon system.

The other possibility is that an AWS might 'deliberately' fire upon a protected target. That is, the AWS's weapon might be directed against a protected target as a result of the AWS targeting code running without error, although not according to its operator's intention. There appear to be three possible causes. First, it may occur if the AWS

[388] For examples and an assessment of the arguments see Krishnan, above n 82, 98.

performs a proportionality calculation, and the result of that calculation is that firing in the 'knowledge' that a civilian will be harmed is justified in the circumstances. Civilian harm attributable to a proportionality calculation performed by an AWS control system does not appear to differ legally from that attributable to a proportionality assessment done by a human combatant. The code which is run to perform that proportionality calculation is, after all, an expression of a human decision process.

Second, an AWS may be maliciously programmed to fire upon a civilian, either by its original developer, an enemy who manages to compromise the targeting system, or some other party. The legal characterisation of that act would, again, not differ from the same act performed against any other weapon system which relies on a guidance system. Depending on the identity of the responsible party it may be an act of sabotage, an indiscriminate attack as per API Article 51(4)(a), or some other failure.

Third, an AWS may be unintentionally misconfigured, either by the developer or by other personnel configuring the weapon system prior to use, such that its targeting system identifies a civilian as a valid target. This type of error is not inherent to AWS, but is properly seen as human error.[389] The failure may be attributed to the person who configured the weapon system, or to a flaw in the weapon review process.

In summary, any failure which might cause an AWS to fire upon a civilian target appears to be equivalent to a failure which can occur in relation to any other type of weapon system, and there does not appear to be any reason to treat the AWS failure differently. Nor are there any types of failure that are unique to AWS or that cannot be addressed by law applicable to weapon systems in general. Malicious acts by any party (a weapon developer, a weapon operator, an enemy saboteur, or any other) do not change their legal character simply by being directed against an AWS as opposed to another type of weapon. Civilian harm which cannot be attributed to a deliberate act by any person amounts to collateral damage.

[389] For a sense of the challenges that would have to be addressed in programming an AWS to target only combatants in complex scenarios, see Jann K Kleffner, 'From 'Belligerents' to 'Fighters' and Civilians Directly Participating in Hostilities – On the Principle of Distinction in Non-International Armed Conflicts One Hundred Years After the Second Hague Peace Conference' (2007) 54 *Netherlands International Law Review* 315.

5.4.3 Superfluous Injury and Unnecessary Suffering

5.4.3.1 Content of the Principle

Article 35(2) of API essentially restates the customary principle:

> It is prohibited to employ weapons, projectiles and material and methods of warfare of a nature to cause superfluous injury or unnecessary suffering.

The principle relates to the physical effects that a weapon has upon its human target, 'injury' referring to the objective, physically observable effects and 'suffering' referring to the subjective psychological effects felt by the target. The intention of the principle is to help ensure that the injury and suffering inflicted on combatants is not more than is necessary to render those combatants *hors de combat* in the circumstances of the attack in which the injury or suffering occurs: 'The object of combat is to disarm the enemy. Therefore it is prohibited to use any means or methods which exceed what is necessary for rendering the enemy *hors de combat*.'[390] Past applications have been to weapons with specific harmful physical effects deemed to go beyond what was militarily necessary:

> The API Commentary concedes that specific applications of the principle have not been numerous. It cites explosive bullets, projectiles filled with glass, 'dum-dum' bullets, bullets of irregular shape or with a hollowed-out nose, poison or poisoned weapons including substances intended to aggravate a wound, asphyxiating gases, and bayonets and lances with serrated or barbed heads.[391]

5.4.3.2 Application to AWS

This principle is understood to place limitations on the nature of what would constitute the weapon sub-system of an AWS, without reference to how that weapon is controlled. It may therefore appear that the principle is of only peripheral relevance to AWS. Indeed, it is unlikely to present much of a practical barrier to AWS development, given that any AWS-mounted weapon which would be prohibited under this principle would already be prohibited independently of AWS involvement.

There is one extra consideration, though. This principle relates to the nature of the weapon, and the legal effect of encoding control instructions into an AWS control

[390] de Preux, above n 2, 400 [1411].
[391] Boothby, above n 64, 60 [citation omitted].

system is to make those instructions part of the weapon's intrinsic behaviour. It is theoretically possible that a control system could cause an otherwise compliant weapon to operate in such a manner as to inflict superfluous injury or unnecessary suffering on a target (however unlikely that might be in practice; it is difficult to envisage any military utility in programming an AWS in such a way). The principle forbidding superfluous injury or unnecessary suffering should therefore be seen as a constraint which must be satisfied when programming an AWS control system: any attached weapon must be operated in such a way as to avoid causing superfluous injury or unnecessary suffering to combatants targeted by the AWS.

5.4.4 Summary

No compelling argument has been presented to support the suggestion that there is some threshold level of autonomous capability beyond which an AWS would be inherently illegal under IHL. Neither is there at present any law specifically directed at AWS. The legality *per se* of any particular AWS must therefore be assessed according to the established principles of distinction and the prevention of superfluous injury and unnecessary suffering.

The novel challenge in applying existing law to AWS is to understand the legal effect of encoding some control activities, such as selecting a target or verifying its status, into an autonomous control system. Those activities are normally understood to be done by combatants and their performance must comply with the rules of targeting law. When done by an AWS, they become part of the intrinsic behaviour of the weapon system and so must comply instead with the rules of weapons law. The targeting law obligations that formerly applied are not transferred to the weapon system; the targeting activities programmed into the weapon system 'switch' legal regimes. The corresponding implications for the content of targeting law obligations are discussed in the next chapter.

An AWS complies with the principle of distinction if, in the circumstances of the attack, the combined effect of its ability to identify targets, via its targeting system, and its ability to strike only the chosen target, enable it to attack specific objectives.

An AWS complies with the principle of avoiding superfluous injury and unnecessary suffering if the weapon it employs does not inherently violate that principle, and the control system does not operate the weapon in a way which would violate the principle.

Chapter 6 Targeting Law

This chapter discusses the restrictions imposed by IHL on the uses to which AWS may be put in armed conflict, and the precautions that must be taken. The chapter is in three sections. The first section explains the activities that define an 'attack' when the weapon used is operating autonomously. The rules of targeting law bind the actions of combatants and their commanders, and attacks are the actions which those rules regulate. When the attack is conducted by a weapon which operates with reduced or no human involvement, the definition of 'attack' must be understood accordingly. The second section steps through a sample targeting process for combatants using AWS to launch an attack. The final section discusses the notion of 'meaningful human control', which has been promoted in the CCW discussions as a guiding principle for design of AWS and interpretation of the law governing their use.

6.1 Autonomous Attacks

The previous chapter made repeated reference to attacks conducted with AWS. Where weaponry is concerned, attacks serve as the nexus between the practice of warfighting and the law that regulates it. They are 'a unit of legal management and tactical action',[392] a means rather than an end; the physical activity of an attack being the means by which military ends are achieved, and the corresponding legal concept being the means by which the humanitarian ends of IHL are achieved. The fundamental principles of IHL regarding distinction and proportionality are embodied in rules governing the conduct of attacks.

By attaching specific legal obligations to the steps taken in planning, preparing for and conducting attacks, IHL achieves its humanitarian aim of limiting the effects of armed conflict. Parties to an armed conflict, when undertaking military operations that constitute attacks under the law governing the conflict, must satisfy a raft of obligations with respect to how the attack is planned, against whom or what it is directed, and how it is conducted. Legal norms, both conventional and customary, oblige an attacking party to take adequate precautions in planning the attack, to ensure that the target of the attack is a military objective, to ensure that the means and methods employed in the

[392] Richard Moyes (Notes, CCW Meeting of Experts on LAWS: Towards a Working Definition of LAWS, April 2016) 1.

attack are permissible, and to ensure that incidental damage to civilians and civilian objects is minimised and not disproportionate to the military advantage expected. An attacker who fails to fulfil those obligations when conducting an attack breaches the law. Operations not amounting to attacks, on the other hand, do not attract such onerous obligations.[393]

Due to the central role of the attack in regulating armed conflict, it is essential that all parties understand the conditions under which an operation amounts to an attack, and the conduct required in relation to the attack. Further, when a significant change is expected in either the nature of weapons to be used in attacks or the ways in which combatants might operate those weapons, it is essential to check those new developments against the law governing attacks to ensure the efficacy of the law is maintained.

The actions amounting to an attack may seem intuitively obvious, but development of increasingly autonomous weapon systems will introduce some complexities which necessitate re-examination of the basic question of what constitutes an attack for the purposes of IHL.

The main issue stems from one implicit intention, and likely consequence, of increasing levels of weapon system autonomy: the distancing of human beings from direct involvement, both physically and causally, in specific acts of violence. That distancing is likely to manifest in at least three ways:

- The possibility of extended deployments, perhaps with limited or no opportunity for human intervention, during the course of which the circumstances in which attacks may be launched could be expected to change. This might include deployments in which it is not known in advance whether any targets might be attacked at all, such as where AWS are used in a defensive capacity.
- The possibility of multiple dissimilar, perhaps unexpected, engagements with adversaries during a single deployment.

[393] For a fuller discussion see, eg, Michael N Schmitt, 'Fault Lines in the Law of Attack' in Susan C Breau and Agnieszka Jachec-Neale (eds), *Testing the Boundaries of International Humanitarian Law* (British Institute of International and Comparative Law, 2006) 277.

- The possibility of direct communication between multiple AWS, or between multiple components of a distributed or networked AWS, rather than information being passed via a human decision-maker.

Distancing humans from combat by enabling weapon systems to select and engage targets without human involvement risks blurring the boundaries of what current law considers an individual attack. That is potentially problematic because of the legal obligations which attach to individual attacks.

One of the primary motivations for States to increase autonomy in weapon systems is to take advantage of their 'potentially unlimited persistent capabilities'.[394] A State may, for example, seek to integrate ISR and strike capabilities on a single platform[395] that can loiter for extended periods in an area of air, sea, land or cyberspace close to enemy operations, responding to threats as they develop without needing extensive, or perhaps any, interaction with human operators. Such a platform would afford the possibility to reduce response times, personnel requirements, costs (potentially) and increase personnel safety. It is reasonable to assume that, as the technology of autonomous systems progresses, developers and operators of AWS will seek to utilise new systems in such operations.

API Article 57(2)(a)(iii) requires a proportionality assessment to be completed in relation to each attack, and that obligation is borne by attack planners and other personnel. Consider how the principle of proportionality might apply to an AWS that is deployed for an extended period, with little or no opportunity for human interaction, such as an autonomous UAV sent to loiter in enemy airspace, or an autonomous UUV set to patrol a length of coastline that is under enemy control. It might not be known in advance whether any enemy military targets will be encountered or, if any are encountered, whether circumstances will favour attacking them. Should each deployment be considered an attack, regardless of its duration or scope, with a proportionality assessment done only at the outset, based on whatever information might be available so far in advance of possible encounters with the enemy? Would that rob the law of its ability to ensure that operations against the enemy effectively balance

[394] Defense Science Board, above n 222, 1.
[395] Ibid 15.

military necessity with humanitarian considerations on a regular basis? Would it risk marginalising the roles of the human bearers of the legal obligations which maintain that balance? Perhaps each separate act of violence during an extended deployment should be considered a separate attack. Would that mean the AWS is the entity to 'plan or decide upon an attack'? If so, how could that be reconciled with the fact that obligations in respect of the attack are to be borne by humans? Would it indicate a requirement that a proportionality assessment must be done in respect of each encounter with an adversary? If so, would that represent a requirement for some minimum level of human intervention, or is it a technical requirement that the AWS must be able to perform its own proportionality assessment?

This section first reviews the legal obligations attached to planning and conducting an attack, and the current understanding of what constitutes an attack under IHL, independently of AWS considerations. It then describes how that law may apply to weapon systems with a capacity for autonomous operation.

6.1.1 Preparing for an Attack

6.1.1.1 Substantive Obligations

The principle of distinction states that parties to the conflict must at all times distinguish between civilians and combatants, and between civilian objects and military objectives. This principle is expressed in several articles of API.[396] Attacks may only be directed against combatants and military objectives, never against civilians or civilian objects (except where a civilian is directly participating in the hostilities).

The principle of proportionality requires that the expected incidental damage to civilians or civilian objects (based on information reasonably available to the personnel planning and conducting the attack) must not be excessive compared to the 'concrete and direct' military advantage anticipated from successfully conducting the attack.[397]

The obligations to ensure that the means and methods employed in the attack are permissible and, more generally, to take adequate precautions in planning the attack, describe the process by which obligations relating to the principles of distinction and proportionality are to be met. This process is expressed in Article 57, although it has

[396] *API* arts 48, 51(2) and 52(2).
[397] *API* arts 51(5)(b) and 57(2)(a)(iii).

been noted that 'to some extent Article 57 reaffirms rules which are already contained explicitly or implicitly in other articles',[398] generally those describing the principles of proportionality and distinction as discussed above. Broadly, Article 57 states that only lawful (military) targets may be made the object of attack, and that specified precautions must be taken in attacking those targets.[399]

6.1.1.2 Standard of Precautions

Many attack-related obligations, notably those in Article 57 of API, are expressed in terms of feasibility: attackers must 'do everything feasible' to verify that targets to be attacked are military objectives,[400] must 'take all feasible precautions' in choosing the means and methods of attack to minimise incidental harm to civilians and civilian objects,[401] and so on.

'Feasibility' functions as both an upper and lower limit on the standard of precautions to be taken. Participants in the 1974-77 Diplomatic Conference explained that the notion is to be taken to mean 'that which is practical or practically possible taking into account all the circumstances at the time, including those relevant to the success of military operations'.[402] That is, the law accepts that some civilian harm will be done during conflict.[403] The purpose of the mandated precautions is to maintain, at the level of each individual attack, the balance between humanity and military necessity that is the objective of IHL. The precise interpretation of 'feasible' necessary to maintain that balance in particular circumstances is 'a matter of common sense and good faith.'[404]

Importantly, it was remarked during negotiation of API that required standards to be applied to the identification of objectives 'depended to a large extent on the technical means of detection available to the belligerents. … For example, some belligerents might have information owing to a modern reconnaissance device, while other belligerents might not have this type of equipment.'[405] Attackers are obliged to make

[398] Pilloud and Pictet, above n 67, 679 [2189].
[399] For a fuller discussion see Henderson, above n 9, ch 7.
[400] *API* art 57(2)(a)(i).
[401] Ibid art 57(2)(a)(ii).
[402] Bothe, Partsch and Solf, above n 68, 405 n 8; Pilloud and Pictet, above n 67, 682 [2198] n 7.
[403] Of course, there are strict limits on the acceptable degree of civilian harm. If, after taking all feasible precautions to minimise it, the expected civilian harm from an attack is deemed disproportionate to the anticipated military advantage, the attack cannot proceed. That is the principle of proportionality.
[404] Pilloud and Pictet, above n 67, 682 [2198].
[405] Ibid 682 [2199].

the best use of the technology and information that is available to them at the time of the attack.[406]

6.1.1.3 Bearers of Obligations

In general, obligations to take precautions in relation to attacks are borne by States, as the parties to the conflict, signatories to API and other conventions, and persons under international law. On the other hand, API assigns the obligation to take precautions in relation to an attack to 'those who plan or decide upon an attack',[407] which is a reference to individuals, specifically military personnel. During the Diplomatic Conference some delegations sought to confine this obligation to senior ranks[408] but no such restriction exists in the final text of API. It is true that '[i]n conventional warfare involving large forces of combined arms these functions are generally performed at higher levels of command',[409] but that does not necessarily apply in all forms of warfare; fighting by guerrilla groups would be one exception. The more natural reading is that 'the obligations apply at whatever level the regulated functions are being performed.'[410] However, there is some support for the view that the 'higher the level [of command] the stricter the required compliance is.'[411]

6.1.2 How is an 'Attack' Defined?

Surprisingly under-explored for such a fundamental concept, the formal definition of 'attack' as it is understood in modern IHL can be traced back only to the mid-1950s.

The Hague Regulations of 1899[412] and 1907[413] set out rules regulating the means and methods of warfare as they were understood at the time, two of which explicitly refer to attacks: Article 25 prohibits '[t]he attack or bombardment, by whatever means, of towns, villages, dwellings, or buildings which are undefended' and Article 26 requires that '[t]he officer in command of an attacking force must, before commencing a

---

[406] Henderson, above n 9, 162 n 27.
[407] *API* art 57(2)(a).
[408] e.g. Switzerland, Austria, New Zealand: Henderson, above n 9, 159-60.
[409] Bothe, Partsch and Solf, above n 68, 405.
[410] Alexandra Boivin, 'The Legal Regime Applicable to Targeting Military Objectives in the Context of Contemporary Warfare' (Research Paper Series No 2/2006, University Centre for International Humanitarian Law, 2006) i.
[411] Jean-Marie Henckaerts and Louise Doswald-Beck, *Customary International Humanitarian Law* (Cambridge University Press, 2005) vol 2, 359 (square brackets in original).
[412] *Hague Convention 1899* annex s II.
[413] *Hague Convention 1907* annex s II.

bombardment, except in cases of assault, do all in his power to warn the authorities.' The regulations do not formally define an 'attack', but make it clear that the word refers to violent actions directed against an adversary. Enormous developments in the means of warfare[414] over the subsequent decades jeopardised observance of the principles set out in those treaties by changing the military practices of States, and resulted in horrendous casualties in conflicts through the twentieth century. At the 1949 Diplomatic Conference leading to formulation of the Fourth Geneva Convention,[415] the Russian delegation attempted to introduce a Draft Resolution prohibiting various weapons, but the Conference resolved that it did not have the authority to set rules relating to 'the means of war, the methods of warfare, or the weapons of war',[416] preferring to leave that task to the United Nations. Consequently, none of the four Geneva Conventions of 1949 formally define an 'attack', although the term is used repeatedly in a sense substantially identical to its use in the Hague Regulations.

In 1956, the International Committee of the Red Cross produced its 'Draft Rules for the Limitation of the Dangers Incurred by the Civilian Population in Time of War'[417] as an attempt to address the problematic disparity between the extant formal rules governing means and methods of warfare, and State practice. Article 3 of those Draft Rules states that

> [t]he present rules shall apply to acts of violence committed against the adverse Party by force of arms, whether in defence or offence. Such acts shall be referred to hereafter as "attacks".

That is the first formal definition of an 'attack' to be found in modern IHL. The Draft Rules were approved by the 19th International Conference of the Red Cross in New Delhi in 1957 but were not subsequently acted on by governments. The definition of 'attack', however, reappeared in substantially the same form in the negotiations leading

---

[414] Principally developments relating to missile technology, artillery and military airpower.

[415] *Geneva Convention Relative to the Protection of Civilian Persons in Time of War*, opened for signature 12 August 1949, 75 UNTS 287 (entered into force 21 October 1950) ('*GCIV*').

[416] *Final Record of the Diplomatic Conference Convened by the Swiss Federal Council for the Establishment of International Conventions for the Protection of War Victims and Held at Geneva from April 21*st *to August 12*th *, 1949* (Federal Political Department, Berne) vol IIB, 498.

[417] *Draft Rules for the Limitation of the Dangers Incurred by the Civilian Population in Time of War* (International Committee of the Red Cross, 1956) <https://www.icrc.org/ihl/INTRO/420?OpenDocument>.

to formulation of the 1977 Additional Protocols to the Geneva Conventions. It appears as Article 44(2) of the Draft Additional Protocols to the Geneva Conventions of 1949:

> These provisions apply to acts of violence committed against the adversary, whether in defence or offence. Such acts are referred to hereafter as "attacks"[418]

and in the proceedings of each part of the 'Conference of Government Experts on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts' held in Geneva in 1971 and 1972:

> This word "attack" is used here in its purely military and technical sense; it means acts of violence perpetrated against the adversary, either defensively or offensively, whatever may be the means or arms employed.[419]

> … the ICRC expert specified that the concept of attack should be understood here in a military and technical sense and not in a politico-legal sense; …[420]

Following those conferences, the conventional definition in general use today appeared, with relatively little controversy, as Article 49(1) of API:

> "Attacks" means acts of violence against the adversary, whether in offence or in defence.

Each of the three parts of the API definition of attack is worth examining in more detail. While the definition still underlies relevant legal practice today, it will be seen that fulfilling the object and purpose of API has required some reinterpretation and extension of the original meanings of its terms.

6.1.2.1 'Acts of Violence'

'Violence' is used in this definition in a physical sense;[421] it refers to a force or influence of a type which can cause injury or death to individuals or damage to objects. Thus, operation of kinetic weapons such as guns and missiles during a military

---

[418] International Committee of the Red Cross, *Draft Additional Protocols to the Geneva Conventions of August 12, 1949 – Commentary* (1973) 54.
[419] International Committee of the Red Cross, *Conference of Government Experts on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts, 24 May -12 June, 1971* (1971) vol III: Protection of the Civilian Population Against Dangers of Hostilities, 21 n 15.
[420] International Committee of the Red Cross, *Conference of Government Experts on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts, Second Session, 3 May -3 June 1972 – Report on the Work of the Conference* (1972) vol I, 148 [3.147].
[421] Bothe, Partsch and Solf, above n 68, 329.

operation constitutes an attack in the right circumstances, as does use of chemical, biological or radiological agents.[422] Operations such as embargoes or dissemination of propaganda, although they may contribute in a general way to inflicting harm on an adversary, do not directly cause physical damage and so do not qualify as attacks in themselves.

Where violence follows immediately from some action of a combatant, such as in the case of attacks conducted by firing guns or launching missiles, it is straightforward to identify the operation as an attack. Undoubtedly, that is the scenario envisaged by the parties involved in negotiations leading to promulgation of API. However, subsequent developments in the means and methods of warfare have necessitated expanding the scope of violence to be considered. In particular, while there must be a causal relationship between the 'act' and the resulting 'violence', the violence need not be part of, or even an immediate consequence of, the act. In some cases, violence may be a second- or higher-order effect; such situations are of particular interest in relation to cyber-attacks where the structure of target systems may require that an attack be conducted in multiple stages and over a protracted period of time:

> The crux of the notion lies in the effects that are caused. Restated, the consequences of an operation, not its nature, are what generally determine the scope of the term 'attack'; 'violence' must be considered in the sense of violent consequences and is not limited to violent acts. For instance, a cyber operation that alters the running of a SCADA [Supervisory Control and Data Acquisition] system controlling an electrical grid and results in a fire qualifies. Since the consequences are destructive, the operation is an attack.[423]

The extension appears to be necessary in order to fulfil the object and purpose of API; the original wording agreed to in Article 49(1) does not anticipate development of technologies which allow combatants to achieve the ends characteristic of an attack while separating themselves from direct commission of what would have been considered an 'act of violence'.

---

[422] *Prosecutor v Tadić (Decision on the Defence Motion for Interlocutory Appeal on Jurisdiction)* (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-94-1, 2 October 1995) [120], [124] ('*Tadić*').

[423] Michael N Schmitt (ed), NATO Cooperative Cyber Defence Centre of Excellence, *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (Cambridge University Press, 2017) 415-6 [3] (rule 92).

It is also noteworthy that the intended violent outcome that would mark an operation as an attack 'is not limited to effects on the targeted cyber system. Rather, it encompasses any reasonably foreseeable consequential damage, destruction, injury, or death.'[424] In other words, unintended but foreseeable consequential harm to civilians or civilian objects would be counted as collateral damage.

Physical damage need not eventuate in order for an operation to qualify as an attack. An act of violence that is defeated by some intervening factor such as the target's own defences, or fails for any other reason, is still an attack for the purposes of IHL.[425] The determining factor is the intended violent outcome, whether that is an intrinsic part of the act or a consequence to follow later.

6.1.2.1.1 Scale

Although the scale of action necessary to constitute an attack is not clearly specified, the Commentary to the Draft Additional Protocols notes that the word is meant in an atomic sense: 'it is related to only one specific military operation, limited in space and time.'[426] In this context, military operations are 'the movements, manoeuvres and actions of any sort, carried out by the armed forces with a view to combat.'[427] In relation to AWS, it is significant that an individual attack is currently considered to coincide with an individual military operation; an attack 'refers simply to the use of armed force to carry out a military operation at the beginning or during the course of armed conflict.'[428] Likewise: 'The author of an attack is he who, whatever his position may be at the outbreak of hostilities, starts a military operation involving the use of arms.'[429]

The 'acts' in 'acts of violence' are to be read as acts attributable to a party to the conflict (a State, via its armed forces, or an organised armed group), each of which may comprise one or more violent acts by members of that party's armed forces. Textual clues in API such as the obligations placed by Article 57(2)(a) on '[t]hose who plan or

---

[424] Ibid 107 [5].
[425] Ibid 110 [15].
[426] International Committee of the Red Cross, *Draft Additional Protocols to the Geneva Conventions of August 12, 1949 – Commentary*, above n 418, 54.
[427] Bruno Zimmermann, 'Protocol I – Article 3 – Beginning and End of Application' in Sandoz, Swinarski and Zimmermann, above n 2, 65, 67 [152]
[428] Claud Pilloud and Jean Pictet, 'Protocol I – Article 49 – Definition of Attacks and Scope of Application' in Sandoz, Swinarski and Zimmermann, above n 2, 601, 603 [1882].
[429] International Committee of the Red Cross, *Draft Additional Protocols to the Geneva Conventions of August 12, 1949 – Commentary*, above n 418, 54.

decide upon an attack' also indicate that some degree of planning and co-ordination by the military hierarchy is required. It has been stated elsewhere that the definition does not exclude 'acts of violence by an individual combatant such as a sniper acting alone, or a single bomber aircraft',[430] but that does not mean that each violent act by an individual combatant would normally constitute an attack in its own right;[431] rather, a set of discrete acts performed in the context of one operation, in pursuit of an objective set by a party to the conflict, constitutes an attack.

Each attack may be linked to one or more objectives, as suggested by the use of plural 'objectives' in API Article 57(2)(a)(i),[432] provided that the set of objectives is pursued 'by a specific military formation engaged in a specific military operation'.[433]

That raises the question of whether there might be an upper limit to the scale of a single attack. For example, could all of the combat action involved in the capture of a city, or a larger region, under the control of the adversary be considered to be one attack for the purposes of API? Commentary to the effect that an attack relates to 'one specific military operation' suggests that an upper limit exists, but nowhere is any formal, absolute limit defined. One can, however, set a *de facto* upper limit for an attack which complies with API by reference to the corresponding legal obligations.

An attack is the military action in respect of which proportionality assessments must be made and other precautions defined in Article 57 must be taken. If those obligations are not fulfilled, the attack cannot proceed in compliance with API. Of course, the ability to assess proportionality and take other precautions is highly situational, dependent on the technologies and other resources available to the attacker as well as the actions of the adversary and other circumstances of the attack. Hypothetically, if a sophisticated attacker is capable of anticipating the combat action involved in a large scale operation in sufficient detail to know the military or civilian status of all possible targets, and support an accurate proportionality assessment, there does not appear to be any *prima facie* barrier to considering the entire operation as one attack for the purposes of API.

---

[430] Bothe, Partsch and Solf, above n 68, 329.
[431] William J Fenrick, 'The Rule of Proportionality and Protocol I in Conventional Warfare' (1982) 98 *Military Law Review* 91, 101–2.
[432] Ibid 102.
[433] Bothe, Partsch and Solf, above n 68, 329.

6.1.2.1.2 Severity

The severity of intended harm needed to qualify an act as an attack is not explicitly specified in API. Given the means and methods of warfare available to armed forces at the time it is unlikely that there was much disagreement about the nature of the harm that was to be subject to regulation. Several articles indicate that the focus is on relatively severe harm, referring to 'loss of civilian life',[434] 'severe losses among the civilian population',[435] 'widespread, long-term and severe damage to the natural environment'[436] and so on. Other articles use less specific terms, such as 'injury to civilians, damage to civilian objects.'[437]

It may be inferred that there is a lower threshold of intended physical harm, below which an operation does not rise to the level of an attack, but that lower threshold is not clearly defined. The question is particularly pertinent in relation to cyber-attacks, wherein the direct harmful effect is less likely to be of a traditional type (kinetic, chemical, etc.) and more likely to involve destruction of data or equipment, which is expected to lead to further harm of another type. It is generally agreed that '*de minimis* damage or destruction does not meet the threshold of harm'[438] required, while death or serious injury of human beings, destruction of buildings and similar degrees of harm do qualify, but the precise location of the lower threshold is not clearly stated. A reasonably conservative view, and one that suits the purposes of this thesis, is that any operation which, if successful, may be reasonably expected to result in death or serious injury to at least one person or destruction or serious damage to at least one target, or equivalent harm, should be considered an 'attack'.

6.1.2.2 'Against the Adversary'

This was the only part of the definition to provoke significant discussion in the Drafting Committee. Some participants felt that restricting attacks to violence against the adverse party was at odds with the purpose of this part of API, being to protect civilians from harm resulting from attacks, including the civilian population of the party launching the

---

[434] *API* art 57.
[435] Ibid art 56(1).
[436] Ibid art 35(3).
[437] Ibid art 51(5)(b).
[438] Schmitt, above n 423, 416 [4] (rule 92).

attack.[439] Although the words were retained, other articles in Part IV of API clarify protection for the civilian population.

Of course, it has become common practice among States, judicial bodies and practitioners to rely on the plain meaning of the word 'attack', in referring to violence directed against civilians, civilian targets and so forth, and consideration of whether an attack has been directed against a legitimate target is common in legal discussions. One could certainly argue that the proviso 'against the adversary' in the definition of attack undermines such consideration: if an action is not directed against the adversary, is it not an attack?

Insofar as 'attack' is the legal concept to which proportionality and superfluous injury / unnecessary suffering considerations relate, it seems correct to omit consideration of attacks against civilians, as those concepts are at issue only where the target of the operation is one permitted by law. Nevertheless, widespread practice relies on a notion of attack that contemplates the possibility that it may be directed against targets other than the adversary.[440]

6.1.2.3 'Whether in Offence or in Defence'

This phrase clarifies that the definition refers to attack as 'use of armed force'[441] or 'combat action',[442] without reference to aggression, the first use of force, or responsibility for initiating a conflict;[443] violent actions intended to repel an aggressor are as much attacks as the aggressive act which prompted them. It has been widely noted that in this sense the definition of 'attack' in API diverges from the generally accepted dictionary definition of the word, and from the usage of the word in the military manuals of many States,[444] but there is little suggestion that such divergence is a source of difficulty.[445]

---

[439] Pilloud and Pictet, above n 428, 602 [1877]; 'Summary Record of the Eleventh Meeting (CDDH/III/SR.11)' in *Official Records of the Diplomatic Conference on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts Geneva (1974–1977)* (1978) vol 14, 85–6.
[440] Schmitt, above n 423, 417 [7] (rule 92).
[441] Pilloud and Pictet, above n 428, 603 [1882].
[442] Ibid 603 [1880].
[443] Ibid 603 [1882].
[444] See, eg, Ibid 603 [1879]–[1880].
[445] But see W Hays Parks, 'Air War and the Law of War' (1990) 32 *Air Force Law Review* 1, 115–6.

6.1.2.4 Application to Delayed-Action Weapons

The negotiations leading to the wording of the definition in API were relatively uncontroversial and occasioned little discussion, with only inclusion of the term 'against the adversary' provoking any significant disagreement at the time. Since finalisation of the treaty though, it has become apparent that the definition is not so easy to apply to attacks with delayed-action weapons; that is, weapons which may be activated some significant time after the last human involvement. Consideration of these questions has led to some extension of the seemingly simple API definition.

6.1.2.4.1 Mines

In 1982, the International Society of Military Law and the Laws of War, at its International Congress in Lausanne, sought the views of delegates from various nations on a range of questions that had been raised about implementation of the new API. One of those questions was 'What is, in your armed forces, the usual meaning of the term "attack"?'[446] There was no great disparity between the views of the delegates, with all giving answers similar to the definition in API, albeit with an emphasis on aggressive action. The response of the United Kingdom representative is typical:

> To attack in common military parlance means, to take offensive military action, but it can include opening fire from a defensive position. In this sense the definition in Protocol I does not appear to cause any difficulty.[447]

However, that response also included the following remark:

> Questions have been raised, however, as to what stage in a mine-laying operation amounts to an attack. Is it when the mine is laid, or when it is armed, when a person is endangered by the mine or when it finally explodes? From a purely legal point of view, the answer must be that the attack occurs when a person is immediately endangered by a mine.[448]

This remark is of some help in identifying the point at which an attack occurs when there is a significant space of time between the last act by a member of the attacking force and infliction of violence on the target. Unfortunately, it also leaves some

[446] International Society of Military Law and the Laws of War, *Armed Forces and the Development of the Law of War (Ninth International Congress, Lausanne, 2–6 September 1982)* (1982) 58.
[447] Ibid 207.
[448] Ibid 207–8.

unanswered questions. No further explanation was given as to the precise circumstances in which a person is 'immediately endangered', despite some disagreement expressed by other delegates.[449] Is a person 'immediately endangered' when they reach some threshold probability of suffering violence? When violence becomes inevitable? Is it related to physical or temporal proximity to the mine exploding? The ICRC later reported that the UK view represented the 'general feeling' of the conference,[450] but without elaboration.

Despite that lack of detail, the answer given at the conference is useful for present purposes in that it indicates that the conditions for an operation to constitute an attack could be fulfilled at some time after the last human involvement in operating a weapon; given a list of alternatives that included the option of placing the attack at the time the mine is laid (the last human involvement), the view of the delegates was that it instead amounts to an attack only later when a person is placed at some immediate risk of violence. Notable also is that a mine attack was said to occur prior to actual infliction of violence, which happens only when, and if, the mine finally explodes.

This notion of a mine attack can be reconciled with the API definition by noting that an operation amounts to an attack when acts of violence are *directed at a target*. The target of an attack is distinguished from other persons who may be in the area by being the focal point of the harmful effects of the weapon used in the attack. Prior to a person being immediately endangered by a mine, there is no single person who can reasonably be considered as a focal point. Stated another way, if an attack with an explosive projectile occurs when the projectile moves towards a specific person, an attack with a stationary explosive occurs when a specific person moves towards the explosive. The identity of the target is an essential element in the legal concept of an attack because, prior to an identifiable focal point emerging, it is not possible to assess whether the attacker has met their obligations regarding distinction, proportionality or other rules which may relate to that particular attack.

[449] See, eg, Ibid 341 (India); Elmar Rauch, 'Intervention' (1983) 22 *Military Law and Law of War Review* 291, 293 (Germany).
[450] Pilloud and Pictet, above n 428, 603 [1881].

6.1.2.4.2 Cyber-Attack

Much more recently, the growing threat of armed conflict involving operations conducted via computer systems or networks has prompted further examination of what actions constitute an attack in the sense of API. While the question is still the subject of some debate, a significant step has been taken in clarifying the law relating to cyber-attacks in the form of the Tallinn Manual on the International Law Applicable to Cyber Warfare [451] and its successor, the Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations.[452]

The International Group of Experts who contributed to drafting the Tallinn Manuals drew on the definition of 'attack' in API, and on the dominant view of how that definition applies to mine attacks, and extended those concepts to the novel environment of a computer network operation:

> The 'general feeling' of the [API] negotiators was that 'there is an attack whenever a person is directly endangered by a mine laid'. By analogy, the introduction of malware or production-level defects that are either time-delayed or activate on the occurrence of a particular event is an attack when the intended consequences meet the requisite threshold of harm. For the majority, this is so irrespective of whether they are activated.[453]

The analogy with mine attacks is not perfect. After drawing that analogy the Tallinn Manual 2.0 notes:

> An attack that is successfully intercepted and does not result in actual harm is still an attack under the law of armed conflict. Thus, a cyber operation that has been defeated by passive cyber defences such as firewalls, anti-virus software, and intrusion detection or prevention systems nevertheless still qualifies as an attack if, absent such defences, it would have been likely to cause the requisite consequences.[454]

Firewalls and intrusion prevention systems typically operate at the border of a network. If they successfully defeat an attempt to place malware inside a target system, then the

---

[451] Michael N Schmitt (ed), NATO Cooperative Cyber Defence Centre of Excellence, *Tallinn Manual on the International Law Applicable to Cyber Warfare* (Cambridge University Press, 2013).
[452] Schmitt, above n 423.
[453] Ibid 419 rule 92 [16] [citation omitted].
[454] Ibid 419 rule 92 [17].

malware is not emplaced, which would be analogous to a failed attempt to lay a mine. The fact that the Tallinn Manual considers such a failed attempt to be an attack suggests that, at least for some cyber-attacks, a more appropriate analogy might be drawn to launching a missile, which constitutes an attack even if the missile is then stopped by a missile defence system.

The Tallinn Manual definition also implicitly assumes that the system into which malware has been introduced is the intended target of the attack. That may not be true if, for example, some further target selection or verification code must run prior to the malware activating. For example, the well-known Stuxnet malware[455] which was used to disable centrifuges at the Natanz nuclear facility in Iran first entered SCADA systems then ran tests to determine whether they were attached to the correct type of centrifuges.[456] Until those tests were run and returned positive results it is not correct to say that the particular SCADA system or centrifuge would be the focal point of an act of violence.

In general, reasoning by analogy between cyberspace and physical space is a hazardous approach; they are highly dissimilar environments in which entities behave in highly dissimilar ways. It is suggested that for legal purposes, one should first seek to identify the elements of an attack (specifically, an attacker, a weapon, and a target) and then, if needed, choose an appropriate analogy to physical space. For example, if malware is to be delivered by embedding it in a web page which people will visit, an attack occurs when a person in fact loads that web page into their browser, thereby marking themselves as the focal point of harm for that malware. If malware is to be delivered by uploading it directly to a target's computer through a compromised firewall, the attack occurs when the malware is transmitted from the attacker's computer because all the elements necessary for a legal assessment of the attack are present at that time.

Just as consideration of mine-laying operations forced an extension of the API definition of 'attack', consideration of cyber-attack forces a further extension. The destructive effects of a mine are inherently confined to a limited region about where the mine is emplaced, so a person who is immediately endangered by approaching the mine

[455] Jackson Maogoto, *Technology and the Law on the Use of Force* (Routledge, 2015) 53.
[456] Thomas M Chen and Saeed Abu-Nimeh, 'Lessons from Stuxnet' (2011) 44(4) *Computer* 91.

marks themselves as the target. Malware moving through a network may transit intermediate networks before arriving at an intended target, and it may be some action by an entity in an intermediate network which enables the malware to travel through and reach its target. This does not alter the principle underlying identification of an attack or its target, but it potentially complicates the process of identifying all the actors involved in an attack.

6.1.3 Applying Current Law to AWS

6.1.3.1 Identifying the Attack

The following discussion begins with consideration of an attack involving only a single AWS selecting and firing on a single target, whether it be an autonomous turret firing on an individual in the Korean DMZ,[457] a ship-mounted CIWS firing at an incoming missile, an autonomous underwater vehicle firing at an enemy submarine, or some other form of AWS selecting and engaging a target. The conclusions are then extended to account for an operation involving an AWS attacking multiple targets in separate incidents.

Discussion of how to apply the law of attack to AWS takes place against a backdrop of two somewhat distinct branches applicable to different types of weapons. In the case of weapons with an instantaneous effect, such as missiles or rifles, an attack occurs at the moment of the last human involvement, when the weapon is activated (such as the rifle's trigger being pulled). In the case of delayed action weapons such as mines, an attack occurs at some time after the last human involvement, when a target is 'immediately endangered' by the weapon. It was proposed above that the two branches may be reconciled by referring to the common point at which a specific target emerges. Can that idea be applied to AWS? Or, given that AWS may employ either instantaneous or delayed action weapons, can a useful analogy be drawn to any existing type of weapon? Or should weapon systems with a significant capacity for autonomous operation be treated as a discrete category, with its own rule?

Boothby has addressed this question in relation to sorties flown by autonomous UAVs:

[457] See, eg, Jean Kumagai, *A Robotic Sentry for Korea's Demilitarized Zone* (1 March 2007) IEEE Spectrum <http://spectrum.ieee.org/robotics/military-robots/a-robotic-sentry-for-koreas-demilitarized-zone>.

> The time of the decision and of execution in the case of an autonomous mission would seem to be the time when the mission is launched, because that is the time when human input to the target recognition systems determines what the machine will attack. After all, for these purposes the time of execution of, e.g., a Tomahawk land attack cruise missile attack must be taken to be the time when the missile is launched, not the time when it impacts and detonates. Where autonomous, self-aware systems of the sort described in [JDN 2/11] are concerned, however, the position may be different. If such future systems replicate a human evaluative decision-making process, it would seem logical that the timing of that sort of attack decision must be taken to be the time when the system logic, which after all *ex hypothesi* replicates human decision-making processes, determines that the attack shall proceed. Similarly, the time of execution must be taken to be the time when the weapon is released by the autonomous platform.[458]

In that argument, Boothby maintains the distinction between two classes of weapons, but instead of speaking in terms of instantaneous weapons and delayed action weapons, he expresses the difference in terms of the 'degree' of autonomous capability:

> The distinction here, therefore, is between what can be described as 'simple autonomy', that is the mechanical implementation of a decision-making process pre-ordained by the personnel initiating the mission, and the more futuristic version of autonomy, or 'complex autonomy', namely the mechanical implementation of an attack decision made by mechanical systems that replicate human evaluative decision-making processes.[459]

Basing the distinction on the complexity of the reasoning performed by the control software is problematic. The analysis presented in Chapter 3 showed that autonomy is a matter of degree rather than a capability that exists in discrete levels. In any case, a level of autonomous capability does not change the nature of a weapon system from a piece of military equipment to a legal decision-maker, and care must be taken in saying that the AWS 'determines that the attack shall proceed' in a legal sense. With the many dimensions along which autonomous capabilities may be expected to vary in weapon systems, it appears difficult and somewhat artificial to attempt to cleanly and

[458] William H Boothby, *Conflict Law: The Influence of New Weapons Technology, Human Rights and Emerging Actors* (Asser, 2014) 111.
[459] Ibid 112.

consistently apply different rules to weapon systems depending on the capabilities they are exhibiting at any particular time.

Instead, one should recognise that 'autonomous' is not a precise description of the behaviour of a weapon system, and a capacity for autonomous operation does not intrinsically support drawing analogies with any other type of weapon. Even today, some weapon systems that exhibit limited autonomy can be analogised easily to missiles[460] while others arguably resemble mines.[461] Cyber weapons with a capacity for autonomous operation may differ again. In the future, it is distinctly possible that more advanced AWS may not be clearly analogous to any one existing weapon.

Lacking an analogy that can be applied to all AWS for a legal analysis, one should seek to identify the elements of an attack in an autonomous operation. As with operations involving other types of weapons, an operation constitutes an attack when all the required elements are present. In particular, when an act of violence is committed against a target. Given that neither API nor any other existing law employs different definitions of attacks for different weapons, it is necessary to place the attack at some common point in respect of weapons that are operated differently. That point is, logically, when all the legal elements are satisfied.

If the time at which an operation amounts to an attack is to differ between AWS and other types of weapons, such as where it lies in relation to the last human involvement, the difference will relate to the functional differences between AWS and other weapons. The principal functional difference that is relevant to the identification of an attack is that some targeting functions are encoded into the AWS control system. As with other weapons, the operation amounts to an attack once a target is selected and preparation for the attack begins; the attack is then executed when the AWS control system activates a weapon. That is the moment at which the operation fulfils all the criteria for being an attack; prior to then there has been no act of violence and no target against which such an act could be directed. The challenge is to identify the moment at which those criteria are fulfilled in relation to the last human involvement in the operation. That moment

[460] See, eg, *AMRAAM*, above n 314; *Harpy*, above n 200.
[461] See, eg, Derick S Hartshorn, *Mk 60 (Captor)* (28 November 2009) Mineman Memories <http://www.hartshorn.us/Navy/navy-mines-10.htm>.

will vary according to the nature of the targeting functions encoded into the AWS control system.

The first possibility is that an AWS is deployed with the intention of having it attack a particular target or type of target, and the AWS is programmed with a description of that target sufficiently complete and specific that it must only compare potential targets to that description in an attempt to find a match. In that case the target was effectively chosen in advance by the weapon system operators and all that remains is to locate it. The attack decision was therefore made when the AWS was deployed and the attack will be executed when (and if) a matching target is located. Existing AWS such as Kongsberg's Joint Strike Missile,[462] or ISI's Harpy loitering munition[463] are of this type.

The second (still hypothetical) possibility is that an AWS is deployed with something less than a complete description of a specific target to be attacked. In this case the AWS control system must do some further processing of potential targets that is more than just asking 'Is this what I was told to attack?' One example may be an AWS used in a sentry capacity, deployed with the intention of assessing whatever incoming threats, unknown at the time of deployment, may be detected, and possibly attacking them. Such an AWS might instead be asking 'Does this look or act like something I should attack?' The attack decision would then be placed at the completion of that processing, provided the answer is 'Yes', when a specific target is selected. Prior to that processing being completed by the control system, there is no entity which could be considered a focal point of an act of violence, or against which adherence to principles of distinction or proportionality could be measured. If the AWS operation completes without any specific target having been selected, then the operation did not amount to an attack.

The other consideration is the time of the last possibility for human intervention before an attack is executed. The discussion above assumes that an AWS is deployed and has no further communication with its human operators prior to executing an attack or completing its operation without attacking a target. However, many AWS in use today and, by most reports, those planned for the near future, will either seek confirmation from a human operator once a target is selected and prior to activating a weapon, or will

---

[462] *Joint Strike Missile (JSM) Brochure*, Kongsberg <https://www.kongsberg.com/~/media/KDS/Files/Products/Missiles/jsm_web_reduced.ashx>.
[463] *Harpy*, above n 200.

at least provide an opportunity for a human operator to override a target selection. In that case the attack decision is placed at the time the human operator confirms the target selection, or decides not to override the AWS's choice of target. If the operator does not allow the attack to proceed against the chosen target, the operation does not amount to an attack.

An attack decision cannot be made before there is a chosen target to attack. Whether target selection is done entirely by human combatants with the final selection being programmed into the AWS, or whether the AWS itself performs some part of the selection process, it is at the end of that process when a specific target is selected that the operation amounts to an attack. Selection of a specific target means the AWS has a description sufficiently complete that the only steps remaining are to locate the target matching that description, complete any required precautionary processes and activate the weapon.

It is suggested that a distinction between 'simple' and 'complex' autonomy is unnecessary. In both cases, the attack should be located as described above. In the case of 'simple' autonomy the personnel initiating the mission make all necessary targeting decisions in the traditional way and employ the AWS only to locate the chosen target and execute the attack. With 'complex' autonomy the personnel initiating the mission have encoded some part of the target selection process into the control system[464] which, together with their decision to launch the AWS at a given time and place, constitutes the same targeting decision. It is inappropriate to say that the control software 'makes a targeting decision' or similar. That would be to anthropomorphise the weapon system and view it as a legal decision-making entity. In respect of AWS, targeting decisions are made by those who activate the weapon system in the knowledge that they will fire on a certain (type of) target, although those decisions do not crystallise until the software is run.

It is therefore unnecessary to draw a distinction between attacks with different types of weapons based on the details of the decision-making process employed; an operation amounts to an attack when a target is selected and preparations for activating a weapon

[464] At the least, those personnel may be assumed to be aware that part of the target selection process is so encoded.

begin, and the attack is executed when the weapon is activated. That rule is consistent with the rule applicable to a simple manually operated weapon system, as the distinction is only in whether or not part of the targeting process is done by executing software in the weapon's control system.

It can also be reconciled with the rule that a mine attack occurs when a target is 'immediately endangered' by recognising that firing a rifle or launching a missile at a target does not necessarily cause harm, just as endangering a target does not necessarily cause harm. The shot might miss, the missile guidance system might malfunction, the target might still take defensive or evasive measures. However, in the ordinary course of events the target will suffer the effects of the attack. Directing a shot at a target immediately endangers the target in the same sense that the target of a mine attack is immediately endangered before the mine actually explodes.

Likewise, where a cyber weapon operates autonomously and has some capacity to select its targets, this finding is consistent with the definition of cyber-attack given in the Tallinn Manual. The Tallinn Manual locates an attack at the time that malware is introduced to the target system; if the malware must seek confirmation from a human operator after selecting a target, whether before or after reaching its destination, then that confirmation must also be given in order for the operation to constitute an attack. For example, Stuxnet has been described as having a degree of autonomous capability:

> Considering that there was very good chance that no Internet connectivity would be available (only access to the internal network), Stuxnet developers put all of its logic in the code without the need of any external communication. As such the Stuxnet was an autonomous goal-oriented intelligent piece of software capable of spreading, communicating, targeting and self-updating; …[465]

If, hypothetically, Stuxnet had been released in the course of an armed conflict, it is likely that the type and degree of damage it was intended to inflict, and in fact inflicted, would have been sufficient to classify the operation as a computer network attack.[466] In that case, the attack would be said to have occurred at a moment after Stuxnet had

[465] Stamatis Karnouskos, 'Stuxnet Worm Impact on Industrial Cyber-Physical System Security' (Paper presented at IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society, Melbourne, Victoria, Australia, 7–10 November 2011) 4492.
[466] Michael Schmitt, 'Classification of Cyber Conflict' (2012) 17 *Journal of Conflict and Security Law* 245, 251-2.

entered the system controlling the centrifuges, and had completed all target selection steps such as confirming that the centrifuge control system was of the correct make and model and that the centrifuges were spinning at the appropriate speed to be attacked;[467] at that point the only step remaining was to issue the command that actually caused damage to the centrifuges.

6.1.3.2 Preparatory Obligations Relate to the Combined Actions of Software and Humans

Some of the functions of the software in an AWS control system, such as target selection, replicate tasks that IHL assumes to be done by humans as part of the planning of, or preparation for, an attack. When those functions are instead made part of the behaviour of an AWS, they are no longer done directly by humans in preparation for activating the weapon system, as the software would generally be executed after the last human involvement in conducting the attack. Legally, though, they still form part of the precautionary measures that the law requires. The obligations that relate to preparation for an attack still fall on 'those who plan or decide upon an attack', but would be fulfilled by running that code rather than by actions done directly by humans. The sum of preparations completed by human operators prior to activating the AWS (such as selecting the time and place of activation), plus those encoded in software in the AWS control system, plus those which may be completed by a human operator or supervisor monitoring the AWS after activation, must be sufficient to satisfy all obligations relating to planning and conducting an attack.

6.1.3.3 Extended Operations

Based on the reasoning in Section 6.1.3.1 above, when an AWS is deployed against an adversary and its control system selects a target and activates a weapon, the operation as a whole can be considered, relatively unambiguously, an 'attack' for legal purposes. That is so whether the attack is directed at one or several objectives, but the legal concept of an attack is normally applied to operations that are limited in space and time, directed at a well-defined objective or set of objectives. Is there a limit to how extensive an operation may be treated as one attack? If an autonomous platform loiters in enemy territory for an extended period, possibly engaging multiple targets on multiple distinct

[467] Karnouskos, above n 465, 4490–1.

occasions, is there a point at which the operation must be 'broken up' into multiple attacks for legal purposes? If so, how is the permissible extent of an attack determined in the context of an AWS operation, and what obligations mark the start and finish of each separate attack?

Given that the written law places no explicit limit on the extent of an attack, it might seem unnecessary to discuss demarcating individual attacks within a long-term operation. However, if the legal concept of an 'attack' is to serve its purpose of constraining combatant actions then the concept itself must be suitably constrained in form and extent; if the concept of 'attack' can be stretched to encompass any arbitrary set of actions by a combatant then it loses its ability to exclude actions which would breach other rules of IHL:

> there has to be some spatial, temporal, or conceptual boundaries to an attack if the law is to function. … If 'attacks' were not conceptualised and subject to legal judgement at the tactical level, but only say the broad strategic level, then a large operation may be determined to be permissible (on the basis of broad anticipated outcomes) whilst containing multiple individual actions that would in themselves be legal violations.[468]

It may seem that spatial or temporal boundaries are most readily applicable, but the spatial and temporal extent of an attack is only incidentally related to its legal role, and is not applicable in any meaningful way to attacks conducted in cyberspace rather than physical space. Instead, a conceptual boundary is required. If the concept of attack is the means by which combatant actions are constrained to comply with the rules and principles of IHL, the conceptual boundaries of attacks are best expressed in terms of those rules and principles.

Many rules of IHL may constrain an attacker's actions in the course of an attack. Some are applicable to every objective attacked, the principal example being the obligation to verify that all objectives of the attack are military objectives and not subject to special protection. Other constraints are applicable to the attack as a whole. The most fundamental 'whole-of-attack' constraints applicable to all attacks are:

[468] Article 36, 'Key Elements of Meaningful Human Control' (Background paper to comments prepared by Richard Moyes, Article 36, for the CCW Meeting of Experts on AWS, April 2016) 3.

- The obligation to direct the attack at the military objective or objectives 'the attack on which may be expected to cause the least danger to civilian lives and to civilian objects',[469] given the purpose of the attack.
- The obligation to '[r]efrain from deciding to launch any attack which 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 and direct military advantage anticipated'.[470] This is supported by the subsidiary obligations to accurately assess both the anticipated 'concrete and direct military advantage' and the expected civilian harm.
- The obligation to '[t]ake all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event to minimizing, incidental loss of civilian life, injury to civilians and damage to civilian objects'.[471]

To fulfil its purpose of confining an attacker's actions to those which are legally permissible, the scope of an attack must be limited such that the conduct of each constituent part of an attack cannot deviate too far from conduct that would be acceptable as an attack in its own right. For example, there may be considerable scope for abuse if it were permissible for an attacker to inflict severe harm on a civilian population in pursuit of one military objective on the understanding that the harm would be 'balanced out' by lesser harm in pursuit of later objectives counted as part of the same attack for the purpose of a proportionality assessment.

There is also an implicit, but perhaps more restrictive in practice, requirement that the attacker must be able to acquire sufficient information prior to the attack to select the appropriate weapon and to support the necessary assessments of military advantage and civilian harm if the attack is executed.

As all of these constraints are linked to an attacker's ability to meet some standard of conduct, the precise extent to which an attack must be limited would logically depend

[469] *API* art 57(3).
[470] *API* art 57(2)(a)(iii).
[471] *API* art 57(2)(a)(ii).

on the circumstances of the attack including, most significantly, the technology available to the attacker.[472]

It seems reasonable that this same proviso would apply to other attack preparations that are done by technical means.

The result of that reasoning is that an 'attack' must be limited to the span of activity which satisfies two conditions:

- The attacker is able to complete all 'whole-of-attack' preparatory obligations to a sufficiently high standard. The limiting factors would include the attacker's technical resources, access to information and any other relevant circumstances of the attack.
- The attacker does not have latitude to inflict excessive harm when attacking individual objectives that would undermine the humanitarian purpose of the law's precautionary obligations.

Where the above constraints would restrict the scope of an attack to a span of activity which cannot encompass an entire operation, it would be necessary to divide the operation, conceptually, into multiple attacks. Each of the attacks would encompass a span of activity over which the attacker is able to perform the 'whole-of-attack' preparatory obligations to a satisfactory standard. The attacks within an operation would then be delimited by the AWS control system performing those preparatory tasks (selecting objectives, estimating civilian harm, and so on) with reference to the next attack.

The legal burden of those obligations, like all others, rests on human beings, not AWS. According to Article 57(2)(a) of API, '[t]hose who plan or decide upon an attack' would be held directly responsible. It is safe to say that the drafters of API would have assumed that those attack planners would have made the relevant decisions 'in person', prior to each attack being launched. It has also been argued by various participants in the AWS discussions that direct human intervention is in fact required in the preparation stage for each individual attack. For example:

[472] Pilloud and Pictet, above n 67, 682 [2199]; Michael N Schmitt, 'Precision Attack and International Humanitarian Law' (2005) 87 *International Review of the Red Cross* 445, 460.

> Recognition that human legal engagement must occur over each attack means that a machine cannot proceed from one attack to another, to another, without human legal judgment being applied in each case, and without capacity for the results of that legal judgment to be acted upon in a timely manner – i.e. through some form of control system. Given that an attack is undertaken, in the law, towards a specific military objective that has been subject to human assessment in the circumstances at the time, it follows that a machine cannot set its own military objective without human authorization based on a human legal judgment.[473]

A literal interpretation of the text of API ('decide upon *an* attack') seems to support that argument. However, the same literal interpretation does not exclude the possibility that encoding an attack planning process in software then causing the software to be executed would also constitute a decision. Arguments for human intervention rely on the premise that a programmatic algorithm, a series of decision-making steps encoded in software, necessarily falls short of being an application of human judgement, even where the algorithm might be specifically designed to replicate a human decision-making process. As detailed in Chapter 3, this thesis disputes that premise. A process for assessing the military status of an attack objective, or a process for estimating incidental civilian harm, or any other process derived from human knowledge and experience, is as much an exercise of human engagement when encoded in software as it is when performed 'live' by an attack planner or other qualified person. The necessity for human engagement in preparation for an individual attack is dependent on the technical capabilities of the AWS control system. If an AWS cannot complete some precautionary task unassisted, such as the difficult and highly contextual proportionality assessment, then direct human intervention in preparation of the attack is needed. However, as long as the AWS is capable of performing all preparatory tasks to such a standard that the resulting attack is conducted in compliance with the law, it is difficult to sustain an argument that the rules of IHL nevertheless require a human to personally intervene.

6.1.3.4 Other Precautions in Extended Operations

API Article 57(2)(c) requires that '[e]ffective advance warning shall be given of attacks which may affect the civilian population, unless circumstances do not permit.' The

[473] Article 36, above n 468, 3.

article does not specify how, when, by whom or to whom the warning must be issued. It is reasonable to say all of those elements would depend on the circumstances of the attack; what is required is that the warning be effective in alerting civilians to the impending attack.[474]

Given that the requirement to issue a warning, or not, relates to each attack, it appears that there are two possible approaches for an AWS on an extended operation involving multiple attacks. Either the AWS must somehow afford a human operator the opportunity to issue a warning prior to commencing an attack, or the AWS must have the ability to: assess whether civilians might be affected by the planned attack, assess whether issuing the warning would unduly threaten the success of the attack or of the extended operation, and then issue the warning in an effective manner prior to commencing the attack if circumstances permit.

Once an attack commences, it 'shall be cancelled or suspended if it becomes apparent that the objective is [protected for various reasons].'[475] On a natural reading of that article, the party to whom the listed problems would 'become apparent' would normally be the human combatant executing the attack. It is understood that where the chosen means of combat does not permit a combatant to directly perceive the target, so that problems would not be apparent, this article 'will be likely to have little to no operation vis-à-vis the person executing the attack.'[476]

However, AWS operations might be conducted in the future where there might be a very considerable separation, in time and intervening events, between the last human involvement and execution of an attack by the AWS. In that situation, the risk of circumstances changing in unforeseen ways would be greatly magnified. A programmed ability for the AWS to cancel or suspend its own attack may become necessary in order to manage the risk of civilian harm consistently with the object and purpose of API.[477]

6.1.3.5 Summary

An operation by an AWS amounts to an attack (the attack decision is made) when a specific target is selected and preparations for activating the weapon begin. Depending

[474] Henderson, above n 9, 188.
[475] *API* art 57(2)(b).
[476] Henderson, above n 9, 183.
[477] As it is in relation to attacks conducted with other weapons: Sassòli, above n 300, 320.

on the capabilities of the AWS and the circumstances of the operation, that target selection may be done by a human or another system and communicated to the AWS, which then sets about locating the chosen target, or it may be done after deployment by the AWS control system, based on whatever target selection criteria are supplied to it. The attack is executed when the AWS control system activates a weapon attached to the system.

Not all AWS operations against an adversary can be encompassed by a single attack. In order to play its role of guiding the behaviour of attacking forces, the legal concept of an 'attack' must be suitably constrained. Some precautionary obligations relate to each objective to be attacked (such as verifying military status) while others relate to the attack as a whole (such as assessing military advantage, and estimating incidental civilian harm). The 'whole-of-attack' obligations constrain the extent of what can be treated by the AWS as an individual attack. If an AWS is not capable of performing those tasks in respect of the entire operation it is to undertake, the operation must be notionally treated as multiple attacks, with the 'whole-of-attack' preparatory tasks performed in preparation for each individual attack. Direct human intervention between, or during, attacks is required where it would be beyond the technical capabilities of the AWS to perform those tasks to the requisite standard without human assistance.

6.2 Legal Use of AWS

As weapon system autonomy means some tasks from the targeting process are taken out of human hands, the central question here is about which tasks, if any, must remain with humans for compliance with the law and what degree of knowledge humans must have about the activities of AWS.[478]

The definitive characteristic of the types of AWS of interest in this thesis is that their control systems replicate some part of a human decision-making process in order to control the attached weapon. That part may include decisions about the military or civilian status of a potential target, the anticipated harm to civilians if a target is attacked, and so on. That is the reason that some commentators have characterised AWS in an anthropomorphic way, as something more akin to combatants than weapons. It is

[478] For another study of human decision-making in AWS targeting processes see Jeffrey S Thurnher, 'Examining Autonomous Weapon Systems from a Law of Armed Conflict Perspective' in Hitoshi Nasu and Robert McLaughlin (eds), *New Technologies and the Law of Armed Conflict* (Asser Press, 2014) 223.

not just the possession of decision-like behaviours which raises questions about the legal character of AWS though; it is also the fact that they employ those behaviours after the last human involvement in launching an attack.

Use of other types of systems with sophisticated capabilities such as decision support systems and ISR platforms is not new, and some may exhibit a degree of autonomous capability. However, they do not raise the types of legal questions that AWS raise because their role is causally prior to a responsible human decision in launching an attack. The outcome of their deployment is information which a human decision-maker draws upon in making the legal judgements required of an attack planner, weapon system operator or other role defined in written IHL.

AWS differ by appearing to delay some part of a human's decision until after deployment, which is some time after the last human involvement. As explained above, the decision is not actually delayed in the sense of the AWS becoming a sort of decision-making entity. The AWS is a computer-based machine, a weapon, like any other sophisticated weapon system. Its control system does not make legal decisions, it carries out decisions already made by its human designers and operators.[479]

But what decision does the operator of an AWS, or their commander, make when deploying an AWS? Is it the decision that the law requires them to make, given that some determinations are, by design, left to the AWS control system to make after being deployed? If an attack planner is required to '[d]o everything feasible to verify that the objectives to be attacked are neither civilians nor civilian objects',[480] but some part of that verification is to be done later by the AWS, has the attack planner fulfilled their obligations in activating the AWS? A survey of current practices indicates that there is no need for an attack planner to complete all of the verification process in respect of each individual target in person prior to activating a weapon system. IAI's Harpy, for example, is used by several States to detect and destroy radar-emitting devices, performing part of the target verification process itself; there has been no suggestion that its use in that capacity is illegal. On the other hand, as a thought experiment (albeit an unrealistic one), one could imagine a hypothetical, highly autonomous weapon

[479] Ibid.
[480] *API* art 57(2)(a)(i).

platform intended to be activated with little, or virtually no, knowledge on the part of its human operators as to what actions it might 'decide' to take on the battlefield. Even if that hypothetical AWS does execute its attacks in accordance with the law, is there some minimum threshold of knowledge, some minimum part of the decision-making process, which legally must be completed by a human being prior to activating a weapon? Or could effectively an entire attack planning and execution process legally be performed within an AWS control system?

This is not a comprehensive guide to all legal aspects of targeting with AWS. Given the complexity of modern conflicts, the level of detail required for such an undertaking would go far beyond the scope of this thesis:

> Targeting decisions and the execution of the use of force in a modern conflict can range from factually very simple to extremely complex. At one end of the spectrum, a soldier may observe a person, conclude that the person is an enemy combatant, and shoot at that person. At the other end of the spectrum, a vast amount of information from various sources may be analysed to determine that an enemy combatant will be in a certain building at a certain time. Further work is then undertaken to determine where that building is geographically located. A pilot, who never sees the information that supported the analysis, is tasked to bomb the building. The attack may be conducted using a laser-guided weapon where the laser designation is provided by yet another person. …[481]

The intent here is merely to explain the ways in which a degree of weapon autonomy affects the legal aspects of planning and conducting a generic attack, and the limits which the law of targeting imposes on the use of AWS. The framework used here is the IHL 6-step targeting process described by Henderson,[482] although the discussion is largely applicable to other targeting processes.[483]

Without wishing to engage in speculation about the circumstances in which future AWS may be employed by armed forces, it seems that the legal significance of autonomous behaviour of weapon systems would be greater in dynamic targeting situations than pre-

[481] Henderson, above n 9, 233.
[482] Ibid 234.
[483] See, eg, Mark Roorda, "NATO's Targeting Process: Ensuring Human Control Over and Lawful Use of 'Autonomous' Weapons" (Research Paper No 2015-06, Amsterdam Center for International Law, 2015).

planned attacks employing deliberate targeting. Where much of the planning for an attack is done far in advance with extensive involvement by commanders, legal officers, intelligence personnel and others, there is likely to be less need for advanced AWS to perform extensive target selection and verification on their own. In those situations, the weapon system would need only to accurately engage the previously designated target using the designated method of attack. It would be in dynamic targeting situations involving unplanned or unanticipated targets that the ability of a weapon system to assess the legal status of a target without human involvement would be of most significance.

The first four phases of the IHL targeting process constitute the planning phase of an attack, broadly covering the legal obligations set out in Article 57(2)(a) of API. The legal burden of this phase is borne by 'those who plan or decide upon an attack'.[484] That burden would not change as a result of executing part of this planning phase in an AWS control system. Just as the weapons law obligations described in Chapter 5 did not alter due to extra functions being encoded into the weapon's control system, the legal burden of making targeting decisions does not change either.

The first two steps relate to Article 57(2)(a)(i), according to which attack planners must

> Do everything feasible to verify that the objectives to be attacked are neither civilians nor civilian objects and are not subject to special protection but are military objectives … and that it is not prohibited by the provisions of this Protocol to attack them

The first task required by this article, to locate and observe the potential target and its surrounding area, does not necessarily raise any particular problems in respect of AWS. It essentially means gathering information, and the use of autonomous platforms rather than other means for ISR purposes would raise novel legal questions only if the platform is able to autonomously 'decide' to process or manipulate the raw data from its sensors in a way that might meaningfully affect the outcome of decisions based on that information.

[484] For a general discussion of the importance of precautionary obligations to use of AWS see Jonathan David Herbach, 'Into the Caves of Steel: Precaution, Cognition and Robotic Weapon Systems Under the International Law of Armed Conflict' (2012) 4(3) *Amsterdam Law Forum* 3.

The second step is to apply the legal tests set out in API to that gathered information, in order to decide upon the legal status of the proposed target: does it meet the criteria for being a military objective? If so, is it subject to special protection? Although this amounts to a legal judgement, it does not necessarily have to be done by a lawyer[485] or even, perhaps, directly by a human being. This step essentially means comparing the observed information about the target and its surroundings to other information describing the characteristics of military objectives. There does not appear to be any legal barrier to performing that comparison in software rather than having it done by an attack planner or commander, although there may be very significant practical challenges in making all of the required information, from whatever source, available to an AWS control system.

Attempting to reconcile that with the remainder of the planning phase is illuminating. The third step relates to Article 57(2)(a)(ii):

> Take all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event to minimizing, incidental loss of civilian life, injury to civilians and damage to civilian objects

The fourth step relates to Article 57(2)(a)(iii):

> Refrain from deciding to launch any attack which 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 and direct military advantage anticipated

On the face of it, there might appear to be a logical inconsistency: an AWS must be chosen as the means of attack before questions of using it to identify targets or perform proportionality assessments arise, but examination of the written law suggests that selection of a means of attack must be based on the information gathered in the previous steps. At the very least, a commander must know that a particular AWS would be capable of successfully attacking the chosen objective before selecting it for the task.

This reveals one limitation, albeit a soft limit, on the decisions which can be delegated to an AWS. An AWS selected for use in an attack must be known:

[485] Ibid 234 n 4.

- to be able to identify the chosen objective in its environment and, equally, to be able to refrain from identifying other persons or objects as objectives;
- to be capable of successfully attacking the desired objective, in terms of the weapons it can employ;
- to be the means which can do so with the least expected harm to civilians.

That means some knowledge of the target and its surroundings must be obtained prior to employing a particular AWS in an attack.[486]

For example, the IAI Harpy is capable of operating entirely unassisted after launch. No human being needs to assist the Harpy in identifying enemy radar installations, either by guiding it to the installations initially or by confirming its target selections before it attacks. The decision to launch the Harpy must be based on prior decisions by humans, though: that destruction of enemy radar installations is militarily necessary, that activity of civilians or protected persons in the area is not such that use of the Harpy would be disproportionate or insufficiently discriminate, and that the Harpy's explosive device is sufficient to destroy the type of radar installations expected without causing excessive damage beyond that. Attack decisions with other AWS must be based on similar knowledge.[487]

Once planning is completed, the attack moves into the execution phase. Step five covers initiating and executing the attack. The legal obligation is given in API art 57(1):

> In the conduct of military operations, constant care shall be taken to spare the civilian population, civilians and civilian objects.

When a manually operated weapon system is employed in an attack, legal responsibility for this phase rests with the combatants operating the weapon system, such as the aircrew of a piloted aircraft. If the role of those combatants is reduced or eliminated by use of an autonomous system, there may be no person who is in a position to take 'constant care' with the conduct of the attack, but the weapons law rule in API Article 51(4) prohibiting indiscriminate attacks still applies. In particular, the rule in Article 51(4)(b) prohibiting attacks 'which employ a method or means of combat which cannot

[486] Rob McLaughlin, 'Unmanned Naval Vehicles at Sea: USVs, UUVs, and the Adequacy of the Law' (2011) 21(2) *Journal of Law, Information and Science* 100, 105.
[487] Thurnher, above n 478, 222.

be directed at a specific military objective', would require the AWS control software to avoid directing an attack at a civilian target. Further, the injunction in Article 57(2)(a)(iii) to refrain from deciding to launch any attack which *may be expected* to cause excessive civilian harm describes an expectation about the capability of the AWS control system to avoid causing civilian harm. That expectation must be based on an understanding of how the AWS is likely to behave in the circumstances of the attack.[488] Taken together, those two provisions seem to have a substantially similar effect on an AWS that Article 57(1) would have on a combatant executing an attack.

Finally, step 6 in the IHL targeting process, relating to API Article 57(2)(b), requires that an attack be cancelled or suspended if circumstances change or new information comes to light. The obligation to cancel or suspend the attack in those circumstances 'is imposed upon anyone who has effective control over the attack. This might be the person conducting the attack or it might be a commander who can issue orders to the person conducting the attack.'[489]

There does not appear to be any explicit obligation on an attacking force to maintain the ability to cancel an attack after it has begun, although there is clearly an assumption that such an ability will be available until very close to the time at which a target is actually harmed:

> It is principally by visual means - in particular, by means of aerial observation - that an attacker will find out that an intended objective is not a military objective, or that it is an object entitled to special protection. Thus, to take a simple example, an airman who has received the order to machine-gun troops travelling along a road, and who finds only children going to school, must abstain from attack.[490]

Nevertheless, the possibility that an attacker's ability to cancel an attack may be curtailed is considered:

> However, with the increased range of weapons, particularly in military operations on land, it may happen that the attacker has no direct view of the objective, either because

[488] Ibid 224.
[489] Ibid 237; Pilloud and Pictet, above n 67, 686 [2220].
[490] Pilloud and Pictet, above n 67, 686 [2221].

> it is very far away, or because the attack takes place at night. In this case, even greater caution is required.[491]

Presumably the need for 'greater caution' would be even more strongly applicable to attacks with AWS in which very significant time delays, and many intervening events, may occur between a platform being activated and a target being engaged. The caution would relate to care in selecting the appropriate AWS for an attack, and the choice of where and when to activate it. That is not to say that there would necessarily be no way of cancelling an autonomous attack; provision of the facility for doing that would be up to the weapon system designers, though, most likely based on an assessment of the military and humanitarian advantages of doing so, as well as the possibilities for the attacking force to maintain contact with the AWS once it is deployed.

6.2.1 Summary

Increasing autonomous capabilities in weapon systems would no doubt have significant effects on many facets of targeting, but the ability to adhere to legal rules would not be greatly affected.

The essential change is in the amount and type of knowledge which attack planners must have in order to meet their obligations. The essence of weapon autonomy is the offloading of specific tasks from humans to the AWS control system; consequently, the information necessary to perform those tasks must be provided to the AWS, but not necessarily to attack planners or commanders or other combatants, in order to conduct the attack. However, attack planners considering use of an AWS to conduct an attack must still have sufficient information about the proposed attack to support a decision to use a particular AWS. Given that the legal burden of the decisions in API Article 57 is still with attack planners and any combatants who can still be considered to be executing the attack, they must also have sufficient information to 'fill in the gaps' around whatever precautionary decisions the AWS is unable to make for itself. For example, humans must constrain the choice of possible targets to a degree where the AWS is able to identify the correct target from among those accessible to it; humans must either perform the proportionality assessment required by API Article 57(2)(a)(iii) or provide sufficient information to the AWS; and so on.

---

[491] Ibid.

The content of the obligations borne by attackers using AWS therefore depends largely on the technical capabilities of the AWS. It is not apparent that there is any fixed upper limit to the level of autonomy that may be legally granted to an AWS other than that humans must be able to justify the decision to use the AWS in the first place.

6.3 Meaningful Human Control

Questions about what constitutes legal use of AWS lead naturally to questions about how to ensure that use is kept within legal limits. Concerns stem from the observation that humans appear to be ceding control over the weapon system to a computer. Accordingly, one of the most prominent features of the AWS debate thus far has been the emergence of the notion of 'meaningful human control' (MHC) over AWS.[492] It refers to the fear that a capacity for autonomous operation threatens to put AWS outside the control of the armed forces that operate them, whether intentionally or not, and consequently their autonomy must be limited in some way in order to ensure they will operate consistently with legal and moral requirements. Although used initially, and most commonly, in the context of objections to increasing degrees of autonomy, the idea has been picked up by many States, academics and NGOs as a sort of framing concept for the debate. This section discusses the place of MHC in the debate, current views on what it entails,[493] and its value in light of the analysis of AWS presented in this thesis.

6.3.1 History

The idea of MHC was first used in relation to AWS by the UK NGO Article 36. In April 2013, Article 36 published a paper arguing for 'a positive obligation in international law for individual attacks to be under meaningful human control.'[494] The paper was a response to broad concerns about increasing military use of remotely controlled and robotic weapon systems, and specifically to statements by the UK

[492] In this section, as in the wider debate, 'meaningful human control' describes a quality that is deemed to be necessary in order for an attack to comply with IHL rules. It does not refer to a particular class of weapon systems which allow or require some minimum level of human control, although it implies that a weapon used in a legally compliant attack would necessarily allow a meaningful level of human control.
[493] For another analysis see Rebecca Crootof, 'A Meaningful Floor for "Meaningful Human Control"' (2016) *Temple International and Comparative Law Journal* 53.
[494] Article 36, 'Killer Robots: UK Government Policy on Fully Autonomous Weapons' (Policy Paper, April 2013) 1 <http://www.article36.org/wp-content/uploads/2013/04/Policy_Paper1.pdf?con=&dom=pscau&src=syndication>.

Ministry of Defence (MoD) in its 2011 Joint Doctrine Note on Unmanned Systems.[495] Despite government commitments that weapons would remain under human control,[496] the MoD indicated that 'attacks without human assessment of the target, or a subsequent human authorization to attack, could still be legal':[497]

> a mission may require an unmanned aircraft to carry out surveillance or monitoring of a given area, looking for a particular target type, before reporting contacts to a supervisor when found. A human-authorised subsequent attack would be no different to that by a manned aircraft and would be fully compliant with [IHL], provided the human believed that, based on the information available, the attack met [IHL] requirements and extant [rules of engagement]. From this position, it would be only a small technical step to enable an unmanned aircraft to fire a weapon based solely on its own sensors, or shared information, and without recourse to higher, human authority. Provided it could be shown that the controlling system appropriately assessed the [IHL] principles (military necessity; humanity; distinction and proportionality) and that [rules of engagement] were satisfied, this would be entirely legal.[498]

As a result, according to Article 36, 'current UK doctrine is confused and there are a number of areas where policy needs further elaboration if it is not to be so ambiguous as to be meaningless.'[499]

Specifically, Article 36 argued that 'it is moral agency that [the rules of proportionality and distinction] require of humans, coupled with the freedom to choose to follow the rules or not, that are the basis for the normative power of the law.'[500] That is, human beings must make conscious, informed decisions about each use of force in a conflict; delegating such decisions to a machine would be inherently unacceptable. Those human decisions should relate to each individual attack:

> Whilst it is recognized that an individual attack may include a number of specific target objects, human control will cease to be meaningful if an [AWS] is undertaking multiple

---

[495] *JDN 2/11*, above n 92.
[496] See, eg, United Kingdom, *Parliamentary Debates*, House of Lords, 26 March 2013, Column 958, 3pm (Lord Astor of Hever) <http://www.publications.parliament.uk/pa/ld201213/ldhansrd/text/130326-0001.htm#st_14>.
[497] Article 36, above n 494, 2.
[498] *JDN 2/11*, above n 92, 5-4 [507].
[499] Article 36, above n 494, 1.
[500] Ibid 2.

> attacks that require specific timely consideration of the target, context and anticipated effects.[501]

The authors acknowledged that some existing weapon systems exhibit a limited capacity for autonomous operation, and are not illegal because of it:

> there are already systems in operation that function in this way - notably ship mounted anti-missile systems and certain 'sensor fuzed' weapon systems. For these weapons, it is the relationship between the human operator's understanding the sensor functioning and human operator's control over the context (the duration and/or location of sensor functioning) that are argued to allow lawful use of the weapons.[502]

Despite that, their conception of problematic 'fully autonomous' weapons, according to an earlier publication, seems to be very broad and could conceivably include calling for a ban on existing weapons:

> Although the relationship between landmines and fully autonomous armed robots may seem stretched, in fact they share essential elements of DNA. Landmines and fully autonomous weapons all provide a capacity to respond with force to an incoming 'signal' (whether the pressure of a foot or a shape on an infra-red sensor). Whether static or mobile, simple or complex, it is the automated violent response to a signal that makes landmines and fully autonomous weapons fundamentally problematic …
>
> [W]e 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 … across to the most complex systems under development.[503]

Nevertheless, based on those concerns, the paper makes three calls on the UK Government. First, they ask the Government to '[c]ommit to, and elaborate, meaningful human control over individual attacks.'[504] Second, '[s]trengthen commitment not to develop fully autonomous weapons and systems that could undertake attacks without

---

[501] Ibid 4.
[502] Ibid 3.
[503] Matthew Bolton, Thomas Nash and Richard Moyes, *Ban Autonomous Armed Robots* (5 March 2012) Article 36 <http://www.article36.org/statements/ban-autonomous-armed-robots/>.
[504] Article 36, above n 494, 3.

meaningful human control.'[505] Finally, '[r]ecognize that an international treaty is needed to clarify and strengthen legal protection from fully autonomous weapons.'[506]

Since 2013, Article 36 has continued to develop the concept of MHC,[507] and it has been taken up by some States and civil society actors. Inevitably, the meaning, rather imprecise to begin with, has changed with use. In particular, some parties have dropped the qualifier 'over individual attacks', introducing some uncertainty about exactly what is to be subject to human control. Does it apply to every discharge of a weapon? Every target selection? Only an attack as a whole? Something else?

Further, each term is open to interpretation:

> The MHC concept could be considered a priori to exclude the use of [AWS]. This is how it is often understood intuitively. However, whether this is in fact the case depends on how each of the words involved is understood. "Meaningful" is an inherently subjective concept … "Human control" may likewise be understood in a variety of ways.[508]

Thoughts about MHC and its implications for development of AWS continue to evolve as the debate continues, but a lack of certainty about the content of the concept has not slowed its adoption. It has been discussed extensively by expert presenters at the CCW meetings on AWS, and many State delegations have referred to it in their statements, generally expressing support or at least a wish to explore the idea in more depth.

At the 2014 Informal Meeting of Experts, Germany spoke of the necessity of MHC in anti-personnel attacks:

> it is indispensable to maintain meaningful human control over the decision to kill another human being. We cannot take humans out of the loop.

---

[505] Ibid 4.
[506] Ibid 5.
[507] Article 36, 'Structuring Debate on Autonomous Weapon Systems' (Memorandum for delegates to the Convention on Certain Conventional Weapons (CCW), November 2013) <http://www.article36.org/wp-content/uploads/2013/11/Autonomous-weapons-memo-for-CCW.pdf>; Article 36, 'Key Areas for Debate on Autonomous Weapon Systems' (Memorandum for delegates at the CCW Meeting of Experts on AWS, May 2014) <http://www.article36.org/wp-content/uploads/2014/05/A36-CCW-May-2014.pdf>.
[508] 'The Weaponization of Increasingly Autonomous Technologies: Considering how Meaningful Human Control Might Move the Discussion Forward' (Discussion Paper, United Nations Institute for Disarmament Research (UNIDIR), 2014) 3.

> We do believe that the principle of human control is already implicitly inherent to [IHL] … And we cannot see any reason why technological developments should all of a sudden suspend the validity of the principle of human control.[509]

Norway explicitly linked 'full' autonomy to a lack of MHC; the delegation expressed concern about the capabilities of autonomous technologies, rather than the principle of delegating decisions on the use of force to an AWS:

> By [AWS] in this context I refer to weapons systems that search for, identify and use lethal force to attack targets, including human beings, without a human operator intervening, and without meaningful human control. … our main concern with the possible development of [AWS] is whether such weapons could be programmed to operate within the limitations set by international law.[510]

The following year, several delegations noted that MHC had become an important element of the discussion:

> [The 2014 CCW Meeting of Experts] led to a broad consensus on the importance of 'meaningful human control' over the critical functions of selecting and engaging targets. … we are wary of fully autonomous weapons systems that remove meaningful human control from the operation loop, due to the risk of malfunctioning, potential accountability gap and ethical concerns.[511]

MHC remained prominent at the 2016 meetings, where there was a widely held view that it was fundamental to understanding and regulating AWS:

> The elements, such as "autonomy" and "meaningful human control (MHC)," which were presented at the last two Informal Meetings are instrumental in deliberating the definition of [AWS].[512]

However, there were questions also emerged about the usefulness of the concept:

> The U.S. Delegation also looks forward to a more in depth discussions [sic] with respect to human-machine interaction and about the phrase "meaningful human control." Turning first to the phrase "meaningful human control," we have heard many

[509] Germany, above n 150, 4.
[510] Norway, above n 151, 1.
[511] Republic of Korea (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2015) 1-2.
[512] Japan (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016) 1-2.

> delegations and experts note that the term is subjective and thus difficult to understand. We have expressed these same concerns about whether "meaningful human control" is a helpful way to advance our discussions.
>
> We view the optimization of the human/machine relationship as a primary technical challenge to developing lethal [AWS] and a key point that needs to be reviewed from the start of any weapon system development. Because this human/machine relationship extends throughout the development and employment of a system and is not limited to the moment of a decision to engage a target, we consider it more useful to talk about "appropriate levels of human judgment."[513]

The idea of MHC over AWS has also been raised outside of a strict IHL context, both at the CCW meetings and elsewhere. For example, the African Commission on Human and Peoples' Rights incorporated MHC into its General Comment No. 3 on the African Charter on Human and Peoples' Rights on the right to life (Article 4) of 2015:

> The use during hostilities of new weapons technologies … should only be envisaged if they strengthen the protection of the right to life of those affected. Any machine autonomy in the selection of human targets or the use of force should be subject to meaningful human control.[514]

For all its prominence, though, the precise content of the MHC concept is still unsettled. The next section surveys the views of various parties.

6.3.2 Meaning

The unsettled content of the MHC concept is perhaps to be expected, as it is not based on a positive conception of something that is required of an AWS. Rather, it is based 'on the idea that concerns regarding growing autonomy are rooted in the human aspect that autonomy removes, and therefore describing that human element is a necessary starting point if we are to evaluate whether current or future technologies challenge that.'[515] That is, the desire to ensure MHC over AWS is based on the recognition that States are embarking on a path of weapon development that promises to reduce direct

---

[513] United States (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016) 2.

[514] *General Comment No. 3 on the African Charter on Human and Peoples' Rights: The Right to Life (Article 4)*, African Commission on Human and Peoples' Rights, 57th ordinary session (18 November 2015) 12 [35] <http://www.achpr.org/instruments/general-comments-right-to-life/>.

[515] Article 36, above n 468, 2.

human participation in conducting attacks,[516] but it is not yet clear how removal of that human element would be accommodated in the legal and ethical decisions that must be made in the course of an armed conflict.

Specifically, Article 36 developed MHC from two premises:

> 1. That a machine applying force and operating without any human control whatsoever is broadly considered unacceptable.
>
> 2. That a human simply pressing a 'fire' button in response to indications from a computer, without cognitive clarity or awareness, is not sufficient to be considered 'human control' in a substantive sense.[517]

The idea is that some form of human control over the use of force is required, and that human control cannot be merely a token or a formality; human influence over acts of violence by a weapon system must be sufficient to ensure that those acts are done only in accordance with human designs and, implicitly, in accordance with legal and ethical constraints. 'Meaningful' is the term chosen to represent that threshold of sufficiency. MHC therefore "represents a space for discussion and negotiation. The word 'meaningful' functions primarily as an indicator that the form or nature of human control necessary requires further definition in policy discourse."[518] Attention should not be focussed too closely on the precise definition of 'meaningful' in this context.

> There are other words that could be used instead of 'meaningful', for example: appropriate, effective, sufficient, necessary. Any one of these terms leaves open the same key question: how will the international community delineate the key elements of human control needed to meet these criteria?[519]

The purpose of discussing MHC is simply 'to delineate the elements of human control that should be considered necessary in the use of force.'[520]

---

[516] For a general discussion of the decline of direct human involvement in combat decision-making see Thomas K Adams, 'Future Warfare and the Decline of Human Decisionmaking' (2001-2) (Winter) *Parameters* 57.
[517] Article 36, above n 468, 2.
[518] Ibid.
[519] Ibid.
[520] Ibid.

In terms of IHL in particular, Article 36 believes that a failure to maintain MHC when employing AWS risks diluting the central role of 'attacks' in regulating the use of weapons in armed conflict.

> it is over individual 'attacks' that certain legal judgments must be applied. So attacks are part of the structure of the law, in that they represent units of military action and of human legal application.[521]

Article 57 of API obliges 'those who plan or decide upon an attack' to take certain precautions. The NGO claims that 'humans must make a legal determination about an attack on a specific military objective based on the circumstances at the time',[522] and the combined effect of Articles 51, 52, and 57 of API is that

> a machine cannot identify and attack a military objective without human legal judgment and control being applied in relation to an attack on that specific military objective at that time … Arguing that this capacity can be programmed into the machine is an abrogation of human legal agency - breaching the 'case-by-case' approach that forms the structure of these legal rules.[523]

Further,

> the drafters' intent at the time was to require humans (those who plan or decide) to utilize their judgment and volition in taking precautionary measures on an attack-by-attack basis. Humans are the agents that a party to a conflict relies upon to engage in hostilities, and are the addressees of the law as written.[524]

Thus, 'the existing legal structure … implies certain boundaries to independent machine operation'.[525] Use of AWS that might be able to initiate attacks on their own, selecting and engaging targets without human intervention, threatens the efficacy of the legal structure:

---

[521] Ibid.
[522] Ibid 3.
[523] Ibid.
[524] Heather M Roff and Richard Moyes, 'Meaningful Human Control, Artificial Intelligence and Autonomous Weapons' (Briefing paper for delegates at the CCW Meeting of Experts on AWS, Article 36, April 2016) 5.
[525] Article 36, above n 468, 3.

> autonomy in certain critical functions of weapons systems might produce an expansion of the concept of 'an attack' away from the granularity of the tactical level, towards the operational and strategic. That is to say, AWS being used in 'attacks' which in their spatial, temporal or conceptual boundaries go significantly beyond the units of military action over which specific legal judgement would currently be expected to be applied.[526]

Whereas:

> By asserting the need for [MHC] over attacks in the context of [AWS], states would be asserting a principle intended to protect the structure of the law, as a framework for application of wider moral principles.[527]

As to the form of human control which would be 'meaningful' in this context, Article 36 proposes four key elements:

- Predictable, reliable and transparent technology: on a technical level, the design of AWS must facilitate human control. 'A technology that is by design unpredictable, unreliable and un-transparent is necessarily more difficult for a human to control in a given situation of use.'[528]
- Accurate information for the user on the outcome sought, the technology, and the context of use: human commanders should be provided with sufficient information 'to assess the validity of a specific military objective at the time of an attack, and to evaluate a proposed attack in the context of the legal rules';[529] to know what types of objects will be targeted, and how kinetic force will be applied; and to understand the environment in which the attack will be conducted.
- Timely human judgement and action, and a potential for timely intervention: human commanders must apply their judgement and choose to activate the AWS. 'For a system that may operate over a longer period of time, some capacity for

[526] Ibid.
[527] Ibid.
[528] Ibid 4.
[529] Ibid.

timely intervention (e.g. to stop the independent operation of a system) may be necessary if it is not to operate outside of the necessary human control.'[530]

- A framework of accountability: structures of accountability should encompass the personnel responsible or specific attacks as well as 'the wider system that produces and maintains the technology, and that produces information on the outcomes being sought and the context of use.'[531]

In summary, Article 36 sees management of individual attacks at the tactical level as the key to regulating the use of force in armed conflict. The law requires legal judgements by the appropriate human personnel in relation to each individual attack, and the design and use of AWS must not exclude those judgements.

Other actors who have taken up the idea of MHC see it somewhat differently and have put forward their own views on the criteria for human control to be 'meaningful'. The International Committee for Robot Arms Control (ICRAC), in its statement on technical issues at the 2014 CCW meetings,[532] expressed concern about the considerable technical challenges facing developers of AWS, and support for MHC as a means of ensuring that humans are able to compensate for those shortcomings:

> Humans need to exercise meaningful control over weapons systems to counter the limitations of automation.
>
> ICRAC hold that the minimum necessary conditions for meaningful control are
>
> First, a human commander (or operator) must have full contextual and situational awareness of the target area and be able to perceive and react to any change or unanticipated situations that may have arisen since planning the attack.
>
> Second, there must be active cognitive participation in the attack and sufficient time for deliberation on the nature of the target, its significance in terms of the necessity and appropriateness of attack, and likely incidental and possible accidental effects of the attack.

---

[530] Ibid.
[531] Ibid.
[532] Frank Sauer, *ICRAC statement on technical issues to the 2014 UN CCW Expert Meeting* (14 May 2014) International Committee for Robot Arms Control <http://icrac.net/2014/05/icrac-statement-on-technical-issues-to-the-un-ccw-expert-meeting/>.

> Third, there must be a means for the rapid suspension or abortion of the attack.[533]

Notably, some of these conditions go beyond the levels of awareness and direct involvement that commanders are able to achieve using some existing weapon systems: 'humans have been employing weapons where they lack perfect, real-time situational awareness of the target area since at least the invention of the catapult.'[534]

At the 2015 CCW meetings, Maya Brehm focussed on control over the harm suffered by persons and objects affected by an attack:

> it is generally expected in present practice that human beings exercise some degree of control over:
>
> – *Who* or *what* is harmed …
> – *When* force is applied / harm is experienced …
> – *Where* force is applied / harm is experienced …
> – *Why* someone or something is targeted / harmed … and *how* armed force is used[535]

According to Brehm, MHC requires that attackers have sufficient information about the effects of an attack to:

> anticipate the reasonably foreseeable consequences of force application. Only if [attackers] can anticipate these consequences, can they make the required legal assessments about the use of force.[536]

Consequently, the degree of autonomy allowed to a weapon system must be limited such that human attackers can be sure of having sufficient information about how the weapon system will behave once it is activated.

According to CNAS, the purpose of MHC should be

> to ensure that human operators and commanders are making conscious decisions about the use of force, and that they have enough information when making those decisions to remain both legally and morally accountable for their actions.[537]

---

[533] Ibid.
[534] Michael C Horowitz and Paul Scharre, 'Meaningful Human Control in Weapon Systems: A Primer' (Working Paper, Center for a New American Security, March 2015) 9.
[535] Maya Brehm (Presentation, CCW Meeting of Experts on LAWS: Characteristics of LAWS, April 2015) 1-2 [emphasis in original].
[536] Ibid 2.

Horowitz and Scharre, also writing in association with CNAS, have summarised the 'two general schools of thought about how to answer the question of why [MHC] is important.'[538]

The first is that MHC is not, and should not be, a standalone requirement, but is

> a principle for the design and use of weapon systems in order to ensure that their use can comply with the laws of war. This … starts from the assumption that the rules that determine whether the use of a weapon is legal are the same whether a human delivers a lethal blow directly, a human launches a weapon from an unmanned system, or a human deploys an [AWS] that selects and engages targets on its own.[539]

The second school of thought positions MHC as an additional legal principle that should be explicitly recognised alongside existing principles of IHL. It states that

> the existing principles under the laws of war are necessary but not sufficient for addressing issues raised by increased autonomy, and that [MHC] is a separate and additional concept. … even if an [AWS] could be used in a way that would comply with existing laws of war, it should be illegal if it could not meet the additional standard of [MHC].[540]

The authors then suggest three essential components of a useful MHC concept:

> 1. Human operators are making informed, conscious decisions about the use of weapons.
>
> 2. Human operators have sufficient information to ensure the lawfulness of the action they are taking, given what they know about the target, the weapon, and the context for action.
>
> 3. The weapon is designed and tested, and human operators are properly trained, to ensure effective control over the use of the weapon.[541]

Geiss offers some more specific suggestions about what may constitute MHC:

---

[537] Center for a New American Security (Text, CCW Meeting of Experts on LAWS: Characteristics of LAWS, April 2015) 1.
[538] Horowitz and Scharre, above n 534, 7.
[539] Ibid.
[540] Ibid.
[541] Ibid 14-15.

> the requisite level of control can refer to several factors: the time-span between the last decision taken by humans and the exertion of force by the machine; the environment in which the machine comes to be deployed, especially with regard to the question of whether civilians are present in that environment; … whether the machine is supposed to engage in defensive or offensive tasks; … whether the machine is set up to apply lethal force; the level of training of the persons tasked with exercising control over the machine; … the extent to which people are in a position to intervene, should the need arise, and to halt the mission; the implementation of safeguards with regard to responsibility …[542]

Horowitz and Scharre also raise the question of the level at which MHC should be exercised. While most commentators focus on commanders responsible for an attack at the tactical level, there are other personnel who are well positioned to ensure that humans remain in control of AWS.

> At the highest level of abstraction, a commander deciding on the rules of engagement for a given use of force is exercising [MHC] over the use of force. Below that, there is an individual commander ordering a particular attack against a particular target … Along a different axis, [MHC] might refer to the way a weapon system is designed in the first place …[543]

6.3.3 Alternatives

Some participants have proposed alternatives to MHC. While not disagreeing with the underlying proposition that human must remain in control of, and accountable for, acts committed via AWS, their view is that attempting to define an objective standard of MHC is not the correct approach.

The United States delegation to the CCW meetings presented the notion of 'appropriate levels of human judgement' being applied to AWS operations, with 'appropriate' being a contextual standard:

> there is no "one-size-fits-all" standard for the correct level of human judgment to be exercised over the use of force with [AWS]. Rather, as a general matter, [AWS] vary greatly depending on their intended use and context. In particular, the level of human judgment over the use of force that is appropriate will vary depending on factors,

[542] Geiss, above n 115, 24-5.
[543] Horowitz and Scharre, above n 534, 15.

> including, the type of functions performed by the weapon system; the interaction between the operator and the weapon system, including the weapon's control measures; particular aspects of the weapon system's operating environment (for example, accounting for the proximity of civilians), the expected fluidity of or changes to the weapon system's operational parameters, the type of risk incurred, and the weapon system's particular mission objective. In addition, engineers and scientists will continue to develop technological innovations, which also counsels for a flexible policy standard that allows for an assessment of the appropriate level of human judgment for specific new technologies.[544]

Measures taken to ensure that appropriate levels of human judgement are applied to AWS operations would then cover the engineering and testing of the weapon systems, training of the users and careful design of the interfaces between weapon systems and users.

Finally, the Polish delegation to the CCW meetings in 2015 preferred to think of State control over AWS, rather than human control:

> What if we accept MHC as a starting point for developing national strategies towards [AWS]? We could view MHC from the standpoint of [S]tate's affairs, goals and consequences of its actions. In that way this concept could also be regarded as the exercise of "meaningful [S]tate control" (MSC). A [S]tate should always be held accountable for what it does, especially for the responsible use of weapons which is delegated to the armed forces. The same goes also for [AWS]. The responsibility of [S]tates for such weapons should also be extended to their development, production, acquisition, handling, storage or international transfers.[545]

6.3.4 Arguments against MHC

The general proposition that humans should maintain close control over the weapons they use is indisputable. Nevertheless, attempts to relate MHC to IHL appear to be generally counterproductive. At most, MHC could reasonably be seen, in Horowitz and Scharre's terms, as 'a principle for the design and use of weapon systems in order to

[544] Michael Meier, 'U.S. Delegation Statement on "Appropriate Levels of Human Judgment"' (Statement to the CCW Informal Meeting of Experts on AWS, 12 April 2016) <https://geneva.usmission.gov/2016/04/12/u-s-delegation-statement-on-appropriate-levels-of-human-judgment/>.
[545] Poland (Text, CCW Meeting of Experts on LAWS: Characteristics of LAWS, April 2015) 1.

ensure that their use can comply with the laws of war'. Even in that respect, though, it is an unnecessary addition; the existing rules of IHL are already sufficient to regulate use of AWS. The principal argument against introducing the idea of MHC into a discussion of AWS and IHL, especially in its more expansive form as a standalone principle or rule, is that it is based on two false premises, one technical and one legal.

The false technical premise underlying a perceived need for MHC is that it assumes that the software and hardware that make up an AWS control system do not themselves constitute an exercise of MHC. One cannot rationally raise the concern that the autonomous capabilities of weapons should be limited in order to ensure humans maintain sufficient control if one understands the weapon's control system as the means by which human control is already maintained.

Yet, the discussion in Chapter 3 showed that machine autonomy is a form of control, not a weakening of control. Weapon developers draw on human operational understanding of how targets are to be selected and attacks are to be conducted, technical understanding of how to operate weapons, and legal understanding of the rules of IHL in programming AWS control systems. Weapon reviews conducted by responsible humans test and verify the behaviour of AWS in the conditions in which they are intended to be used, ensuring they comply with all relevant rules of weapons law. Attack planners and commanders are required by existing rules of IHL to '[t]ake all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event to minimizing, incidental loss of civilian life, injury to civilians and damage to civilian objects;'[546] that means, at a minimum, selecting an AWS that has been shown to operate successfully in the circumstances of the attack at hand. After an AWS is activated, its control system, tested by humans, controls the weapon system in the circumstances for which it has been tested, just as the control systems of existing weapons do. It is difficult to see how any part of that process can be interpreted as constituting a lack of human control.

Concerns about maintaining human control over AWS might best be understood as fears about events that might occur after the weapon system is activated in the course of an attack; fears that it might perform some proscribed act, such as firing upon a civilian

[546] *API* art 57(2)(a)(ii).

target. The classification of civilian harm by an AWS was discussed earlier in Section 5.4.2.2.1. If such an unfortunate event were to occur, it would be the result of either an intentional act by a human, a malfunction by the AWS, or unavoidable collateral damage. None of those concerns are unique to AWS, and all are considered in existing law; no new notion of MHC is required.

The false legal premise underlying MHC is that it assumes that existing rules of IHL do not ensure a level of human control over AWS sufficient to achieve the aims of the law. Examination of current targeting law shows that is not the case. It does not appear possible for a weapon system to be beyond human control without its use necessarily violating an existing rule. If attack planners cannot foresee that an AWS will engage only legal targets then they cannot meet their obligations under the principle of distinction.[547] If they cannot ensure that civilian harm will be minimised and that the AWS will refrain from attacking some objective if the civilian harm would be excessive then they cannot meet their obligations under the principle of proportionality.[548] If they cannot ensure that the AWS will cancel or suspend an attack if conditions change, they also fail to meet their obligations.[549]

There seems to have been some confusion on this point. Human Rights Watch has cited the bans on several existing weapons as evidence of a need for recognition of MHC:

> Although the specific term [MHC] has not appeared in international arms treaties, the idea of human control is not new in disarmament law. Recognition of the need for human control is present in prohibitions of mines and chemical and biological weapons, which were motivated in part by concern about the inability to dictate whom they engage and when. After a victim-activated mine is deployed, a human operator cannot determine at what moment it will detonate or whom it will injure or kill. Although a human can choose the moment and initial target of a biological or chemical weapons attack, the weapons' effects after release are uncontrollable and can extend across space and time causing unintended casualties. The bans on mines and chemical and biological

[547] *API* art 57(2)(a)(i).
[548] *API* art 57(2)(a)(iii).
[549] *API* art 57(2)(b).

> weapons provide precedent for prohibiting weapons over which there is inadequate human control.[550]

Examination of the prohibitions on mines,[551] biological[552] and chemical[553] weapons shows they were each prohibited for violating fundamental rules and principles that long predate any notion of MHC as a standalone concept. Insofar as one may view indiscriminate behaviour of a weapon or substance as evidence of an inability to exercise control, then the bans could be attributed to a lack of control, but in that case the idea of MHC seems to add nothing to the existing principle of distinction. Mines are strictly regulated because a simple pressure switch is a very imprecise means of identifying a combatant; biological and chemical weapons employ harmful agents the effects of which are indiscriminate, independently of how the weapon system itself is controlled.

Beyond those two main concerns, recognising MHC as a limitation on development of new control system technologies risks interfering with advances which might improve an attacker's ability to engage military objectives with greater precision, and less risk of civilian harm, than is currently possible. Human Rights Watch has previously recognised the value of precision weapons in attacking targets in densely populated areas;[554] it seems implausible to suggest that further advances in selecting and assessing potential targets on board a weapon system after activation will not create further opportunities for avoiding civilian casualties.

Finally, even if fears about a likely path of weapon development are seen as a valid basis for regulation, it is not clear exactly what development path proponents of MHC are concerned about: is it that AWS will be too 'smart', or not 'smart' enough? Fears that AWS will be too smart amount to fears that humans will be unable to predict their behaviour in the complex and chaotic circumstances of an attack. Fears that AWS will not be smart enough amount to fears that they will fail in a more predictable way,

---

[550] Human Rights Watch, 'Killer Robots and the Concept of Meaningful Human Control' (Memorandum to CCW Delegates, April 2016) 10 <https://www.hrw.org/sites/default/files/supporting_resources/robots_meaningful_human_control_final.pdf>.
[551] *Ottawa Convention*, Preamble.
[552] Henckaerts and Doswald-Beck, above n 19, 256.
[553] Ibid 259.
[554] Human Rights Watch, *Off Target: The Conduct of the War and Civilian Casualties in Iraq* (2003) Summary and Recommendations <https://www.hrw.org/reports/2003/usa1203/usa1203_sumrecs.pdf>.

whether it be in selecting legitimate targets or another failure mode. In either case, using a weapon which is the object of such concerns would breach existing precautionary obligations.

6.3.5 Controllability

Existing IHL does not contemplate any significant level of autonomous capability in weapon systems. It implicitly assumes that each action of a weapon will be initiated by a human being and that after completion of that action the weapon will cease operating until a human initiates some other action. If there is a failure in use of a weapon, such that a rule of IHL is broken, it is assumed to be either a human failure (further assuming that the weapon used is not inherently illegal), or a failure of the weapon which would be immediately known to its human operator. Generally, facilities would be available to prevent that failure from continuing uncontrolled.

If an AWS fails after being activated, in circumstances in which a human cannot quickly intervene, its failure will be in the nature of a machine rather than a human. The possibility of runaway failure is often mentioned by opponents of AWS development. Horowitz and Scharre mention it in arguing for 'controllability' as an essential element of MHC:

> Militaries generally have no interest in developing weapon systems that they cannot control. However, a military's tolerance for risk could vary considerably … The desire for a battlefield advantage could push militaries to build weapons with high degrees of autonomy that diminish human control … While any weapon has the potential for failure and accidents, [AWS] arguably add a new dimension, since a failure could, in theory, lead to the weapon system selecting and engaging a large number of targets inappropriately. Thus, one potential concern is the development of weapons that are legal when functioning properly, but that are unsafe and have the potential to cause a great deal of harm if they malfunction or face unanticipated situations on the battlefield.[555]

Use of AWS in situations where a human is not able to quickly intervene, such as on long operations or in contested environments, may change the nature of the risk borne by non-combatants.

[555] Horowitz and Scharre, above n 534, 8.

Controllability, as described by Horowitz and Scharre, could be seen as no different to the requirement for any weapon to be capable of being directed at a specific military objective, and malfunctions are similarly a risk which accompanies all weapon systems. To an extent the different type of risk that accompanies failure of an AWS is simply a factor that must be considered by attack planners in their precautionary obligations. However, if that risk acts to prevent the advantages of AWS from being realised, then one possible response might be to recognise a requirement for a failsafe, whether that be a facility for human intervention, or some other form:

> Although some systems may be designed to operate at levels faster than human capacity, there should be some feature for timely intervention by either another system, process, or human.[556]

6.3.6 Summary

A desire to maintain MHC over the operations of AWS is a response to the perception that some human element would be removed from military operations by increasing the autonomous capabilities of weapon systems. The idea that a formal requirement for MHC may be identified in, or added to, existing IHL was originated by civil society actors and is being taken up by an increasing number of States participating in the CCW discussions on AWS.

Although the precise definition of MHC is yet to be agreed upon, it appears to be conceptually flawed. It relies on the mistaken premise that autonomous technologies constitute a lack of human control, and on a mistaken understanding that IHL does not already mandate adequate human control over weapon systems.

[556] Roff and Moyes, above n 524, 3.

Chapter 7 Accountability

The focus of the rules of IHL is, rightly, prevention of acts which would undermine the goals of containing the effects of armed conflict and alleviating the suffering of those affected by it. Prevention alone is not sufficient, though. In the chaotic environment of the battlefield, violations of the law are inevitable. If the law is to have normative power, those violations must bear legal consequences for the entity which has transgressed the rules. Such consequences can be applied only where an entity can justly be held accountable for violations. A robust system of accountability supporting enforcement of the law is therefore essential.

Mechanisms of accountability for violations of IHL have evolved considerably since the advent of modern IHL, and are still evolving. When the first Geneva Convention concerning the treatment of wounded soldiers was signed in 1864,[557] the measures available for holding parties to account were minimal. Gustave Moynier, one of the instigators of the 1864 Convention and a driving force behind establishment of the ICRC, was initially of the view that the pressure of public opinion would be sufficient to ensure compliance, and that States Parties would voluntarily enact municipal legislation imposing penalties for violations.[558] Unfortunately, neither proved to be the case. The Franco-Prussian War of 1870-71, in which both belligerents were parties to the Geneva Convention of 1864, saw widespread atrocities in violation of the Convention, inflamed rather than restrained by the press and public opinion. Following that war, Moynier began working for a system of enforcement of the rules of IHL.[559]

Today there are several regimes of international accountability which may be brought to bear on violations of IHL. This chapter begins with a general statement on how accountability regimes apply to AWS behaviour, then discusses the application of two established modes of legal responsibility to situations involving AWS. The law of State responsibility attempts to restrain violations of rules of international law by States and their agents. International criminal law attempts to restrain intentional violations of IHL by imposing penalties on the individuals who commit those offences. Those two

[557] *Convention for the Amelioration of the Condition of the Wounded in Armies in the Field*, opened for signature 22 August 1864, 22 Stat 940 (entered into force 22 June 1865).
[558] See, eg, Christopher Keith Hall, 'The First Proposal for an International Criminal Court' (1998) 38 *International Review of the Red Cross* 57, 58-9.
[559] Ibid 57.

sections focus narrowly on the effects that use of AWS is likely to have on the efficacy of existing law and the responses, if any, which might be needed in order to maintain the effectiveness of the law. Following that is a broader discussion of how to restrain violations of IHL in the face of increasing reliance on highly autonomous systems in armed conflict. Some alternative and emerging international accountability regimes are surveyed and their usefulness in relation to AWS is assessed.

7.1 General Statement on Accountability for AWS Behaviour

Perhaps the most common objection to further development of AWS is that a sufficient level of autonomous capability would create an 'accountability gap'.[560] That is, if an AWS is capable of 'making its own decisions', thereby determining its own behaviour, no person or organisation could be held accountable for the actions of the weapon system, due to a lack of control over, or perhaps even knowledge of, those decisions and actions. According to this view, if an AWS can decide for itself whether or not to fire upon some target, that decision would belong to the AWS rather than to a human in the chain of command, or any other entity which has a status or role which would allow them to be held accountable. The 'gap' thus created between the AWS and the State and armed forces who wield it would potentially rob the law of its normative power.[561]

The fear of an accountability gap implies a perception that the targeting system built into an AWS is, in terms of accountability, the ultimate cause of the harm done to a target, rather than just a proximate cause as, for example, the guidance system of a fire-and-forget missile might be viewed. As explained in Section 4.1, this thesis rejects any such suggestion that an AWS, no matter how advanced, is anything more than a weapon operating in accordance with human designs. Members of armed forces who select, configure and deploy AWS bear the same primary obligations that they would bear in relation to any other weapon system: to direct attacks only at military targets, to minimise the expected harm to civilians, to take all feasible precautions in selecting the appropriate weapon system for that purpose, and so on. They are also subject to the

[560] See generally, Human Rights Watch, *Mind the Gap: The Lack of Accountability for Killer Robots*, above n 299; Specific examples of this objection are given in the sections relating to each accountability regime.

[561] For a rebuttal of Human Rights Watch's position see Charles J Dunlap, 'Accountability and Autonomous Weapons: Much Ado about Nothing?' (2016) 30 *Temple International and Comparative Law Journal* 63.

same accountability regimes. Undeniably, armed forces which operate advanced AWS will be taking a step back from the process of engaging individual targets, but they will still be required to do everything that is practically possible to ensure that attacks done by AWS comply with the State's IHL obligations.

The act of activating a given AWS at a given time and place is itself an exercise of control over the eventual behaviour of the weapon. Even in the operation of today's AWS such as CIWS systems or IAI's Harpy, it may be the only human control applied in the normal course of events, yet no argument is made that there is an accountability gap in relation to those existing weapons.

Nor does placing a more complex computer-based control system in direct control of a weapon sever the connection between the actions of the weapon and the conduct of the entities which deployed it in such a way that those entities could not be held accountable for harm caused by the weapon. This is so regardless of the complexity of the control system's operation, the foreseeability of its 'decisions', or the ability or inability of a human to assume direct control of the weapon. Existing law accounts for the behaviour of existing weapon systems that are largely governed by computer control systems, without reference to the nature of the software employed in those control systems. The legal constraints that apply to weapon use, such as the requirement to ensure that attacks are directed only against military targets, are independent of the form of control employed (although the specific measures required of attack planners and weapon system operators would, of course, differ accordingly).

Consider, as a thought experiment, an AWS which is known to select targets randomly: that is, according to some random process which makes no attempt to comply with any legal obligations such that future targeting choices cannot be feasibly predicted by the weapon system's operators. Is there an accountability gap which would prevent any person or other entity from being held accountable for harm caused by the weapon? The precise effect of a computer-based weapon control system on each accountability regime will be discussed in the relevant sections below. In general, though, it seems that even in the extreme scenario of a weapon acting according to a random process, the actions of the weapon system's operators in deploying the weapon system are sufficient to attract accountability for any harm caused. In the more realistic case of a control

system which follows some non-random process, although possibly based on probabilistic reasoning and with some margin for error, the case for holding its operators accountable is even stronger.

As an aside, it appears that, even if one were to accept that an accountability gap could exist, the circumstances in which it would be at issue are unlikely to arise. Accountability for targeting decisions and actions taken in attacking the chosen targets becomes an issue only during the course of an attack, after the precautions required of an attack planner have been taken. Saying that no person is accountable for those targeting and attack processes, due to lacking control over them, is saying that the AWS was deployed on that attack despite no person being in a position to take the required precautions to ensure that only legal targets are attacked. Utilising a weapon system in an attack when it is not possible to take sufficient precautions to ensure that only legal targets are struck is an illegal act in itself, in violation of the principle of distinction.[562] If those precautions are taken, they would amount to an exercise of control over the AWS which would bridge any accountability gap.

One could object that neither API nor customary law dictates a precise level of certainty which must be reached regarding the status of the targets to be attacked, but at the very least an attack planner must be sufficiently certain about the targets which an AWS is likely to engage to be confident that the chosen AWS is a better choice (in legal terms) than another means or method of attack. Of course, complete certainty about the actions of a complex AWS may not be achievable in practice, but neither is it required by law. No weapon allows an attacker to be completely certain about the outcome of using it in an attack. One would not say there is an accountability gap because of the possibility that a missile might strike a civilian target that is within its circular error probable[563] just because no person selects that target for attack. Neither is there a gap because operators of an AWS cannot foresee with complete precision everything the AWS will do once activated.

[562] See, eg, Henckaerts and Doswald-Beck, above n 19, 55.
[563] Nelson, above n 384.

Accordingly, this chapter applies the existing regimes of accountability for violations of IHL to attacks involving AWS, just as they would be applied to attacks using other weapon systems.

7.2 State Responsibility

The law of State responsibility is a 'cardinal institution of international law.'[564] States are the main bearers of obligations under international law, and the law of State responsibility sets out the secondary rules by which those obligations are upheld. Although State responsibility has been supplemented (and, some would argue, obscured[565]) by other forms of international legal responsibility, it remains the frame of reference for consideration of those other forms.[566]

The principal reference point for the law of State responsibility today was promulgated by the International Law Commission and adopted in 2001.[567] The Draft Articles on the Responsibility of States for Internationally Wrongful Acts[568] codify and develop existing law, setting out the rules of State responsibility: the conditions for States to be considered responsible for wrongful acts and omissions, and the resulting legal consequences.[569] They do not define the content of international obligations, only the conditions and consequences of breaches of those obligations. The Draft Articles do not have the status of a treaty and are not formally binding on States,[570] but they are widely

---

[564] James R Crawford, 'State Responsibility' in *Max Planck Encyclopedia of Public International Law* (September 2006) [1] <http://opil.ouplaw.com/home/EPIL>.
[565] Rebecca Crootof, 'War Torts: Accountability for Autonomous Weapons' (2016) 164 *University of Pennsylvania Law Review* 1347, 1364.
[566] Crawford, above n 564, [1].
[567] *Report of the International Law Commission*, UN GAOR, 56th sess, Supp No 10, UN Doc A/56/10 (2001) ('*ILC Report*'). For a general discussion of State responsibility for violations of IHL prior to that time see Frits Kalshoven, 'State Responsibility for Warlike Acts of the Armed Forces' (1991) 40 *International and Comparative Law Quarterly* 827; see also Dieter Fleck, 'Individual and State Responsibility for Violations of the Ius in Bello: An Imperfect Balance' in Wolff Heintschel von Heinegg and Volker Epping (eds), *International Humanitarian Law Facing New Challenges* (Springer, 2007) 171, 171 n 2. For one application of the law to a violation of IHL prior to 2001 see Markus Rau, 'State Liability for Violations of International Humanitarian Law — The *Distomo* Case before the German Federal Constitutional Court' (2006) 7 *German Law Journal* 701.
[568] *ILC Report* 43 ('*Draft Articles*').
[569] James Crawford, *Articles on Responsibility of States for Internationally Wrongful Acts: Historical Background and Development of Codification*, United Nations Audiovisual Library of International Law <http://legal.un.org/avl/ha/rsiwa/rsiwa.html>.
[570] See, eg, *Prosecutor v Nikolić (Decision on Defence Motion Challenging the Exercise of Jurisdiction by the Tribunal)* (International Criminal Tribunal for the Former Yugoslavia, Trial Chamber II, Case No IT–94–2–PT, 9 October 2002) [60].

seen as a codification of customary law.[571] Since 2001 the Draft Articles have been commended to the attention of governments by the United Nations General Assembly in a series of resolutions[572] and an early draft text was applied by the ICJ in the case *Gabčíkovo-Nagyamaros Project (Hungary/Slovakia)*.[573]

State responsibility is enlivened by any 'internationally wrongful act' of a State, being any act or omission which is attributable to the State and which breaches an international obligation of the State.[574] Obligations within the scope of the law include any treaty and customary obligations, whether owed to another State, a non-State entity or the international community as a whole. Generally, the question of whether an act or omission breaches an obligation of the State is independent of the severity of any harm caused or the intention of the entity which caused the harm, unless the primary rule at issue specifies either of those conditions.[575]

The question of which acts and omissions are attributable to the State is more complicated. A State cannot act of itself. Acts 'of the State' are acts of persons or groups whose conduct can, in the circumstances, be attributed to the State. While there is no exhaustive and precise specification of whose conduct may be attributed to a State, 'the general rule is that the only conduct attributed to the State at the international level is that of its organs of government, or of others who have acted under the direction, instigation or control of those organs, i.e. as agents of the State.'[576] For the purposes of determining responsibility, '[a]n organ includes any person or entity which has that status in accordance with the internal law of the State',[577] even if acting in excess of

---

[571] See, eg, *Noble Ventures Inc v Romania (Award)* (ICSID Arbitral Tribunal, Case No ARB/01/11, 12 October 2005) [69].

[572] *Responsibility of States for Internationally Wrongful Acts*, UN GAOR, 6th Comm, 56th sess, 85th plen mtg, Agenda Item 162, UN Doc A/RES/56/83 (12 December 2001); *Responsibility of States for Internationally Wrongful Acts*, UN GAOR, 6th Comm, 59th sess, 65th plen mtg, Agenda Item 139, UN Doc A/RES/59/35 (2 December 2004); *Responsibility of States for Internationally Wrongful Acts*, UN GAOR, 6th Comm, 62nd sess, 62nd plen mtg, Agenda Item 78, UN Doc A/RES/62/61 of 6 December 2007; *Responsibility of States for Internationally Wrongful Acts*, UN GAOR, 6th Comm, 65th sess, 57th plen mtg, Agenda Item 75, UN Doc A/RES/65/19 (6 December 2010).

[573] *Gabčíkovo-Nagyamaros Project (Hungary v Slovakia) (Judgment)*, [1997] ICJ Rep 7, 35 [47].

[574] *Draft Articles* arts 1, 2; also *ILC Report* 68-9 [1]-[2].

[575] For example, the Convention on the Prevention and Punishment of the Crime of Genocide defines genocide as '… any of the following acts committed with intent to destroy, in whole or in part, a national, ethnical, racial or religious group, …'.

[576] *Materials on the Responsibility of States for Internationally Wrongful Acts*, United Nations Legislative Series Book 25, UN Doc ST/LEG/SER.B/25, (2012) 27 [2] ('*Materials on the Responsibility of States for Internationally Wrongful Acts*').

[577] *Draft Articles* art 4; see also art 5 for entities exercising governmental authority.

authority or contrary to instructions.[578] A State is normally not responsible for the conduct of private parties, but 'a State may be responsible for the effects of the conduct of private parties, if it failed to take necessary measures to prevent those effects.'[579]

IHL imposes many positive and negative obligations on States through their organs, both in armed conflict and in peacetime.[580] The obligations of interest for this discussion are those incumbent on armed forces personnel in planning and executing attacks, in particular to take precautions in verifying that targets may be legally attacked and to ensure that the chosen means and methods of warfare used in the attack are appropriate.

The Draft Articles function as a set of default rules, and can be overridden by special secondary rules applicable to specific areas of law. Article 55, entitled 'Lex specialis' provides that the Draft Articles 'do not apply where and to the extent that the conditions for the existence of an internationally wrongful act or the content or implementation of the international responsibility of a State are governed by special rules of international law.'

It is universally acknowledged that the Draft Articles are generally applicable to violations of IHL: the ILC commentary to the Draft Articles makes numerous references to violations of IHL as examples of their implementation, and international tribunals have applied some of the general principles contained in the Draft Articles when attributing responsibility for violations of IHL to a given State.[581]

IHL implements its own accountability mechanisms in relation to some primary rules, but in the context of this thesis they do not conflict in any significant way with the rules of the Draft Articles.[582] State armed forces are a State organ and are acting in that capacity when engaged in conflict on behalf of the State.[583] In specific instances there may be doubt about whether a member of the armed forces is acting in their official capacity in performing an act, but in general the use of weapons in combat is certainly

---

[578] *Draft Articles* art 7.
[579] *Materials on the Responsibility of States for Internationally Wrongful Acts*, 28 [4].
[580] Crootof, above n 565, 1356.
[581] Marco Sassòli, 'State Responsibility for Violations of International Humanitarian Law' (2002) 84 *International Review of the Red Cross* 401, 407 n 19, 408 n 20.
[582] See generally Ibid 404ff.
[583] Their actions are, therefore, within the scope of the *Draft Articles*: *ILC Report* 91 [13].

attributable to the State. The *lex specialis* of Article 91 of API 'exemplifies'[584] that principle:[585] 'A Party to the conflict … shall be responsible for all acts committed by persons forming part of its armed forces', where '[t]he armed forces of a Party to a conflict consist of all organized armed forces, groups and units which are under a command responsible to that Party for the conduct of its subordinates'[586] Thus, if a member of a State's armed forces breaches a rule of IHL in the course of an armed conflict, there is no impediment to attributing that breach to the State.[587]

7.2.1 Applying the Law of State Responsibility to AWS

The starting point for applying the law of State responsibility to AWS operations is Article 2 of the Draft Articles:

> There is an internationally wrongful act of a State when conduct consisting of an action or omission:
> (a) is attributable to the State under international law; and
> (b) constitutes a breach of an international obligation of the State.

The wrongful acts which are relevant here are breaches of a State's IHL obligations resulting from the use of weapons. The events which would underlie such breaches are taken to be common to both autonomous and manually operated weapons: for example, a State might fail to take adequate precautions in planning an attack or identifying a target, resulting in civilian deaths. As autonomy relates to how a weapon is controlled rather than the nature of the weapon itself, there is no reason to suppose the violent act, such as discharging a weapon, would differ between manual and autonomous weapons. The AWS activating its weapon would be the immediate cause of that violence, just as a soldier firing a rifle would be an immediate cause of other violent conduct.

The conduct of interest in assessing a State's responsibility for harm caused to a targeted person or object is the conduct which leads to the weapon being discharged. Attributing that conduct to the State is the step which is most at issue when an AWS is used in an attack. Perceived difficulties in attributing AWS actions to the State, or to

---

[584] Sassòli, above n 581, 405.
[585] *Hague Convention 1907* art 3 is similar.
[586] *API* art 43(1).
[587] State responsibility may also extend beyond State armed forces. For a general discussion of State responsibility for armed forces see Kalshoven, above n 567; also *Draft Articles* arts 5, 8, 11.

individuals, are the basis of the 'accountability gap' which opponents of AWS development often cite as support for prohibiting or restricting their use.

The process of attributing AWS behaviour to the State which deploys the weapon system begins with the finding, noted above, that States are responsible for the acts of members of their armed forces in the course of an armed conflict. Thus, if a member of a State's armed forces breaches the law by intentionally killing a civilian with a rifle or other manual weapon in the course of an armed conflict, there is no impediment to attributing that breach to the State.

The question is whether and how that situation might change if a weapon system used by State armed forces possesses a degree of autonomous capability. If a State organ, such as a member of the armed forces, or another party whose actions can be attributed to the State, causes an AWS to inflict harm in violation of the relevant law, then the State can be held responsible just as for a manual weapon system. If no State organ can be said to have so acted, the State is not responsible. It has already been shown in Section 7.2 that under existing IHL the State is responsible for the acts of its armed forces in selecting a specific AWS for use in an attack and activating the AWS at a given time and place. If the resulting harm is not to be attributable to the State then the chain of responsibility must be interrupted by some aspect of the AWS. Although it is not widely argued, some authors suggest that future AWS operating with very high degrees of autonomy may present challenges to the process of establishing State responsibility for AWS actions.

7.2.1.1 Lack of Human Judgement

Beard argues 'that the elusive search for individual culpability for the actions of [AWS] foreshadows fundamental problems in assigning responsibility to [S]tates for the actions of these machines.'[588] Both State and individual responsibility, according to Beard, are based on the exercise of human judgement in relation to an attack:

> Legal and political accountability flows from the judgments of commanders [in meeting the requirements of IHL]. Not only is it necessary that a human make this judgment, but

[588] Beard, above n 365, 617.

> increasingly (under the applicable policies of [S]tates), a human at the highest levels of authority must make this judgment …[589]

According to this argument, where an AWS would operate with such a degree of autonomy that no organ of the State would be applying their human judgement to control the progress of the attack, the connection required by the law of State responsibility is broken.

Beard's view appears to be based on the flawed premise that the direct control of an engagement that would be exercised by a combatant using manual weapons is the only basis of State responsibility. As Schmitt observes, 'The mere fact that a human might not be in control of a particular engagement does not mean that no human is responsible for the actions of the [AWS].'[590] The actions of attack planners, weapon developers and other personnel who contribute to an AWS being used in an attack all provide a link to the State for establishing responsibility for AWS-caused harm. In addition, the reference to human judgement mistakenly implies that there is necessarily a mental element to State responsibility. However, the ILC has stated, in the commentary to Draft Article 2, that: '[i]n the absence of any specific requirement of a mental element in terms of the primary obligation, it is only the act of a State that matters, independently of any intention.'[591] None of the obligations borne by armed forces personnel in relation to weapon use, or planning or executing an attack, contains specific requirements of a mental element, so it may be concluded that State responsibility is established by an action being done via an AWS, independently of any exercise of human judgement. The international law of State responsibility is, in this sense, a law of strict liability.

Beard's view also differs in other respects from the argument presented in this thesis. He appears to implicitly rely on a conception of advanced AWS as being something more than weapons, in suggesting that decisions about adherence to the law may be made by the AWS. He also disregards the role of human commanders in the decision to deploy an AWS at a given time and place, and the legal burden that is attached to that decision.

---

[589] Ibid 675.
[590] Schmitt, above n 305, 33.
[591] *ILC Report* 73 [10].

7.2.1.2 AWS as Agent

Other authors argue that responsibility could stop with the AWS itself. This is the position taken, implicitly or explicitly, in arguments that view an advanced AWS as something more than just a weapon: for example, an agent of the State, or other entity which would displace human combatants in a legal or moral sense as well as a physical or operational sense. This argument does not feature prominently in the legal debate, but is sometimes raised in ethical and other philosophical discussions.[592] It is generally made in reference to hypothetical advanced AWS which would exhibit some human-like intelligence and would operate with a very high degree of independence such that human operators could not reasonably be seen as exercising control over the machine.

As explained in Chapter 3, AWS are physically nothing more than weapon systems under the control of (particularly complex and capable) computer systems, not fundamentally different from other computer-controlled weapon systems. The natural course is to place them in the same legal category as other weapon systems, which means that accountability for harm done by the weapon rests with its operator. Treating a computer-controlled weapon as something which could bear legal responsibility for its actions would require an argument for an unprecedented step in the development of IHL. As Chapters 5 and 6 explain, today's IHL framework is adequate for regulating the use of AWS now and in the foreseeable future. No argument has been made which demonstrates the necessity of categorising AWS as bearers of legal obligations.

7.2.2 Sufficiency of the Law of State Responsibility

The law of State responsibility is as effective at attaching consequences to use of AWS as it is to use of other weapons in armed conflict. However, that does not mean that it is a sufficient means of ensuring accountability for violations of IHL done via AWS.[593] While the existence of autonomous capabilities in weapon systems does not reduce the efficacy of the law of State responsibility, neither does it alleviate any of the limitations which are equally applicable to incidents involving AWS and to incidents in armed conflict in general. Hammond describes the doctrinal and practical challenges faced by individuals seeking to access the institutions of State responsibility in relation to harm

[592] See, eg, John Sullins, 'When is a Robot a Moral Agent?' (2006) 6 *International Review of Information Ethics* 23.
[593] Nor is it unnecessary: Crootof, above n 565, 1364.

caused by AWS,[594] although the points made are applicable much more widely to other violations of IHL. Crootof notes that development of State responsibility institutions appears to have stagnated, and the challenges that AWS development may pose for international criminal responsibility draws attention to that limitation:

> Because they may take unpredictable action … [AWS] break the causal chain necessary for individual criminal liability and thereby highlight the relative lack of institutional means of holding [S]tates responsible for serious violations of [IHL].[595]

The extent to which AWS might inhibit processes of international criminal responsibility will be examined in the next section.

7.3 Individual Criminal Responsibility

The principle of individual criminal responsibility for war crimes has been recognised since the advent of modern IHL. It was represented in the Lieber Code of 1863[596] and the Oxford Manual of 1880. In the aftermath of the First World War, the Allied Commission on Responsibility of the Authors of the War and on Enforcement of Penalties enumerated a list of offences for further investigation and subsequent prosecution despite the lack of explicit criminalisation of serious violations of IHL in any of the relevant treaties.[597] The proposed Allied High Tribunals for the prosecution of defendants from the defeated Central Powers failed to eventuate, but not because of any principled objection to the notion of individual criminal responsibility.[598] The principle was first explicitly articulated in the Nuremberg Charter[599] of 1945 followed by the Tokyo Charter[600] in 1946. It first appeared in a multilateral treaty in the Genocide

---

[594] Daniel N Hammond, 'Autonomous Weapons and the Problem of State Accountability' (2015) 15 *Chicago Journal of International Law* 652.
[595] Crootof, above n 565, 1366.
[596] Francis Lieber, *Instructions for the Government of Armies of the United States in the Field* (Government Printing Office, 1898) [originally issued as: General Orders No 100, Adjutant General's Office, 1863].
[597] Commission on Responsibility of the Authors of the War and on Enforcement of Penalties, *Report Presented to the Preliminary Peace Conference* (29 March 1919), reprinted in (1920) 14 *American Journal of International Law* 95, 114–15.
[598] For more detail on the proposed Allied High Tribunals, see Gary J Bass, *Stay the Hand of Vengeance: The Politics of War Crimes Tribunals* (Princeton University Press, 2000); Timothy L H McCormack, "*Their* Atrocities and *Our* Misdemeanours: The Reticence of States to Try Their 'Own Nationals' for International Crimes" in Mark Lattimer and Philippe Sands (eds), *Justice for Crimes Against Humanity* (Hart Publishing, 2003) 107.
[599] *Trial of the Major War Criminals before the International Military Tribunal*, above n 51, 10.
[600] *Special Proclamation Establishing an International Military Tribunal for the Far East and Charter of the International Military Tribunal for the Far East (as amended)* (19 January 1946, amended 26 April

Convention of 1948[601] and then in the grave breaches regime common to all four Geneva Conventions of 1949.[602] States Parties to those four Conventions are obliged to criminalise the specific breaches identified as 'grave' and to investigate and either prosecute or extradite for trial those allegedly responsible for perpetration of any of the grave breaches. Many other treaties affirm the principle and it is recognised as customary law.[603] It was the basis of the war crimes prosecutions at Nuremberg and Tokyo after World War 2 as well as today in the ICTY and the ICC.

The reality of warfighting has changed considerably over that time. Where armed conflicts were once fought mostly between sovereign States and often far from major civilian centres, now non-international conflict is becoming the norm with civilians all too often caught in the fighting, or participating willingly. The means and methods of warfighting have evolved too, sometimes driving and sometimes driven by those other changes.[604] The law governing warfare has been forced to develop in response, with the rise of international criminal law being perhaps the most significant change:

> Since the laws of war are for the most part still premised on the concept of classic international armed conflict, it proved difficult to fit this law into 'modern' war crimes trials dealing with crimes committed during non-international armed conflicts. The criminal law process has therefore 'updated' the laws of war. The international criminal judge has brought the realities of modern warfare into line with the purpose of the laws of war (the prevention of unnecessary suffering and the enforcement of 'fair play').[605]

1946) in Burns H Weston and Jonathan C Carlson, *International Law & World Order: Weston's & Carlson's Basic Documents* (Brill, 2012) II.E.2.

[601] *Convention on the Prevention and Punishment of the Crime of Genocide*, opened for signature 9 December 1948, 78 UNTS 277 (entered into force 12 January 1951) art 4.

[602] *Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field*, opened for signature 12 August 1949, 75 UNTS 31 (entered into force 21 October 1950) art. 50 ('*GCI*'); *Geneva Convention for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces at Sea*, opened for signature 12 August 1949, 75 UNTS 85 (entered into force 21 October 1950) art. 51 ('*GCII*'); *Geneva Convention Relative to the Treatment of Prisoners of War*, opened for signature 12 August 1949, 75 UNTS 135 (entered into force 21 October 1950) art 130 ('*GCIII*'); *GCIV* art 147.

[603] Henckaerts and Doswald-Beck, above n 19, 551.

[604] For a discussion of how the acceleration of battlefield technology may be a response to the rise of international criminal law see Kenneth Anderson, 'The Rise of International Criminal Law: Intended and Unintended Consequences' (2009) 20 *European Journal of International Law* 331, 344.

[605] Elies van Sliedregt, *Individual Criminal Responsibility in International Law* (Oxford University Press, 2012) 3.

This section discusses the implications of the latest major development in warfighting, brought by advances in the technologies of autonomous systems.

The law covered here is that of the ICC, and in particular the focus is on war crimes as defined in Article 8(2) of the Rome Statute. It is not suggested that the ICC is the only forum in which prosecutions for offences involving AWS would be conducted. Indeed, restrictions on the Court's jurisdiction and its rules on admissibility of cases are likely to redirect many relevant cases elsewhere.[606] In particular, the ICC can try a case only if a State is 'unwilling or unable genuinely to carry out the investigation or prosecution'[607] and, in respect of war crimes, 'in particular when committed as part of a plan or policy or as part of a large-scale commission of such crimes.'[608] However, it is not feasible to survey the domestic law of each State which may be in a position to prosecute relevant crimes.

This is not a general discussion of international criminal law, or even of prosecution of war crimes in the ICC. The aim is merely to assess the effect, if any, that development of AWS will have on the efficacy of international criminal law as a means of restraining breaches of IHL. The focus is therefore on the points at which weapon autonomy may affect the progress of a prosecution.

AWS are perhaps unique among weapons in their effect on accountability mechanisms because, unlike other classes of weapons, they do not represent a particular type of physical harm which must be regulated or prohibited, as is the case with chemical weapons or blinding lasers, for example. Rather, the significance of AWS is in how they will change the behaviour of personnel belonging to the State which wields them: attack planners, weapon system operators, perhaps weapon developers, and so forth. The physical and mental steps required to plan and conduct an attack with a highly autonomous weapon are likely to differ from those required to conduct the same attack with manually operated weapon systems. Given the specificity with which the ICC's Elements of Crimes[609] are set out, and the requirement in Article 22(2) of the Rome Statute that '[t]he definition of a crime shall be strictly construed and shall not be

[606] *Rome Statute* arts 5, 17.
[607] Ibid art 17(1)(a).
[608] Ibid art 8(1).
[609] International Criminal Court, *Elements of Crimes*, Doc No ICC-ASP/1/3 and Corr.1 and RC/11 (adopted 11 June 2010) ('*Elements of Crimes*').

extended by analogy', it is prudent to determine whether the existing law adequately describes events as they are likely to occur when AWS are used in conflict.

There are two matters to address. First, would a capacity for autonomous behaviour in weapon systems threaten the ICC's ability to successfully prosecute relevant offences defined in the Rome Statute? The primary problem identified is that the requisite mental elements may be harder to satisfy when the weapon system used in perpetrating an offence is capable of operating in a highly autonomous manner.

The second question is about how to respond to that challenge. The possibilities of altering the requisite mental element and tracing accountability back to developers of AWS are discussed.

7.3.1 Prosecuting an Offence Involving an AWS

The precise effect of AWS capabilities on the ability to prosecute an offence obviously depends on the nature of the offence. Offences not directly involving the use of weapons are unlikely to be affected at all. Interest must be focussed on offences which are likely to involve use of weapon systems, of which there are many within the jurisdiction of the ICC.

Given the generality of the problem (being to assess the implications of a broad path of weapon development), it is understandable that most of the concerns that have been expressed so far relate to violation of the basic principles of IHL, notably the principles of distinction and proportionality. As it is not feasible to thoroughly examine all potentially relevant offences in a work of this scope, this section examines one relevant offence in detail.

Given that the primary focus of the thesis is the law represented in API, this part of the discussion will focus on war crimes defined in:

- Article 8(2)(b) of the Rome Statute, being '[o]ther serious violations of the laws and customs applicable in international armed conflict'.[610] These crimes include

[610] *Rome Statute* art 8(2)(b).

violations of API as well as the 1899 and 1907 Hague Conventions, the 1925 Geneva Gas Protocol and others.[611]

- Article 8(2)(e) of the Rome Statute, being '[o]ther serious violations of the laws and customs applicable in armed conflicts not of an international character'.[612]

'Other' in this case means violations of rules other than those of the 'grave breaches' of the four 1949 Geneva Conventions, which are defined in Article 8(2)(a) for international armed conflicts and of the serious violations of Article 3 common to the four 1949 Geneva Conventions defined in Article 8(2)(c) for non-international armed conflicts respectively.

The most persistent objections to development of AWS are based on the fear of their causing harm to civilians, so the offence chosen for close examination here is that defined in Articles 8(2)(b)(i) (for international armed conflict) and 8(2)(e)(i) (for non-international armed conflict), the 'War crime of intentionally directing an attack against civilians':

> Intentionally directing attacks against the civilian population as such or against individual civilians not taking direct part in hostilities;

The international offence relates to violations of the rule of civilian immunity in armed conflict, given in API Article 51(2):[613]

> The civilian population as such, as well as individual civilians, shall not be the object of attack.

It correlates to the grave breach defined in API Article 85(3)(a):

> the following acts shall be regarded as grave breaches of this Protocol, when committed wilfully, in violation of the relevant provisions of this Protocol, and causing death or serious injury to body or health:
> (a) Making the civilian population or individual civilians the object of attack;
> …

[611] Antonio Cassese, Paola Gaeta and John R W D Jones, *The Rome Statute of the International Criminal Court: A Commentary* (Oxford University Press, 2002) vol 1, 395.
[612] *Rome Statute* art 8(2)(e).
[613] Cassese, Gaeta and Jones, above n 611, 397. The equivalent provision in relation to non-international armed conflict is given in *APII* art 13(2).

Unlike the grave breach offence, the ICC offence has no result requirement. That is, the attack is not required to result in 'death or serious injury to body or health'.[614] It is sufficient that a qualifying attack is launched, regardless of how it then proceeds.

One must be wary of drawing inferences about the progress of particular cases based on general and hypothetical discussions such as the one in this thesis about the likely effects of a broad class of weapon systems. The facts of a particular situation may undermine any conclusions reached here. However, a general understanding of the interaction between the relevant law and weapon autonomy as a capability, free from the constraints of any particular implementation or application, could perhaps be usefully applied to cases involving AWS alongside consideration of case-specific facts that might alter these conclusions. To that end, the reader is invited to consider a generic example of a relevant factual situation while reading the following: for example, an autonomous UAV might be deployed on a sortie during which it targets and fires upon a group of civilians.

The next sections discuss the physical and mental elements of the offence defined in Article 8(2)(b)(i) and how their adjudication might be affected where a case concerns such an incident involving a highly autonomous weapon system.

7.3.1.1 Material Elements

The material elements of the offence 'war crime of attacking civilians' are:

1. The perpetrator directed an attack.
2. The object of the attack was a civilian population as such or individual civilians not taking direct part in hostilities.
3. The perpetrator intended the civilian population as such or individual civilians not taking direct part in hostilities to be the object of the attack.
4. The conduct took place in the context of and was associated with an international armed conflict.
5. The perpetrator was aware of factual circumstances that established the existence of an armed conflict.[615]

[614] Knut Dörmann, International Committee of the Red Cross, *Elements of War Crimes under the Rome Statute of the International Criminal Court* (Cambridge University Press, 2003) 130.

The last two elements, common to all crimes under Article 8(2)(b), are unlikely to raise any novel questions in relation to AWS. They could, hypothetically, be at issue if a weapon developer, working outside of an armed conflict, were to be charged with a war crime resulting from the activities of an AWS for which that developer is responsible. However, the possibility of that occurring is discussed and dismissed in Section 7.3.3.

The first material element to consider in detail is that the perpetrator must have 'directed an attack'. 'Attack' refers to the activity defined in Article 49(1) of API:[616] 'acts of violence against the adversary, whether in offence or in defence.'[617] There is no reason to suggest that 'directing an attack' has any different meaning in relation to attacks involving AWS than it does in relation to any other attack. In past cases, persons in a position of command over a force which used weapons against civilian targets have been convicted for acts which were consistent with their roles as commanders in relation to that attack and which will to continue to be done by commanders of forces using AWS in conflict. For example, Germain Katanga, a senior commander of the FRPI militia in the Democratic Republic of the Congo, was convicted for his role in planning an attack on civilians in the village of Bogoro and for issuing orders to that effect.[618] If some of the tasks that form part of directing or conducting the attack are encoded in software in an AWS rather than performed directly by a human, responsibility for those tasks still lies with the commander responsible for deploying the AWS. Generally, the attack planning and preparatory activities that lead to a decision to activate a weapon system at a specific time and place will be done by the same people in the same circumstances regardless of a capacity for autonomous operation in the chosen weapon systems.

The second material element is that 'the object of the attack was a civilian population as such or individual civilians not taking direct part in hostilities.' In reality, the offence covers a broad range of attacks, not just those directed solely at a civilian target. In

---

[615] *Elements of Crimes* 18.

[616] Dörmann, above n 614, 134; *Prosecutor v Kordić (Appeal Judgement)* (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-95-14/2-A, 17 December 2004) [47]; *Prosecutor v Strugar (Trial Judgement)* (International Criminal Tribunal for the Former Yugoslavia, Trial Chamber II, Case No IT-01-42-T, 31 January 2005) [282].

[617] See Section 6.1.3 for a detailed discussion of how this definition applies to AWS.

[618] *Prosecutor v Katanga (Trial Judgment)* (International Criminal Court, Trial Chamber II, Case No ICC-01/04-01/07, 7 March 2014) ('*Katanga*').

relation to the offence as committed in a non-international context, the Trial Chamber in *Katanga* found:

> that the crime may be established even if the military operation also targeted a legitimate military objective. It is important, however, to establish that the primary object of the attack was the civilian population or individual civilians. … It must be noted that indiscriminate attacks … may qualify as intentional attacks against the civilian population or individual civilians, especially where the damage caused to civilians is so great that it appears to the Chamber that the perpetrator meant to target civilian objectives. … The Chamber notes in this regard that an indiscriminate attack does not, however, automatically constitute an attack against the civilian population under article 8(2)(e)(i), as the subjective element is decisive in respect of the second case.[619]

Dörmann, writing about the international offence, also states that the scope is broad:

> It essentially encompasses attacks that are not directed against a specific military objective or combatants or attacks employing indiscriminate weapons or attacks effectuated without taking necessary precautions to spare the civilian population or individual civilians, especially *failing to seek precise information* on the objects or persons to be attacked.[620]

Dörmann also states that the same scope 'might' apply to the non-international offence.[621]

A plain language reading of 'directing' an attack may seem to preclude merely employing an indiscriminate weapon in an attack without some further qualification. It seems to imply a mental element of intent or wilfulness. In this offence, 'directing' is to be read in a strictly physical sense; the intent of the perpetrator is captured in a separate element of the crime. As Dörmann points out,[622] the ICJ made a similar statement in its Nuclear Weapons Advisory Opinion:

---

[619] *Katanga* [802] [citations omitted].
[620] Dörmann, above n 614, 131-2.
[621] Ibid 443-5.
[622] Ibid 137.

> States must never make civilians the object of attack and must consequently never use weapons that are incapable of distinguishing between civilian and military targets.[623]

Given that interpretation, the offence would relate to, relevantly, attacks prohibited by Articles 51(4)(a) and (b) of API: [624]

> Indiscriminate attacks are:
>
> (a) Those which are not directed at a specific military objective;
>
> (b) Those which employ a method or means of combat which cannot be directed at a specific military objective;
>
> …

It is straightforward to apply this element to attacks with AWS. It was noted in Section 5.4.2.2 that target selection code which runs in an AWS control system is an aspect of the nature or design of the weapon rather than a step in its use. It follows that employing an AWS which runs target selection code that cannot direct the system's weapon at a specific military objective in the circumstances of the attack amounts to attacking with an indiscriminate weapon as per the description in Article 51(4)(b) of API.

That is not to say that every attack with such a weapon system would violate a rule of IHL. A person in a position to be accused of directing an attack contrary to the law would often bear precautionary obligations under Article 57 of API as one of those who 'plan or decide upon an attack'. Included in those is the obligation to 'take all feasible precautions in the choice of means and methods of attack', where feasibility has been interpreted (in other provisions of API) as doing everything that is practically possible in the circumstances of the attack (in this case, to select weapons that will minimise expected harm to civilians).[625] If an AWS selects a civilian target despite all feasible precautions having been taken to prevent that from occurring, the offence of attacking civilians is unlikely to be made out. In a strict physical sense a civilian target would still be the object of the attack, but taking all feasible precautions indicates that the third element, that the perpetrator *intended* the civilian population as such or individual

---

[623] *Legality of the Threat or Use of Nuclear Weapons (Advisory Opinion)* [1996] ICJ Rep 226, 257 [78].
[624] Article 51(4)(c) also prohibits attacks 'which employ a method or means of combat the effects of which cannot be limited as required by this Protocol;' as autonomy relates to how a weapon is controlled rather than the type of physical harm it causes, discussion of that point is omitted.
[625] Pilloud and Pictet, above n 67, 681 [2198].

civilians not taking direct part in hostilities to be the object of the attack, would not apply.

On that third element, Dörmann notes that it 'seems to be unnecessary, but it was justified in particular by the fact that the term "intentionally" is contained in the Statute and the insertion would add more clarity.'[626] The Trial Chamber in *Katanga* also notes that it was needed 'to make a clear distinction between this crime, which proscribes violations of the principle of distinction, and acts violating the principles of proportionality and/or precaution.'[627]

7.3.1.2 Mental Elements

The Trial Chamber in *Katanga* set out the mental element of the war crime of attacking civilians in a non-international armed conflict:

> the perpetrator must have (1) intentionally directed an attack; (2) intended the civilian population or individual civilians to be the object of the attack; (3) been aware of the civilian character of the population or of civilians not taking part in hostilities; and (4) been aware of the factual circumstances that established the existence of an armed conflict.[628]

This thesis does not dispute the finding in that case, but will re-examine the mental element in order to identify points at which it might differ when the weapons used in an attack have some autonomous capability.

The default mental element applicable to crimes under the Rome Statute is given in Article 30(1):

> Unless otherwise provided, a person shall be criminally responsible and liable for punishment for a crime within the jurisdiction of the Court only if the material elements are committed with intent and knowledge.

The elements of the war crime of attacking civilians provides that the perpetrator must have '*intended* the civilian population as such or individual civilians' (emphasis added) to be the object of attack, and 'drafters of the Elements … regarded references in the Rome Statute to intent and wilfulness as consistent with the default rule in Article

[626] Dörmann, above n 614, 131.
[627] *Katanga* 298 [806] n 1851.
[628] *Katanga* 299 [808].

30.'[629] Given that Article 9(3) of the Rome Statute requires the Elements of Crimes to be consistent with the Statute, this discussion takes the conservative view that the intent specified in the Elements of Crimes for the crime of attacking civilians is expressing the same general intent as that referenced in Article 30(1).[630] There must therefore be both volitional and cognitive elements to the perpetrator's mental state, each of which, in relation to this offence, may apply to different elements of the crime.

The material element of directing an attack is an example of conduct. According to Article 30(2)(a) of the Rome Statute, 'a person has intent where … [i]n relation to conduct, that person means to engage in the conduct'.[631] This element requires merely that directing the attack is, for the perpetrator, 'voluntary conduct without regard to its purpose or result.'[632] If directing an attack may comprise typical command activities such as planning the attack, issuing orders and similar tasks, it is not problematic to apply existing law to situations involving AWS. Some complexity arises in the case of extended AWS operations involving multiple separate attacks as discussed in Section 6.1.3.3, where it may not be obvious that one individual directed each attack. In that case the Court would need to either identify the portion of the perpetrator's activities which constituted directing the particular attack in question, or consider how the element of 'directing an attack' may be extended to apply to such extended operations.

According to the Trial Chamber in *Katanga*, 'the second element of the Elements of Crimes, … "[t]he object of the attack was a civilian population as such or individual civilians not taking direct part in hostilities", must be regarded as conduct.'[633] To satisfy that element, it is only necessary for the perpetrator to have meant to make civilians the object of the attack and to have done so voluntarily, regardless of the purpose of the attack or the actual result. If that view holds true for incidents involving AWS then there are no novel challenges in applying the same rule.

---

[629] Leena Grover, 'A Call to Arms: Fundamental Dilemmas Confronting the Interpretation of Crimes in the Rome Statute of the International Criminal Court' (2010) *European Journal of International Law* 543, 577.
[630] Cassese, Gaeta and Jones, above n 611, 901; for a discussion of other views see ibid 577ff & notes.
[631] *Katanga* [805] confirms that this applies to directing an attack in the context of this crime.
[632] Cassese, Gaeta and Jones, above n 611, 913.
[633] *Katanga* 299 [806]

It is plausible, though, that the second element may be seen as a circumstance rather than as conduct in some attacks involving AWS. In the context of the Rome Statute, 'circumstances' are a broad class:

> circumstances can be any objective or subjective facts, qualities or motives with regard to the subject of the crime …, the object of the crime (such as the victim or other impaired interests), or any other modalities of the crime …[634]

Where a manually operated weapon is used, it is clear that the person or organisation directing the attack selects the object of the attack; the weapon itself plays no role until after the object has been selected. Where the weapon has some ability to select targets autonomously, the picture is not so clear. The entity directing the attack (possibly acting through subordinates) plays some part in selecting the object by activating a specific weapon system with specific targeting behaviour (the operation of which may or may not be thoroughly understood by the perpetrator) at a specific time and place. After the last human involvement ends though, the AWS targeting software takes over, identifying particular targets to be attacked and discharging the weapon. It was noted in Section 6.2 that failure to select a weapon system which would minimise the expected civilian harm from an attack is a failure to perform the precautionary obligation of API Article 57(2)(a)(ii). That would still satisfy the material element as discussed in Section 7.3.1.2, and may support a finding that the attack is indiscriminate.[635]

However, a court may decide that the choice of the particular target that is attacked was more properly attributed to the immediate cause of target selection, the AWS targeting software running after the last human involvement, than to any intentional act of the perpetrator. That software would normally have been designed, written and installed outside of the control of the perpetrator of an offence. In that case, the second element might be seen as a circumstance rather than as conduct: specifically, the circumstance would be that the AWS targeting software would cause the weapon system to select and fire upon a civilian target in the conditions under which the attack occurred. The cognitive mental component of 'knowledge' would then be applicable to the second material element. According to Article 30(3) of the Rome Statute, knowledge in relation

[634] Cassese, Gaeta and Jones, above n 611, 919.
[635] Dörmann, above n 614, 132.

to a circumstance means 'awareness that a circumstance exists'. The standard of knowledge that is required is high, appearing to require positive knowledge of the relevant circumstance:

> Under strict interpretation, Article 30(3) first sentence first alternative ICC Statute seems to suggest that the concept of intent in relation to attendant circumstances applicable and binding within the jurisdiction of the ICC should be limited to situations where the agent was actually aware of the relevant facts.[636]

If the Court were to adopt a less strict interpretation of Article 30(3) there is a possibility that it may be interpreted to cover situations of wilful blindness. In that case the knowledge requirement may be satisfied 'where the accused realised the high probability that the circumstance existed, but purposely refrained from obtaining the final confirmation because he or she wanted to be able to deny knowledge.'[637]

Actual knowledge of the operation of targeting software on an AWS would be very difficult to prove, given the intrinsic complexity of such software and the difficulty, even in a laboratory setting, of determining exactly how it will behave in a specific situation on the battlefield. Proving that knowledge would mean showing conclusively that the perpetrator had knowledge of how the software selects targets, the environment in which targets would be encountered, the civilian status of those targets and the fact that all three were such that civilians would be attacked. It seems that extending the scope to include wilful blindness would not be of much assistance if the perpetrator's understanding of those three factors still had to yield a 'high probability' of civilians being attacked.

7.3.2 Altering the Mental Element

Whether or not the Court would take that particular view of the second element of the offence of attacking civilians, it is probably true to say that, in general, the current default cognitive mental element may be difficult to satisfy when an incident involves use of a highly autonomous weapon system. Elements of crimes that amount to a

---

[636] Mohamed Elewa Badar and Sara Porro, 'Rethinking the Mental Elements in the Jurisprudence of the ICC' in Carsten Stahn (ed), *The Law and Practice of the International Criminal Court* (Oxford University Press, 2015) 665; Sarah Finnin, 'Mental Elements under Article 30 of the Rome Statute of the International Criminal Court: A Comparative Analysis' (2012) 61 *International and Comparative Law Quarterly* 325, 349.

[637] Finnin, above n 636, 350.

requirement to have precise knowledge of how an AWS will behave in a particular situation may need to be revised or reinterpreted in order to maintain their effectiveness with incidents involving AWS. At the very least, it seems there would be considerable scope for an accused to deny knowledge of how the weapon system would behave once activated (except, perhaps, for cases where it could be shown that the accused had previously operated a very similar AWS in a very similar situation; although, even that exception may not apply where the AWS is one which learns online after being activated).

Indeed, the stringent standard currently applied to the knowledge element may not accurately reflect the process by which offences might be committed when the weapon system involved operates with a high degree of autonomy. Unlike manually operated weapons, much of the behaviour of an AWS would be determined by parties other than the individual who employs it on the battlefield. Care must be taken to ensure that the elements of crimes involving the weapon system accurately reflect the involvement of the person who is the perpetrator of the offence. The purpose of autonomy in a weapon system is to have the system manage itself in some way once it is activated, and much of a highly autonomous system's behaviour would depend on events that occur after the system is activated and human involvement may have ended. The focus of its operators (such as a commander who may be in a position to 'direct an attack' against civilians) should therefore be on the precautions to take leading to activation of the weapon system, in recognition of the fact that after activation their ability to intervene to alter or prevent specific actions of the weapon may be limited. That is, their actions should focus on managing the risk associated with activating the weapon system.

To capture that change in how weapons would operate, and to retain a reasonable possibility of securing a conviction, the required standard of the mental element associated with activating an AWS could be lowered to something akin to recklessness (which is a common law concept, broadly equivalent to *dolus eventualis* in civil law systems). The ICTY has applied similarly relaxed standards in its jurisprudence:

> Since the landmark *Galić* case, the Yugoslavia Tribunal has consistently applied to the war crime of 'attacks against civilians' a lower form of intent which a reader may

> perceive … as akin to the civil law concept of *dolus eventualis* or to the common law concept of recklessness.[638]

That standard of knowledge could reasonably be identified in someone who is generally familiar with the operation of a class of AWS and with the environment of an attack. It involves someone taking 'an unreasonable and unjustifiable risk, of which he or she is aware.'[639] The cognitive component would amount to an awareness that there is some significant (and unjustifiable) risk that the AWS would engage civilian targets, and the volitional component would entail voluntary acceptance of that risk. There is room for discussion about whether the volitional component should be kept high or lowered.[640] The important step would be to reduce the cognitive component to a level that reasonably reflects the understanding that a person in charge of directing an attack would have about the behaviour of a complex AWS in the environment of the attack.

However, such a change would require a significant departure from the existing practice of the ICC regarding the mental element of crimes such as attacking civilians. The notion of *dolus eventualis* was discussed by the Preparatory Committee and appeared in an early draft of the Statute but was discarded.[641] While no precise explanation of the reasons seems to be available, it appears the intent of the drafters was to confine the mental element of Rome Statute offences to higher levels of intent.

7.3.3 Developer Responsibility

As noted previously, much of the behaviour of an AWS would depend on work done prior to the AWS being used in an attack. Development of AWS, like any modern weapon system, is a complex undertaking; large teams of people and organisations are involved at all stages of the process, from creation of the concept, through the various stages of development (with extensive interaction between military decision makers providing specifications, producers responding to those specifications, and those responsible for testing and approval for production) to production and ultimately deployment. 'Developers,' for the purposes of this section, refer broadly to people who play some significant role in defining the behaviour of an AWS, as opposed to

[638] Badar and Porro, above n 636, 650.
[639] Finnin, above n 636, 334.
[640] For a different view see Jens David Ohlin, 'The Combatant's Stance: Autonomous Weapons on the Battlefield' (2016) 92 *International Law Studies* 1.
[641] Finnin, above n 636, 345.

‘operators,’ which refers to those responsible for utilising the system in some situation during armed conflict.

Unlike developers of other weapon systems though, developers of AWS will play a role arguably not contemplated by, or not entirely consistent with, the current legal framework. Developers of AWS will exert some control over not only the range of actions the weapon system is capable of performing, but the specific actions that it in fact performs after being deployed. That influence stems from the fact that the weapon’s control system ‘steps into the shoes’ of an operator to some significant extent, reducing the need and perhaps the opportunity for intervention by weapon system operators, and placing the weapon developer somewhat ‘between’ the soldier and activation of the weapon. In general, the higher the level of autonomous operation exhibited by a weapon system, the less control will be exercised by a human operator in the field and the more control will be exercised by developers of the control software (that has reduced the level of control over the weapon system exercised by that human operator).

If the level of autonomous operation can vary from complete human control through various degrees and types of shared control to complete computer control, it is necessary to ask: in what circumstances, if any, does a system’s inbuilt capacity for autonomous operation start to mark developer activities as possibly affecting the allocation of criminal responsibility for serious violations of IHL? One requirement is, of course, that the function under computer control be the subject of legal regulation. The second indication is that the degree of control over the regulated action exercised by the system itself, and, by extension, by its developers, has increased to the point where it is not reasonable to say that a human operator alone exercises effective control; that is, the individual who would normally decide to ‘pull the trigger’ is no longer solely responsible for that decision. At that point, some part of the effective human control over the behaviour of the weapon system is built-in to the system during development rather than after deployment or at the time of the actual use of force. It would generally be the case that such development would occur prior to commencement of an armed conflict in which the weapon system is deployed.

The next question then is whether these circumstances have yet arisen, or are within the scope of development plans that have so far been made public. Weapon systems in use

today generally exhibit capabilities for autonomous operation far below what would be required to raise serious questions about developer responsibility, although many of the component technologies which may one day comprise such a weapon system either exist already or are under development. The degree of autonomy needed to raise questions about the influence and responsibility of weapon developers may not be particularly high if one considers the possibility of accessorial modes of liability (which are discussed below). Provided that other threshold requirements for criminal responsibility are met, a significant possibility exists that weapon developers may exercise a sufficient degree of control over proscribed acts by the systems they create such that they could be held liable for serious violations of the law. The proviso that other threshold requirements are met is important here. The next sections consider some of the requisite elements for criminal responsibility in the context of AWS which are relevant to the question of developer responsibility.

7.3.3.1 Threshold Requirement of the Existence of an Armed Conflict

It is axiomatic that a war crime can only be perpetrated in the context of an armed conflict. The application of IHL is only triggered when an armed conflict exists, so it follows *ab initio* that there cannot be a serious violation of that law in the absence of the specific context in which the law applies. In many war crimes trials, it is obvious that the charges against the accused involved the perpetration of alleged acts in the context of an armed conflict. In other trials, the prosecution has a more challenging burden of establishing the existence of an armed conflict. The canonical formulation of the threshold requirement, repeatedly cited by international criminal courts and tribunals as well as in national criminal trial proceedings, was articulated by the Appeals Chamber of the ICTY in that Tribunal's very first trial against Duško Tadić when it held that:

> an armed conflict exists whenever there is a resort to armed force between States or protracted armed violence between governmental authorities and organized armed groups or between such groups within a State. [IHL] applies from the initiation of such armed conflicts and extends beyond the cessation of hostilities until a general conclusion of peace is reached; or, in the case of internal conflicts, a peaceful settlement is achieved. Until that moment, [IHL] continues to apply …[642]

[642] *Tadić* [70].

The ICTY Appeals Chamber decision, reflecting both treaty and customary international law, distinguished between international and non-international armed conflicts and articulated a different threshold test for these two types of armed conflict, with the latter having a significantly higher threshold. Considerable effort is often expended in international criminal trials to determine whether or not an armed conflict exists and, if so, whether that armed conflict is international or non-international. In cases involving war crimes under the Rome Statute, the Prosecutor must prove the existence either of an international or of a non-international armed conflict as a requisite threshold element of any particular alleged war crime.

The critical issue in relation to developer responsibility arises from the fact that much weapon development work is done outside of the context of an armed conflict, whether in a commercial context by a corporate entity or by a governmental entity during times of peace, perhaps derived from more fundamental research done in an academic context by university researchers.

It is possible that, in a protracted armed conflict lasting many years, a weapon system may be developed and deployed during the conduct of the armed conflict. Given the long lead time for development of AWS, that scenario, while possible (US forces have already been deployed in Afghanistan for approximately 15 years, for example), will be the exception rather than the norm. In exceptional circumstances then, where all other threshold requirements are met, it may well be that developers can be prosecuted for war crimes because the threshold contextual element will be satisfied. But what of the more likely scenario: where the development of the system has concluded well before the armed conflict in which the weapon system is deployed has commenced, such that the temporal distance of the developers' conduct from the context of armed conflict results in a serious obstacle to individual criminal liability?

The Elements of Crimes includes for every war crime a requirement that the alleged conduct 'took place in the context of and was associated with an [international or non-international, depending upon the precise provision of the Statute,] armed conflict.'[643] *Prima facie*, conduct done prior to the commencement of an armed conflict fails the 'context of' test and cannot constitute a war crime, despite that development of a

[643] *Elements of Crimes* 13.

weapon system in preparation for deployment in an armed conflict when and if the need arises would be very likely to satisfy the 'associated with' test.[644]

However, one could ask whether developer activity is properly confined to the development phase, or whether, to the extent that it cannot be modified later by an operator, it is best located where the AWS's programmed behaviour crystallises during an armed conflict. While that property of AWS may support a general notion of developer accountability for AWS behaviour, it does not bypass IHL's threshold requirement that the developer's acts must have been done in the context of an armed conflict. At best, it is an argument that developers may be held responsible under a body of law which is active at the time of the development activity, such as a relevant product liability regime in the domestic law of the developers' State.

7.3.3.2 Physical Perpetration and Accessorial Modes of Liability

In addition to proving all the constituent elements of a particular charge, the prosecution must also prove the criminal liability of the accused by reference to one or other of the enumerated grounds of individual criminal responsibility in Article 25 of the Rome Statute. In this respect, the nature of development activities is significant and several observations can be made.

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

[644] Knut Dörmann, 'Preparatory Commission for the International Criminal Court: The Elements of War Crimes' (2000) 82 *International Review of the Red Cross* 771, 779-80.

Another hypothetical scenario, as applicable to conventional weapons and to weapons of mass destruction as it will be to AWS, should also be mentioned. Where evidence exists of an explicit common criminal purpose in which weapon developers intentionally develop a weapon system for a known illegal purpose, the fact that certain actions were undertaken prior to the commencement of an armed conflict in which the weapon was subsequently deployed and the intended crimes perpetrated will not preclude prosecution of the developers' conduct. Such scenarios will, hopefully, be rare but the international community has witnessed ruthless political regimes willing to utilise all the resources available to them, including the co-opting of scientific and technical expertise, to produce a capability which they have not hesitated to use, including against their own civilian populations. The doctrine of common criminal purpose as a mode of individual criminal liability clearly envisages acts of preparation to contribute to or to further an agreed criminal activity or criminal purpose.[645]

Second, to the extent that weapon developers may be said to instigate or control some action of an AWS, they do so through the control software and other subsystems that they design, and the degree of control exercised by a developer depends on the degree of autonomy exhibited by the weapon system in respect of an action. Insofar as it shapes the relationship between operators and weapon systems and excludes operators from intervening to exercise direct physical control over proscribed acts, the degree of system autonomy would also determine the potential for developers to be held accountable for actions instigated by those systems and is the likely mode for finding liability.

As explained above, as the degree of autonomy exhibited by a system increases, operators and developers move through various forms of shared control of and responsibility for the behaviour of the system. As long as an operator remains able to intervene in the functioning of the weapon system in some significant way, control is shared between developer and operator. It seems that at the very least an operator will always remain able to select the time, place and other circumstances in which an AWS would be activated. Hypothetically though, at some point yet to be defined it may be that system developers could occupy a position of effective control over the actions taken by a system in some situation encountered during a conflict, such that soldiers and

[645] *Rome Statute* art 25(3)(d).

their commanders may be effectively excluded and cannot be reasonably said to instigate those actions or be in a position to intervene. It therefore seems necessary to examine two modes of liability: where a developer via an AWS's control software plays some contributory, but not decisive, role in instigating and carrying out a course of action, and where the system exercises such control that soldiers in the field cannot be said to have meaningfully contributed. Perhaps the most likely and appropriate alleged ground of liability where any degree of shared control is present would be pursuant to Article 25(3)(c) as an accessory who 'aids, abets or otherwise assists in [the] commission [of the crime] … including providing the means for its commission.'

7.3.3.3 International Jurisprudence on *Ex Ante* Aiding and Abetting

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

There does not appear to be any jurisprudence from international war crimes trials to support the notion of individual criminal responsibility for war crimes where an accused's acts have occurred prior to the commencement of an armed conflict. At best, there are broad general statements as to hypothetical possibilities of accessorial liability for acts committed prior to the principal offence. By way of *obiter dicta*, the Appeals Chamber of the ICTY in the *Blaškić* judgment affirmed the possibility that: '[T]he *actus reus* of aiding and abetting a crime may occur before, during, or after the principal crime has been perpetrated, and that the location at which the *actus reus* takes place may be removed from the location of the principal crime.'[646] However, that was a broad statement of principle and one not tested by the facts in the Blaškić trial (it was decided that aiding and abetting was not the appropriate ground of criminal liability for the charges against Blaškić). At any rate, the nature of the allegations against the accused

[646] *Prosecutor v Blaškić (Judgement)* (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-95-14-A, 29 July 2004) [48] ('*Blaškić*').

did not involve acts of preparation prior to the outbreak of hostilities in the Lašva Valley region of Bosnia-Herzegovina.

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

> The Accused may aid and abet at one or more of the "planning, preparation or execution" stages of the crime or underlying offence. The lending of practical assistance, encouragement or moral support may occur before, during or after the crime or underlying offence occurs. The *actus reus* of aiding and abetting does not require "specific direction." No evidence of a plan or agreement between the aider and abettor and the perpetrator is required, except in cases of *ex post facto* aiding and abetting where "at the time of planning or execution of the crime, a prior agreement exists between the principal and the person who subsequently aids and abets the principal."[647]

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

In the absence of jurisprudence considering actual *ex ante* conduct, one can only speculate about the hypothetical determinations of a court. But given the formulation in the Elements of Crimes of the requisite contextual existence of an armed conflict for all war crimes; and Article 22(2) of the Rome Statute which requires that 'the definition of a crime shall be strictly construed and shall not be extended by analogy' and that 'in case of ambiguity, the definition shall be interpreted in favour of the person being

[647] *Prosecutor v Taylor (Judgement)* (Special Court for Sierra Leone, Trial Chamber II, Case No SCSL-03-01-T, 18 May 2012) [484] (citations omitted) ('*Taylor*').

investigated, prosecuted or convicted,' any chamber of the ICC asked to decide the question in the future is likely to read the requisite elements narrowly.

7.3.3.4 *Mens Rea* Requirements

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

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

> [I]t is not necessary that the aider and abettor should know the precise crime that was intended and which in the event was committed. If he is aware that one of a number of crimes will probably be committed, and one of those crimes is in fact committed, he has intended to facilitate the commission of that crime, and is guilty as an aider and abettor.[648]

The Appeals Chamber of the ICTY in *Blaškić* later concurred with that statement,[649] which seems to indicate that a developer need not be a party to detailed and specific plans to employ an AWS in a criminal act in order to be found guilty of aiding and abetting. That condition would likely be readily satisfied by a competent developer who imparts behaviour to a weapon system which a reasonable person would be expected to see as criminal in nature. However, there would be two significant challenges for prosecutors to overcome.

First, existing case law relates only to situations in which the accused is found to be aware of a criminal intent which is held by the physical perpetrator *at the time that the accused engages in conduct which assists the perpetrator*. The members of Anto Furundžija's anti-terrorist unit possessed criminal intent and were committing criminal acts while Furundžija was present and rendering assistance. Similarly Charles Taylor

---

[648] *Prosecutor v Furundžija (Judgement)* (International Criminal Tribunal for the Former Yugoslavia, Trial Chamber, Case No IT-95-17/1-T, 10 December 1998) [246].
[649] *Blaškić* [50].

knew of a criminal intent which had been formed (and acted on) by his soldiers at the time he was providing supplies. Earlier, in the post-Second World War *Zyklon B* case,[650] German industrialists were found guilty of war crimes for supplying poison gas which was used to exterminate Allied nationals interned in concentration camps, in the knowledge that the gas was to be so used, in violation of Article 46 of the *Hague Convention 1907*. In that case, the intent of the SS officers who exterminated the internees had been formed at the time the gas was supplied to them, and the industrialists who supplied it were aware of that intent.[651] There is no suggestion in the jurisprudence that the requisite mental state may extend to a developer's expectation that their work would be of assistance to a future perpetrator who may later form a criminal intent which did not exist at the time the developer completed their work.

Second, Article 25(3)(c) of the Rome Statute also specifies a special mental element which applies in relation to aiding and abetting. This is in addition to the general requirements of Article 30 that a developer's conduct would need to be done 'for the purpose of facilitating the commission' of a crime within the jurisdiction of the Court. Prosecutors would be required to demonstrate that a weapon developer acts not only with awareness of the eventual physical perpetrator's intention to commit the crime and with the knowledge that his or her conduct would assist in perpetration of the offence,[652] but also that s/he acts for the purpose of facilitating the crime, rather than for the sole purpose of, for example, selling weapon systems for profit[653] or achieving another legal military objective.

Equally formidable challenges would also face a prosecutor seeking to show that a developer of an AWS committed a crime 'through another person, regardless of whether that other person is criminally responsible,' though the picture is less clear. Despite

---

[650] 'Case No 9: The Zyklon B Case: Trial of Bruno Tesch and Two Others (British Military Court, Hamburg)' in United Nations War Crimes Commission, *Law Reports of Trials of War Criminals* (His Majesty's Stationery Office, 1947) vol 1, 93.

[651] In a similar vein, *The I.G Farben Trial* dealt with the prosecution of the leaders of the corporation I.G. Farben Industrie A.G. for 'the use of poison gas, supplied by Farben, in the extermination of inmates of concentration camps' as a crime against humanity: 'Case No 57: The I G Farben Trial: Trial of Carl Krauch and Twenty-Two Others (United States Military Tribunal, Nuremberg)' in United Nations War Crimes Commission, *Law Reports of Trials of War Criminals* (His Majesty's Stationery Office, 1949) vol 10, 5.

[652] *Taylor* [487].

[653] Robert Cryer et al, *An Introduction to International Criminal Law and Procedure* (Cambridge University Press, 2nd ed, 2010) 312.

recent signs of judicial opposition to the concept,[654] the ICC has generally favoured the 'control of the crime' approach to adjudicating on commission through another person. If that approach is adhered to, it would seem that a developer of an AWS could, with criminal intent, produce a weapon system which performs a proscribed act without an instruction to that effect from its operator. However, similar problems exist here as in relation to aiding and abetting: the crime which the developer purportedly intends to commit, or knows will be committed in the ordinary course of events, can only occur later during an armed conflict. At a minimum, prosecutors would need to demonstrate that the developer understood that the weapon system would behave in an illegal manner, for example, to be incapable of distinguishing between military personnel and civilians by targeting and killing as yet unidentified civilian victims. Evidence of a developer intentionally building in to the weapon system a mechanism for targeting a specific individual known by the developer to be a civilian immune from attack would, of course, satisfy the *mens rea* requirements, but that type of evidence of specific intent will likely be rare.

7.3.3.5 Attribution of Responsibility

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

7.3.3.6 Conclusions on Responsibility of Developers

The likely scenario that developers will complete their involvement in a weapon system prior to the commencement of an armed conflict raises real concerns about the satisfaction of various requisite elements of any alleged war crime. The first potential

[654] See, e.g., *Prosecutor v Chui (Judgment Pursuant to Article 74 of the Statute)* (International Criminal Court, Trial Chamber II, Case No ICC-01/04–02/12, 18 December 2012) [4]–[30] (Concurring Opinion of Judge Christine Van den Wyngaert).

obstacle is the threshold requirement that the alleged acts take place 'in the context of an armed conflict.' That particular threshold requirement may amount to giving de facto immunity to weapon developers for subsequent war crimes. A prosecutor may seek to argue that the culmination of acts perpetrated prior to the commencement of an armed conflict 'in the [subsequent] context of an armed conflict' is sufficient to satisfy the threshold element, but that is a novel argument and there is no guarantee that it would be endorsed by judicial decision makers. It is true that international jurisprudence on the scope of aiding and abetting as a mode of liability envisages the sufficiency of *ex ante* acts of preparation. However, currently there does not appear to be any international jurisprudence to substantiate the proposition that acts of preparation to aid and abet a subsequent war crime can occur prior to the commencement of an armed conflict.

Another obstacle to weapon developer liability is the requisite *mens rea* requirement for aiding and abetting. A weapon developer accused of aiding and abetting would typically render their assistance prior to an armed conflict, whereas a criminal intent on the part of the physical perpetrator would typically emerge during the conflict. Bridging that temporal gap would require significant development of the law beyond the current jurisprudence.

Even if each of these requisite elements could be established in respect of a weapon developer, there is the further practical, perhaps practically insurmountable, obstacle that to speak of 'the developer' is to misrepresent the realities of the development process. A sophisticated AWS will not be developed by a single individual, but by many teams of developers in many organisations working on a multitude of subsystems with complex interdependencies. Attempting to identify an individual most responsible for subsequent behaviour of a deployed weapon system that constitutes a war crime may simply be too difficult for the purposes of initiating trial proceedings.

The existence of those obstacles may point to a need to develop the current normative framework to better accommodate the expanded role of weapon developers. The threshold requirement that relevant acts occur 'in the context of an armed conflict' could be amended to explicitly (or could be interpreted to implicitly) include acts of preparation prior to the commencement of the armed conflict provided that the completion of the crime occurred in the relevant context. *Mens rea* requirements for

developers to aid and abet subsequent offences could also be clarified to cover the scenario under consideration here.

7.3.4 Command or Superior Responsibility

The notion that military commanders may be held responsible for transgressions by their subordinates has a long history,[655] but the first clear treaty reference is in API. Articles 86 and 87 require commanders to prevent and suppress breaches of IHL, and commanders may be held responsible 'if they knew, or had information which should have enabled them to conclude in the circumstances at the time,'[656] that the breach was being, or would be, committed. Article 28(a) of the Rome Statute extends individual criminal responsibility to commanders who fail in their duties to prevent and suppress breaches. The Trial Chamber in the *Čelebići* case also affirmed that 'the principle of individual criminal responsibility of superiors for failure to prevent or repress the crimes committed by subordinates forms part of customary international law'.[657]

Kastan posits that command responsibility will take on renewed importance in respect of breaches of IHL involving AWS.[658] However, his argument is based on the premise that, when a weapon system operates with a sufficiently high level of autonomy, there is nobody in the position of 'operator' of the weapon system, and a commander would interact directly with the AWS as with a subordinate (human).[659] That view conflicts with the position taken in this thesis, of an AWS as merely a weapon system, albeit one with extensive capabilities, operated by human forces.

Command responsibility requires first that a crime be committed by a subordinate over whom the commander exercises effective control.[660] The novel effect of autonomy in a weapon system is how it alters the relationship between the weapon and its direct operators, and consequently the novel accountability questions raised by use of AWS focus on the criminality of actions by the operator. It is not apparent that the operation of the law of command responsibility would be directly impacted by development of

---

[655] Cryer, above n 653, 387-9.
[656] *API* art 86(2).
[657] *Prosecutor v Mucić (Judgement)* (International Criminal Tribunal for the Former Yugoslavia, Trial Chamber, Case No IT-96-21-T, 16 November 1998) [343] ('*Čelebići*').
[658] Kastan, above n 327.
[659] Kastan, above n 327, 66.
[660] *Prosecutor v Bemba Gombo (Judgment)* (International Criminal Court, Trial Chamber III, Case No ICC-01/05-01/08, 21 March 2016) [170], [175] n 389.

highly autonomous weapons. The relationship between commander and subordinates is unlikely to change, and the IHL obligations borne by both commander and subordinate are unaltered, albeit that they may require different steps to fulfil.

In situations of very high levels of weapon autonomy, the operator who deploys the weapon system may well be excused of individual responsibility for lack of effective control over the behaviour of the weapon system. However, the commander who has called for the deployment of the weapon system will be on notice once the system behaves in such a manner as to have resulted in a serious violation of IHL. At that point of the commander's awareness, the doctrine of command responsibility requires the commander to take 'all necessary and reasonable measures' to prevent further crimes, as per Rome Statute Article 28(a)(ii). Such measures may well extend to ordering the standing down of the system pending re-programming/adjustment/tweaking to prevent recurrent offences.

7.4 Other Forms of Accountability

The above discussions of State responsibility and individual criminal responsibility for violations of IHL show that those existing regimes will continue to function relatively untroubled by increasing military use of autonomous systems, at least according to the development plans that have been made public so far. That is not to say they are entirely sufficient to reign in violations of IHL, though. Broadly speaking, accountability regimes are instruments to secure control of public power and protect the rule of law, and as such they must evolve along with shifts in governance and public authority.[661] In the relatively narrow and specialised context of an armed conflict, the available regimes must be matched to the exercise of military power and the use of force. Where that matching fails, accountability gaps emerge, and with them the risks of unchecked exercise of power.[662] The fear of an accountability gap due to automation of combat operations is one example of that concern, representing significant changes in the practice and risks of warfighting due to:

- The influence that weapon developers may have over actions taken on the battlefield;

[661] Deirdre Curtin and André Nollkaemper, 'Conceptualizing Accountability in Public International Law' (2005) XXXVI *Netherlands Yearbook of International Law* 3, 6, 9.
[662] Ibid 6.

- The dual-use nature of many of the technologies involved and their accessibility to non-State actors;
- Perhaps, the suitability of AWS for engaging in persistent low-level conflict;[663]
- The risk of (further) alienating local populations through automating the infliction of violence wherein combatants are unable to confront their enemy.[664]

Increasing use of AWS may also be seen as one of many changes in the practice of warfare over recent decades that have called into question the efficacy of existing accountability regimes, and which underscore the need to examine accountability more broadly. This section surveys some alternative means of ensuring accountability for the use of force in armed conflict, in particular force applied via AWS. The survey is divided into two sections. The first deals with retrospective accountability: an approach that is designed to address past wrongs and thereby, hopefully, guide future behaviour. The second section deals with forms of prospective accountability, being forward-looking, preventative systems that 'inform and remind individual and institutional actors that they are expected to behave in certain ways.'[665]

7.4.1 Retrospective Accountability

If serious accountability problems are to eventually arise through continued development of AWS, they are likely to stem from the unique ways in which weapon system operators, developers and commanders interact with and exert control over the weapon systems. Two factors suggest that collective modes of accountability may appropriately account for breaches with AWS: that control over the behaviour of weapons may be split in ways not contemplated thus far in the development of IHL, and the possible difficulty of proving a sufficiently high level of culpability in any one individual in relation to a breach. That is not to say that large scale forms of collective responsibility such as system criminality,[666] put forward as the form of criminality

[663] George W Casey Jr, "America's Army In an Era Of Persistent Conflict" (2008) 58 *Army* 19.
[664] Douglas A Pryer, 'The Rise of the Machines: Why Increasingly "Perfect" Weapons Help to Perpetuate our Wars and Endanger Our Nation' [2013] (March-April) *Military Review* 14.
[665] Deborah G Johnson, 'Software Agents, Anticipatory Ethics, and Accountability' in Gary E Marchant, Braden R Allenby and Joseph R Herkert (eds), *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight: The Pacing Problem* (Springer, 2011) 61, 70.
[666] van Sliedregt, above n 605, 20; B V A Röling, 'Aspects of the Criminal Responsibility for Violations of the Laws of War' in Antonio Cassese (ed), *The New Humanitarian Law of Armed Conflict* (Napoli: Editoriale Scientifica, 1979) 138, 203; Mark A Drumbl, 'Accountability for System Criminality' (2010) 8 *Santa Clara Journal of International Law* 373.

behind mass atrocities, would necessarily be at issue with a development as 'low level' as a new weapon technology. It does, however, seem prudent to allow for the possibility that a proscribed action by an AWS in armed conflict would be likely to result from acts or omissions of several people, possibly with a mixture of motivations and levels of culpability.

7.4.1.1 War Torts

Crootof has proposed the recognition of 'war torts', being 'serious violations of [IHL] that give rise to [S]tate responsibility'.[667] The proposal is a recognition that States are the entities best placed to ensure that AWS are designed and used in compliance with IHL, and generally have the resources to make reparations for harm caused by AWS, but the existing law of State responsibility is 'rarely discussed or enforced with regard to serious violations of [IHL]'.[668] Individual criminal responsibility, on the other hand, is based on the premise that serious violations of IHL flow from wilful acts on individuals. That condition may be missing from situations involving AWS, where the harm may be traced back to a failure of very complex code operating unpredictably in a chaotic environment rather than to an intentional act by an individual.[669] What is needed, according to Crootof, is 'a separate term, a different name, for violations where there is fault, regardless of whether or not there is also guilt.'[670] A war torts regime would be something akin to the United States Alien Tort Statute,[671] under which individuals and other entities (but generally not States) can be held accountable for international torts, except it would allow States to be held accountable in the international sphere.

While such a regime would address some potential problems, such as the difficulty of establishing individual criminal intent, it is premised on a strong view of the degree of independence that may be enjoyed by highly autonomous weapon systems. It also assumes the AWS may become 'uniquely unpredictable': 'the sheer complexity of an [AWS]'s program may make it impossible for human beings to predict how it will act with complete accuracy — or even reconstruct why it acted a certain way after the

[667] Crootof, above n 565, 1348.
[668] Ibid 1364.
[669] Ibid 1377, 1385.
[670] Ibid 1386.
[671] *Alien Tort Statute*, 28 USC § 135 (2012).

fact.'[672] As discussed in Chapter 6, this thesis takes a stricter view, that where the legality of the behaviour of an AWS cannot be predicted with a high degree of confidence, the AWS cannot be legally deployed, and legal sanctions would apply to the (human) act of deploying the weapon system.

7.4.2 Prospective Accountability

There is perhaps more scope, and more potential benefit, in promoting measures supporting prospective accountability in relation to use of AWS than in focussing on redressing past wrongs. When weapons may operate in circumstances where no human is able to intervene at short notice to correct problems such as selection of illegal targets, prospective accountability is critical. When the complexity of the weapon system and its behaviour may inhibit efforts to identify the cause of a fault, be it human or technical, the focus should be on measures to ensure that problems are prevented from occurring at all.

While accountability is often understood as a retrospective process, that view is being challenged by calls for more forward-looking, participative processes.[673] In this matter, IHL is perhaps more progressive than many other areas of international law. The Geneva Conventions already provide a range of prospective accountability measures, applicable in peace time as well as armed conflict, designed to repress abuses and infractions of IHL.[674]

Most broadly, States Parties must disseminate the text of the Conventions (and the Additional Protocols) 'as widely as possible' and 'include the study thereof in their programmes of military instruction'.[675] This is fundamental to preventing any breach of IHL; members of armed forces must know the law if they are to uphold it. To the extent that legally regulated activities (for example, target selection) are encoded in AWS control systems, the purpose of the obligation suggests it would extend to disseminating

---

[672] Crootof, above n 565, 1373.
[673] Curtin and Nollkaemper, above n 661, 8.
[674] For a longer discussion of these and other accountability measures see: Helen Durham and Eve Massingham, 'Moving from the Mechanics of Accountability to a Culture of Accountability: What More Can be Done in Addition to Prosecuting War Crimes?' in Jadranka Petrovic (ed), *Accountability for Violations of International Humanitarian Law: Essays in Honour of Tim McCormack* (Routledge, 2016) 267.
[675] *GCI* art 47; *GCII* art 48; *GCIII* art 127; *GCIV* art 144; *API* art 83; *APII* art 19.

relevant law to weapon developers and others who are in a position to define the behaviour of AWS.

Building on that, States Parties must 'ensure that legal advisers are available, when necessary, to advise military commanders at the appropriate level on the application of the Conventions and this Protocol'.[676] This requirement recognises the fact that 'even when limited to uncontroversial rules directly applicable by military commanders in the field, the law of armed conflict is becoming increasingly complex, detailed and extensive.'[677]

Two other measures are more directly relevant to weapon use. First, States must ensure that precautions are taken in planning and conducting attacks. Those precautions were discussed in detail earlier, principally in Chapter 6, and will not be covered again here.

7.4.2.1 Weapon Reviews

The other major preventative measure is the conventional, and perhaps, to an extent, customary, requirement to review new weapon systems for compatibility with a State's legal obligations. That requirement seeks to suppress breaches of IHL by preventing the development and use of weapons which would be incompatible with rules governing the nature and use of weapons. Neither API nor general practice by non-party States mandate any particular process for conducting weapon reviews, instead leaving States to develop their own internal procedures.[678] It is not clear how many States have done so,[679] and the extent of compliance with this provision is, unfortunately, open to question,[680] despite its importance in light of the leaps made in weapon technologies in recent years.

---

[676] *API* art 82

[677] Jean de Preux, 'Protocol I – Article 82 – Legal Advisers in Armed Forces' in Sandoz, Swinarski and Zimmermann, above n 2, 947, 948 [3342].

[678] de Preux, above n 345, 424 [1470], 428 [1482].

[679] For a list of States known to have weapon review procedures: International Committee of the Red Cross, above n 270, 934 n 8; Vincent Boulanin, 'Implementing Article 36 Weapon Reviews in the Light of Increasing Autonomy in Weapon Systems' (SIPRI Insights on Peace and Security No 2015/1, Stockholm International Peace Research Institute, November 2015) Annex A; Boothby, above n 64, 341..

[680] Isabelle Daoust, Robin Coupland and Rikke Ishoey, 'New Wars, New Weapons? The Obligation of States to Assess the Legality of Means and Methods of Warfare' (2002) 84 *International Review of the Red Cross* 345, 354.

The importance, and difficulty, of reviewing AWS is prominent among the concerns being discussed at the CCW meetings.[681] Reviews of AWS, like other weapons, are critical for legal compliance. In light of the attention that the matter receives at the CCW discussions, it may seem that deeper discussion is required here.[682] However, the novel challenges of conducting those reviews are more technical than legal, although they arguably require lawyers to gain some understanding of the engineering issues involved.[683] Section 4.4.1.1 discussed the difficulty of understanding the operation of very complex software systems, in particular where that software will run unattended in a changing and perhaps poorly understood combat environment.[684]

Access to the appropriate expertise to conduct such a test is therefore likely to be an impediment for some States. In particular, States which intend to purchase AWS from exporting States, and do not possess the relevant expertise themselves, may face challenges in conducting a review that thoroughly tests an AWS for compatibility with their own legal obligations. Some States have made progress towards establishing standards applicable to their internal processes: for example, the US Department of Defense 'Unmanned Systems Safety Guide for DoD Acquisition'[685] sets out detailed safety-related considerations applicable when purchasing unmanned systems, although it is not a detailed guide to evaluation of AWS. By far the greatest problem, though, is in testing the software that will run in AWS control systems. In theory, one would wish to verify that an AWS control system will issue the correct instructions to its attached weapon in all scenarios in which it will be deployed on the battlefield. That would mean identifying and testing all possible combinations of variables which the AWS would encounter: target appearances, types and behaviours; enemy actions in defence and offence; environmental conditions including weather, terrain, vegetation and artificial structures; appearance and behaviour of other weapon systems; all the possible

[681] See Section 2.4.2 for an outline.
[682] For an example of a hypothetical review procedure applied to AWS see Michael W Meier, 'Lethal Autonomous Weapons Systems (LAWS): Conducting a Comprehensive Weapons Review' (2016) 30 *Temple International and Comparative Law Journal* 119, 126-32.
[683] Backstrom and Henderson, above n 132, 507.
[684] For a discussion of challenges facing reviewers of weapons based on new technologies in general see Damian P Copeland, 'Legal Review of New Technology Weapons' in Nasu and McLaughlin, above n 478, 43.
[685] 'Unmanned Systems Safety Guide for DoD Acquisition' (United States Department of Defense, 27 June 2007) <https://acc.dau.mil/adl/en-US/683704/file/75176/Unmanned%20Systems%20Safety%20Guide%20forDOD%20Acquisition,27June%202007.pdf>.

instructions the AWS might receive from its operators; and so on. For even a moderately complex AWS operating in a moderately complex environment, the number of combinations to test will be infeasibly high. Nor is this problem specific to AWS; it is a well-known challenge in software development and means that software testing can only minimise, not entirely eliminate, errors:

> exhaustive input testing is impossible. Two implications of this are that (1) you cannot test a program to guarantee that it is error free and (2) a fundamental consideration in program testing is one of economics. That is, since exhaustive testing is out of the question, the objective should be to maximize the yield on the testing investment by maximizing the number of errors found by a finite number of test cases.[686]

There are three questions which States wishing to acquire and use AWS will have to address. The first is somewhat urgent, given that software sufficiently complex to be beyond exhaustive testing already exists: if a weapon system is too complex to thoroughly review for compliance with IHL, can it legally be used in armed conflict? The answer must be a clear 'no'; if a State is 'under an obligation to determine' the legality of weapons, surely a high level of certainty about their behaviour is required. Existing weapon systems also run software which is certainly too complex to have been exhaustively tested, which indicates that the problem can perhaps be overcome. However, existing weapons are generally designed to perform only one or a small range of narrow, well-defined operations. As such, it may be that the number of combinations of input variables they are expected to face is small enough to be manageable for testing purposes. As the range of capabilities built into an AWS expands, the number of combinations of inputs to be tested would increase exponentially and quickly become large enough to make exhaustive testing infeasible.

That leads into the second question: given that complete certainty about the operation of complex software cannot feasibly be achieved, how can an acceptable level of certainty be defined for testing purposes? Neither the Geneva Conventions nor the Additional Protocols mandate an objective minimum acceptable level of certainty about weapon system operation; there is no particular 'passing score' which must be achieved. However, the fundamental importance of the principles at issue in operating weapon

[686] Glenford J Myers, *The Art of Software Testing* (John Wiley & Sons, 2nd ed, 2004) 13.

systems (principally, adherence to distinction and proportionality requirements) and the strong terms in which they are couched in the Conventions and Protocols indicate that a high level of reliability is needed and, of equal importance, it must be known that a high level of reliability has been achieved.

IHL has developed in an environment where the role of a weapon is simple and contained, and one can be confident that the way the weapon behaves during development and testing is the way it will behave when used in a conflict. Where a weapon must respond only to a known set of inputs that are well understood (the operator selects a target for the weapon, the operator pulls the trigger, etc.), the resultant uncertainty about the weapon's behaviour depends only on the degree of understanding about those known inputs. That is, the variables which affect the weapon's behaviour are known, although there may be uncertainty about their values. That uncertainty can be measured, and a level of confidence about the weapon's behaviour reached. If unexpected situations arise during conflict, it is assumed that the human soldiers who operate weapons will directly face that new, unknown factor and manage whatever risks it presents; the weapon's role is to carry out the human's instructions, the full range of which can be known and tested in advance.

Where an AWS is expected to operate unattended in a combat situation, it must directly respond to any such unexpected situations without human assistance. That changes not only the degree of uncertainty which the weapon system faces, but the type of uncertainty. Where the variable factors that may affect a weapon system's operation are poorly understood, such as would be expected when an AWS is operating unattended in a chaotic, contested combat environment, the sources of uncertainty are themselves uncertain. That is, the variables which affect the weapon's behaviour are not well understood. For example, an adversary might attempt to interfere with the AWS by unanticipated means, or a civilian within range of the weapon might behave in an unexpected manner, or some other entirely 'unknown unknown' might occur. Consequently, those variables cannot be studied and it is much more difficult to reach a level of confidence about how the weapon will behave. Not only is there a margin of error in how the weapon system might respond to previously known inputs, but there is the possibility of unexpected and therefore untested influences on an AWS's behaviour.

The implication of that for weapon reviewers is that it becomes very difficult to attach a level of confidence to a weapon system's behaviour when the environment in which it is to operate is not well controlled. Kalmanovitz, at the 2016 CCW Meeting of Experts, described the problem of uncertainty in AWS behaviour in relation to proportionality judgments:

> As many experts have argued, proportionality requires that this uncertainty be bounded — it must be probabilistic, not second-order uncertainty (not uncertainty about the uncertainty). It must be possible to determine probabilistically the machine's range of action, because only on that basis it is possible to make judgments of proportionality.[687]

The third question is about how to review AWS that are designed to undertake *online learning*.[688] That is, AWS that, by design, alter their behaviour over time based on experience in order to make themselves more effective. This is perhaps more of a long-term concern than an immediate problem, based on what is known of current weapons and development plans. It is generally accepted that weapons which are significantly modified after their initial fielding require a new legal review.[689] It can be argued that a weapon system which alters its own operation during use would, in effect, become a different weapon after each alteration, and would accordingly require a new review.[690] Continuous reviews of a weapon would, of course, be infeasible in practice. Another approach would be to review the learning algorithm itself rather than the precise state of the weapon's control system at any one time. If one can establish the ways in which the learning algorithm is likely to alter the AWS's behaviour, or at least the limits of those changes, then perhaps only one initial review is needed, as with non-learning weapon systems.

One possible means of alleviating the problems caused by the difficulty of testing would be to formulate and enforce generally applicable standards of IHL compliance which would govern the design of AWS control system software. In a 2007 report, Arkin provided what may be the beginning of an enforceable standard when he described a set

---

[687] Pablo Kalmanovitz (Notes, CCW Meeting of Experts on LAWS: Human Rights and Ethical Issues, April 2016) 2.
[688] Discussed briefly in Section 3.1.2.
[689] See, eg, Jeffrey S Thurnher, 'The Law That Applies to Autonomous Weapon Systems' (2013) 17(4) *ASIL Insights* 1.
[690] See, eg, Giacca, above n 179.

of 'design criteria, recommendations, and test scenarios to design and construct an autonomous robotic system architecture capable of the ethical use of lethal force'.[691] No concrete standard has yet emerged, though. Indeed, the prospects are slim, even if a suitable and detailed standard could be agreed upon. At the very least, given the vast research and development costs, and resulting commercial value, that effective AWS control software would entail, it is likely that weapon developers would resist efforts to open their software for examination.

7.5 Conclusion

Given the development plans that have been made public, it seems that there is no grave, immediate threat to existing accountability regimes for violations of IHL via AWS. For all their complexity and capability, AWS will still be weapon systems operated by armed forces who are subject to the same obligations they bear when operating other weapon systems. Armed forces personnel are accountable for the precautionary measures they take in preparation for activating an AWS and for whatever supervisory role they play during the course of an attack. Harm caused by the AWS after it is activated will fall within one of the recognised legal categories outlined in Section 5.4.2.2.1, all of which are consistent with current accountability mechanisms.

Although fears of an 'accountability gap' are unwarranted, it is possible that some development of the law may be required to allow for the different ways in which weapon operators interact with AWS. Hypothetically, very high levels of autonomous capability, if they ever find their way into weapon systems, might force a stronger reliance on prospective and collective accountability regimes.

In particular, weapon reviews will take on renewed importance, particularly in relation to AWS which will operate without the possibility of human intervention for extended periods of time and/or in particularly unpredictable environments.

[691] Arkin, above n 81, 98.

Chapter 8 Conclusions and Recommendations

This thesis finds that, overall, IHL presents no impenetrable barriers to continued development and use of AWS. Existing restrictions on the nature and use of weapons continue to apply to AWS as they do to other types of weapons. States which abide by those restrictions in developing, acquiring and operating increasingly autonomous weapons will remain in compliance with their IHL obligations. The main legal challenge is to understand how to interpret existing legal rules in light of increasing weapon autonomy.

Given the significant controversy surrounding the matter, that finding may be seen by some as an endorsement of continued development. That is not the case. This thesis neither endorses nor opposes development or use of AWS. The aim has been simply to map out part of the legal landscape that will be encountered if development plans are pursued, with a view to assisting States whose interests are at stake. The intention has not been to argue for or against pursuit of any particular plans.

The decision about whether or not to undertake a path of weapon development must be based on considerations far beyond IHL: moral and ethical considerations, strategic goals, security concerns, political interests and broader legal concerns, notably those of human rights law, must all be taken into account. Those matters cannot all be addressed in a project of this scope and must be studied elsewhere.

Two recommendations can be made based on the research presented here. First, a description of 'autonomy' in weapon systems which focusses on the aspects of legal significance and supports a constructive debate. Second, a tentative proposal for a regulatory approach which may address the main concerns raised thus far.

8.1 Defining AWS

Much of the debate within the CCW process and elsewhere has become polarised around, and paralysed by, competing understandings of the meaning of 'autonomous' as it applies to weapon systems. States with an interest in the debate should urgently seek to establish an understanding of weapon autonomy that would support a thorough analysis of its legal and other effects. The debate about regulation of AWS is fundamentally a response to a set of technologies and their collective impact on the

threat, nature and consequences of armed conflict. States cannot hope to reach agreement on a meaningful response without first agreeing on an understanding of the change that must be addressed. Specifically, an understanding of weapon autonomy should include the points summarised below:

- A weapon system's autonomy is simply its ability to operate without needing human interaction.
- Autonomy describes a means of controlling weapons. AWS are not a specific type of weapon with a specific physical effect, like mustard gas or blinding lasers. In fact, autonomy does not relate directly to the 'weapon' part of a weapon system at all, only to how the weapon is controlled. The technologies that make a weapon system autonomous would generally be part of a weapon platform, which may or may not be directly attached to the weapon being controlled. An AWS may even draw on a multitude of sensors and other devices spread over a wide area via an armed force's communications network.
- A weapon system's capacity to operate autonomously is a matter of degree, and is variable. That is, a weapon system may be able to operate with a high degree of autonomy in respect of some functions (e.g. navigation) but lower levels with respect to others (e.g. identifying targets), and those levels of autonomy might vary in different phases of an operation or in response to other changes in circumstances. Weapon systems with limited capacities for autonomous operation have been in use by some armed forces for decades and their legal status is not contentious.
- The functions of most interest in studying AWS are the critical functions of selecting and engaging targets.
- Autonomy does not mean true independence or a lack of human control, although there may, in some cases, be little or no opportunity for a human operator to intervene in an AWS operation after it the AWS has been activated. Instead, humans control AWS by designing software which operates the weapon system on their behalf. That software steps into the shoes of a human operator to control whichever functions of the weapon system are to be made autonomous.
- Anthropomorphic images of AWS are misleading and unhelpful in a legal discussion. They imply that an AWS is something other than a weapon, perhaps

something akin to a combatant. The legal rules applicable to use of AWS are the same rules that apply to use of other weapons.

- Legal concerns about AWS may arise from the functional scope of its autonomous capabilities (what it can do autonomously), the facility for human oversight and intervention, and the circumstances in which it is permitted to operate autonomously.

8.2 Regulating AWS

IHL-based arguments for a complete ban on AWS development are not supported either by the status of current AWS under IHL or by publically available development plans:[692] there is nothing to suggest that AWS are or will become inherently indiscriminate, nor that they are of a nature to inflict superfluous injury or unnecessary suffering, nor that their use will necessarily violate the law in any other way. Certainly every weapon system will have some limitation to its ability to select and engage targets in a legally compliant manner. However, if a particular weapon system cannot be used to legally attack a target in a particular situation it is difficult to see how that limitation can be generalised beyond a prohibition on using that weapon system in that situation.

There is also a definitional difficulty to overcome in order to implement a ban. Autonomy being variable and a matter of degree, AWS are not a sharply defined class of weapons with specific effects on targets as are, for example, blinding laser weapons, which were subject to a successful pre-emptive ban. Attempts to define AWS in a way which allows any given weapon system to be unambiguously classified as either 'autonomous' or 'not autonomous' have so far suffered from one of three problems. Some are too narrow, describing futuristic weapons with human-like capabilities that go far beyond published development plans.[693] Others, primarily technical definitions which recognise autonomy as simply a capacity for self-management present in many existing weapon systems, do not delineate the legally significant aspects.[694] Definitions that focus on one aspect of AWS deemed to be problematic, such as learning ability,[695]

[692] Again, no comment is made about whether or not there might be ethical or other grounds for a ban.
[693] See, eg, *JDN 2/11*, above n 92, 2-3 [206]: AWS 'must be capable of achieving the same level of situational understanding as a human.'
[694] See the various definitions of technical origin given in Section 3.1.
[695] See, eg, Roff, above n 288, 212.

risk becoming outdated as technology advances, or excluding weapon systems that may require regulation for other reasons.

Indeed, a ban may be counterproductive from a humanitarian viewpoint. It is widely acknowledged, including by opponents of AWS, that greater precision in attacks yields humanitarian benefits as well as military utility.[696] Precision through superior sensing, superior control over the quantum of force applied to a target and other advanced capabilities are among the benefits promised by further AWS development.

On the other hand, there are obvious risks in pursuing a complex new technology of great military utility, where the drive to gain the benefits of that utility may, absent appropriate regulation, lead developers and their customers to act incautiously. As the CCW discussions, not to mention the wider public debate, show, there is widespread discomfort with the idea of humans 'delegating' lethal decisions to a machine, even if the machine is just implementing human instructions. Whether that discomfort stems from moral, legal or other reservations, there is a reasonably widely held perception that development of AWS is unacceptable. That perception of risk may in itself provide a basis for regulation:

> All existing prohibitions on conventional armaments stem from the conviction that the use of the kinds of conventional weapons described bears the unacceptable risk of violating the most fundamental principles of humanitarian law'[697]

If it is deemed desirable to apply special regulatory measures to development and use of AWS, those measures should be cautiously applied. They should be based on a clear understanding of the nature of autonomous machines and should carefully address specific, identifiable risks while minimising the risk of derailing genuinely useful innovations. As Anderson and Waxman have noted:

> sensible regulation of these systems as they emerge is both possible and desirable. But regulation has to emerge along with the technologies themselves, and against the backdrop of a world that will likely come to adopt, over coming decades, technologies of autonomy for self-driving vehicles, or advanced nursing or elder-care robots, or any number of other technologies that evolve from being increasingly automated to

[696] Human Rights Watch, above n 554.
[697] den Dekker, above n 171, 79.

> performing some functions with genuine machine autonomy. With many of these technologies, however, the machine will take actions with potentially lethal consequences — and it will happen largely because people conclude over successive decades that machines can sometimes do certain tasks better than humans can.[698]

Despite the reference to a distinction between automation and 'genuine machine autonomy', with which this thesis disagrees, Anderson and Waxman capture the spirit of caution with which regulatory measures should be applied.[699]

Given that a process has already begun within the context of the CCW, inspiration may be found by examining the content of *CCW Protocol II*. That Protocol has several features which suggest it would provide a useful model for regulating AWS.

First, *CCW Protocol II* recalls general IHL principles, reinforces that they continue to apply to use of mines and, where helpful, provides specific rules of application: art 3(3) on superfluous injury and unnecessary suffering; arts 3(8)(a), 3(8)(b) and 3(9) on distinction; art 3(8)(c) on proportionality; art 3(10) on precautions; and art 3(11) on warnings.

Second, the Protocol identifies specific risks arising from use of mines and mandates specific measures to address those risks. To address the risk that civilians might inadvertently enter a mined area, Article 5(2)(a) provides that mines must be 'placed within a perimeter-marked area which is monitored by military personnel and protected by fencing or other means, to ensure the effective exclusion of civilians from the area.'

Third, the Protocol mandates positive measures to maximise its protective value. Under Article 11 each State Party undertakes to provide technological cooperation and assistance to other parties in implementing the Protocol. Such measures would include provision of equipment, information and expertise.

Finally, the Protocol provides a Technical Annex containing detailed technical information which supports implementation of the primary rules.

---

[698] Anderson and Waxman, above n 212, 2-3.

[699] For another argument in favour of regulation of AWS rather than a ban, see Kenneth Anderson, Daniel Reisner and Matthew Waxman, 'Adapting the Law of Armed Conflict to Autonomous Weapon Systems' (2014) 90 *International Law Studies* 386, 406.

It is too early to specify the precise content of the rules of such a Protocol applicable to AWS. States have not yet agreed on whether specific regulation is needed at all. If it is decided that such regulation is needed, it will be necessary to first establish the overall aim of regulating AWS (whether it is to guide, restrict or prohibit development or use) and define the set of weapon systems or technologies that must be regulated before substantive rules can be formulated.

This is a novel area of law and AWS are at a very early stage of development, so their nature is still something of a moving target. There is sufficient information available to make it clear that endowing lethal weapons with a capacity for autonomous operation will raise difficult questions, but not yet enough information to conclusively answer all of those questions. More research is needed into all the aspects of conflict and international security that would be affected by development of AWS. It is hoped that this thesis has provided some answers to questions about AWS and IHL.

Nevertheless, this thesis has addressed only the first few steps of a long process of legal development so it is difficult to find a conclusive point on which to end. Perhaps the best one would be simply an exhortation to States and lawyers to continue engaging closely with developers and future users of AWS, to understand the challenges and develop solutions which help to achieve and maintain international peace and security.

Bibliography

Articles / Books / Reports

Adams, Thomas K, 'Future Warfare and the Decline of Human Decisionmaking' (2001-2) (Winter) *Parameters* 57

Alberts, David S, John J Garstka and Frederick P Stein, *Network Centric Warfare* (Department of Defense Command and Control Research Program, 2nd revised ed, 1999)

Allen, Craig H, 'The Seabots are ~~Coming~~ Here: Should they be Treated as "Vessels"?' (2012) 65 *The Journal of Navigation* 749

Allenby, Braden R, 'Are New Technologies Undermining the Laws of War?' (2014) 70(1) *Bulletin of the Atomic Scientists* 21

Alston, Philip, 'Lethal Robotic Technologies: The Implications for Human Rights and International Humanitarian Law' (2011) 21(2) *Journal of Law, Information and Science* 35

Anderson, Kenneth, 'The Rise of International Criminal Law: Intended and Unintended Consequences' (2009) 20 *European Journal of International Law* 331

Anderson, Kenneth and Matthew C Waxman, "Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can" (Hoover Institution, 9 April 2013)

Anderson, Kenneth, Daniel Reisner and Matthew Waxman, 'Adapting the Law of Armed Conflict to Autonomous Weapon Systems' (2014) 90 *International Law Studies* 386

Antsaklis, Panos J, Kevin M Passino and S J Wang, 'An Introduction to Autonomous Control Systems' (1991) 11(4) *IEEE Control Systems* 5

Arkin, Ronald C, 'Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture' (Technical Report No GIT-GVU-07-11, Mobile Robot Laboratory, College of Computing, Georgia Institute of Technology, 2007)

Arkin, Ronald C, *Governing Lethal Behavior in Autonomous Robots* (Chapman and Hall, 2009)

Article 36, 'Killer Robots: UK Government Policy on Fully Autonomous Weapons' (Policy Paper, April 2013) <http://www.article36.org/wp-content/uploads/2013/04/Policy_Paper1.pdf?con=&dom=pscau&src=syndication>

Article 36, 'Structuring Debate on Autonomous Weapon Systems' (Memorandum for delegates to the Convention on Certain Conventional Weapons (CCW), November 2013) <http://www.article36.org/wp-content/uploads/2013/11/Autonomous-weapons-memo-for-CCW.pdf>

Article 36, 'Key Areas for Debate on Autonomous Weapon Systems' (Memorandum for delegates at the CCW Meeting of Experts on AWS, May 2014) <http://www.article36.org/wp-content/uploads/2014/05/A36-CCW-May-2014.pdf>

Article 36, 'Key Elements of Meaningful Human Control' (Background paper to comments prepared by Richard Moyes, Article 36, for the CCW Meeting of Experts on AWS, April 2016)

Asaro, Peter, 'On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making' (2012) 94 *International Review of the Red Cross* 687

Aspray, William, 'The Stored Program Concept' (1990) 27(9) *IEEE Spectrum* 51

Backstrom, Alan and Ian Henderson, 'New Capabilities in Warfare: An Overview of Contemporary Technological Developments and the Associated Legal and Engineering Issues in Article 36 Weapons Reviews' (2012) 94 *International Review of the Red Cross* 483

Badar, Mohamed Elewa and Sara Porro, 'Rethinking the Mental Elements in the Jurisprudence of the ICC' in Carsten Stahn (ed), *The Law and Practice of the International Criminal Court* (Oxford University Press, 2015) 665

Barak, Eitan (ed), *Deadly Metal Rain: The Legality of Flechette Weapons in International Law* (Martinus Nijhoff, 2011)

Bass, Gary J, *Stay the Hand of Vengeance: The Politics of War Crimes Tribunals* (Princeton University Press, 2000)

Beard, Jack M, 'Autonomous Weapons and Human Responsibilities' (2014) 45 *Georgetown Journal of International Law* 617

Beizer, Boris, *Software Testing Techniques* (Dreamtech, 2003)

Bekey, George A, *Autonomous Robots: From Biological Inspiration to Implementation and Control* (MIT Press, 2005)

Bergen, Peter and Katherine Tiedemann, 'Washington's Phantom War: The Effects of the US Drone Programs in Pakistan' (2011) 90(4) *Foreign Affairs* 12

Bialek, William, Ilya Nemenman and Naftali Tishby, 'Predictability, Complexity, and Learning' (2001) 13 *Neural Computation* 2409

Boivin, Alexandra, 'The Legal Regime Applicable to Targeting Military Objectives in the Context of Contemporary Warfare' (Research Paper Series No 2/2006, University Centre for International Humanitarian Law, 2006)

Bolton, Matthew, 'From Minefields to Minespace: An Archeology of the Changing Architecture of Autonomous Killing in US Army Field Manuals on Landmines, Booby Traps and IEDs' (2015) 46 *Political Geography* 41

Boothby, William H, *Weapons and the Law of Armed Conflict* (Oxford University Press, 2009)

Boothby, William, 'How Far Will the Law Allow Unmanned Targeting to Go?' in Saxon, Dan (ed), *International Humanitarian Law and the Changing Technology of War* (Martinus Nijhoff, 2013) 56

Boothby, William H, *Conflict Law: The Influence of New Weapons Technology, Human Rights and Emerging Actors* (Asser, 2014)

Borenstein, Jason, 'The Ethics of Autonomous Military Robots' (2008) 2(1) *Studies in Ethics, Law and Technology* 13

Bothe, Michael, Karl Josef Partsch and Waldemar A Solf, *New Rules for Victims of Armed Conflicts* (Martinus Nijhoff, 2013)

Boulanin, Vincent, 'Implementing Article 36 Weapon Reviews in the Light of Increasing Autonomy in Weapon Systems' (SIPRI Insights on Peace and Security No 2015/1, Stockholm International Peace Research Institute, November 2015)

Boylan, Patrick J, *Review of the Convention for the Protection of Cultural Property in the Event of Armed Conflict* (UNESCO, 1993) <http://unesdoc.unesco.org/images/0010/001001/100159eo.pdf>

Bubnicki, Zdzislaw, *Modern Control Theory* (Springer, 2005)

Byrnes, Michael W, 'Nightfall: Machine Autonomy in Air-to-Air Combat' (2014) (May-June) *Air & Space Power Journal* 48

Carey, John, William V Dunlap and R John Pritchard, *International Humanitarian Law: Origins* (Brill, 2003)

Carnahan, Burris M, 'The Law of Land Mine Warfare: Protocol II to the United Nations Convention on Certain Conventional Weapons' (1984) 105 *Military Law Review* 73

Casey, George W Jr, "America's Army in an Era of Persistent Conflict" (2008) 58 *Army* 19

Casey-Maslen, Stuart and Sharon Weill, 'The Use of Weapons in Armed Conflict' in Stuart Casey-Maslen (ed), *Weapons Under International Human Rights Law* (Cambridge University Press, 2014) 248

Cassese, Antonio, Paola Gaeta and John R W D Jones, *The Rome Statute of the International Criminal Court: A Commentary* (Oxford University Press, 2002) vol 1

Chen, Thomas M and Saeed Abu-Nimeh, 'Lessons from Stuxnet' (2011) 44(4) *Computer* 91

Citron, Danielle Keats, 'Technological Due Process' (2008) 85 *Washington University Law Review* 1249

Commission on Responsibility of the Authors of the War and on Enforcement of Penalties, *Report Presented to the Preliminary Peace Conference* (29 March 1919), reprinted in (1920) 14 *American Journal of International Law* 95

Copeland, Damian P, 'Legal Review of New Technology Weapons' in Hitoshi Nasu and Robert McLaughlin (eds), *New Technologies and the Law of Armed Conflict* (Asser Press, 2014) 43

Cordeschi, Roberto, 'Automatic Decision-Making and Reliability in Robotic Systems: Some Implications in the Case of Robot Weapons' (2013) 28 *AI & Societ*y 431

'Country Views on Killer Robots' (Campaign to Stop Killer Robots, 13 December 2016) <http://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC_CountryViews_13Dec2016.pdf>

Crawford, James R, 'State Responsibility' in *Max Planck Encyclopedia of Public International Law* (September 2006) <http://opil.ouplaw.com/home/EPIL>

Crootof, Rebecca, 'The Killer Robots Are Here: Legal and Policy Implications' (2015) 36 *Cardozo Law Review* 1837

Crootof, Rebecca, 'War Torts: Accountability for Autonomous Weapons' (2016) 164 *University of Pennsylvania Law Review* 1347

Crootof, Rebecca, 'A Meaningful Floor for "Meaningful Human Control"' (2016) *Temple International and Comparative Law Journal* 53

Cryer, Robert et al, *An Introduction to International Criminal Law and Procedure* (Cambridge University Press, 2nd ed, 2010)

Cummings, Mary L, 'Automation and Accountability in Decision Support System Interface Design' (2006) XXXII *The Journal of Technology Studies* 23

Curtin, Deirdre and André Nollkaemper, 'Conceptualizing Accountability in Public International Law' (2005) XXXVI *Netherlands Yearbook of International Law* 3

Daoust, Isabelle, Robin Coupland and Rikke Ishoey, 'New Wars, New Weapons? The Obligation of States to Assess the Legality of Means and Methods of Warfare' (2002) 84 *International Review of the Red Cross* 345

de Preux, Jean, 'Protocol I – Article 35 – Basic Rules' in Yves Sandoz, Christophe Swinarski and Bruno Zimmermann (eds), *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (Martinus Nijhoff, 1987) 389

de Preux, Jean, 'Protocol I – Article 36 – New Weapons' in Yves Sandoz, Christophe Swinarski and Bruno Zimmermann (eds), *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (Martinus Nijhoff, 1987) 421

de Preux, Jean, 'Protocol I – Article 82 – Legal Advisers in Armed Forces' in Yves Sandoz, Christophe Swinarski and Bruno Zimmermann (eds), *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (Martinus Nijhoff, 1987) 947

den Dekker, Guido, *The Law of Arms Control: International Supervision and Enforcement* (Martinus Nijhoff, 2001)

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

Dörmann, Knut, 'Preparatory Commission for the International Criminal Court: The Elements of War Crimes' (2000) 82 *International Review of the Red Cross* 771

Dörmann, Knut, *Elements of War Crimes under the Rome Statute of the International Criminal Court* (Cambridge University Press, 2003)

Drumbl, Mark A, 'Accountability for System Criminality' (2010) 8 *Santa Clara Journal of International Law* 373

Dunlap, Charles J, 'Accountability and Autonomous Weapons: Much Ado about Nothing?' (2016) 30 *Temple International and Comparative Law Journal* 63

Durham, Helen and Eve Massingham, 'Moving from the Mechanics of Accountability to a Culture of Accountability: What More Can be Done in Addition to Prosecuting War Crimes?' in Jadranka Petrovic (ed), *Accountability for Violations of International Humanitarian Law: Essays in Honour of Tim McCormack* (Routledge, 2016) 267

Ekberg, Peter J, 'Remotely Delivered Landmines and International Law' (1995) 33 *Columbia Journal of Transnational Law* 149

Eyffinger, Arthur, *The 1899 Hague Peace Conference: The Parliament of Man, the Federation of the World* (Martinus Nijhoff, 1999)

Fenrick, William J, 'The Rule of Proportionality and Protocol I in Conventional Warfare' (1982) 98 *Military Law Review* 91

*Final Record of the Diplomatic Conference Convened by the Swiss Federal Council for the Establishment of International Conventions for the Protection of War Victims and Held at Geneva from April 21*st *to August 12*th *, 1949* (Federal Political Department, Berne) vol IIB

Finn, Anthony and Steve Scheding, *Developments and Challenges for Autonomous Unmanned Vehicles* (Springer, 2010)

Finnin, Sarah, 'Mental Elements under Article 30 of the Rome Statute of the International Criminal Court: A Comparative Analysis' (2012) 61 *International and Comparative Law Quarterly* 325

Fleck, Dieter, 'Individual and State Responsibility for Violations of the Ius in Bello: An Imperfect Balance' in Wolff Heintschel von Heinegg and Volker Epping (eds), *International Humanitarian Law Facing New Challenges* (Springer, 2007)

François, Charles (ed), *International Encyclopedia of Systems and Cybernetics* (K G Saur, 2nd ed, 2004) vol 1

Fussell, Susan R et al, 'How People Anthropomorphize Robots' (2008) *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction* 145

Gasser, Hans-Peter, 'A Look at the Declaration of St. Petersburg of 1868' (1993) 33 *International Review of the Red Cross* 511

Geiss, Robin, 'The International-Law Dimension of Autonomous Weapons Systems' (Study, Friedrich-Ebert-Stiftung, International Policy Analysis, October 2015)

Gillespie, Tony and Robin West, 'Requirements for Autonomous Unmanned Air Systems set by Legal Issues' (2010) 4(2) *The International C2 Journal* 1

Ginsberg, Matt, *Essentials of Artificial Intelligence* (Newnes, 2012)

Goodrich, Michael A and Alan C Schultz, 'Human-Robot Interaction: A Survey' (2007) 1 *Foundations and Trends in Human-Computer Interaction* 203

Greenwood, Christopher, 'Customary Law Status of the 1977 Geneva Protocols' in Astrid J M Delissen and Gerard J Tanja (eds), *Humanitarian Law of Armed Conflict: Challenges Ahead: Essays in Honour of Frits Kalshoven* (Martinus Nijhoff, 1991) 111

Grimmelmann, James, 'Regulation by Software' (2005) 114 *Yale Law Journal* 1719

Grover, Leena, 'A Call to Arms: Fundamental Dilemmas Confronting the Interpretation of Crimes in the Rome Statute of the International Criminal Court' (2010) *European Journal of International Law* 543

Grut, Chantal, 'The Challenge of Autonomous Lethal Robotics to International Humanitarian Law' (2013) 18(1) *Journal of Conflict & Security Law* 5

Hall, Christopher Keith, 'The First Proposal for an International Criminal Court' (1998) 38 *International Review of the Red Cross* 57

Hammond, Daniel N, 'Autonomous Weapons and the Problem of State Accountability' (2015) 15 *Chicago Journal of International Law* 652

Hancock, Peter A and Stephen F Scallen, 'Allocating Functions in Human–Machine Systems' in Robert R Hoffman, Michael F Sherrick and Joel S Warm (eds), *Viewing Psychology as a Whole: The Integrative Science of William N Dember* (American Psychological Association, 1998) 521

Hangos, Katalin M, Rozália Lakner and Miklós Gerzson, *Intelligent Control Systems: An Introduction with Examples* (Kluwer Academic Publishers, 2004)

Hayashi, Nobuo, 'Military Necessity as Normative Indifference' (2013) 44 *Georgetown Journal of International Law* 675

Hazen, H L, 'Theory of Servo-Mechanisms' (1934) 218(3) *Journal of the Franklin Institute* 279

Hearst, Marti A, 'Trends & Controversies: Mixed-Initiative Interaction' (1999) 14(5) *IEEE Intelligent Systems* 14

Henckaerts, Jean-Marie and Louise Doswald-Beck, *Customary International Humanitarian Law* (Cambridge University Press, 2005) vols 1 and 2

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

Hennessy, Michael A and B J C McKercher (eds), *War in the Twentieth Century: Reflections at Century's End* (Praeger, 2003)

Herbach, Jonathan David, 'Into the Caves of Steel: Precaution, Cognition and Robotic Weapon Systems Under the International Law of Armed Conflict' (2012) 4(3) *Amsterdam Law Forum* 3

Heyns, Christof, 'Human Rights and the Use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement' (2016) 38 *Human Rights Quarterly* 350

Horowitz, Michael C and Paul Scharre, 'Meaningful Human Control in Weapon Systems: A Primer' (Working Paper, Center for a New American Security, March 2015)

Human Rights Watch, *Off Target: The Conduct of the War and Civilian Casualties in Iraq* (2003) Summary and Recommendations <https://www.hrw.org/reports/2003/usa1203/usa1203_sumrecs.pdf>

Human Rights Watch, *Losing Humanity: The Case Against Killer Robots* (November 2012)

Human Rights Watch, *Shaking the Foundations: The Human Rights Implications of Killer Robots* (2014)

Human Rights Watch, *Mind the Gap: The Lack of Accountability for Killer Robots* (April 2015)

Human Rights Watch, 'Killer Robots and the Concept of Meaningful Human Control' (Memorandum to CCW Delegates, April 2016) 10 < https://www.hrw.org/sites/default/files/supporting_resources/robots_meaningful_human_control_final.pdf>

Hunter, Dan, 'Reason is Too Large: Analogy and Precedent in Law' (2001) 50 *Emory Law Journal* 1197

International Committee of the Red Cross, *Conference of Government Experts on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts, 24 May -12 June, 1971* (1971) vol III: Protection of the Civilian Population Against Dangers of Hostilities

International Committee of the Red Cross, *Conference of Government Experts on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts, Second Session, 3 May -3 June 1972 – Report on the Work of the Conference* (1972) vol I

International Committee of the Red Cross, *Draft Additional Protocols to the Geneva Conventions of August 12, 1949 – Commentary* (1973)

International Committee of the Red Cross, *Conference of Government Experts on the Use of Certain Conventional Weapons (Lucerne, 24.9-18.10.1974): Report* (1975)

International Committee of the Red Cross, *Conference of Government Experts on the Use of Certain Conventional Weapons (Second Session - Lugano, 28.1-26.2.1976): Report* (1976)

International Committee of the Red Cross, 'A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977' (2006) 88 *International Review of the Red Cross* 931

International Committee of the Red Cross, 'Autonomous Weapon Systems: Implications Of Increasing Autonomy In The Critical Functions Of Weapons' (Report on Expert Meeting, Versoix, Switzerland, 15-16 March 2016)

'International Criminal Court: Making the Right Choices – Part V: Recommendations to the Diplomatic Conference' (Position Paper No IOR 40/010/1998, Amnesty International, 30 April 1998)

International Society of Military Law and the Laws of War, *Armed Forces and the Development of the Law of War (Ninth International Congress, Lausanne, 2–6 September 1982)* (1982)

Jenks, Chris, 'False Rubicons, Moral Panic, & Conceptual Cul-De-Sacs: Critiquing & Reframing the Call to Ban Lethal Autonomous Weapons' (2016) 44(1) *Pepperdine Law Review* (Article 2)

Johnson, Deborah G, 'Software Agents, Anticipatory Ethics, and Accountability' in Gary E Marchant, Braden R Allenby and Joseph R Herkert (eds), *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight: The Pacing Problem* (Springer, 2011) 61

Kalshoven, Frits, 'State Responsibility for Warlike Acts of the Armed Forces' (1991) 40 *International and Comparative Law Quarterly* 827

Karnow, Curtis E A, 'The Application of Traditional Tort Theory to Embodied Machine Intelligence' in Calo, Ryan, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar, 2016) 54

Kastan, Benjamin, 'Autonomous Weapons Systems: A Coming Legal "Singularity"?' [2013] *Journal of Law, Technology & Policy* 45

Kleffner, Jann K, 'From 'Belligerents' to 'Fighters' and Civilians Directly Participating in Hostilities – On the Principle of Distinction in Non-International Armed Conflicts One Hundred Years After the Second Hague Peace Conference' (2007) 54 *Netherlands International Law Review* 315

Kleffner, Jann, 'Sources of the Law of Armed Conflict' in Rain Liivoja and Tim McCormack (eds), Routledge Handbook of the Law of Armed Conflict (Routledge, 2016) ch 4

Kreps, Sarah, 'Flying Under the Radar: A Study of Public Attitudes Towards Unmanned Aerial Vehicles' (2014) 1(1) *Research & Politics*

Krishnan, Armin, *Killer Robots: Legality and Ethicality of Autonomous Weapons* (Ashgate, 2009)

Landau, I D, R Lozano and M M'Saad, *Adaptive Control* (Springer, 1998)

Leigh, James R, *Control Theory: A Guided Tour* (Institution of Engineering and Technology, 2012)

Lieber, Francis, *Instructions for the Government of Armies of the United States in the Field* (Government Printing Office, 1898) [originally issued as: General Orders No 100, Adjutant General's Office, 1863]

Lieblich, Eliav and Eyal Benvenisti, 'The Obligation to Exercise Discretion in Warfare: Why Autonomous Weapons Systems are Unlawful' in Nehal Bhuta et al (eds),

Autonomous Weapons Systems: Law, Ethics, Policy (Cambridge University Press, 2016) 245

Liivoja, Rain, Kobi Leins and Tim McCormack, 'Emerging Technologies of Warfare' in Rain Liivoja and Tim McCormack (eds), Routledge Handbook of the Law of Armed Conflict (Routledge, 2016) ch 35

Linkens, D A and H O Nyongesa, 'Learning Systems in Intelligent Control: An Appraisal of Fuzzy, Neural and Genetic Algorithm Control Applications' (2002) 143(4) *IEE Proceedings - Control Theory and Applications* 367

Liu, Hin-Yan, 'Categorization and Legality of Autonomous and Remote Weapons Systems' (2012) 94 *International Review of the Red Cross* 627

Maogoto, Jackson, *Technology and the Law on the Use of Force* (Routledge, 2015)

Marchant, Gary E et al, 'International Governance of Autonomous Military Robots' (2011) XII *Columbia Science and Technology Law Review* 272

McClelland, Justin, 'The Review of Weapons in Accordance with Article 36 of Additional Protocol I' (2003) 85 *International Review of the Red Cross* 397

McCormack, Timothy L H, '*Their* Atrocities and *Our* Misdemeanours: The Reticence of States to Try Their 'Own Nationals' for International Crimes' in Mark Lattimer and Philippe Sands (eds), *Justice for Crimes Against Humanity* (Hart Publishing, 2003) 107.

McFarland, Tim and Tim McCormack, 'Mind the Gap: Can Developers of Autonomous Weapons Systems be Liable for War Crimes?' (2014) 90 *International Law Studies* 361

McLaughlin, Rob, 'Unmanned Naval Vehicles at Sea: USVs, UUVs, and the Adequacy of the Law' (2011) 21(2) *Journal of Law, Information and Science* 100

Meier, Michael W, 'Lethal Autonomous Weapons Systems (LAWS): Conducting a Comprehensive Weapons Review' (2016) 30 *Temple International and Comparative Law Journal* 119

Myers, Glenford J, *The Art of Software Testing* (John Wiley & Sons, 2nd ed, 2004)

NIST Engineering Laboratory, 'Autonomy Levels for Unmanned Systems' (National Institute of Standards and Technology, 6 June 2010) <http://www.nist.gov/el/isd/ks/autonomy_levels.cfm>

Official Records of the Diplomatic Conference on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts Geneva (1974–1977) (1978) vol 1

Ohlin, Jens David, 'The Combatant's Stance: Autonomous Weapons on the Battlefield' (2016) 92 *International Law Studies* 1

Parasuraman, Raja, Thomas B Sheridan and Christopher D Wickens, 'A Model for Types and Levels of Human Interaction with Automation' (2000) 30(3) *IEEE Transactions on Systems, Man and Cybernetics* 288

Parks, W Hays, 'Air War and the Law of War' (1990) 32 *Air Force Law Review* 1

Parks, W Hays, 'Conventional Weapons and Weapons Reviews' (2005) 8 *Yearbook of International Humanitarian Law* 55

Pilloud, Claud and Jean Pictet, 'Protocol I – Article 49 – Definition of Attacks and Scope of Application' in Yves Sandoz, Christophe Swinarski and Bruno Zimmermann (eds), *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (Martinus Nijhoff, 1987) 601

Pilloud, Claude and Jean Pictet, 'Protocol I – Article 51 – Protection of the Civilian Population' in Yves Sandoz, Christophe Swinarski and Bruno Zimmermann (eds), *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (Martinus Nijhoff, 1987) 613

Pilloud, Claude and Jean Pictet, 'Protocol I – Article 57 – Precautions in Attack' in Yves Sandoz, Christophe Swinarski and Bruno Zimmermann (eds), *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (Martinus Nijhoff, 1987) 677

Pocar, Fausto, 'To What Extent Is Protocol I Customary International Law?' (2002) 78 *International Law Studies* 337

Pryer, Douglas A, 'The Rise of the Machines: Why Increasingly "Perfect" Weapons Help to Perpetuate our Wars and Endanger Our Nation' [2013] (March-April) *Military Review* 14

Rau, Markus, 'State Liability for Violations of International Humanitarian Law — The *Distomo* Case before the German Federal Constitutional Court' (2006) 7 *German Law Journal* 701

Rauch, Elmar, 'Intervention' (1983) 22 *Military Law and Law of War Review* 291

Richards, Neil M and William D Smart, 'How Should the Law Think about Robots?' in Calo, Ryan, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar, 2016) 17

Roff, Heather, 'The Strategic Robot Problem: Lethal Autonomous Weapons in War' (2014) 13(3) *Journal of Military Ethics* 211

Roff, Heather M and Richard Moyes, 'Meaningful Human Control, Artificial Intelligence and Autonomous Weapons' (Briefing paper for delegates at the CCW Meeting of Experts on AWS, Article 36, April 2016)

Rogers, A P V, 'A Commentary on the Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices' (1987) 26 *Military Law and Law of War Review* 185

Röling, B V A, 'Aspects of the Criminal Responsibility for Violations of the Laws of War' in Antonio Cassese (ed), *The New Humanitarian Law of Armed Conflict* (Napoli: Editoriale Scientifica, 1979) 138

Roorda, Mark, "NATO's Targeting Process: Ensuring Human Control Over and Lawful Use of 'Autonomous' Weapons" (Research Paper No 2015-06, Amsterdam Center for International Law, 2015)

Russell, Stuart and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Prentice Hall, 3rd ed, 2009)

Sassòli, Marco, 'State Responsibility for Violations of International Humanitarian Law' (2002) 84 *International Review of the Red Cross* 401

Sassòli, Marco, 'Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified' (2014) 90 *International Law Studies* 308

Scharre, Paul, 'Autonomous Weapons and Operational Risk' (Center for a New American Security, Ethical Autonomy Project, February 2016)

Schindler, Dietrich and Jiri Toman, *The Laws of Armed Conflicts: A Collection of Conventions, Resolutions and Other Documents* (Martinus Nijhoff, 4th rev ed, 2004)

Schmitt, Michael N, 'The Principle of Discrimination in 21st Century Warfare' (1999) 2 *Yale Human Rights and Development Journal* 143

Schmitt, M N, 'Future War and the Principle of Discrimination' (1999) 28 *Israel Yearbook on Human Rights* 55

Schmitt, Michael N, 'Precision Attack and International Humanitarian Law' (2005) 87 *International Review of the Red Cross* 445

Schmitt, Michael N, 'Fault Lines in the Law of Attack' in Susan C Breau and Agnieszka Jachec-Neale (eds), *Testing the Boundaries of International Humanitarian Law* (British Institute of International and Comparative Law, 2006) 277

Schmitt, Michael N, 'Military Necessity and Humanity in International Humanitarian Law: Preserving the Delicate Balance' (2010) 50 *Virginia Journal of International Law* 796

Schmitt, Michael, 'Classification of Cyber Conflict' (2012) 17 *Journal of Conflict and Security Law* 245

Schmitt, Michael N, *Essays on Law and War at the Fault Lines* (T M C Asser Press, 2012)

Schmitt, Michael N, 'Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics' (2013) *Harvard National Security Journal Features* 4

Schmitt, Michael N and Jeffrey S Thurnher, '"Out of the Loop": Autonomous Weapon Systems and the Law of Armed Conflict' (2013) 4 *Harvard National Security Journal* 231

Schmitt, Michael N (ed), NATO Cooperative Cyber Defence Centre of Excellence, *Tallinn Manual on the International Law Applicable to Cyber Warfare* (Cambridge University Press, 2013)

Schmitt, Michael N (ed), NATO Cooperative Cyber Defence Centre of Excellence, *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (Cambridge University Press, 2017)

Sharkey, Noel, 'Cassandra or False Prophet of Doom: AI Robots and War' [2008] (July/August) *IEEE Intelligent Systems* 14

Sharkey, Noel, 'Grounds for Discrimination: Autonomous Robot Weapons' [2008] (October) *RUSI Defence Systems* 86

Sheridan, Thomas B and William L Verplank, 'Human and Computer Control of Undersea Teleoperators' (Technical Report, Massachusetts Institute of Technology, 14 July 1978)

Singer, Peter W, *Wired for War: The Robotics Revolution and Conflict in the 21st Century* (Penguin Press, 2009)

Smith, Philip J, C Elaine McCoy and Charles Layton, 'Brittleness in the Design of Cooperative Problem-Solving Systems: The Effects on User Performance' (1997) 27(3) *IEEE Transactions on Systems, Man and Cybernetics* 360

Solum, Lawrence B, 'Legal Personhood for Artificial Intelligences' (1992) 70 *North Carolina Law Review* 1231

Strachan, Hew and Sibylle Scheipers (eds), *The Changing Character of War* (Oxford University Press, 2011)

Sukman, Daniel, 'Lethal Autonomous Systems and the Future of Warfare' (2015) 16 *Canadian Military Journal* 44

Sullins, John, 'When is a Robot a Moral Agent?' (2006) 6 *International Review of Information Ethics* 23

'Summary Record of the Eleventh Meeting (CDDH/III/SR.11)' in *Official Records of the Diplomatic Conference on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts Geneva (1974–1977)* (1978) vol 14, 85

'The Weaponization of Increasingly Autonomous Technologies: Considering how Meaningful Human Control Might Move the Discussion Forward' (Discussion Paper, United Nations Institute for Disarmament Research (UNIDIR), 2014)

Thurnher, Jeffrey S, 'The Law That Applies to Autonomous Weapon Systems' (2013) 17(4) *ASIL Insights* 1

Thurnher, Jeffrey S, 'Examining Autonomous Weapon Systems from a Law of Armed Conflict Perspective' in Hitoshi Nasu and Robert McLaughlin (eds), *New Technologies and the Law of Armed Conflict* (Asser Press, 2014) 223

Trial of the Major War Criminals before the International Military Tribunal, Nuremberg, 14 November 1945 – 1 October 1946 (1947) vol 1

Truszkowski, Walt et al, 'NASA's Swarm Missions: The Challenge of Building Autonomous Software' (2004) 6(5) *IT Professional Magazine* 47

Tucker, Spencer C, *Instruments of War: Weapons and Technologies That Have Changed History* (ABC-CLIO, 2015)

van Sliedregt, Elies, *Individual Criminal Responsibility in International Law* (Oxford University Press, 2012)

Wagner, Markus, 'Taking Humans Out of the Loop: Implications for International Humanitarian Law' (2011) 21(2) *Journal of Law, Information and Science* 155

Wagner, Markus, 'Autonomy in the Battlespace: Independently Operating Weapon Systems and the Law of Armed Conflict' in Saxon, Dan (ed), *International Humanitarian Law and the Changing Technology of War* (Martinus Nijhoff, 2013) 105

Wagner, Markus, 'The Dehumanization of International Humanitarian Law: Legal, Ethical, and Political Implications of Autonomous Weapon Systems' (2014) 47 *Vanderbilt Journal of Transnational Law* 1371

Wallach, Wendell and Colin Allen, *Moral Machines: Teaching Robots Right from Wrong* (Oxford University Press, 2009)

Wareham, Mary, "Report on Outreach on the UN Report on 'Lethal Autonomous Robotics'" (Report, Campaign to Stop Killer Robots, 31 July 2013) <http://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC_ReportHeynsUN_Jul2013.pdf>

Werrell, Kenneth P, *The Evolution of the Cruise Missile* (Air University Press, 1985)

Weston, Burns H and Jonathan C Carlson, *International Law & World Order: Weston's & Carlson's Basic Documents* (Brill, 2012)

Whittaker, James A, 'What is software testing? And why is it so hard?' (2000) 17(1) *IEEE Software* 70

Zimmermann, Bruno, 'Protocol I – Article 3 – Beginning and End of Application' in Yves Sandoz, Christophe Swinarski and Bruno Zimmermann (eds), *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (Martinus Nijhoff, 1987) 65

Armed Forces Documents

Cummings, Mary and Angelo Collins, 'Autonomous Aerial Cargo/Utility System (AACUS)' (Concept of Operations, Office of Naval Research) <http://www.onr.navy.mil/~/media/Files/Funding-Announcements/BAA/2012/12-004-CONOPS.ashx>

Defense Science Board, 'Patriot System Performance' (Report Summary, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, January 2005) <http://www.dtic.mil/get-tr-doc/pdf?AD=ADA435837>

Defense Science Board, 'The Role of Autonomy in DoD Systems' (Task Force Report, US Department of Defense, July 2012)

Department of Defense, 'Autonomy Research Pilot Initiative (ARPI)' (Invitation for Proposals, November 2012) 1 <http://auvac.org/uploads/publication_pdf/Autonomy%20Research%20Pilot%20Initiative.pdf>

Dotterway, Kristen A, *Systematic Analysis of Complex Dynamic Systems: The Case of the USS Vincennes* (Masters Thesis, Naval Postgraduate School, 1992) <http://www.dtic.mil/get-tr-doc/pdf?AD=ADA260260>

Fogarty, William, 'Formal Investigation into the Circumstances Surrounding the Downing of Iran Air Flight 655 on 3 July 1988' (Investigation Report, 19 August 1988) <http://www.dod.mil/pubs/foi/Reading_Room/International_Security_Affairs/172.pdf>

'Legal Review of New Weapons' (Defence Instruction (General) OPS 44-1, Australian Department of Defence, 2 June 2005)

Lin, Patrick, George Bekey, and Keith Abney, 'Autonomous Military Robotics: Risk, Ethics, and Design' (Report No ADA534697, US Department of Navy, Office of Naval Research, 20 December 2008) <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA534697>

McKenna, Tim et al., 'Science & Technology for Australian Network-Centric Warfare: Function, Form and Fit' (Australian Government Department of Defence, 2005) <http://docsfiles.com/view.php?view=http://www.dsto.defence.gov.au/attachments/ST%20for%20Australian%20NCW.pdf&keyword=the%20australian%20approach%20to%20warfare&count=>

Ministry of Defence, 'The UK Approach to Unmanned Aircraft Systems' (Joint Doctrine Note 2/11, 30 March 2011)

Nelson, William, 'Use of Circular Error Probability in Target Detection' (Technical Report, MITRE Corporation, May 1988) <http://www.dtic.mil/dtic/tr/fulltext/u2/a199190.pdf>

'Network Enabled Capability' (JSP 777 ed 1, United Kingdom Ministry of Defence, 2005) <http://webarchive.nationalarchives.gov.uk/20121026065214/http://www.mod.uk/NR/rdonlyres/E1403E7F-96FA-4550-AE14-4C7FF610FE3E/0/nec_jsp777.pdf>

Office of the Secretary of Defense, 'Unmanned Aircraft Systems Roadmap, 2005 – 2030' (US Department of Defense, 4 August 2005) <https://fas.org/irp/program/collect/uav_roadmap2005.pdf>

Proud, Ryan W, Jeremy J Hart and Richard B Mrozinski, 'Methods for Determining the Level of Autonomy to Design into a Human Spaceflight Vehicle: A Function Specific Approach' (Report, NASA Johnson Space Center, September 2003) 4 <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA515467>

Secretary of Defense, 'DoD Policy on Cluster Munitions and Unintended Harm to Civilians' (Memorandum, United States Department of Defense, 19 June 2008) <http://www.acq.osd.mil/tc/treaties/ccwapl/DoD%20Policy%20on%20Cluster%20Munitions.pdf>

Strategic Technology Office, Defense Advanced Research Projects Agency, 'Military Imaging and Surveillance Technology — Long Range (MIST-LR) Phase 2' (Broad Agency Announcement No DARPA-BAA-13-27, 12 March 2013) <https://www.fbo.gov/index?s=opportunity&mode=form&id=78b0ddbf382678fa9ace985380108f89&tab=core&_cview=0>

Tank-Automotive Research, Development, and Engineering Center, 'Robotics Strategy White Paper' (Army Capabilities Integration Center, US Army, 19 March 2009) (http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA496734>

United States Air Force, 'United States Air Force Unmanned Aircraft Systems Flight Plan, 2009-2047' (Report No ADA505168, Headquarters, United States Air Force, 18 May 2009) <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA505168>

'Unmanned Systems Safety Guide for DoD Acquisition' (United States Department of Defense, 27 June 2007) <https://acc.dau.mil/adl/en-US/683704/file/75176/Unmanned%20Systems%20Safety%20Guide%20forDOD%20Acquisition,27June%202007.pdf>

US Army Science Board, 'Ad-hoc Study on Human Robot Interface Issues' (September 2002) 16 <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA411834>

US Department of Defense, 'Unmanned Systems Integrated Roadmap FY2011-2036' (No 11-S-3613)

US Department of Defense, 'Autonomy in Weapon Systems' (Directive No 3000.09, Office of the Under Secretary of Defense for Policy, 21 November 2012)

Cases

'Case No 9: The Zyklon B Case: Trial of Bruno Tesch and Two Others (British Military Court, Hamburg)' in United Nations War Crimes Commission, *Law Reports of Trials of War Criminals* (His Majesty's Stationery Office, 1947) vol 1, 93

'Case No 57: The I G Farben Trial: Trial of Carl Krauch and Twenty-Two Others (United States Military Tribunal, Nuremberg)' in United Nations War Crimes Commission, *Law Reports of Trials of War Criminals* (His Majesty's Stationery Office, 1949) vol 10, 5

Fisheries Jurisdiction (Spain v Canada) (Judgment) [1998] ICJ Rep 432

Gabčíkovo-Nagyamaros Project (Hungary v Slovakia) (Judgment), [1997] ICJ Rep 7

Katz v United States 389 US 347

Legality of the Threat or Use of Nuclear Weapons (Advisory Opinion) [1996] ICJ Rep 226

Noble Ventures Inc v Romania (Award) (ICSID Arbitral Tribunal, Case No ARB/01/11, 12 October 2005)

Olmstead v United States 277 US 438, 466 (1928)

Prosecutor v Bemba Gombo (Judgment) (International Criminal Court, Trial Chamber III, Case No ICC-01/05-01/08, 21 March 2016)

Prosecutor v Blaškić (Judgement) (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-95-14-A, 29 July 2004)

Prosecutor v Chui (Judgment Pursuant to Article 74 of the Statute) (International Criminal Court, Trial Chamber II, Case No ICC-01/04–02/12, 18 December 2012)

Prosecutor v Furundžija (Judgement) (International Criminal Tribunal for the Former Yugoslavia, Trial Chamber, Case No IT-95-17/1-T, 10 December 1998)

Prosecutor v Katanga (Trial Judgment) (International Criminal Court, Trial Chamber II, Case No ICC-01/04-01/07, 7 March 2014)

Prosecutor v Kordić (Appeal Judgement) (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-95-14/2-A, 17 December 2004)

Prosecutor v Kupreškić (Judgement) (International Criminal Tribunal for the Former Yugoslavia, Trial Chamber, Case No IT-95-16, 14 January 2000)

Prosecutor v Mucić (Judgement) (International Criminal Tribunal for the Former Yugoslavia, Trial Chamber, Case No IT-96-21-T, 16 November 1998)

Prosecutor v Nikolić (Decision on Defence Motion Challenging the Exercise of Jurisdiction by the Tribunal) (International Criminal Tribunal for the Former Yugoslavia, Trial Chamber II, Case No IT–94–2–PT, 9 October 2002)

Prosecutor v Strugar (Trial Judgement) (International Criminal Tribunal for the Former Yugoslavia, Trial Chamber II, Case No IT-01-42-T, 31 January 2005)

Prosecutor v Tadić (Decision on the Defence Motion for Interlocutory Appeal on Jurisdiction) (International Criminal Tribunal for the Former Yugoslavia, Appeals Chamber, Case No IT-94-1-1, 2 October 1995)

Prosecutor v Taylor (Judgement) (Special Court for Sierra Leone, Trial Chamber II, Case No SCSL-03-01-T, 18 May 2012)

Legislation

Alien Tort Statute, 28 USC § 135 (2012)

National Prohibition Act 41 Stat 305, ch 85

Treaties

Additional Protocol 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 (Protocol IV, entitled Protocol on Blinding Laser Weapons), opened for signature 13 October 1995, 2024 UNTS 163 (entered into force 30 July 1998)

Convention for the Amelioration of the Condition of the Wounded in Armies in the Field, opened for signature 22 August 1864, 22 Stat 940 (entered into force 22 June 1865)

Convention on Cluster Munitions, opened for signature 30 May 2008, 2688 UNTS 39 (entered into force 1 August 2010)

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 October 1980, 1342 UNTS 137 (entered into force 2 December 1983)

Convention on the Prevention and Punishment of the Crime of Genocide, opened for signature 9 December 1948, 78 UNTS 277 (entered into force 12 January 1951)

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)

Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on their Destruction, opened for signature 13 January 1993, 1974 UNTS 45 (entered into force 29 April 1997)

Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on their Destruction, opened for signature 18 September 1997, 2056 UNTS 211 (entered into force 1 March 1999)

Convention respecting the Laws and Customs of War on Land and its Annex: Regulations concerning the Laws and Customs of War on Land, signed 18 October 1907 [1910] ATS 8 (entered into force 26 January 1910)

Convention (V) Respecting the Rights and Duties of Neutral Powers and Persons in Case of War on Land, signed 18 October 1907, 205 CTS 299 (entered into force 26 January 1910)

'Declaration IV, 1 to Prohibit for the Term of Five Years the Launching of Projectiles and Explosives from Balloons, and Other New Methods of a Similar Nature' in James Scott Brown (ed), *The Hague Conventions and Declarations of 1899 and 1907* (Oxford University Press, 2nd ed, 1915) 220

'Declaration IV, 2 Concerning Asphyxiating Gases' in James Scott Brown (ed), *The Hague Conventions and Declarations of 1899 and 1907* (Oxford University Press, 2nd ed, 1915) 225

'Declaration IV, 3 Concerning Expanding Bullets' in James Scott Brown (ed), *The Hague Conventions and Declarations of 1899 and 1907* (Oxford University Press, 2nd ed, 1915) 227

'Declaration XIV Prohibiting the Discharge of Projectiles and Explosives from Balloons' in James Scott Brown (ed), *The Hague Conventions and Declarations of 1899 and 1907* (Oxford University Press, 2nd ed, 1915) 220

Declaration Renouncing the Use, in Time of War, of Explosive Projectiles under 400 Grams Weight, [1901] ATS 125 (signed and entered into force 11 December 1868)

Draft Rules for the Limitation of the Dangers Incurred by the Civilian Population in Time of War (International Committee of the Red Cross, 1956) <https://www.icrc.org/ihl/INTRO/420?OpenDocument>

Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field, opened for signature 12 August 1949, 75 UNTS 31 (entered into force 21 October 1950)

Geneva Convention for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces at Sea, opened for signature 12 August 1949, 75 UNTS 85 (entered into force 21 October 1950)

Geneva Convention Relative to the Treatment of Prisoners of War, opened for signature 12 August 1949, 75 UNTS 135 (entered into force 21 October 1950)

Geneva Convention Relative to the Protection of Civilian Persons in Time of War, opened for signature 12 August 1949, 75 UNTS 287 (entered into force 21 October 1950)

International Convention for Adapting to Maritime Warfare the Principles of the Geneva Convention of 22 August 1864, signed 29 July 1899, [1901] ATS 132 (entered into force 4 September 1900)

International Convention for the Pacific Settlement of International Disputes, signed 29 July 1899, [1901] ATS 130 (entered into force 4 September 1900)

International Convention relative to the Laying of Automatic Submarine Contact Mines, signed 18 October 1907, [1910] ATS 11 (entered into force 26 January 1910)

International Convention with respect to the Laws and Customs of War on Land, signed 29 July 1899, [1901] ATS 131 (entered into force 4 September 1900)

International Covenant on Civil and Political Rights, opened for signature 16 December 1966, 999 UNTS 171 (entered into force 23 March 1976)

Project of an International Declaration concerning the Laws and Customs of War, signed 27 August 1874 (reprinted in: *The Proceedings of the Hague Peace Conferences: The Conference of 1899* (James Brown Scott trans, William S Hein & Co, 2000) 564 [trans of: *Conférence internationale de la paix. La Haye, 18 mai-29 juillet 1899.* (first published 1907)])

Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), opened for signature 8 June 1977, 1125 UNTS 3 (entered into force 7 December 1978)

Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), opened for signature 8 June 1977, 1125 UNTS 609 (entered into force 7 December 1979)

Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous or other Gases, and of Bacteriological Methods of Warfare, opened for signature 17 June 1925, [1930] ATS 6 (entered into force 8 February 1928)

Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices as amended on 3 May 1996 (Protocol II, as amended on 3 May 1996) annexed 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 3 May 1996, 2048 UNTS 93 (entered into force 3 December 1998)

Rome Statute of the International Criminal Court, opened for signature 17 July 1998, 2187 UNTS 90 (entered into force 1 July 2002)

The Laws of War on Land: Manual Published by the Institute of International Law, adopted 9 September 1880 (reprinted in: *Resolutions of the Institute of International*

Law Dealing with the Law of Nations (James Brown Scott trans, William S Hein & Co, 2003)

United Nations Documents

Alston, Philip, *Interim Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, UN GAOR, 65th sess, Agenda Item 69(b), UN Doc A/65/321 (23 August 2010)

Final Report of the Meeting of the High Contracting Parties 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, 2013 sess, Agenda Item 13, UN Doc CCW/MSP/2013/10 (16 December 2013)

Heyns, Christof, *Statement by Mr Christof Heyns, Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, UN GAOR, 3rd Comm, 65th sess, Agenda Item 68(a) (22 October 2010)

Heyns, Christof, *Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, Human Rights Council, 23rd sess, Agenda Item 3, UN Doc A/HRC/23/47 (9 April 2013)

Materials on the Responsibility of States for Internationally Wrongful Acts, United Nations Legislative Series Book 25, UN Doc ST/LEG/SER.B/25, (2012)

Recommendations to the 2016 Review Conference 1 <http://www.unog.ch/80256EDD006B8954/(httpAssets)/6BB8A498B0A12A03C1257FDB00382863/$file/Recommendations_LAWS_2016_AdvancedVersion+(4+paras)+.pdf>

Report of the 2014 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS), 2014 sess, Provisional Agenda Item 8, UN Doc CCW/MSP/2014/3 (11 June 2014)

Report of the 2015 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS), 2015 sess, Provisional Agenda Item 8, UN Doc CCW/MSP/2015/3 (2 June 2015)

Report of the 2016 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS) (Advanced Version) <http://www.unog.ch/80256EDD006B8954/(httpAssets)/DDC13B243BA863E6C1257FDB00380A88/$file/ReportLAWS_2016_AdvancedVersion.pdf>

Report of the International Law Commission, UN GAOR, 56th sess, Supp No 10, UN Doc A/56/10 (2001)

Responsibility of States for Internationally Wrongful Acts, UN GAOR, 6th Comm, 56th sess, 85th plen mtg, Agenda Item 162, UN Doc A/RES/56/83 (12 December 2001)

Responsibility of States for Internationally Wrongful Acts, UN GAOR, 6th Comm, 59th sess, 65th plen mtg, Agenda Item 139, UN Doc A/RES/59/35 (2 December 2004)

Responsibility of States for Internationally Wrongful Acts, UN GAOR, 6th Comm, 62nd sess, 62nd plen mtg, Agenda Item 78, UN Doc A/RES/62/61 of 6 December 2007

Responsibility of States for Internationally Wrongful Acts, UN GAOR, 6th Comm, 65th sess, 57th plen mtg, Agenda Item 75, UN Doc A/RES/65/19 (6 December 2010)

Work of the Advisory Board on Disarmament Matters: Report of the Secretary-General, UN GAOR, 68th sess, Provisional Agenda Item 101, UN Doc A/68/206 (26 July 2013)

Websites

Adams, William Lee, *Brief History: Wiretapping* (11 October 2010) Time Magazine <http://content.time.com/time/magazine/article/0,9171,2022653,00.html>

Advanced Medium Range Air-to-Air Missile (AMRAAM), Naval Air Systems Command <http://www.navair.navy.mil/index.cfm?fuseaction=home.display&key=D3FAC4AB-2E9F-4150-9664-6AFBC83F203E>

Aegis Combat System, Lockheed Martin <http://www.lockheedmartin.com.au/us/products/aegis.html>

Bolton, Matthew, Thomas Nash and Richard Moyes, *Ban Autonomous Armed Robots* (5 March 2012) Article 36 <http://www.article36.org/statements/ban-autonomous-armed-robots/>

Crawford, James, *Articles on Responsibility of States for Internationally Wrongful Acts: Historical Background and Development of Codification*, United Nations Audiovisual Library of International Law <http://legal.un.org/avl/ha/rsiwa/rsiwa.html>

DARPA Tactical Technology Office, Defense Advanced Research Projects Agency, *Hydra* <http://www.darpa.mil/Our_Work/TTO/Programs/Hydra.aspx>

Defence Science and Technology Organisation, *Network-Centric Warfare* (3 September 2007) Australian Government Department of Defence <http://www.dsto.defence.gov.au/research/4051/>

Federal Business Opportunities, *Hydra* (22 August 2013) <https://www.fbo.gov/index?s=opportunity&mode=form&id=4cc32f06144bd6f3eba18655135d6155&tab=core&_cview=1>

Federal Research Division, *Official Records of the Diplomatic Conference on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts* (16 July 2010) Library of Congress <https://www.loc.gov/rr/frd/Military_Law/RC-dipl-conference-records.html>

First Committee, *Disarmament and International Security - Documents of the 68th Session*, General Assembly of the United Nations <http://www.un.org/en/ga/first/68/documentation.shtml>

France convenes Seminar at the UN (26 September 2013) Campaign to Stop Killer Robots <http://www.stopkillerrobots.org/2013/09/france-seminar/>

Global Patriot Solutions, Raytheon <http://www.raytheon.com/capabilities/products/patriot/>

Harpy Air Defense Suppression System (4 March 2006) Defense Update: International Online Defense Magazine <http://defense-update.com/directory/harpy.htm>

Harpy NG, Israel Aerospace Industries <http://www.iai.co.il/2013/36694-16153-en/IAI.aspx>

Hartshorn, Derick S, *Mk 60 (Captor)* (28 November 2009) Mineman Memories <http://www.hartshorn.us/Navy/navy-mines-10.htm>

Iron Dome Weapon System, Raytheon <http://www.raytheon.com/capabilities/products/irondome/>

Joint Strike Missile (JSM) Brochure, Kongsberg <https://www.kongsberg.com/~/media/KDS/Files/Products/Missiles/jsm_web_reduced.ashx>

Keller, John, *DARPA Considers Unmanned Submersible Mothership Designed to Deploy UAVs and UUVs* (23 July 2013) Military & Aerospace Electronics <http://www.militaryaerospace.com/articles/2013/07/darpa-uuv-mothership.html>

Kumagai, Jean, *A Robotic Sentry for Korea's Demilitarized Zone* (1 March 2007) IEEE Spectrum <http://spectrum.ieee.org/robotics/military-robots/a-robotic-sentry-for-koreas-demilitarized-zone>

Lockheed Martin Corporation, *K-MAX* <http://www.lockheedmartin.com/us/products/kmax.html>

Meier, Michael, 'U.S. Delegation Statement on "Appropriate Levels of Human Judgment"' (Statement to the CCW Informal Meeting of Experts on AWS, 12 April 2016) <https://geneva.usmission.gov/2016/04/12/u-s-delegation-statement-on-appropriate-levels-of-human-judgment/>

Micro Autonomous Systems and Technology (MAST), *Research Thrusts* <https://alliance.seas.upenn.edu/~mastwiki/wiki/index.php?n=Main.Objectives>

NATO Network Enabled Capability (27 October 2010) North Atlantic Treaty Organization <http://www.nato.int/cps/de/SID-815535E4-57782C82/natolive/topics_54644.htm>

Office of Naval Research, A*utonomous Aerial Cargo/Utility System Program* (27 September 2013) <http://www.onr.navy.mil/en/Science-Technology/Departments/Code-35/All-Programs/aerospace-research-351/Autonomous-Aerial-Cargo-Utility-AACUS.aspx>

Phalanx Close-In Weapon System: Last Line of Defense for Air, Land and Sea, Raytheon Australia <http://www.raytheon.com.au/capabilities/products/phalanx/>

Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the protection of victims of international armed conflicts (Protocol I), United Nations Treaty Collection <https://treaties.un.org/Pages/showDetails.aspx?objid=08000002800f3586>

Sauer, Frank, *ICRAC statement on technical issues to the 2014 UN CCW Expert Meeting* (14 May 2014) International Committee for Robot Arms Control <http://icrac.net/2014/05/icrac-statement-on-technical-issues-to-the-un-ccw-expert-meeting/>

Sauer, Frank, *Banning Lethal Autonomous Weapon Systems (LAWS): The Way Forward* (13 June 2014) International Committee for Robot Arms Control <http://icrac.net/2014/06/banning-lethal-autonomous-weapon-systems-laws-the-way-forward/>

Scharre, Paul, *Between a Roomba and a Terminator: What is Autonomy?* (18 February 2015) War on the Rocks <https://warontherocks.com/2015/02/between-a-roomba-and-a-terminator-what-is-autonomy/?singlepage=1>

United States Army Research Laboratory, *Micro Autonomous Systems and Technology (MAST)* (25 February 2011) <http://www.arl.army.mil/www/default.cfm?page=332>

Other

Amnesty International (Statement, CCW Meeting of Experts on LAWS: Human Rights and Ethical Issues, April 2016)

Belgium (Presentation, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016)

Borrie, John (Presentation, CCW Meeting of Experts on LAWS: Security Issues, April 2016)

Bourcier, Daniele (Presentation, CCW Meeting of Experts on LAWS: Human Rights and Ethical Issues, April 2016)

Brehm, Maya (Presentation, CCW Meeting of Experts on LAWS: Characteristics of LAWS, April 2015)

Canada (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016)

Center for a New American Security (Text, CCW Meeting of Experts on LAWS: Characteristics of LAWS, April 2015)

Center for a New American Security (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016)

Chatila, Raja (Text, CCW Meeting of Experts on LAWS: Technical Issues, May 2014)

Dahlmann, Anja (Presentation, CCW Meeting of Experts on LAWS: Towards a Working Definition of LAWS, April 2016)

'Final Act of the International Peace Conference' in James Scott Brown (ed), *The Hague Conventions and Declarations of 1899 and 1907* (Oxford University Press, 2nd ed, 1915) 1

'Final Act of the Second International Peace Conference' in James Scott Brown (ed), *The Hague Conventions and Declarations of 1899 and 1907* (Oxford University Press, 2nd ed, 1915) 1

France, 'Characterization of a LAWS' (Working Paper, CCW Meeting of Experts on LAWS, April 2016)

Galliott, Jai (Presentation, CCW Meeting of Experts on LAWS: Security Issues, April 2016)

General Comment No. 3 on the African Charter on Human and Peoples' Rights: The Right to Life (Article 4), African Commission on Human and Peoples' Rights, 57th ordinary session (18 November 2015) 12 [35] <http://www.achpr.org/instruments/general-comments-right-to-life/>

Germany (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, May 2014)

Germany (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016)

Germany (Statement, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016)

Giacca, Gilles (Notes, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016)

Harpy NG Anti-Radiation Loitering Weapon System, Israel Aerospace Industries, MBT Missiles Division <http://www.iai.co.il/Sip_Storage//FILES/5/41655.pdf>

Heyns, Christof (Comment, CCW Meeting of Experts on LAWS: Overarching Issues, April 2015)

Heyns, Christof (Note, CCW Meeting of Experts on LAWS: Human Rights and Ethical Issues, April 2016)

Holy See (Text, CCW Meeting of Experts on LAWS: Overarching Issues, April 2015)

Human Rights Watch (Statement, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016)

International Criminal Court, *Elements of Crimes*, Doc No ICC-ASP/1/3 and Corr.1 and RC/11 (adopted 11 June 2010)

Israel (Text, CCW Meeting of Experts on LAWS: Characteristics of LAWS, April 2015)

Israel (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016)

Israel (Statement, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016)

Japan (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016)

Jenks, Chris (Notes, CCW Meeting of Experts on LAWS: Towards a Working Definition of LAWS, April 2016)

Kalmanovitz, Pablo (Notes, CCW Meeting of Experts on LAWS: Human Rights and Ethical Issues, April 2016)

Karnouskos, Stamatis, 'Stuxnet Worm Impact on Industrial Cyber-Physical System Security' (Paper presented at IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society, Melbourne, Victoria, Australia, 7–10 November 2011)

Kester, Leon (Notes, CCW Meeting of Experts on LAWS: Mapping Autonomy, April 2016)

Koh, Collin S L (Note, CCW Meeting of Experts on LAWS: Security Issues, April 2016)

Kozyulin, Vadim (Presentation, CCW Meeting of Experts on LAWS: Security Issues, April 2016)

Lieblich, Eliav (Note, CCW Meeting of Experts on LAWS: Human Rights and Ethical Issues, April 2016)

Lin, Patrick (Text, CCW Meeting of Experts on LAWS: Overarching Issues, April 2015)

Mines Action Canada (Statement, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016)

Moyes, Richard (Notes, CCW Meeting of Experts on LAWS: Towards a Working Definition of LAWS, April 2016)

Netherlands (Statement, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016)

Norway (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, May 2014)

Pakistan (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016)

Poland (Text, CCW Meeting of Experts on LAWS: Characteristics of LAWS, April 2015)

Republic of Korea (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2015)

Richter, Wolfgang, 'Utility and Limitations of the Use of LAWS in Military Operations' (Text, CCW Meeting of Experts on LAWS: Operational and Military Aspects, May 2014)

Rickli, Jean-Marc (Text, CCW Meeting of Experts on LAWS: Overarching Issues, April 2015)

'Russian Circular Note Proposing the Program of the First Conference' in James Scott Brown (ed), *The Hague Conventions and Declarations of 1899 and 1907* (Oxford University Press, 2nd ed, 1915) xvii

Scharre, Paul (Text, CCW Meeting of Experts on LAWS: Technical Issues, May 2014)

Scholtz, Jean, 'Theory and Evaluation of Human Robot Interactions' (Paper presented at 36th Annual Hawaii International Conference on System Sciences, Hawaii, 6–9 January 2003)

Stojanovic, Milica, 'Underwater Acoustic Communications: Design Considerations on the Physical Layer' (Paper presented at Fifth Annual Conference on Wireless on Demand Network Systems and Services, Garmisch-Partenkirchen, Germany, 23-25 January 2008)

Sweden (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, May 2014)

Sweden (Statement, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016)

Switzerland (Statement, CCW Meeting of Experts on LAWS: Towards a Working Definition of LAWS, April 2016)

Switzerland (Statement, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016)

United Kingdom, *Parliamentary Debates*, House of Lords, 26 March 2013, Column 958, 3pm (Lord Astor of Hever) <http://www.publications.parliament.uk/pa/ld201213/ldhansrd/text/130326-0001.htm#st_14>

United Kingdom (Statement, CCW Meeting of Experts on LAWS: Challenges to International Humanitarian Law, April 2016)

United States (Opening Statement, CCW Meeting of Experts on LAWS: General Exchange, April 2016)

van Bezooijen, B J A, P J M D Essens and A L W Vogelaar, 'Military Self-Synchronization: An Exploration of the Concept' (Paper presented at 11th International Command and Control Research and Technology Symposium, Cambridge, UK, 26–28 September 2006) <http://www.dodccrp.org/events/11th_ICCRTS/html/papers/065.pdf>

Zawieska, Karolina (Presentation & Text, CCW Meeting of Experts on LAWS: Overarching Issues, April 2015)