Forthcoming in Rain Liivoja and Ann Väljataga (eds), *Autonomous Cyber Capabilities Under International Law* (NATO Cooperative Cyber Defence Centre 2021).

Autonomous cyber capabilities and individual criminal responsibility for war crimes

Abhimanyu George Jain*

I. Introduction

Weaponization of the pervasive cyber infrastructure and of nascent autonomous technologies poses difficult challenges for individual criminal responsibility for war crimes arising from the military use of these technologies. For cyber technologies, the problem is identifying the perpetrator of the conduct which corresponds to the *actus reus* of the relevant war crime. For autonomous technologies, there is the problem of the 'responsibility gap': if the impugned conduct is effectuated by an algorithm or by a human relying on an algorithm,[1] there is no human being with the *mens rea* required for the crime.[2] Autonomous cyber capabilities (**ACC**) compound these two very significant and quite different challenges.

This chapter analyses the challenges posed by ACC to criminal responsibility for war crimes. It does so by considering the practicalities of war crimes prosecution and how these challenges might be addressed in practice. On this basis, it presents two arguments. First, the practical impact of the challenge of identifying perpetrators and of the responsibility gap on individual criminal responsibility may be mitigated in some cases by the practice of charging and adjudicating war crimes. Second, for the remaining cases, the impossibility of criminal responsibility should not be seen as diminishing the enforcement of international humanitarian law (**IHL**) but instead, as indicating the preferability of enforcing IHL in these cases by invoking the parallel responsibility of belligerents.

This chapter does not deny the significance of the difficulty of identification or of the responsibility gap and nor does it take on the quixotic burden of resolving them. Instead, it

* This chapter has benefitted greatly from comments and suggestions from Paola Gaeta, Andrew Clapham, Dorothea Endres, Shri Singh, Alessandra Spadaro and David Stewart, and from the participants at the NATO CCDCOE conference on autonomous cyber capabilities and international law. I retain sole responsibility for any errors.

[1] This chapter adopts the definition of 'autonomy' proposed in Tim McFarland, 'The Concept of Autonomy', this volume, ch. 2. However, even recognising that autonomy is a form of control rather than the absence of control (ibid 22, ¶ 52) does not necessarily or directly address the responsibility gap in the form discussed here.

[2] The characterisation of this problem in terms of 'responsibility' demands consideration. Existing analyses refer variously to gaps in responsibility and accountability, and on occasion these terms seem to have been used interchangeably in the broader criminal law literature. This chapter takes the view that 'responsibility' refers to the substantive capacity to comply with legal obligations, to be bound by them and to be liable for breach. Liability for breach of the rule presumes responsibility under the rule. 'Accountability' refers to the procedures of enforcing legal obligations upon obligors. See, Renée SB Kool, '(Crime) Victims' Compensation: The Emergence of Convergence' (2014) 10 Utrecht Law Review 14, 16–20; HLA Hart, *Punishment and Responsibility: Essays in the Philosophy of Law* (2nd edn, Oxford University Press 2008) 196–7. Given that what is at issue here is not only the enforcement of the obligation but also the more fundamental question of applicability of the obligation and existence of an obligor, the term 'responsibility gap' is preferred to 'accountability gap'. However, while the problem is better articulated in terms of responsibility than of accountability, from a strictly technical perspective, the problem is narrower than responsibility. The crux of the problem is that of culpability (guilty mind or *culpa*) which, along with actionable conduct, is a necessary requirement for criminal responsibility. Notwithstanding the greater technical accuracy of 'culpability gap', the term 'responsibility gap' is preferred here because the culpability gap necessarily produces a responsibility gap. Moreover, the significance and consequences of the technical problem of culpability become much clearer when seen from the perspective of a gap in responsibility: the absence of a human who can be held criminally responsible for breaches of international humanitarian law, and the consequent gap in the criminal enforcement of international humanitarian law.

argues for destabilising the unitary and fixed nature of these problems, and for seeing them as variable challenges which may manifest differently in different cases. This differentiated perspective allows for the recognition that in some cases these challenges do not pose insurmountable barriers to prosecution and allows for the refocussing of attention on the residual cases. In turn, the resilience of these challenges in 'residual' cases draws attention to the possibility that criminal responsibility is simply inappropriate in these cases, and that these cases are better addressed in terms of the underlying responsibility of belligerents.

The scope of this chapter is limited in several ways. To begin with it is restricted to the prosecution of conduct of hostilities war crimes, which present significant evidentiary challenges even in kinetic contexts, and within that, for reasons of space, to war crimes corresponding to the IHL rule of distinction. However, the analysis presented here may be extended *mutatis mutandis* to other conduct of hostilities war crimes. Further, only prosecution at the International Criminal Court (**ICC**) is considered here, although a fuller discussion of this issue would also take account of the prospects of prosecution at the national level. Finally, this chapter assumes the existence of an armed conflict, the applicability of IHL and that the impugned use of ACC constitutes a cyber-attack.[3]

Sections II and III discuss the challenge of identifying perpetrators and the responsibility gap respectively. Section IV concludes with a discussion of the preferability of addressing residual cases through belligerents' responsibility under IHL.

II. Identifying perpetrators

The difficulties of tracing cyber-attacks and identifying the perpetrators are well-recognised, particularly in the context of state responsibility for breaches of international law rules on uses of force, sovereignty and interventions in states' internal affairs.[4] The crux of the problem lies in identifying the source of the cyber operation and, frequently, in attributing the actions of non-state actors to states in accordance with the international law of responsibility.

In this context, there is increasing recognition that attribution is not a unique, technical problem with a definite answer: it is an art not a science.[5] It is a political process which necessarily involves subjective assessments, and its nature and results vary depending on the purposes of attribution (public or not), standard of proof, timeframes, attributing agency (political or

[3] An account of the questions implicated here is set out in Michael N Schmitt (ed), *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (2nd edn, Cambridge University Press 2017) (rr. 80-5, 92 and accompanying commentary); Kai Ambos, 'International Criminal Responsibility in Cyberspace' in Nicholas Tsagourias and Russell Buchan (eds), *Research Handbook on International Law and Cyberspace* (Edward Elgar Publishing 2015).

[4] In addition to the chapters by Michael N Schmitt and Samuli Haataja in this volume, recent contributions to the literature on the challenges of attribution include Nicholas Tsagourias and Michael Farrell, 'Cyber Attribution: Technical and Legal Approaches and Challenges' [2020] European Journal of International Law; William Banks, 'Who Did It? Attribution of Cyber Intrusions and the Jus in Bello' in Major Ronald TP Alcala and Eric Talbot Jensen (eds), *The Impact of Emerging Technologies on the Law of Armed Conflict* (Oxford University Press 2019); Hans-Georg Dederer and Tassilo Singer, 'Adverse Cyber Operations: Causality, Attribution, Evidence, and Due Diligence' (2019) 95 International Law Studies 430.

[5] Clement Guitton, *Inside the Enemy's Computer: Identifying Cyber-Attackers* (Hurst & Company 2017). See also, Thomas Rid and Ben Buchanan, 'Attributing Cyber Attacks' (2015) 38 Journal of Strategic Studies 4. These and similar approaches are endorsed in legal analyses by Tsagourias and Farrell (n 4) 4–11; Banks (n 4) 248; Dan Efrony and Yuval Shany, 'A Rule Book on the Shelf? Tallinn Manual 2.0 on Cyberoperations and Subsequent State Practice' (2018) 112 American Journal of International Law 583, 636.

judicial), etc.[6] In short, attribution as a technical process cannot produce absolute certainty. This rationalisation of expectations has highlighted the related but distinct technical and legal aspects of attribution,[7] and focussed attention on adaptation of legal requirements to technical limitations, including reliance on a 'preponderance of evidence' standard of proof.[8]

When cyber-attacks correspond to the *actus reus* of a war crime, the difficulties of attribution for the purposes of state responsibility are translated into the challenges of identifying the perpetrator of the attack for the purposes of criminal responsibility. In this context, recognising the limitations of technical attribution and turning to a preponderance of evidence standard of proof may not be feasible. For criminal responsibility a presumption of innocence operates until guilt is established 'beyond reasonable doubt'.[9] In the context of both criminal and state responsibility, the common problem is that of identifying the perpetrators, but in the criminal responsibility context this identification must satisfy the requirements of the criminal standard of proof.

This is the problem of identifying the perpetrators of ACC-related IHL breaches. This articulation of the challenge relies on a particular, fixed and rigid idea of 'beyond reasonable doubt' and assumes that this is a very stringent and rigorous standard. This assumption is not supported by the practice of criminal law at either the national or international levels.

Though the generally applicable[10] criminal standard of proof beyond reasonable doubt is frequently abbreviated to certainty or near certainty,[11] in fact it simply requires the elimination of all alternative possibilities that are reasonable or plausible.[12] There is an extensive body of literature discussing the difficulty of defining and applying the standard, pointing to significantly lower thresholds for conviction in practice in both national and international criminal law.

For instance, in the American context, interpretations of reasonable doubt vary widely,[13] to the extent that definitions deemed acceptable by one court are found by other courts to violate the

---

[6] Guitton (n 5) 11.

[7] Thus, for instance, Rid and Buchanan propose three levels of attribution: tactical (how the attack was conducted), operational (what it entailed) and strategic (who and why). None of the three levels necessarily yields certain answers, but the level of uncertainty increases from the tactical (technical) to the strategic level. See, Rid and Buchanan (n 5).

[8] See, e.g., Tsagourias and Farrell (n 4).

[9] This is the requirement under Art. 66 of the Rome Statute of the ICC.

[10] It is true that civil law systems prefer the '*intime conviction du juge*' standard, but the reasonable doubt standard has a long history in international criminal law, having been adopted and applied by the Yugoslavia and Rwanda tribunals: Salvatore Zappalà, 'The Rights of the Accused' in Antonio Cassese, Paola Gaeta and John RWD Jones (eds), *The Rome Statute of the International Criminal Court: A Commentary* (Oxford University Press 2002) 1346–7. Moreover, the reasonable doubt standard finds some recognition in international human rights law: Otto Triffterer and Kai Ambos (eds), *The Rome Statute of the International Criminal Court: A Commentary* (3rd edn, CH Beck 2016) 1643.

[11] Quantifications of the standard generally characterise it as requiring 90-95% certainty. See, e.g., Stephen Wilkinson, 'Standards of Proof in International Humanitarian and Human Rights Fact-Finding and Inquiry Missions' (Geneva Academy of International Humanitarian Law and Human Rights and Geneva Call) 17 <https://www.geneva-academy.ch/joomlatools-files/docman-files/Standards%20of%20Proof%20in%20Fact-Finding.pdf> accessed 7 October 2020.

[12] Triffterer and Ambos (n 10) 1645; Zappalà (n 10) 1347.

[13] Larry Laudan, *Truth, Error, and Criminal Law: An Essay in Legal Epistemology* (Cambridge University Press 2006) 32–47. He also notes the increasing practice of simply not instructing juries on the meaning of the standard: ibid 47–51. To similar effect, an official training document for criminal trial judges in the UK provides: "*It is unwise to elaborate on the standard of proof…although if an advocate has referred to "beyond reasonable*

constitutional rights of the defendant.[14] Empirical research has shown that juries frequently enter convictions based on a perceived probability of guilt ranging from 50-75%.[15]

In the international context, Combs has undertaken an extensive review of trial transcripts at the International Criminal Tribunal for Rwanda, the Special Court for Sierra Leone and the Special Panels in the Dili District Court in East Timor. She highlights a large number of significant infirmities in the evidence relied on for conviction, including extensive reliance on organisational affiliation as evidentiary proxy,[16] ultimately questioning these tribunals' adherence to the requirements of proof beyond reasonable doubt.[17] Her explanation of the handful of acquittals (six) at the Rwanda tribunal is particularly disturbing: "*the inclination of these Trial Chambers to conduct a more searching inquiry into testimonial deficiencies was driven primarily by their sense that the defendant did not generally support the genocide.*"[18]

Another aspect of the indeterminacy of the reasonable doubt standard is reflected in the frame of assessment for evidence relating to specific facts. This question is at the heart of an ongoing and unsettled debate in the case law of the ICC.[19] One side suggests that each piece of evidence relating to a fact should be individually assessed for evidentiary value (e.g., reliability) and then all eligible pieces of evidence should be considered together to determine whether they establish the fact in question.[20] Contrasted with this 'atomistic' or 'fragmentary' approach is a 'holistic' approach which proposes collective assessment of all pieces of evidence pertaining to a fact to determine whether, as a whole, they establish the fact in question.[21] An illustration

---

doubt", the jury should be told that this means the same thing as being sure." See, UK Judicial College, 'The Crown Court Compendium - Part I: Jury and Trial Management and Summing Up' 5–1 <https://www.judiciary.uk/wp-content/uploads/2020/07/Crown-Court-Compendium-Part-I-July-2020-09.10.20.pdf> accessed 7 October 2020.

[14] Laudan (n 13) 47.

[15] Nancy Amoury Combs, *Fact-Finding Without Facts: The Uncertain Evidentiary Foundations of International Criminal Convictions* (Cambridge University Press 2014) 350.

[16] ibid 235–72.

[17] ibid 189–223.

[18] ibid 254.

[19] For academic commentary, see, Mark Klamberg, 'Epistemological Controversies and Evaluation of Evidence in International Criminal Trials' [2019] Stockholm Faculty of Law Research Paper Series 65 <https://ssrn.com/abstract=3313509> accessed 7 October 2020; Darryl Robinson, 'The Other Poisoned Chalice: Unprecedented Evidentiary Standards in the Gbagbo Case? (Part 1)' (*EJIL: Talk!*, 5 November 2019) <https://www.ejiltalk.org/the-other-poisoned-chalice-unprecedented-evidentiary-standards-in-the-gbagbo-case-part-1/> accessed 7 October 2020; Yvonne McDermott, 'Strengthening the Evaluation of Evidence in International Criminal Trials' (2017) 17 International Criminal Law Review 682. As these authors note, this issue has also been discussed at length in the jurisprudence of other international criminal tribunals.

[20] As per the separate opinions referred to *infra* note 21, this approach has been followed in *The Prosecutor v Laurent Gbagbo and Blé Goudé (TC I) (No Case to Answer Decision)* [2019] International Criminal Court ICC-02/11-01/15-1263; *The Prosecutor v Mathieu Ngudjolo Chui (AC) (Appeal Judgment)* [2015] International Criminal Court ICC-01/04-02/12-271-Corr; *The Prosecutor v Germain Katanga (TC II) (Trial Judgment) (Minority Opinion of Judge Christine van den Wyngaert)* [2014] International Criminal Court ICC-01/04-01/07-3436-AnxI.

[21] This approach has been endorsed in *The Prosecutor v Laurent Gbagbo and Blé Goudé (TC I) (No Case to Answer Decision) (Dissenting Opinion of Judge Herrera Carbuccia)* [2019] International Criminal Court ICC-02/11-01/15-1263-AnxC-Red [5, 26–51]; *The Prosecutor v Mathieu Ngudjolo Chui (AC) (Appeal Judgment) (Joint Dissenting Opinion of Judges Ekaterina Trendafilova and Cuno Tarfusser)* [2015] International Criminal Court ICC-01/04-02/12-271-AnxA [31–51]; *The Prosecutor v Germain Katanga (TC II) (Trial Judgment) (Concurring Opinion of Judges Fatoumata Diarra and Bruno Cotte)* [2014] International Criminal Court ICC-01/04-01/07-3436-AnxII-tEng [4–5]. This approach also finds support in *The Prosecutor v Thomas Lubanga Dyilo (AC) (Appeal Judgment)* [2014] International Criminal Court ICC-01/04-01/06-3121-Red [22]: "*In the view of the Appeals Chamber, when determining whether* [the reasonable doubt standard] *has been met, the Trial*

of the difference between these two approaches lies in their treatment of contradictory evidence. An atomistic approach might reject two pieces of evidence altogether based on their mutual inconsistency; a holistic approach might advocate reconciling the inconsistency by reference to the broader evidence adduced in relation to the fact in question.[22]

This difference has the reasonable doubt standard at its centre. Compliance with the standard is much more difficult through the atomistic approach than through the holistic approach.[23] This difference has not yet been resolved, and arguably it never will be. The inherent subjectivity of the idea of 'reasonable doubt', and divisions along the axes of public international law / criminal law and civil law / common law which characterise even the narrow epistemic community of ICC judges,[24] make the resolution of this difference difficult. For present purposes, the existence of the debate is more interesting than its resolution: the persistence of these differences in the interpretation of a long-standing and well-established standard embellishes its inherent variability and context-specificity.

In sum, shorn of rhetoric and mythology, the reasonable doubt standard is indeterminate, subjective and dependent on context, including the nature of the crime charged and the decider's perception of the defendant.[25] As Lord Justice Denning has noted, "*In criminal cases the charge must be proved beyond reasonable doubt, but there may be degrees of proof within that standard."*[26] And as noted by Damaška, "*it seems psychologically naive to assume that sufficiency of proof requirements do not change in the process of decision-making*."[27]

Nor is this indeterminacy and variability, in and of itself, necessarily problematic. A standard of proof is not a historical, legal or moral necessity. It represents a socio-political decision as to the appropriate allocation of the burden of legal error: a high standard of proof beyond reasonable doubt suggests that acquitting the guilty is seen as far preferable to convicting the innocent.[28] A low standard of proof, conversely, suggests that convicting the innocent is a lesser concern than acquitting the guilty. In the socio-political context of international criminal law, there are a large number of reasons supporting a shifting of this allocation of the burden of legal error. The conviction of the innocent may be seen as less costly than the acquittal of the guilty by reference to, *inter alia*:[29] the investigatory challenges of international prosecutions;[30] the likelihood that a defendant whose case has reached this far bears some

Chamber is required to carry out a holistic evaluation and weighing of ***all the evidence taken together*** *in relation to the fact at issue. Indeed, it would be incorrect for a finder of fact to do otherwise.*" (emphasis in the original)

[22] This example is drawn from *Ngudjolo Chui appeal judgment (Trendafilova and Tarfusser)* (n 21) paras 47–51.

[23] This is particularly clear in ibid 31–41. Consider, for instance, at para 31: "*The Chamber assessed in isolation individual items of evidence and failed to properly consider the evidence in its entirety. As a result of this approach, the Trial Chamber disregarded trustworthy, coherent and vital evidence which, when pieced together with other relevant and credible evidence, would have provided a solid basis for the determination of the truth."*

[24] See, e.g., the qualifications for ICC judges set out in Rome Statute Art. 36.

[25] Combs (n 15) 344–50; Laudan (n 13) 32–51. A well-recognised example is that men tend to be more concerned than women about wrongful conviction for sexual crimes: Combs (n 15) 350.

[26] Bater v. Bater (1951), *cited in* Combs (n 15) 348.

[27] Mirjan Damaška, 'Evidentiary Barriers to Conviction and Two Models of Criminal Procedure: A Comparative Study' (1973) 121 University of Pennsylvania Law Review 506, 542.

[28] Laudan (n 13) 1–2.

[29] Combs (n 15) 350–9; Fergal Gaynor and others, 'Law of Evidence' in Göran Sluiter and others (eds), *International Criminal Procedure: Principles and Rules* (Oxford University Press 2013) 1148. The impracticability and undesirability of excessively fastidious application of the reasonable doubt standard in international criminal law features in criticism of the ICC's recent decision that Laurent Gbagbo and Blé Goudé had no case to answer, by the dissenting judge and academics alike. See, *Gbagbo and Goudé no case (Carbuccia)* (n 21) paras 6–7; Robinson (n 19).

[30] See, e.g., *Katanga trial judgment (Diarra and Cotte)* (n 21) para 5.

responsibility;[31] and, the accountability, deterrence and historical aspects of international criminal trials.

None of this implies a rejection of the reasonable doubt standard, and nor should it be interpreted to support a dilution of the standard for international crimes in general or for ACC-related war crimes in particular. The objective of the foregoing analysis is simply to draw attention to the well-recognised variability and context-specificity of the reasonable doubt standard. The mythologies of certainty surrounding the standard should not obfuscate the inherent and unavoidable contingency of any factual determination.[32]

Against this rationalised understanding of the reasonable doubt standard and its operation in practice, the difficulties of identifying perpetrators of ACC-related war crimes may not be as significant or as uniform as they seem at first sight.

III. Responsibility gap

This section discusses the challenge of the responsibility gap and proceeds in three sub-parts. Section III.A introduces the responsibility gap. Sections III.B and III.C discuss two features of international war crimes prosecution which may operate to mitigate the responsibility gap to varying degrees – the practicalities of proving *mens rea* (III.B) and the in-built seriousness requirement in the ICC's jurisdiction (III.C).

A. Introducing the responsibility gap

The responsibility gap presents a conceptual problem for criminal responsibility for ACC-related war crimes, based on the impossibility of a culpable human i.e., a human who has the *mens rea* necessary for criminal responsibility.

Under the Rome Statute of the ICC the default *mens rea* or mental element requirement is intention and knowledge.[33] The Rome Statute war crimes relating to the IHL rule of distinction refer to "*Intentionally directing attacks against the civilian population as such or against individual civilians not taking direct part in hostilities*" in international and non-international

[31] See, e.g., *supra* note 18 and accompanying text.

[32] The contingent nature of determinations of 'fact' is well-recognised in the rich literature on the epistemological philosophy of criminal law and fact-finding but seemingly under-recognised in the practice and promise of criminal law. For a review of this literature, see Simon de Smet, 'Justified Belief in the Unbelievable' in Morten Bergsmo and Carsten Stahn (eds), *Quality Control in Fact-Finding* (Torkel Opsahl Academic EPublisher 2020). For a persuasive account of the role of judges in 'constructing' facts in international adjudication, providing an explanation for the deficiencies identified by Combs in the fact-finding practice in international criminal tribunals (*supra* note 15), see, Ana Luísa Bernardino, 'The Discursive Construction of Facts in International Adjudication' (2020) 11 Journal of International Dispute Settlement 175.

[33] Rome Statute Art. 30. The rather confusing structure of Art. 30 necessitates a brief explanation of its requirements. The requirement of 'intent and knowledge' in Art. 30(1) does not mean that both together constitute the default, but instead that the general mental element (*mens rea* or *dolus*) under the Rome Statute has both volitional (intent or purpose or wanting) and cognitive (knowing or awareness) elements. One or both may constitute the requirement for specific war crimes, and Arts. 30(2) and 30(3) go on to define what each means. Thus, Art. 30 envisages neither a default separation of intent and knowledge nor a default conjunction. See, Elements of Crimes of the International Criminal Court (General Introduction, para 2); Triffterer and Ambos (n 10) 1117; Elies van Sliedregt, *Individual Criminal Responsibility in International Law* (Oxford University Press 2012) 46; Albin Eser, 'Mental Elements—Mistake of Fact and Mistake of Law' in Antonio Cassese, Paola Gaeta and John RWD Jones (eds), *The Rome Statute of the International Criminal Court: A Commentary* (Oxford University Press 2002) 904–8.

armed conflicts and "*Intentionally directing attacks against civilian objects, that is, objects which are not military objectives"* in international armed conflicts.[34] These crimes of attacking civilian targets have been interpreted as crimes of conduct which do not require a specific result,[35] and as requiring the deliberate launching of an attack against a target known to be civilian in nature.[36]

In the event of ACC-related breaches of IHL the stringent *mens rea* requirement of the Rome Statute may prove a barrier to war crimes prosecutions and criminal responsibility.[37] The soldier who deploys or relies on ACC knowing that civilian targets will be attacked and intending to attack them satisfies the *mens rea* requirement and may be criminally responsible if the *actus reus* requirements are met.[38] The soldier who has no reason to doubt the prospective IHL-compliance of ACC but is implicated in a breach of IHL is not culpable and need not concern us further. It is the soldier who has reason short of certainty to doubt the deployment of or reliance on ACC i.e., the soldier who acts negligently or recklessly or with *dolus eventualis*, who is difficult to accommodate within the mental element required by the Rome Statute. It is impossible that this soldier *knew* that a civilian target would be attacked and *intended* that attack, and it is this soldier who bestrides the responsibility gap, rather like a colossus.[39]

This responsibility gap is not simply a theoretical proposition. The nature of ACC means that they act and respond to their environments independently, and human involvement is restricted to controlling or supervising very sophisticated and possibly unpredictable technologies.[40] In this context there is a transfer of agency which diminishes the possibility of culpability of the soldier.

---

[34] Rome Statute Arts. 8(2)(b)(i), 8(2)(b)(ii) and 8(2)(e)(i). See also, the elements of these crimes in Elements of Crimes of the International Criminal Court. Given the similar structure and elements of these crimes, in the following analysis they will be treated as co-extensive and grouped under the common rubric of 'war crimes of attacks against civilian targets'.

[35] *The Prosecutor v Bosco Ntaganda (TC VI) (Trial Judgment)* [2019] International Criminal Court ICC-01/04-02/06-2359 [904]; *The Prosecutor v Germain Katanga (TC II) (Trial Judgment)* [2014] International Criminal Court ICC-01/04-01/07-3436-tENG [799].

[36] *Ntaganda trial judgment* (n 35) paras 903, 917, 921; *Katanga trial judgment* (n 35) para 808. Though these war crimes explicitly require 'intentionally directing attacks', the *Katanga* trial chamber has confirmed that the reference to 'intentionally' does not amount to a specific *mens rea* requirement distinct from the Art. 30 default: *ibid 806.*

[37] As noted *supra* note 1, recognising that autonomy is a form of control rather than its absence, and that ACC are developed by humans and operate within human-defined parameters, does not necessarily mitigate the responsibility gap in relation to soldiers deploying or relying on ACC in active hostilities.

[38] It is possible that in cases of deployment of ACC there may also be questions as to whether the *actus reus* component of the war crime has been satisfied. This possibility arises from the *Ntaganda* trial judgment where the ICC interpreted the requirement of 'directing attacks' as "*selecting the intended target and deciding on the attack*". See, *Ntaganda trial judgment* (n 35) para 917.

[39] This formulation of the problem is drawn from Shakespeare's Julius Caesar, and is deployed to reflect the vast consternation and analysis the problem has spawned. The substance of the underlying problem is drawn from Neha Jain, 'Autonomous Weapons Systems: New Frameworks for Individual Responsibility' in Nehal Bhuta and others (eds), *Autonomous Weapons Systems: Law, Ethics, Policy* (Cambridge University Press 2016) 315; Jens David Ohlin, 'The Combatant's Stance: Autonomous Weapons on the Battlefield' (2016) 92 International Law Studies 1, 21–2.

[40] Vincent Boulanin and others, 'Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control' (Stockholm International Peace Research Institute and International Committee of the Red Cross 2020) ix <https://www.icrc.org/en/document/limits-autonomous-weapons> accessed 7 October 2020.

Early articulations of the responsibility gap hypothesised weapon systems which could autonomously make and execute attack decisions.[41] More recent literature points instead to the application of autonomy at multiple stages of the targeting process in support of human decision-making.[42] Thus, for instance, autonomous technologies might be used for intelligence analysis, for generating possible targets, for target-identification, for assessing collateral damage, etc.

The shift from deploying autonomous technologies to relying on them for human decision-making changes the responsibility gap: it makes it less obvious but no less significant. In cases of deployment, the conduct which constitutes a breach of IHL (and consequently, possibly, the *actus reus* for the corresponding war crime) is effectuated by the weapon system and the only proximate human conduct is the decision of deployment. The culpability of this deployment is a necessary precondition for criminal responsibility. In cases of reliance, the conduct which constitutes a breach of IHL is effectuated by a human, but in reliance on autonomous technologies. This reliance may be unwarranted or unjustified or compromised by cognitive biases such as over or under-reliance and cognitive overloading.[43] In these cases, criminal responsibility requires establishing the culpability of this reliance.

Thus, the shift from deployment to reliance on autonomous technologies shifts the locus of the culpability assessment.[44] But while this shifts the source of the responsibility gap, it does not change it. In both cases the key underlying premise of the responsibility gap – the questionable culpability of the soldier who is negligent or reckless in their deployment of or reliance on autonomous technologies remains the same.[45]

For present purposes, it is not necessary to assess the exact nature and scope of the change in the responsibility gap across deployment and reliance, and it suffices to note that these may constitute two different but related challenges. Indeed, in the context of ACC both versions of the responsibility gap (deployment and reliance) are relevant given the existing state of the technology and expected trajectories of development.[46] The singular exception to the emerging

[41] In addition to the sources *supra* note 39, see also, Rebecca Crootof, 'War Torts: Accountability for Autonomous Weapons' (2016) 164 University of Pennsylvania Law Review 1347.

[42] Merel AC Ekelhof, 'Lifting the Fog of Targeting: "Autonomous Weapons" and Human Control through the Lens of Military Targeting' (2018) 71 Naval War College Review.

[43] Marta Bo, 'The Human-Weapon Relationship in the Age of Autonomous Weapons and the Attribution of Criminal Responsibility for War Crimes' (2019) <https://robots.law.miami.edu/2019/wp-content/uploads/2019/03/Bo_Human-Weapon-Relationship.pdf> accessed 7 October 2020. The problem posed by biases should be treated with some caution. It is undeniably true that the question of biases is particularly significant in the context of human-machine interaction and teaming, and all the more so in relation to autonomous technologies where the interaction poses existential challenges for the very meaning of human agency and control. However, the fact remains that biases are an unavoidable (and in their role as heuristics, possibly necessary) aspect of human cognition which can never be eliminated but can only be accounted for and managed.

[44] As to the relationship between culpability and the responsibility gap, see the discussion *supra* note 2.

[45] The shift from deployment to reliance may change the responsibility gap in one more way, through multiplication in the instances of human-machine interaction, and proliferation of consequent questions of culpability and responsibility in relation to a single attack. However, determining whether this constitutes a change, in what way and to what extent depends on the specificities of the autonomous weapons being deployed and the autonomous technologies being relied on, and cannot be assessed further in the abstract.

[46] Rain Liivoja, Maarja Naagel and Ann Väljataga, 'Autonomous Cyber Capabilities Under International Law' (NATO Cooperative Cyber Defence Centre of Excellence 2019) 11–13 <https://ccdcoe.org/uploads/2019/07/Autonomy-in-Cyber-Capabilities-under-International-Law_260619-002.pdf> accessed 7 October 2020.

consensus[47] regarding the requirement of meaningful human control for autonomous weapons systems relates to defensive applications of ACC.[48] In this context, the original responsibility gap thesis, centred around the increasingly obsolete trope of deployment of 'killer robots' without a human in or on the 'loop', regains prominence and applies in parallel to the revised responsibility gap thesis relating to reliance on ACC.

This brief introduction to the nature of the responsibility gap explains the furore it has generated. The soldier who deploys or relies on ACC with reason short of certainty to doubt the IHL compatibility of the ensuing actions exposes a gap in the criminal enforcement of IHL.[49] Though there is no conceptual solution to the responsibility gap, its practical significance may be mitigated by two features of the practice of charging and adjudicating war crimes at the ICC. It is to the first of these two features that the next sub-section turns.

B. Mens rea *in probative practice*

The first feature of the practice of war crimes prosecution which may mitigate the responsibility gap concerns the modalities of proving mental elements or *mentes reae*. It will be argued here that the common practice of inferring intent from conduct and circumstances provides limited mitigation of the responsibility gap by shifting focus from what the soldier actually knew and intended to what they *must have* known and *therefore* intended.

Criminal law, both national and international, relies on a strict application of Cartesian dualism – the distinction between body and mind, according to which criminal responsibility requires the conjunction of *actus reus* and *mens rea*.[50] The distinction has been the subject of criticism and critique in psychology and neuroscience,[51] and in criminal theory,[52] but it endures in criminal law.[53] However, the insistence on a guilty mind in addition to proscribed conduct raises evidentiary challenges because "*substantive rules regarding the mental element require the actual occurrence of a subjective mental state, whereas the law of evidence can provide only an assumption that the required state may have occurred.*"[54]

---

[47] 'Report of the 2019 Session' (Group of Governmental Experts 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 2019) CCW/GGE.1/2019/3 paras 21–2 <https://undocs.org/en/CCW/GGE.1/2019/3> accessed 7 October 2020.

[48] Tanel Tammet, 'Autonomous Cyber Defence Capabilities', this volume, ch. 3; Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (2018) ch 14.

[49] Academic analyses have turned instead to theories of indirect perpetration, including command responsibility. See, e.g., Russell Buchan and Nicholas Tsagourias, 'Command Responsibility and Autonomous Cyber Weapons', this volume, ch. 13; Jain (n 39); Ohlin, 'The Combatant's Stance: Autonomous Weapons on the Battlefield' (n 39).

[50] Jeroen Blomsma, *Mens Rea and Defences in European Criminal Law* (Intersentia 2012) 41. This is expressed in the Latin phrase: "*actus non facit reum nisi mens sit rea*".

[51] Dov Fox and Alex Stein, 'Dualism and Doctrine' in Dennis Patterson and Michael S Pardo (eds), *Philosophical Foundations of Law and Neuroscience* (Oxford University Press 2016).

[52] Antony Duff, *Intention, Agency and Criminal Liability: Philosophy of Action and the Criminal Law* (Blackwell 1990).

[53] On this gap between philosophy and criminal law, see George P Fletcher, *Rethinking Criminal Law* (Oxford University Press 2000) 451–2.

[54] Keren Shapira-Ettinger, 'The Conundrum of Mental States: Substantive Rules and Evidence Combined' (2007) 28 Cardozo Law Review 2577, 2685. See also, Fletcher (n 53) 120. Hart neatly encapsulates the resonance of this concern and its rejection by juxtaposing the 15th century dictum of Chief Justice Bryan – "*The thought of man is not triable; the devil alone knoweth the thought of man*" – with that of Lord Justice Bowen in the 19th century – "*the state of a man's mind is as much a fact as the state of his digestion.*": Hart (n 2) 188.

In practice, this hurdle is overcome by inferring *mens rea* from conduct and circumstances.[55] It is uncontroversial, for instance, that the intent to murder can be inferred from the act and context of stabbing the victim in the stomach.[56] This constraint posed by the law of evidence upon the ideals of the criminal law is not a rejection of the *mens rea* requirement.[57] It is simply an acknowledgment that there are acts and circumstances (e.g., stabbing somebody in the stomach) for which a particular mental state (intention) is the only possible, *though still rebuttable*,[58] conclusion.[59]

This probative practice – inferring mental element from conduct and circumstances which allow for no reasonable alternative explanation – is also well-established in international war crimes prosecutions.

The International Criminal Tribunal for the Former Yugoslavia (**ICTY**) has consistently held that intent for the war crime of attacking civilians:[60]

> "*can be inferred from many factors, including the means and method used in the course of the attack, the status and number of the victims, the nature of the crimes committed, the extent to which the attacking force may be said to have complied or attempted to*

[55] Jens David Ohlin, *Criminal Law: Doctrine, Application, and Practice* (Wolters Kluwer 2016) 167 (American law); AP Simester and others, *Simester and Sullivan's Criminal Law: Theory and Doctrine* (6th edn, Hart Publishing 2016) 147–8 (English law); Michael Bohlander, *Principles of German Criminal Law* (Hart Pub 2009) 65 (German law); *State of Maharashtra v Mohd Yakub s/o Abdul Hamid & Ors* [1980] SCR (2) 1158 (Supreme Court of India) 1163–4 (Indian law); *X und Y gegen Staatsanwaltschaft des Kantons Luzern sowie Obergericht des Kantons Luzern* (Urteil des Kassationshofes) [8.4] (Swiss law); Blomsma (n 50) 54–8 (Dutch, English, German and European law); Thomas Weigend, 'Subjective Elements of Criminal Liability' in Markus D Dubber and Tatjana Hörnle (eds), *The Oxford Handbook of Criminal Law* (Oxford University Press 2014) 508 (generally). By way of example, s. 8 of the UK Criminal Justice Act 1967 provides: "*Proof of criminal intent. A court or jury, in determining whether a person has committed an offence,— (a) shall not be bound in law to infer that he intended or foresaw a result of his actions by reason only of its being a natural and probable consequence of those actions; but (b) shall decide whether he did intend or foresee that result by reference to all the evidence, drawing such inferences from the evidence as appear proper in the circumstances.*" On this issue more generally, see, Shapira-Ettinger (n 54).

[56] In itself this practical reality represents the endorsement of Wittgenstein's observation: "*An 'inner process' stands in need of outward criteria*.": Ludwig Wittgenstein, *Philosophical investigations* (PMS Hacker and Joachim Schulte eds, GEM Anscombe, PMS Hacker and Joachim Schulte trs, 4th edn, Wiley-Blackwell 2009) s 580. On the challenges and difficulties of defining and divining intention from the perspective of analytical philosophy, see, GEM Anscombe, *Intention* (2nd edn, Harvard University Press 2000).

[57] The separate but intertwined nature of *mens rea* and *actus reus*, and the possibility of divining one from the other finds a parallel in the relationship between state practice and *opinio juris* for the determination of customary international law. For instance, though the International Law Commission's draft conclusions on the identification of customary international law emphasise the distinct but conjunctive requirements of state practice and *opinio juris*, they also recognise the possibility of inferring *opinio juris* from state practice. See, International Law Commission, 'Draft Conclusions on Identification of Customary International Law, With Commentaries' (Conclusion 2, Commentary para 1; Conclusion 10, Commentary para 3).

[58] Compare, for instance, ss. 8(a) and 8(b) of the UK Criminal Justice Act 1967, quoted *supra* note 55.

[59] This is in keeping with the observation, *supra* note 12 and accompanying text, that the requirement of proof beyond reasonable doubt does not require absolute certainty, but merely the elimination of all plausible alternatives.

[60] *Prosecutor v Dragomir Milošević (TC)* [2007] International Criminal Tribunal for the Former Yugoslavia IT-98-29/1-T [948]. See also, *Prosecutor v Stanislav Galić (AC)* [2006] International Criminal Tribunal for the Former Yugoslavia IT-98-29-A [132]; *Prosecutor v Tihomir Blaškić (TC)* [2000] International Criminal Tribunal for the Former Yugoslavia IT-95-14-T [501–12]; *Prosecutor v Zoran Kupreškić et al (TC)* [2000] International Criminal Tribunal for the Former Yugoslavia IT-95-16-T [513]; Héctor Olásolo, *Unlawful Attacks in Combat Situations: From the ICTY's Case Law to the Rome Statute* (Martinus Nijhoff Publishers 2008) 76–8.

> *comply with the precautionary requirements of the laws of war and the indiscriminate nature of the weapon used.*"

Consider, for instance, the *Galić* trial chamber's discussion of 'Scheduled Shelling 5', a shell-strike on Markale open-air market on 5 February 1994. It engaged at great length with a range of expert evidence to determine the source and direction of the attack, and concluded that the shell in question was fired from territory controlled by the Sarajevo Romanija Corps (a unit of the Bosnian-Serb Army commanded by General Galić) and was aimed at the market.[61] Its subsequent discussion of the legal characterisation of the attack referred simply to the absence of military targets in the vicinity,[62] and on this basis it found that civilians had been made the object of attack intentionally or recklessly.[63] In other words, from a seeming lack of military justification for attacks, the ICTY has been willing to infer the *mens rea* of the war crime of attacking civilian objectives.[64]

The 'elementary proposition'[65] that intent can be inferred from conduct and circumstances is also well-recognised at the ICC. For instance, the General Introduction to the Elements of Crimes of the International Criminal Court expressly provides: "*Existence of intent and knowledge can be inferred from relevant facts and circumstances.*"[66]

Similarly, the 'Means of Proof Digest' of the ICC's Case Matrix expressly endorses the possibility of inferring intent to attack civilian targets from conduct and context.[67] For proving that the perpetrator intended to make civilians the object of attack in relation to Rome Statute Arts. 8(2)(b)(i) and 8(2)(e)(i), the digest refers to means of proving the 'Knowledge of the

---

[61] *Prosecutor v Stanislav Galić (TC)* [2003] International Criminal Tribunal for the Former Yugoslavia IT-98-29-T [438–94]. Judge Nieto-Navia disagreed with the majority on this finding: *Prosecutor v Stanislav Galić (TC) (Separate and Partially Dissenting Opinion of Judge Nieto-Navia)* [2003] International Criminal Tribunal for the Former Yugoslavia IT-98-29-T [71–97]. The appeals chamber upheld the trial chamber's decision: *Galić appeal judgment* (n 60) paras 314–35.

[62] *Galić trial judgment* (n 61) paras 495–6. This was affirmed by the appeals chamber: *Galić appeal judgment* (n 60) paras 334–5.

[63] *Galić trial judgment* (n 61) para 596.

[64] To be clear, the jurisprudence of the ICTY has no formal significance before the ICC. Indeed, given the that the war crime of attacking civilians and civilian objects requires intent or recklessness in the ICTY's case law and intent under the Rome Statute, the ICTY's jurisprudence is technically irrelevant to the interpretation of the mental element of the war crimes defined in Rome Statute Arts. 8(2)(b)(i), 8(2)(e)(i) and 8(2)(e)(ii). Compare the ICC's interpretation of the war crimes of attacking civilian targets, *supra* notes 35-36 and accompanying text, with that of the ICTY in, e.g., *Dragomir Milošević trial judgment* (n 60) para 951; *Galić appeal judgment* (n 60) para 140. Nonetheless, the ICC has referred extensively to the jurisprudence of the ICTY, including in relation to these particular war crimes and specifically in relation to proof of mental elements. See, e.g., *Ntaganda trial judgment* (n 35) para 921; *Katanga trial judgment* (n 35) para 807. Consequently, the probative practice of the ICTY in relation to *mens rea* has been referred to here, first, because of its influence upon the probative practice of the ICC, and second, to demonstrate the pedigree of the specific probative practice of inferring intent from conduct and circumstances in international war crimes prosecutions.

[65] Triffterer and Ambos (n 10) 1117.

[66] Elements of Crimes of the International Criminal Court (General Introduction para 3).

[67] 'Means of Proof Digest of the International Criminal Court' <https://cilrap-lexsitus.org/means-proof-digest> accessed 7 October 2020. The Means of Proof digest is not a formal ICC publication, and nor does it have any authority before the ICC. However, the Means of Proof Digest and the Case Matrix Network and Legal Tools, of which it forms a part, were developed at the ICC, though the updating and maintenance of these tools have now been outsourced and the ICC disclaims any responsibility for their content. Thus, the Means of Proof Digest is not an authoritative source. However, it does provide a valuable guide to understanding how specific questions of law and procedure have been addressed in international criminal trials, at the ICC and also at other international criminal tribunals. It is in this capacity that it is referred to here.

perpetrator about the civilian status of the object of the attack' and means of proving the 'Intent of the perpetrator to target civilians'. In relation to the latter, it refers to:

> "*(a) evidence of the absence of military objects and/or military activity in the vicinity of the attacked area; (b) evidence showing that no military objects, real or believed, in the attacked area were targeted; (c) evidence of the extensive targeting of non-military objects in and around the attacked area concerned; (d) evidence of repeated shooting on civilians; (e) evidence of the indiscriminate nature of a weapon employed; (f) evidence of failure to take all necessary precautions to avoid injury, loss or damage to the civilian population; (g) evidence disproving accident caused by stray or ricocheting bullet; and, (h) evidence showing that the market area being attacked drew large number of people.*"[68]

None of these elements directly establishes an intention to attack civilians but instead, infers it from conduct and circumstances. Individually or collectively, these elements operate to discard alternative explanations for attacks on civilian targets, leaving intention to do so as the only plausible remaining possibility.

This probative practice – inferring mental element from conduct and circumstances which allow for no reasonable alternative explanation – is also evident in the case law of the ICC. Trial chambers have repeatedly endorsed the probative value of circumstantial evidence.[69] In relation to the war crime of attacking civilians in a non-international armed conflict, echoing the ICTY,[70] the *Katanga* trial chamber said that the intent to make civilians the object of attack:[71]

> "*may be inferred from various factors establishing that civilians not taking part in the hostilities were the object of the attack, such as the means and methods used during the attack, the number and status of the victims, the discriminatory nature of the attack or, as the case may be, the nature of the act constituting the attack.*"

Similarly, the *Ntaganda* trial chamber has held that in relation to attacks against co-located civilian and military targets, lack of discrimination or precaution in attack may constitute an attack against civilian targets.[72]

It must be emphasised that notwithstanding these broad formulations of proof of intent to attack civilian targets in the jurisprudence of the ICTY and the ICC, in practice both courts have usually been able to rely on far more specific and concrete evidence. Thus, for instance, the *Ntaganda* trial chamber relied on the use of the phrase 'kupiga na kuchaji' by the defendant in

---

[68] Similarly, to prove that the perpetrator intended to make civilian objects the object of attack under Rome Statute Art. 8(2)(b)(ii), the digest refers to: "*(a) evidence with regard to ability to target with precision; (b) evidence inferred from indiscriminate targeting; (c) evidence of the non-existence of military objects in the attacked area; (d) evidence of the demarcation etc being obvious to the perpetrator at the time of the attack; (e) evidence of the extensive targeting of non-military objects in and around the attacked area concerned.*"

[69] *Ntaganda trial judgment* (n 35) paras 69–70; *The Prosecutor v Jean-Pierre Bemba Gombo (TC III) (Trial Judgment)* [2016] International Criminal Court ICC-01/05-01/08-3343 [239]; *Katanga trial judgment* (n 35) para 109; *The Prosecutor v Thomas Lubanga Dyilo (TC I) (Trial Judgment)* [2012] International Criminal Court ICC-01/04-01/06-2842 [111].

[70] *Supra* note 60 and accompanying text.

[71] *Katanga trial judgment* (n 35) para 807.

[72] *Ntaganda trial judgment* (n 35) para 921.

ordering attacks, a term which the chamber interpreted as an exhortation to attack the entire Lendu community without distinction as to civilian and combatant.[73]

Most cases involving conduct of hostilities war crimes which have been tried before international courts and tribunals so far have featured inter-ethnic strife where entire communities are targeted, regardless of civilian or combatant status. In these contexts, it is unsurprising that trial chambers have frequently been able to rely on more specific evidence of intent to attack civilian targets. This should not, however, detract from the significance of broad endorsements of the possibility of inferring intent from conduct and circumstances. Indeed, the consistency of the ICC (and the ICTY) in maintaining the validity of this probative practice despite its limited utility in the specific cases before them embellishes its significance and pedigree.[74]

In sum, though intent can be established through an insight into the perpetrator's mind – e.g., through a confession or witness testimony, it can also be inferred from conduct and circumstances which do not admit of alternative explanation.

Once we acknowledge this probative practice it becomes evident that the responsibility gap thesis rests on a false premise. It focusses only on the subjective state of mind of the deploying or relying soldier and ignores the possibility of inferring intent from the manner and context of deployment or reliance.

The probative practice outlined here makes it possible to shift the focus of the *mens rea* analysis from the subjective state of mind of the soldier to the more objective manner and context of deployment or reliance of ACC. This shift in the focus of the inquiry produces a shift in the nature of the inquiry: from an inquiry into what the soldier knew and intended to what the soldier must have known and consequently intended. The former posits an ambitious inquiry into the actual knowledge and intent of the soldier, which is invariably impossible in the absence of a confession or verifiable declaration of intent.[75] The latter resolves this evidentiary difficulty by invoking a standard of reasonableness to infer constructive knowledge and on this basis presuming intent, subject to rebuttal.

The mechanics of this shift in the *mens rea* analysis can be seen in the example of the person who stabs another in the stomach. As discussed, it is uncontroversial that it is possible to infer intent to murder from the act of stabbing somebody in the stomach. The chain of reasoning here may be broken down as follows. First it is necessary to posit that a reasonable person would recognise the fatal consequences of stabbing another person in the stomach. On the basis of this standard of reasonableness we can assume, subject to rebuttal, that the perpetrator recognised the fatal consequences of their action. On the basis of this constructive knowledge we can assume, again subject to rebuttal, that the perpetrator intended the consequences of their

[73] *ibid* 415, 484, 922, 1181. See also, *Katanga trial judgment* (n 35) paras 850–5. The extent of the *Ntaganda* trial chamber's reliance on the use of the phrase 'kupiga na kuchaji' and the questionable correctness of its interpretation is highlighted in *The Prosecutor v Bosco Ntaganda (AC) (Defence Appeals Brief - Part II)* [2020] International Criminal Court ICC-01/04-02/06-2465-Red-Corr [75–90].

[74] For instance, compare para 865 of the *Katanga* judgment in which the chamber summarises its factual findings in relation to one part of the impugned attack, and the prior summary of evidence it relies on - § VIII(B)(2)(b) of the judgment. The former is phrased far more generally (even allowing for the limitations of summary) than the latter.

[75] In this regard, see note 54 *supra,* and accompanying text.

action.[76] In this manner we can infer intent to murder from the act of stabbing somebody in the stomach.

In this way the possibility of inferring intent from conduct and circumstances effectively skirts the responsibility gap. Whether the soldier knew that their deployment of or reliance on ACC would result in an attack on a civilian target and whether they intended that attack is still useful, but it is not determinative of the criminal responsibility of that soldier. Criminal responsibility can equally be based on what the soldier must have known and therefore what they can be presumed to have intended. What they must have known can be derived from the manner and circumstances of deployment of ACC, including what was known about the performance of the ACC, what its operational abilities and constraints were, what precautions were taken,[77] the context of deployment or reliance, etc.[78] If the soldier must have known that their deployment of or reliance on ACC would result in an attack on a civilian target, it may be presumed, subject to rebuttal, that the attack was intended, and would implicate the criminal responsibility of the soldier for the war crime of attacking civilian targets.

This approach to the war crimes of attacking civilian targets finds support in the reasoning of the *Ntaganda* trial chamber. The trial chamber broke the crime of attacking civilians into two requirements: directing attacks; against civilians. It defined the first requirement as "*selecting the target and deciding on the attack*".[79] Turning to the second requirement, it went on to say:[80]

> *"As the burden of proof lies with the Prosecution, it must be established that in the circumstances at the time,* ***a reasonable person could not have believed*** *that the individual or group he or she attacked was a fighter or directly participating in hostilities."* (emphasis supplied)

In effect, the trial chamber is using the standard of what a reasonable person must have known to determine knowledge of civilian status, and from a deliberate and otherwise unjustified attack against this target, it is willing to infer intent to attack civilians.

A similar willingness to replace knowledge with constructive knowledge is evident in the *Katanga* trial chamber's interpretation of awareness that a consequence 'will occur in the ordinary course of events' in Rome Statute Art. 30(2)(b) as 'virtual certainty' of occurrence.[81] It went on to describe virtual certainty in the following terms: "*it is nigh on impossible for him or her to envisage that the consequence will not occur.*"[82]

---

[76] On the well-established nature of both these assumptions – of foresight of consequences and of their intendment, see, Hart (n 2) 175.

[77] In his contribution to this volume Eric Talbot Jensen argues that the obligation to take precautions can be fulfilled by autonomous weapons themselves, but "*Such systems, before being employed, and presumably throughout the process, should be subject to a very rigorous weapons review process and potentially revisit that process as autonomous systems "learn" from their circumstances.*" See, Eric Talbot Jensen, 'Precautions and Autonomy in the Law of Armed Conflict', this volume, ch. 9. If this argument is correct, 'what must have been known' may still be derived from failure to conduct the required ongoing reviews or shortcomings in the reviews.

[78] In this regard, see the dicta of chambers of the ICTY and ICC quoted *supra* notes 60 and 70-72 and accompanying text.

[79] *Ntaganda trial judgment* (n 35) para 917.

[80] ibid 921.

[81] *Katanga trial judgment* (n 35) para 776.

[82] ibid 777. The original French text of the judgment provides: « *il lui est à peu près impossible d'envisager que la conséquence ne surviendra pas.* »

In summary, absent a confession or other declaration of intent, probative limitations are *constitutive* of the *mens rea* requirement. The *mens rea* of the soldier deploying or relying on ACC may, if possible, be determined by what they actually knew and therefore intended. But it can equally be determined by reference to what they must have known and intended. It is not necessary for the prosecution to establish knowledge of civilian status; it suffices to demonstrate that it was impossible not to have known of civilian status.

The responsibility gap thesis ignores the possibility of this shift from actual to constructive knowledge, from knowing to the impossibility of not knowing. As demonstrated here, this probative practice mitigates the responsibility gap to some extent. The scope and extent of this mitigation is, however, subject to three important restrictions and clarifications.

First, presumptions of what must have been known and consequently intended i.e., presumptions of recognition of consequences and thereby of intendment, are rebuttable. Thus, if what must have been known was not in fact known, the inference of intent may be rebuttable.[83]

Second, a clarification is necessary as to the role of the reasonableness standard here.

The standard of reasonableness provides a perspective for assessment, it does not determine the substance of the assessment. The role of the reasonable person is to provide a benchmark of comparison. Whether the substance of the comparison is what the reasonable person *should* have known or what the reasonable person *must* have known depends on the underlying rule. In relation to the war crimes of attacking civilian targets in the jurisprudence of the ICC, the reasonableness standard is deployed to determine what a reasonable person in similar circumstances *must* have known.[84] This is a significantly more stringent requirement than that of what a reasonable person should have known.[85]

It bears emphasis that the very contingent[86] idea of the 'reasonable person' in this case refers to the reasonable military commander. Compliance with the IHL rule of distinction is determined by reference to the standard of the reasonable commander,[87] and it stands to reason that the same reasonable commander would provide the benchmark of reasonableness for the purposes of the corresponding war crimes. This conclusion is also supported by the explicit connection drawn by the ICC between the war crimes of attacking civilian targets and the IHL rule of distinction.[88]

---

[83] In this regard, it is interesting to note that the particular phrasing of s. 8 of the UK Criminal Justice Act 1967, cited *supra* note 55 as support for the probative possibility of inferring intent from conduct and circumstances, was specifically intended to preserve the possibility of this inference, while ensuring that it remained rebuttable. See, Hart (n 2) 175.

[84] *Supra* notes 80-82 and accompanying text.

[85] The jurisprudence of the ICTY is inconsistent on this point. Some cases have used a 'should have known' standard: *Dragomir Milošević trial judgment* (n 60) para 952; *Galić trial judgment* (n 61) para 55. Others have used an 'impossibility of not knowing' standard: *Prosecutor v Pavle Strugar (TC)* [2005] International Criminal Tribunal for the Former Yugoslavia IT-01-42-T [280]; *Blaškić trial judgment* (n 60) para 180.

[86] Hart (n 2) 171: "*the judgment of the reasonable man very often is a mere projected shadow, cast by the judge's own moral views or those of his own social class.*"

[87] Sigrid Redse Johansen, *The Military Commander's Necessity: The Law of Armed Conflict and Its Limits* (Cambridge University Press 2019) 77.

[88] *Ntaganda trial judgment* (n 35) para 916; *Katanga trial judgment* (n 35) para 797.

In effect then, the mitigating influence of this probative practice on the responsibility gap will be limited to particularly egregious cases. The soldier in the responsibility gap will be deemed to have the required *mens rea* only when it is impossible for a reasonable commander in their position not to have known that deployment of or reliance on ACC would result in an attack on a civilian target. This is an important and significant limitation.

Third, the foregoing analysis raises undeniable concerns regarding the conflation of intent and recklessness or *dolus eventualis*.[89]

The simple answer to this concern is to acknowledge it. Inferring intent from conduct and circumstances, in shifting focus from what the soldier knew and intended to what must have been known and intended, assimilates the most egregious cases of recklessness or *dolus eventualis* into intent. This concern is valid, but it is also not new.[90] Moreover, it is mitigated to varying degrees by, first, the inherent limitation to the most egregious cases of recklessness or *dolus eventualis*; and, second, the continuing possibility of rebutting the inference of intent. In this regard, it bears emphasis that this chapter does not propose this conflation but instead, draws attention to its longstanding vintage in international war crimes prosecutions.

However, another answer to this concern might question its premises. A concern as to the conflation of intent and recklessness or *dolus eventualis* seems to assume a strict distinction between them, based on stable contours of the concept of intent and a bright line difference between intent and recklessness or *dolus eventualis*. This is a questionable assumption.

Intention and recklessness or *dolus eventualis* are inherently indeterminate concepts and the boundary between them is semantic and constructed rather than natural and immutable.[91] Consider the soldier in the responsibility gap. If they are certain that their deployment of or reliance on ACC will result in an attack on a civilian target and they intend this attack, they have intent. If they are not certain of this consequence, they lack intent. But what of the soldier who is 99% certain, or 95% or 90%? It is definitely possible to deny the (conceptual) intention of this latter soldier to attack a civilian target, but only by invoking a rigidly doctrinaire conception of intent which would sit uncomfortably with the social and political objectives of criminal responsibility.[92] Acknowledging the questionable distinction between 100% and 95%

[89] The difference between recklessness and *dolus eventualis* may be summarised as the difference between being culpably indifferent to risks and culpably accepting risks: Blomsma (n 50) 134.

[90] The same concerns have been raised in relation to the jurisprudence of the ICTY. See, e.g., Jens David Ohlin, 'Targeting and the Concept of Intent' (2013) 35 Michigan Journal of International Law 79. These concerns may implicate a broader tussle between IHL and war crimes law: whether war crimes are simply means for enforcing IHL through criminal sanction, or whether war crimes independently secure the same values as the corresponding IHL rules.

[91] Hart (n 2) 117.

[92] An interesting example here is the position of certain forms of wilful blindness in English law. In cases where the defendant intentionally chooses not to inquire into the truth of something because they have no doubt as to the answer, or because they don't want to know the answer, English law assumes knowledge on the part of the defendant, even while recognising the conceptual impossibility of knowledge: Simester and others (n 55) 157–9. See also, more generally, Weigend (n 55) 497–8. Incidentally, it is worth noting that the possibility of wilful blindness has featured extensively in concerns regarding the exclusion of recklessness and *dolus eventualis* from Rome Statute Art. 30: Eser (n 33) 931–2. It has been argued that the exclusion of wilful blindness cannot have been in the contemplation of the drafters of the Rome Statute: Knut Dörmann, Louise Doswald-Beck and Robert Kolb, *Elements of War Crimes Under the Rome Statute of the International Criminal Court: Sources and Commentary* (Cambridge University Press 2003) 131–2, 137–40, 145–7.

in this case forces the recognition that the line between intent and recklessness or *dolus eventualis* is necessarily fluid and contingent.[93]

Stated differently, if *mens rea* is inferred from conduct and circumstances then there cannot be a clear and definite boundary between intent and recklessness or *dolus eventualis*. It ceases to matter whether the defendant was certain or only 90% confident that stabbing the victim in the stomach would prove fatal; or whether the soldier was certain or only 95% confident that their deployment of or reliance on ACC would result in an attack on civilian targets. In both cases, the assessment of *mens rea* will focus on what the defendant must have known rather than what they did know.[94]

The argument that has been presented here may be summarised as follows.

The responsibility gap is concerned with the impossibility of intent to commit a war crime in the soldier who has reason short of certainty to believe that their deployment of or reliance on ACC might lead to an attack against civilian targets. This statement of the responsibility gap ignores the practicalities of proving *mens rea* in war crimes prosecutions (and criminal prosecutions more generally), where intent can be and is inferred from conduct and circumstances subject to elimination of plausible alternative explanations. This probative practice means that successful prosecution will not require establishment of what the soldier knew or intended, which may well fall short of intent to commit the war crime. Instead, it is sufficient to establish that the soldier must have known that civilian targets would be attacked, and that therefore, the soldier must have intended that attack. Consequently, the impossibility of the careless or uncertain soldier knowing that civilian targets would be attacked and of intending the attack is not an absolute bar to the criminal responsibility of that soldier. Based on available information as to the context and manner of deployment of or reliance on the ACC, constructive knowledge and consequently intent can be imputed to the soldier if it seems impossible that the soldier did not recognise the virtual certainty of attacking civilian targets. In effect, the probative practice of inferring intent from conduct and circumstances allows for the assimilation of egregious cases of recklessness or *dolus eventualis* into the category of intent, notwithstanding the conceptual impossibility of intent.

C. The restricted focus of the ICC on the most serious international crimes

This leads neatly to the second feature of international war crimes prosecution which mitigates the responsibility gap: the restricted focus of the ICC on those most responsible for the most serious crimes. This means that the defendants who are likely to attract the attention of the ICC

[93] The contingency of this line represents a socially rooted classification of degrees of culpability. This gives rise to the practical possibility that the line varies depending on the nature and circumstances of the crime, as is the case, discussed *supra* note 25, in relation to the standard of proof beyond reasonable doubt. Indeed, given the role of judicial interpretation in defining these concepts, it may be possible to point to an iterative process of judicial definition, social response and judicial redefinition. An example of this process relevant to the present context might be the revision of the *Gotovina* trial judgment by the ICTY appeals chamber following stakeholder responses. See, e.g., Gary D Solis, 'The Gotovina Acquittal: A Sound Appellate Course Correction' (2013) 215 Military Law Review 78.

[94] A further extension of this answer to the concern of conflating intention and recklessness might recognise that mental states, like emotions, are not (only) psychological states but socio-cultural practices. See, Sara Ahmed, *The Cultural Politics of Emotion* (2nd edn, Edinburgh University Press 2014) 8–9. This is not to say that a defendant is not intending 'something'. However, the meaning of 'intention' and the classification of the mental state of the defendant are contingent socio-cultural - and in relation to criminal responsibility, political – practices which themselves play a constitutive role in defining the defendant's mental state.

are precisely those whose deployment of or reliance on ACC was so egregious that they must have known of the virtual certainty of attacking civilian targets. And consequently, through the argument set out above, they may be presumed, subject to rebuttal, to have the requisite *mens rea*.

Rome Statute Art. 5 provides that "*The jurisdiction of the Court shall be limited to the most serious crimes of concern to the international community as a whole*." It goes on to indicate that war crimes generally are an example of such crimes. Art. 8(1) then provides that "*The Court shall have jurisdiction 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.*" Neither of these provisions amounts to a concrete restriction of the ICC's war crimes jurisdiction by reference to criteria of seriousness, plans or policies, or scale, but they suggest the prioritisation of war crimes which bear these features, and the de-prioritisation of isolated instances.[95]

The prioritisation suggested by Rome Statute Arts. 5 and 8(1) is mandated by Art. 17(1)(d) which posits the case not being "*of sufficient gravity to justify further action by the Court*" as a ground of inadmissibility.[96] The ICC has described the gravity requirement as a mandatory rarefaction of the already restricted (on the basis of seriousness in Art. 5) material jurisdiction of the court.[97] In its Art. 15 decision on the *Kenya* situation, the ICC described the gravity requirement in terms of restricting focus to those who bear the greatest responsibility for the gravest crimes,[98] and listed the following factors as 'useful guidance' for assessing gravity:[99]

> "*(i) the scale of the alleged crimes (including assessment of geographical and temporal intensity); (ii) the nature of the unlawful behaviour or of the crimes allegedly committed; (iii) the employed means for the execution of the crimes (i.e., the manner of their commission); and (iv) the impact of the crimes and the harm caused to victims and their families…*"

This approach to the gravity of crimes has been adopted in the policies of the Office of the Prosecutor,[100] as well as in its practice.[101] The interpretation of these criteria has been the subject of disagreement between the Prosecutor and the Court,[102] but the charging practice of the Prosecutor reflects a continued adherence to a strict interpretation of these criteria by

---

[95] Triffterer and Ambos (n 10) 321–2; Michael Bothe, 'War Crimes' in Antonio Cassese, Paola Gaeta and John RWD Jones (eds), *The Rome Statute of the International Criminal Court: A Commentary* (Oxford University Press 2002) 380–1.

[96] Triffterer and Ambos (n 10) 811–6.

[97] *Situation in the Republic of Kenya (PTC II) (Art 15 Decision)* [2010] International Criminal Court ICC-01/09-19 [56–7]; *Situation in the Democratic Republic of Congo (PTC I) (Arrest Warrant Decision)* [2006] International Criminal Court ICC-01/04-520-Anx2 [44, 46].

[98] *Kenya Art. 15* (n 97) para 59. See also, *Situation in the Republic of Côte d'Ivoire (PTC III) (Art 15 Decision)* [2011] International Criminal Court ICC-02/11-14-Corr [204]; *The Prosecutor v Bahar Idriss Abu Garda (PTC I) (Confirmation of Charges Decision)* [2010] International Criminal Court ICC-02/05-02/09-243-Red [30–2].

[99] *Kenya Art. 15* (n 97) para 62. See also, *Situation in Georgia (PTC I) (Art 15 Decision)* [2016] International Criminal Court ICC-01/15-12 [51–7]; *Côte d'Ivoire Art. 15* (n 98) para 204.

[100] Office of the Prosecutor of the International Criminal Court, 'Policy Paper on Case Selection and Prioritisation' 35–41 <https://www.icc-cpi.int/itemsDocuments/20160915_OTP-Policy_Case-Selection_Eng.pdf> accessed 7 October 2020.

[101] *Situation on Registered Vessels of Comoros, Greece and Cambodia: Article 53(1) Report* (Office of the Prosecutor of the International Criminal Court) [133–48].

[102] Compare ibid; *Situation on the Registered Vessels of Comoros, Greece and Cambodia (PTC I) (Decision on Review of Prosecutor's Decision Not to Initiate an Investigation)* [2015] International Criminal Court ICC-01/13-34 [20–50]. See also, Triffterer and Ambos (n 10) 816.

reference, *inter alia*, to requirements of scale and systemic nature.[103] Moreover, most cases of conduct of hostilities war crimes which have been tried at the ICC (and the ICTY) featured large-scale and systematic violations of IHL.[104]

In other words, war crimes prosecutions relating to ACC at the ICC will likely involve large-scale and systematic violations. These are precisely the sort of violations where the probative practice identified in the previous sub-section could be most significant in mitigating the responsibility gap.

IV. Conclusion: belligerents' responsibility

This chapter has examined two challenges to the prosecution of ACC-related IHL breaches as war crimes under the Rome Statute of the ICC. First, there is the difficulty of identifying the perpetrator of the conduct corresponding to the *actus reus* of the war crime, in accordance with the criminal standard of proof beyond reasonable doubt. Second, there is the difficulty of accommodating the actual mental state of the concerned human – inevitably negligence, recklessness or *dolus eventualis*, within the stringent *mens rea* requirement of intent and knowledge under the Rome Statute.

It has been argued here that the practical realities of charging and adjudicating war crimes may, in some cases, mitigate these challenges.

The challenge of identifying perpetrators relies on a fixed and rigid understanding of the reasonable doubt standard. In principle, the standard does not require absolute certainty but merely the elimination of reasonable alternatives. In practice, it is a variable and context-specific standard which, particularly in the context of international crimes, may not pose a very exacting threshold. Shorn of its mythologies of certainty, the reasonable doubt standard may not prove to be an insurmountable hurdle to identification and prosecution.

Similarly, the challenge of the responsibility gap relies on the conceptual impossibility of the actual mental state of the alleged perpetrator corresponding to the Rome Statute requirement of intent and knowledge. This framing of the problem ignores the universal probative practice

---

[103] Pre-Trial Chamber I's request to the Prosecutor to reconsider the decision not to initiate an investigation into the situation referred by Comoros, etc. was challenged and finally rejected by the Prosecutor. See, *Registered vessels of Comoros, Greece and Cambodia review* (n 102); Office of the Prosecutor of the International Criminal Court, 'Report on Preliminary Examination Activities 2017 - Registered Vessels of Comoros, Greece and Cambodia' <https://www.icc-cpi.int/itemsDocuments/2017-PE-rep/2017-otp-rep-PE-Comoros_ENG.pdf> accessed 7 October 2020. In relation to the preliminary examination into the conduct of UK forces in Iraq, the Prosecutor has noted: "*In the present situation, while there is a significant body of allegations, in light of the circumstances in which some of such allegations were collected, it remains unclear whether the crimes alleged were committed on the scale alleged by communication senders. Additionally, while several failings in army leadership, planning, and training, leading to prisoners' abuses were reported especially in the early phases of Op. Telic, the Office is seeking to assess the gravity of the role of other military or civilian personnel who may bear responsibility as an accessory or as a commander/superior.*" See, Office of the Prosecutor of the International Criminal Court, 'Report on Preliminary Examination Activities 2018' para 208 <https://www.icc-cpi.int/itemsDocuments/181205-rep-otp-PE-ENG.pdf> accessed 7 October 2020.

[104] A notable exception may be the *Abu Garda* case in which charges were based on a single attack against UN peacekeeping personnel resulting in 12 deaths (and eight further attempted killings) and damage to and appropriation of UN property: *Abu Garda confirmation* (n 98) paras 21–4. The Prosecutor has justified the gravity of the impugned conduct in this case by reference to the interests implicated – the security of peacekeeping personnel in the context of the role they play in maintaining the collective security order: *Registered vessels of Comoros, Greece and Cambodia Art. 53(1)* (n 101) para 145.

of inferring intent from conduct and circumstances i.e., shifting the frame of analysis from what the perpetrator actually knew and intended to what they *must have known* and *therefore intended.* This effectively means that in cases where the manner and mode of deployment of or reliance on ACC suggest that it was impossible that a reasonable commander in the position of the perpetrator would not have recognised the virtual certainty of an attack upon a civilian target, intent to attack civilian targets may be presumed from this knowledge. These particularly egregious cases are precisely the putative ACC-related war crimes most likely to satisfy the gravity requirement and attract the attention of the ICC. In other words, the practical realities of proving *mens rea* in war crimes prosecutions mitigate some of the challenges of the responsibility gap, at least in the most egregious cases.

These mitigating effects of the practice of charging and adjudicating war crimes are subject to two important clarifications.

First, it is necessary to emphasise that this chapter has not suggested the dilution of the criminal standard of proof or the conflation of intent and recklessness or *dolus eventualis*. Instead, it has drawn attention to the inherent indeterminacy of the reasonable doubt standard and the possibilities provided thereby for the variation or dilution of the standard in practice. And it has noted the ubiquity of inferring subjective mental states from objective physical indicators and has argued that this necessarily entails the replacement of actual knowledge with constructive knowledge, by reference to a benchmark of reasonableness. Neither of these well-recognised features of war crimes prosecutions (or prosecution more generally) is endorsed here, and nor can it be denied that they raise significant concerns for a body of law that is already vulnerable to withering critique on grounds of fairness and legitimacy.[105] That said, insofar as these practices exist,[106] they do provide some amelioration for the difficulties of prosecuting ACC-related IHL breaches as war crimes.

Second, the practical realities of proof beyond reasonable doubt and of establishing *mens rea* do not eliminate the difficulty of identifying perpetrators or resolve the responsibility gap. They operate in some cases to ameliorate these challenges and facilitate prosecution, for instance, in cases where the perpetrator of an ACC-related IHL breach can be identified with some degree of certainty, or in cases of egregious recklessness of *dolus eventualis*. This is not an insignificant argument, because it opens the door to destabilising the unitary and fixed nature of perpetrator identification and the responsibility gap and for seeing them as difficulties which may manifest differently in different cases. However, it undeniably leaves many residual cases where these factors operate to hinder the prosecution of ACC-related IHL breaches as war crimes.

---

[105] See, e.g., Frédéric Mégret, 'International Criminal Justice: A Critical Research Agenda' in Christine EJ Schwöbel (ed), *Critical Approaches to International Criminal Law: An Introduction* (Routledge 2014). Indeed, it may be possible to recast the problem of criminal responsibility for ACC-related war crimes as one of prosecution rather than conviction. The history of international criminal trials (and their discourse – *supra* notes 29-31 and accompanying text) suggests that international criminal courts and tribunals usually find a way around legal barriers to the conviction of those prosecuted before them. A more significant challenge to criminal responsibility may lie in the difficulty of prosecuting members of armed forces and citizens of technologically advanced states which are leading the race to develop these technologies. If that barrier is overcome, the technical problems posed by the standard of proof the responsibility gap may prove (relatively) easier to resolve.

[106] It is worth highlighting the possibility that the assessment of criminal practice presented here is simply 'moronic', in the sense that Bilbo describes in *Foucault's Pendulum* by Umberto Eco. "*Morons never do the wrong thing. They get their reasoning wrong. Like the fellow who says all dogs are pets and all dogs bark, and cats are pets, too, and therefore cats bark…Morons will occasionally say something that's right, but they say it for the wrong reason…"* See, Umberto Eco, *Foucault's Pendulum* (William Weaver tr, Vintage Books 2001) 65.

In concluding this chapter, it is useful to consider these residual cases briefly.

To begin with, it must be emphasised that while the challenges of individual criminal responsibility in these cases suggest a gap in the criminal enforcement of IHL, they do not imply a gap in the enforcement of IHL.

War crimes are serious breaches of IHL which implicate individual criminal responsibility.[107] But war crimes and the criminal responsibility they entail are only one part of the enforcement infrastructure of IHL. The broader and indeed primary part of IHL's enforcement infrastructure draws on the responsibility of belligerents for breaches of the rules of IHL. In recent years, the difficulties of enforcing the rules of IHL against states and non-state actors alike, the comparative successes of international criminal courts and tribunals, and the lure of ending impunity have combined to privilege individual criminal responsibility over belligerents' responsibility under IHL.[108] But even if subordinated, belligerents' responsibility persists in relation to IHL rules, and applies equally to ACC-related IHL breaches.[109]

Indeed, it may be easier to invoke the responsibility of belligerents for breaches of IHL than to prosecute those breaches as war crimes. Recklessness and negligence suffice for triggering responsibility under IHL provided they result in unreasonable errors in attack,[110] resolving the

---

[107] *Prosecutor v Duško Tadić (AC) (Decision on the Defence Motion for Interlocutory Appeal on Jurisdiction)* [1995] International Criminal Tribunal for the Former Yugoslavia IT-94-1-A [94]; International Committee of the Red Cross, 'Explanatory Note: What Are "Serious Violations of International Humanitarian Law"?' <https://www.icrc.org/en/doc/assets/files/2012/att-what-are-serious-violations-of-ihl-icrc.pdf> accessed 7 October 2020.

[108] This privileging of individual criminal responsibility has raised systemic challenges for IHL, including the de-prioritisation of those IHL norms which are not capable of individualisation and criminalisation, and the frequent misinterpretation of IHL norms. This argument is developed in greater detail in Paola Gaeta and Abhimanyu George Jain, 'Individualisation of IHL Rules Through Criminal Responsibility for War Crimes and Some (Un)Intended Consequences' in Dapo Akande and Jennifer Welsh (eds), *The Individualisation of War* (Oxford University Press 2021). See also, Paola Gaeta, 'Autonomous Weapon Systems and the Alleged Responsibility Gap' (International Committee of the Red Cross 2016) <https://www.icrc.org/en/publication/4283-autonomous-weapons-systems> accessed 7 October 2020.

[109] United States of America, 'Working Paper on Autonomy in Weapon Systems' (Group of Governmental Experts 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 2017) CCW/GGE.1/2017/WP.6 para 24 <https://www.unog.ch/80256EDD006B8954/(httpAssets)/20092911F6495FA7C125830E003F9A5B/$file/2018_GGE+LAWS_Final+Report.pdf> accessed 7 October 2020.

[110] IHL conduct of hostilities rules such as the rule of distinction do not guarantee the protection of civilians, they guarantee that civilians will not be made the object of attack and tolerate the possibility of an erroneous attack on civilians. This begs the question of which errors are permissible and the answer to that question relies on the standard of the reasonable commander. Consequently, deliberate attacks on civilians constitute a breach of the rule of distinction, as do negligent and reckless attacks (assessed in accordance with the standard of the reasonable commander). This approach to the requirements of the rule of distinction is reflected in, e.g., 'U.S. Department of Defense Law of War Manual' s 5.3 <https://dod.defense.gov/Portals/1/Documents/pubs/DoD%20Law%20of%20War%20Manual%20-%20June%202015%20Updated%20Dec%202016.pdf?ver=2016-12-13-172036-190> accessed 7 October 2020; 'Situation of Human Rights in Yemen, Including Violations and Abuses Since September 2014' (Human Rights Council 2018) A/HRC/39/43 7 <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/252/79/PDF/G1825279.pdf?OpenElement> accessed 7 October 2020; *Partial Award: Central Front - Ethiopia's Claim 2* (Eritrea-Ethiopia Claims Commission) [101–13]; 'Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia' (International Criminal Tribunal for the Former Yugoslavia 2000) paras 80–5

issue of the responsibility gap. As to the difficulty of identifying perpetrators, the transition from individual criminal responsibility to belligerents' responsibility entails less stringent burdens and standards of proof,[111] as well as a shift from identifying specific perpetrators to identifying the responsible belligerent and attributing the acts to it.

In relation to the difficulty of criminal responsibility for ACC-related IHL breaches, turning to the responsibility of belligerents is not only possible and easier, it may be more appropriate.

Consider the residual responsibility gap.[112] It includes soldiers whose mental state in deploying or relying on ACC falls below the principled intent and knowledge requirements of the ICC, and also falls below the practical 'must have known and therefore intended' threshold. In other words, these soldiers fall well below the *mens rea* requirement set out in the Rome Statute for the war crimes of attacking civilian targets.

A particular *mens rea* requirement represents a socio-political determination that a particular degree of culpability is required for criminal responsibility.[113] Breach of the rule is necessary but not sufficient to trigger criminal sanction: the breach must be accompanied by a specified culpable state of mind.[114] In other words, the impossibility of accommodating the soldier in the residual responsibility gap within the principled or practical *mens rea* requirements of the Rome Statute suggests the inadequate culpability of the soldier. The soldier may not have been culpable at all – they may have been justifiably unaware of a risk of an attack on a civilian target. Or the soldier may not have been sufficiently culpable – they may have been unjustifiably unaware of or accepted a risk of attacking civilian targets, but in either case it cannot be said that they must have known of the virtual certainty of attacking civilian targets.[115]

By conceptualising the mental state of the solider in the residual responsibility gap in terms of absent or insufficient culpability it becomes possible to reconceptualise the residual responsibility gap. The difficulty of satisfying the *mens rea* requirement of the Rome Statute does not only indicate the impossibility of individual criminal responsibility, it also indicates its inappropriateness. This reconceptualization of the residual responsibility gap recognises that IHL breaches are less frequently the result of individual deviance and more often arise from institutional factors including systemic interpretation and implementation of IHL.[116]

<https://www.icty.org/en/press/final-report-prosecutor-committee-established-review-nato-bombing-campaign-against-federal> accessed 7 October 2020.

[111] See, e.g., Banks (n 4) 247–8; Dederer and Singer (n 4) 439–45.

[112] The residual cases also include those where the perpetrator of the attack cannot be identified in accordance with even the variable and operational standard of proof beyond reasonable doubt. This is a 'practical' problem rather than the 'conceptual' problem posed by the responsibility gap and so it is not addressed separately here. However, the argument as to the inappropriateness of criminal responsibility for residual cases applies equally in this context: continuing and significant doubts as to the identity of the perpetrator should not be seen as an impediment to criminal responsibility but instead as an indication of the impropriety of criminal responsibility.

[113] It is also possible that particular conduct may be criminalised even in the absence of a culpable state of mind, as is the case with so-called 'strict liability' crimes.

[114] See, e.g., Hart (n 2) 160.

[115] It is worth emphasising that the insufficiency of culpability in these cases is only by reference to the specific and contingent standard of the Rome Statute. The inadequate culpability of negligence, recklessness or *dolus eventualis* in the context of lethal force and civilian lives is a socio-political choice and not an immutable fact.

[116] See, e.g., Matthew Talbert and Jessica Wolfendale, *War Crimes: Causes, Excuses, and Blame* (Oxford University Press 2018). This also reflects the broader idea that international crimes necessitate difficult distinctions between individual conduct and systemic criminality: e.g., Mégret (n 105) 28–30; Neha Jain, 'Individual Responsibility for Mass Atrocity: In Search of a Concept of Perpetration' (2013) 61 American Journal of Comparative Law 831, 831–2.

Reconceptualising the residual responsibility gap in terms of the inappropriateness of individual criminal responsibility rather than its impossibility does not deny that there has been an attack against a civilian target which may constitute a breach of IHL. It does not deny the importance of criminal responsibility in the enforcement of IHL.[117] And finally, it does not deny the broader significance and uses of criminal responsibility, from the perspectives of victims, offenders and society more broadly. But it does wonder whether in cases where attacks against civilian targets result from inadequate training, flaws in the development or use of ACC, or systemic misinterpretation or disregarding of IHL, criminal responsibility is the appropriate means of IHL enforcement. In these cases, it seems evident that the responsibility of belligerents should be the focus of enforcement efforts, and individual criminal responsibility is not only inapplicable but also inappropriate.[118]

Clearly, the appropriateness of criminal responsibility for cases in the residual responsibility gap cannot be determined in the abstract and must be considered on a case-by-case basis. But it would seem reasonable to assume that cases in the residual responsibility gap would largely arise from systemic factors rather than individual deviance, particularly given the exclusion of cases meeting the 'must have known and therefore intended' threshold.[119]

The possibility of systemic factors in IHL breaches relating to ACC (and lethal autonomy more generally) is particularly significant given the radical changes in the very nature of armed conflict which is facilitated by these technologies. The tactical, operational and strategic possibilities created by ACC may prove difficult to accommodate within the existing IHL framework and technological change may spur legal change. Military applications of ACC may challenge the binary of IHL compliance and breach, imperilling the prospect of prosecuting IHL breaches as war crimes. Put another way, cases in the residual responsibility gap may reflect disagreement as to the military use and manner of use of ACC. Those disagreements are entirely legitimate and indeed, necessary, but their resolution through the individual criminal responsibility of the implicated soldier is inappropriate.

The early years of the drone debates provide a fitting analogue here. In that context there were similar concerns about IHL breaches resulting from drone strikes and about the possibility of establishing criminal responsibility.[120] These concerns stemmed from disagreement as to how to assimilate the new military possibilities enabled by remote warfare within the requirements of IHL. Drone strikes have not resulted in significant war crimes prosecutions or convictions. Concerns about the manner in which drone strikes are conceptualised and conducted have been

[117] Although, as argued above, there are concerns as to the primacy of individual criminal responsibility in the enforcement of IHL and as to the exclusion of belligerents' responsibility.

[118] An interesting example here is the mistaken American strike on the Chinese embassy in Belgrade during the NATO intervention in Kosovo. In its final report to the prosecutor, the committee established to review the bombing campaign noted that the strike was erroneous and that the error resulted from misidentification of the Chinese embassy and inadequacies in the targeting process which prevented discovery of the error. But though the report deemed the strike to constitute a breach of the rule of distinction, it did not consider it appropriate to invoke the criminal responsibility of the pilots and senior military leaders on account of the systemic source of the error. See, 'Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia' (n 110) paras 80–5.

[119] Of course, there remains the possibility of a core residual responsibility gap comprising of cases featuring un-prosecutable individual deviance.

[120] See, e.g., Kevin Jon Heller, '"One Hell of a Killing Machine": Signature Strikes and International Law' (2013) 11 Journal of International Criminal Justice 89.

discussed largely within the framework of IHL, producing a slow and incomplete but discernible process of reconciliation between the novel practice of drone strikes and IHL.

The concerns raised by drones, like the concerns raised by the residual category of ACC-related IHL breaches, implicate the interpretation and application of substantive IHL norms at the level of belligerents rather than the actions of individual soldiers. Their resolution through the prism of the criminal enforcement mechanism of IHL would be inappropriate and unfair.