Hendrik Huelss
School of Politics and International Relations
University of Kent
h.c.huelss@kent.ac.uk

Author's post-print manuscript

Norms are what machines make of them: Autonomous Weapons Systems and the normative implications of human-machine interactions

Abstract

The emergence of autonomous weapons systems (AWS) is increasingly in the academic and public focus. Research largely focuses on the legal and ethical implications of AWS as a new weapons category set to revolutionise the use of force. However, the debate on AWS neglects the question what introducing these weapons systems could mean for how decisions are made. Pursuing this from a theoretical-conceptual perspective, the article critically analyses what impact AWS can have on norms as standards of appropriate action. The article draws on the Foucauldian "apparatus of security" to develop a concept that accommodates the role of security technologies for the conceptualisation of norms guiding the use of force. It discusses to what extent a technologically mediated construction of a normal reality emerges in the interplay of machinic and human agency and how this leads to the development of norms. The article argues that AWS provide a specific construction of reality in their operation and thereby define procedural norms that tend to replace the deliberative, normative-political decision on when, how, and why to use force. The article is a theoretical-conceptual contribution to the question of why AWS matter and why we should further consider the implications of new arrangements of human-machine interactions in IR.

Introduction

Autonomous weapons systems (AWS) are on the rise. While the international community has set up institutional procedures for a comprehensive, substantial discussion on the implications, possible regulation or prohibition of AWS, for example in the framework of the UN's Convention on Certain Conventional Weapons (UN-CCW) in Geneva, the development and deployment of AWS is accelerating. As of now, the potential large-scale usage of AWS is rather a strategic vision than the reality of contemporary security policy. But the crucial challenges of this development in terms of ethical, legal, and legitimacy aspects have not gone unnoticed. A central aspect are implications of AWS for international law governing the use of force. At the same time, studies at the intersection of law and International Relations (IR) have only started to respond to the emergence of AWS and slowly move beyond the existing, substantial literature on drone warfare as the first generation of unmanned weapons. Autonomous weapons as systems that 'can select and engage targets without further human intervention' (Heyns 2016, 4) differ from remotely-controlled drones in having varying degrees of autonomy as a main feature.

Autonomy in terms of technological, machinic agency is also the reason why the international debate focuses on AWS, while the technical features of drones have remained largely unconsidered. However, the example of drone warfare shows that the influence of international law regulating this type of the use of force is limited. In terms of *jus ad bellum*, the right to war, the deployment of drones in general is not considered as illegal if used in declared theatres of war. However, outside declared theatres of war, such as exemplified by the U.S. use of drones for the purpose of targeted killing in Pakistan, Somalia, or Yemen, drones operate in a grey area of international law, for example where states are considered as "unable or unwilling" (Bode 2017; Warren and Bode 2014) to counteract "terrorism". In terms of *jus in bello*, governing the conduct of war, drones are measured against principles such as discrimination and proportionality. Constituting so-called "precision" weapons, they even promise and are often considered as being better equipped to comply with these principles (Ekelhof 2017, 314; Walsh 2015; DeJonge Shulman 2018), but more importantly, law seems to lack leverage when it comes to the use of force executed by (new) technologies of warfare. This is particularly the case with regard to AWS as so far vague and unregulated weapons category.

The practice of warfare in the form of small-scale interventions and "precision"-strikes arguably changes our understanding of what appropriate use of force is. Notwithstanding that *jus ad bellum* and *jus in bello* apply to AWS just like to any other form of weapons system, this article makes two additional, novel arguments. First, the basic indeterminacy of law makes it important to shift the focus from law to norms as standards of appropriate action, which are to be differentiated from more narrowly defined legal rules. Norms are therefore decisive for how understandings of how to use force "appropriately" materialise in legal grey areas. Second and consequently, the article argues that we require a closer examination of the way norms are conventionally conceptualised as related to law in terms of normativity, as well as a greater awareness of how norms can be the products of technologically translated normality perceptions.

The article seeks to conceptualise the process of norm-emergence in the interaction of human actors and increasing technological autonomy. While conventional studies on the power of norms focus on how norms shape actors' behaviour, often linked to a pre-defined normativity specifying "good" or desirable actions, the article is interested in the role of normality, how technological autonomy can shape normality, and how understandings of normality have an impact on the emergence of norms and normativity. The article contributes to the debate on AWS by emphasising the importance of human-technology interactions for understanding the

role of autonomy in weapons systems. In other words, the decisive aspect to be considered is not the extent to which weapons systems operate autonomously in technical terms but how humans interact with technological systems and how these systems can construct a specific normality perception. This requires defocusing from the alleged uniqueness and importance of "drones" as the academically dominant materialisation of contemporary use of force policy by conceptually accommodating the complexity of arrangements comprising human and non-human agency as well as material and social structures.[1]

The law-focused political debate tends to employ rather clear-cut concepts of automation and autonomy, of human agents and of material objects as self-standing elements. Likewise, the rather misleading images of a "Terminator" or of "killer robots", which are points of reference in the public debate and also considered by different, critical perspectives (Garcia 2015; Campaign to Stop Killer Robots 2016; Arkin 2015; Garcia 2014; Gubrud 2015; Sparrow 2007; Wareham 2017), invoke the emergence of humanoid killer machines vested with sophisticated artificial intelligence (AI) as a new form of completely independent machinic agency, which does not exist yet nor is source of the central political-normative problem. Rather, emerging autonomous features in weapons systems require rethinking the meaning of agency as a relational concept, which is however largely neglected in relevant literature. Instead, the political debate in the framework of the UN-CCW remains focused on sounding options for how human actors can exert and keep control over machines, which is decisive for defining the legality of weapons systems. However, human agency, understood in this debate as the ability to control machines, is increasingly influenced or compromised by technologies. This is not even a novel phenomenon. The debate should therefore re-conceptualise its analytical approach to what constitutes agency towards a relational arrangement of humans and machines. We should understand "agency not as an attribute of either humans or machines, but rather as an effect of particular human-machine configurations" (Suchman and Weber 2016, 100).

Questions of autonomy and agency that are central for these reflections are closely linked to the issue of decision-making in terms of who decides and based on what standards. While the main premises of structuration theory posit the theoretically influential mutual constitution of agency and structure – reproduced by an assumed normative structure maintained by the "dual quality" (Wiener 2007, 48) of norms as constitutive and constituted elements – norms

[1] The reference to material and non-material or to technical and social qualities is not meant to establish a dichotomy between these dimensions. For example, elaborating on the question of how to categorise software or algorithms, which have a material component but are only meaningful in their social function, is beyond the scope of this article.

research in IR predominantly considers the impact of norms on actors while neglecting to problematise normativity and normality. This means that the implementation of norms, possible transformations therein and whether the initial meaning ascribed to norms is, in fact, translated into practice receives very little attention (see Huelss 2017; Bode and Karlsrud 2018). The structural quality of instruments taken as material structures, for instance, are examined from different perspectives (Fowler and Harris 2015; Lundborg and Vaughan-Williams 2015; Walters 2014; Barry 2013, 2001; Lemke 2015) and the role of "algorithms" (Zarsky 2016) as synonym of technological autonomy is increasingly in the focus of research. However, the question of how technologies implement and transform normativity, or shape our understanding of normality as the basis of norms, remains understudied.

The article draws on Michel Foucault's "apparatus of security" to approach the relationship between the normativity and normality of norms. Foucault provides for an innovative conceptual understanding of how norms emerge in assessing a social reality perceived as normality. Hence, the objective of this article is a theoretical reflection and conceptualisation of norms between notions of normativity and normality, using the increasing technological autonomy of systems such as drones and AWS as an empirical illustration.

The remainder of the article is structured as follows: the first section introduces the research problem by outlining the emergence of AWS and initial thoughts on how normality and norms interrelate. The second section discusses the usefulness of Foucault's apparatus of security as a theoretical perspective on the constitutive qualities of AWS, conceptualising the interplay of human and machinic agency. The third section provides a theoretical exploration of weapons systems as apparatuses of security and outlines how norms may emerge and are shaped by the increasing degree of technological autonomy. This is followed by the conclusion, arguing that the implications of the human-machine collaboration in the context of AWS and security technologies so far remain a theoretically under-conceptualised and understudied issue of IR scholarship.

AWS and research on norms in International Relations

The meaning of the term of AWS is contested.[2] Highlighting autonomous features, it represents a broad and heterogenous understanding of what falls into this weapons category. A widely reproduced concept of autonomy and control differentiates between "in the loop" systems that have humans in manual control, which mainly refers to remotely-controlled systems such as drones; "on the loop" systems, where humans are overseeing operations with the option to interfere, and "out of the loop" systems, where humans are completely absent from the operational scenario (see Heyns 2016, 14). Still, this only provides a broad and ambiguous distinction of systems because decision-making and human interference can take place at different levels, from steering actions to the programming of command sequences outside of operational settings. Others, suggesting the concept of "functional autonomy" (Boulanin and Verbruggen 2017, 18), underline the fragmented character of autonomy and the fact that autonomy can define certain functional aspects of a weapons systems only, such as targeting or navigation. This emphasises that autonomy is a platform-bound technological feature but not an essential quality. In a similar vein, Sharkey defines AWS as "systems that, once activated, can track, identify and attack targets with violent force without human intervention" (Sharkey 2016b, 23), thereby focussing on specific operational qualities of AWS.

Reflecting on automation and autonomy, roboticists emphasise that autonomous systems differ from automated systems, which are "running through a fixed pre-programmed sequence of action" (Winfield 2012, 12–13), because their "actions are determined by its sensory inputs, rather than where it is in a preprogramed sequence" (Winfield 2012, 12–13). In this regard, AWS in their very basic form can "decide" how to act based on sensor input and are vested with an action sequence within a pre-programmed range of possible actions. It could be argued that AWS are different from automated weapons in terms of having access to a greater range of action alternatives based on sensor input and, ultimately, their ability to learn and develop new action sequences, which will be discussed further below. Automated systems such as landmines or cruise missiles currently in operation lack this ability and depend on human updates to change their range of actions or even course and location.

Nevertheless, reactive automation and pro-active autonomy short of the ability to "learn", often overlap and are difficult to define in a way that satisfies the requirements of the legal-political debate. This is also one of the main reasons why the question of human control has

[2] Due to the given limitations, the article refrains from providing detailed examples of AWS in development or use. See (Adams 2001; Boulanin and Verbruggen 2017; Human Rights Watch 2012; Sharkey 2016a; Williams 2015) for empirical illustrations of AWS.

moved into the focus in recent years. Turning away from finding a clear-cut definition of AWS, the analytically more decisive question pertains to agency and the capacity to increase the number of possible actions/decisions from binary choices (see Suchman and Weber 2016, 102). The novel type of "decision" weapons that can develop and establish a new form of decision-making without human intervention and control because human actors might not even be able to retrace how decisions were made, is therefore clearly different from indiscriminate weapons such as landmines, but also from so-called, long-established "precision" weapons such as guided missiles.

In normative regards, the introduction of novel weapons systems such as chemical and nuclear weapons or even submarines has led to practices that have influenced how the appropriate use of force is defined. The political response to the use of anti-personnel landmines (Ottawa Treaty drafted in 1997) shows how weapons can trigger legal-normative changes in international relations. But this normative effect of weapons, usually linked to understandings of what is morally "right", depends on a political decision to establish international relations norms regulating or banning such weapons. This means that normativity is deliberatively created in a process that is supposed to shape the normality of the use of force. My research is, however, interested in reversing this perspective and investigating how perceptions of "normal" use of force influence norms and normativity in the case of AWS.

How does the political arena deal with AWS? The international community is considering the emergence and legal status of AWS in the framework of the UN-CCW since 2014. Initially, the topic was discussed at three CCW review conferences and three informal meetings of experts. In December 2016, CCW state parties agreed to formalise the deliberations by establishing a Group of Governmental Experts (GGE), holding an initial meeting in November 2017. Currently, twenty-eight countries have explicitly called for a ban on lethal autonomous weapons systems (Campaign to Stop Killer Robots 2018). In political terms, the decisive and controversial aspect of AWS is related to the fundamental legal, ethical, and philosophical question regarding whether, when, and how AWS should be involved in ending human life. States and NGOs advocating a ban of AWS hold a position that "fundamentally objects to permitting machines to take a human life on the battlefield or in policing, border control, and other circumstances" (Campaign to Stop Killer Robots 2017). Their ultimate

objective is a comprehensive, pre-emptive ban of AWS as a weapons category, which goes beyond merely regulating their usage.

Research on this topic has therefore centred on questions of autonomy, international legality, and the regulation of AWS (Roff 2015; Hammond 2014; Kastan 2013; Grut 2013; Garcia 2016; Sartor and Omicini 2016). This intersects with ethical considerations (Altmann 2013; Leveringhaus 2016; Lin 2010; Lokhorst and Van den Hoven 2012; Purves, Jenkins, and Strawser 2015; Schwarz 2018; Sharkey 2008; Asaro 2012). But current AWS research in IR is characterised by two shortcomings: first, technological developments are typically assessed against pre-defined and fixed legal or ethical standards. This underestimates the extent to which a broad spectrum of features of technological autonomy can influence perceptions of normality and the normal conduct of the use of force. Second, the focus on how to define autonomy, assuming the emergence of an independent machinic agency, risks overemphasising the importance of technological sophistication. Even relatively simple weapons systems meet the definition of basic technological autonomy. These systems are already at an operational stage of development or will approach this stage very soon. Moreover, even drones currently deployed deliver a technologically processed interpretation of reality that can be decisive for perceptions of normality and strongly influence manifestations of human agency. This means that the level and quality of technological sophistication is not decisive for the impact of AWS on norms. Far below the threshold of "killer robots", security technologies with a limited range of capabilities used as surveillance, reconnaissance, or defensive technologies have an impact on how humans perceive normality. We therefore require a perspective that accommodates "new models of human–machine collaborations" (Boulanin and Verbruggen 2017, 64) in terms of a new, complex understanding of agency. This agency is not a "capacity intrinsic to singular actors (human or artefactual)" but "an effect of subject/object relations that are distributed and always contingently enacted" (Suchman and Weber 2016, 99). To approach this issue further, I will briefly discuss how the role of norms is conventionally conceptualised in IR research before turning to a closer investigation of Foucault's apparatus of security.

Disentangling normality, norms and normativity

IR research on norms has studied their impact on actors for about three decades in the context of the constructivist agenda, for example inspired by the influential "logic of appropriateness" (March and Olsen 1989). While the function of norms in terms of impact is elaborated in various studies and approaches – not least contributing to the important debate on theories or

logics of actions in the 1990s and 2000s and their implications for IR, scholars tend to neglect the origin of norms in terms of sources and ways of emergence.

First, accepting the implicit sequence of emergence, function, and change of norms as an ontological reality of conventional constructivism has inspired research to consider norm emergence as being separate from the perceived function and change of norms. This has led to overemphasising law as the main source of norms, but also of the stability and relevance of a normative/legal structure for shaping actions. Furthermore, law as the main source of norms is often not problematised, which is part of the reason why studies on liberal "fundamental norms" (Wiener 2016, 23) such as human rights, democracy, liberty, or rule of law have dominated the academic debate. The motivations behind why norms are established or why actors follow specific norms might not always be highly relevant when researching the impact of norms. But the prevalent model of norm-emergence based on formal norm-setting in terms of deliberation works with an understanding of normativity as the "right" thing to do.

More importantly, research has been dominated by the concept of a sequence, which presupposes the existence of a defined normativity, emerging from deliberation, enshrined by formalisation, and implemented in practices. These deliberative practices are considered as shaping normality of international relations. In my conceptualisation, I contrast this sequence and the associated emergence of "deliberative norms" with a reverse effect in which perceptions of normality emerging in practices shape the development of norms. These norms might be formalised but are initially only locally relevant. This perspective also questions the existence as well as the stability of normative substance. While legal rules based on international law certainly play an important role in the governing of weapons – examples comprise conventions on nuclear weapons or chemical weapons where norms also have socialising effects (Tannenwald 1999; Jefferson 2014) – international law is often indeterminate (Koskenniemi 2011). This leads to two general questions: first, does a detailed normative substance exist and, second, should a legal structure be equated with a normative structure? These questions are particularly relevant for studies working with the "principle of contestedness" (Wiener 2014, 2017), promoting the influential concept of "contestation". Is there a concrete, stable and powerful normative structure to be contested and is the contestation of macro-structural norms more relevant than the interpretation, "localization" (Acharya 2004) or diverging implementation of norms?

Still, the questions of how norms are implemented and how and whether pre-defined norms influence practices remain widely neglected. This is analytically unfortunate, as these concerns are particularly relevant in complex settings of international politics/international

organisations, where the formal decision-making arena is spatially and temporally separated from implementation processes that are theoretically only expected to "translate" normative meaning. The abovementioned concepts of localization and contestation accommodate a second layer of decision-making agency at the level of norm-implementation. However, analysing the local emergence of norms should exceed studying the impact of implementation. It entails rethinking the very way in which norms come into existence: based on a simple concept of norms as standards of appropriate action, it can be assumed that norms emerge in the context of a specific interpretation of reality. This reality or understanding of normality is de-linked from normativity.

In human-machine collaborations, this would mean that the reality or normality conception offered by the machine informs the norm-perception of the human actor – this type of *procedural* norm (Bode and Huelss 2018) is different from political-normative norms. But the research focus on a supposedly stable normative structure established by political actors in specific deliberations or law-making settings leaves alternative ways of norm-emergence unconsidered and overemphasises the importance of deliberative norms in establishing normativity. It neglects discussing concepts such as normality, normal, or normalisation sufficiently, which are related to the origin and function of norms, and the role of specific practices in this.

A perspective on practices as defining standards of appropriate action beyond the clear conceptual link to law is neither particularly novel nor contested. This perspective argues that the use of AWS, due to their perceived advantages in terms of efficiency and effectiveness will change how and when the use of force is considered as appropriate – depending on the degree of autonomy, examples of such advantages comprise technological supremacy in terms of the high-speed processing of large amounts of quantitative data, persistence, or reliability (see Ekelhof 2017, 313). Research on drone warfare also confirms how practices can change use of force policies outside of explicit international law (Bode 2016; Warren and Bode 2015). Here, practices as patterned, structured and repeated ways of doings things in a social context (Leander 2008, 18) have led to the adoption of new, implicit standards of appropriate use of force in international relations. While it might serve as an interesting case for exploring the agency-structure model, in which agency has constituted the normative structure that in turn constitutes future actions, the emergence of normative substance outside of the deliberate norm-setting context remains unclear. In other words, how the use of specific weapons

systems and the technical qualities of these weapons systems influence perceptions of what normality is, requires further conceptual accommodation.

Moving towards reconsidering the role of technology in this context is related to earlier studies on the constitutive or "performative" (Leander 2013) quality of weapons systems. This entails unpacking the complex relation between human actors, non-human agency, and materiality embedded in a specific socio-political environment. The emergence of technological autonomy linked to advances in AI and represented by the growing research interest in algorithms, introduces a novel, potentially game changing element for the mature models of human agency and material/non-material structure dominant in IR theory. In this regard, the novelty of "deliberative autonomy" (Boulanin and Verbruggen 2017, 24) and its implications for deliberative norms, for instance, clearly goes beyond the performative, structural effects of materiality.

While the development and use of AI-based weapons systems should not overshadow the current influence of practices involving "stupid machines" (Bode and Huelss 2017), future scenarios could see a much stronger role of machinic autonomy. From analytical but also operational and political viewpoints, the main challenge is the systemic complexity of AWS featuring machine learning algorithms for example, which can make investigating and predicting the exact operation of algorithms analytically infeasible. An algorithm, basically defined as a "systematic procedure that produces – in a finite number of steps – the answer to a question or the solution of a problem" (Encyclopaedia Britannica 2006) is only meaningful in the interplay of programme (code) executing the algorithm, software platform and material configurations, to name a few components. While an algorithm is basically only a sequence of steps that provide an unambiguous set of instructions, it is crucial to understand an algorithm "as a running system, running in a particular place, on a particular computer, connected to a particular network, with a particular hardware configuration" (Dourish 2016, 5). Moreover, the importance of algorithms in the context of weapons systems lies in their increasing sophistication, particularly regarding machine learning algorithms used to improve the performance of a specific task. Without going into more technical detail here, supervised or semi-supervised learning algorithms work with labelled training data in the form of input variable and desired output variables. They can learn a mapping function that predicts the outputs for inputs even in cases where the input was not part of the training data (semi-supervised). The learning process is supervised when the correct output is known. In the case of unsupervised learning, the algorithms work with a set of data that only consists of input and learn about underlying structures and patterns in the input data, while the data has not been

labelled as in the case of supervised learning and correct answers do not exist (see Brownlee 2016). The algorithm is autonomous in finding "interesting" patterns in the data and can improve its performance in the process of learning. In the practice of security policy, pattern recognition based on machine learning plays an experimental role, for example, in predictive policing and offender risks assessment (Babuta, Oswald, and Rinik 2018; Kaufmann, Egbert, and Leese 2019) or in the identification of military targets (Boulanin and Verbruggen 2017, 24–25; Sharkey 2010).

While the technical complexity of AWS is challenging, we require an analytical perspective that considers a comprehensive, relational security technology of human-machine interaction, regardless of what specific technical qualities it might have. In the following, the article proceeds with conceptualising AWS as part of a Foucauldian apparatus of security, thereby, first, discussing how technology affects the constitution of normality, and, second, contextualising this in considering the possible role of algorithms and machine learning.

Mechanisms of governing: The apparatus of security and AWS

As part of the lecture series "Security, Territory, Population" held during 1977 and 1978, Foucault outlined a comprehensive perspective on the historical conditions and emergence of contemporary governing, widely known as "governmentality". While the depth and amplitude of the different concepts and themes addressed in these lectures make an adequate but necessarily de-contextualised presentation of key concepts challenging, I highlight in the following that Foucault introduced a novel perspective on the emergence and function of norms, as well as offering a redefined perspective on the meaning of security. I will first outline three mechanisms of governing which Foucault presented in these lectures. This will enable me to clarify the meaning and status of the "apparatus of security" in Foucault's perspective and as an analytical concept relevant for this article.

Approaching governing as "a regime of practice" (Foucault 1991, 75), Foucault differentiates between the mechanism of discipline and the apparatus of security. Both systems are defined by the opposing role norms play therein: in the mechanism of the discipline, the law codifies a norm (Foucault 2007, 56). Once established, the legal structure equals a normative structure and is the basis for an "optimal" model, differentiating the normal from the abnormal: "In the disciplines, one started from a norm, and it was in relation to the training carried out with reference to the norm that the normal could be distinguished from the abnormal" (Foucault

2007, 63). Foucault considers this process as "normation" rather than normalisation (Foucault 2007, 57), the meaning and importance of normalisation will be outlined further below.

This model resembles the predominant status of international law as the source of fundamental norms (normative structure) in constructivist approaches. In broad terms, the normality of international relations, understood as the definition of the range of normal and hence permitted and beneficial behaviour, often linked to normativity, is decided at the level of the normative structure. In contrast to this concept of static, fixed norms being constituted in deliberative processes of norm-setting or law-making, the apparatus of security is based on the idea that statistical methods and calculations are decisive for determining norms based on an empirical normality. Fundamentally, the mechanism of security in the Foucauldian sense primarily derives an understanding of normality from statistical distributions of the average normal, such as the "Bell Curve", and "these distributions will serve as the norm" (Foucault 2007, 63). In other words, Foucault differentiates between the fixed, pre-defined norm (discipline), which seeks to bring reality in line with it; and the construction of norms based on statistical perceptions of a normal reality (security).

This differentiation works with normalisation as a decisive process in establishing and promoting norms. Again, the apparatus of security is contrasted with the opposite concept of the legal mechanism: Foucault noted that "if it is true that the law refers to a norm, and that the role and function of the law therefore – the very operation of the law – is to codify a norm (…) the problem that I am trying to mark out is how techniques of normalization develop from and below a system of law, in its margins and maybe even against it" (Foucault 2007, 56). Foucault, therefore, highlights that law is not necessarily linked to the emergence of norms, while normalisation as a process is different from what would conventionally be considered as "making normal". The mechanism of discipline in contrast

> "divides the normal from the abnormal. Disciplinary normalization consists first of all in positing a model, an optimal model that is constructed in terms of a certain result, and the operation of disciplinary normalization consists in trying to get people, movements, and actions to conform to this model (…) it is not the normal and the abnormal that is fundamental and primary in disciplinary normalization, it is the norm" (Foucault 2007, 57).

In introducing the apparatus of security, Foucault summarised that "we have then a system that is, I believe, exactly the opposite of the one we have seen with the disciplines", which act with reference to the norm that distinguishes the normal from the abnormal. The apparatus of security is based on "different curves of normality, and the operation of normalization

consists in establishing an interplay between these different distributions of normality (…) these distributions will serve as the norm" (Foucault 2007, 63).

The apparatus of security does not, however, stand for a specific, material security technology. Rather, the apparatus is a social relation comprising different, interrelated elements – for example, actors, material technologies, and techniques, which together build a constitutive complex. In this apparatus, normality is assessed, the normal is specified, and the process of normalisation specifies norms. Furthermore, for Foucault, security is a constructed, constitutive process instead of a relatively stable state or condition. Security is an administrative operation that seeks to assess, monitor, regulate, and stabilise developments in providing expectations for processes in specific situations.

Overall, Foucault points to the emergence of what Amoore calls the "mobile norm" (Amoore 2011, 31), which can be contrasted with the fixed norm central for the idea of a normative structure both, guiding decision-making but also being shaped in deliberations. The apparatus of security responds to the fluidity of an empirical reality to be assessed and measured in considering normalities. The effect is a dislocation of space and time in terms of two different, not necessarily linked levels of decision-making: first, formal deliberation and norm-setting and second, procedural norm-construction. However, what is assessed as the average normal changes in different situations and the norm informed by "curves of normality" changes accordingly. This now takes place in the context of machinic agency that becomes increasingly influential in the process of normalisation.

Machine learning and the control problem

In order to approach the impact of AWS on the normative dimension further, I recall the main, controversial questions of the current debate: what does the autonomy in weapons systems mean and to what extent can there be "meaningful human control" (Crootof, 2016; Roff and Moyes, 2016)? Discussions on these aspects in the political and academic community are still emerging but one set of scholars argues that some sort of human control is always present – be it direct interference in operations or indirect steering by programming AWS to perform a clearly defined sequence and set of actions (Noel Sharkey 2016a; Roff and Moyes 2016). However, the practice of drone warfare already suggests that remote-controlled vehicles *inter alia* represent a technologically mediated reality in human-machine interactions. In other words, drones, which were initially designed as surveillance and reconnaissance instruments, are gathering, filtering and processing data, thereby offering a specific numerically mediated translation or visualisation of reality to the operator. The higher

the degree of technological sophistication, the more difficult the control of operations, even if a human is still in or on the loop. The operational advantage of increasing technical autonomy lies in processing huge amounts of data in a comparatively short time. Therefore, AWS can be superior in dealing with a specific numerical complexity while performing pre-defined tasks but their capacity to deal with social abstraction and interpretative issues is insufficient. This complexity in terms of speed and quantity is not meaningfully controllable by humans in operational scenarios and therefore poses a major challenge for the regulation and control of AWS.

In an abstract sense and considered from the viewpoint of mechanisms of discipline outlined above, remote-controlled weapons systems (such as drones) are still in line with the conventional perspective on a codified, pre-defined norm translating normativity that distinguishes the normal from the abnormal: these weapons are supposed to be used according to clear norms and objectives defining procedures. While remote-control does not rule out the emergence of a specific, technologically created normality, the difference to AWS is that the latter are governed by machinic agency, which is at heart of autonomy. This has two major implications: first, political oversight and public scrutiny of AWS' actions are impeded by the lack of open and transparent information on the nature of algorithms, source code, as well as software and hardware architecture. The basic rationality informing the algorithm remains unclear, actions are unpredictable, and it is difficult to assess whether the algorithm is correctly implemented by code (Dourish 2016, 4). This problem is vital when considering machine learning by, for example, deep, neural networks where the input and output data might be known but the process of reasoning is not (Boulanin and Verbruggen 2017, 17). Basically, "[a] neural network is an interconnected assembly of simple processing elements, units or nodes, whose functionality is loosely based on the animal neuron. The processing ability of the network is stored in the interunit connection strengths, or weights, obtained by a process of adaptation to, or learning from, a set of training patterns" (Gurney 2004, 13). The artificial neural network hence consists of connections which link the output of one neuron to the input of another neuron, governed by a learning rule such as an algorithm.

The increasing importance of machine learning in the U.S. military for instance, makes this control problem particularly crucial. The U.S. Defense Advanced Research Projects Agency (DARPA) announced in September 2018 a multi-year investment of more than $2 billion in new and existing programs called the "AI Next" campaign" (DARPA 2018), which also focuses on "Next Generation AI", particularly machine learning. Previously, the U.S. Department of Defense (DoD) had announced the "Establishment of an Algorithmic Warfare

Cross-Functional Team (Project Maven)" (U.S. Department of Defense 2017) in April 2017. The project "will initially provide computer vision algorithms for object detection, classification, and alerts" (U.S. Department of Defense 2017) based on full-motion video sequences and involve deep learning solutions in analysing data.

The current focus on deep learning in defence research suggests that this dimension will play an increasing role in the development of autonomous functionality in weapons systems. It is precisely here that questions of meaningful human control become most pertinent because autonomous learning and problem-solving is clearly different from following a pre-defined, programme sequence and deciding on different pre-defined options. Machine learning represents a level of autonomy that fundamentally alters the ways machines are governed and actions are informed, which has significant implications for accountability, responsibility, and human control. It elevates the role of algorithms from directional, sequential (quantitative) problem solving mechanisms to administrative, deliberative operators. The requirements of control and predictability and the emergence of acceptable technical solutions, however, will also influence to what extent machine learning will become part of AWS beyond technological feasibility: in the military context, "there is little room for algorithmic mystery, and the Department of Defense has identified explainability as a key stumbling block" (Knight 2017). But the translated, filtered and transformed depiction of reality offered by weapons systems already in operation, such as the surveillance functions of drones, can have profound effects in representing normality, measuring what normal is, and offering particular understandings of norms. I suggest broadening the understanding of agency to include the reality shaping capacities of machines in depicting normality, which links the operation of AWS but also of automated systems to the concept of apparatus of security. In this regard, the apparatus of security can serve as a conceptual approach to the study of AWS, covering different degrees of autonomy.

In the following, I outline the constitution of normality by weapons system referring to drones in particular, which gives an initial understanding of the relevance of sophisticated weapons systems for the normative dimension in international relations. These reflections are meant to encourage further theoretical conceptualisations as well as subsequent empirical studies of AWS.

AWS and normalisation: the normal as the norm

In order to investigate the constitutive quality of AWS as part of an apparatus of security, it is useful to reconsider the "distinction between security and discipline" (Foucault 2007, 55),

which is mirrored in the transition from remotely-controlled weapons to autonomous, machine learning systems at the end of the spectrum. The mechanism of discipline is fundamentally based on "optimal sequences or co-ordinations" (Foucault 2007, 57), which correspond to the ideal type of an instrument accurately translating the user's intention into action. This understanding represents the idea of a normative structure shaping actors and actions in an unaltered way in its purest form. It is what Foucault called "a normation (*normation*) rather than normalization" (Foucault 2007, 57). "Technologies of security", in contrast, are relevant in the "rationalization of chance and probabilities" (Foucault 2007, 59), establishing a norm based on assessment and calculations defining different distributions of normality. My point is that AWS in the broad sense can constitute norms as standards of appropriate procedures, which are derived from a representation of a normal reality. The process of normalisation is crucial because it produces a filtered understanding of normality. In other words, in contrast to a simple form of automation, which steers a system by running through a pre-programmed and fixed sequence of action, autonomous systems have the ability to adapt and to select a specific way of operating (out of a range of options) based on external data input from sensors and potentially the ability to develop new options independently (learning algorithm).

How does this normalisation unfold in the practice of the use of force? Aradau and Blanke (Aradau and Blanke 2018) have studied the effects of drone warfare with regard to the technological construction of subjects of security and self/other relations. They introduce the concept of anomaly as a supplement to abnormality in the Foucauldian sense. While the authors rightly describe the capacity of anomaly detection as the core reason for why these security technologies are deployed, how does this practice link to the political dimensions of normativity and norm-construction as standards of appropriate action? The authors outline that calculating anomaly is based on modelling what is similar and dissimilar. However, we should also consider that the tendency to develop an understanding of an optimum, of deviance, and of ways of bringing the deviant in line with the normal does not disappear. The technological analysis of anomaly patterns is transferred into a model of normal and abnormal and provides an image of normality with political consequences for the use of force. While it is argued that "[a]lgorithmic security has not only relinquished the desire for normalising the 'other', but calculations of spatial, temporal, and topological similarity seemingly bypass the negative polarity of racialised and gendered other" (Aradau and Blanke 2018, 20), the detection of anomaly is only an initial step in an action sequence that leads to a process which

could be abstractly labelled “normation”: a norm based on differentiating normal from abnormal in an analysis of anomalies still emerges, and this norm not only defines statistically what appropriate behaviour is in the sense of the average but also how to act appropriately on this behaviour. Any deviation from the average will be counteracted.

Moreover, the data anomaly and its visualisation constitute a projection, “produced from fragments of data, from isolated elements that are selected, differentiated and reintegrated to give the appearance of a whole” (Amoore 2011, 29). It is decisive that the norm is constructed based on projections as post-human reality, which, however, must be transferred to the dimension of political or moral decision-making if human agency is still involved, contributing to the spatial and temporal fragmentation of decisions. In the case of AWS, this fragmentation might mean that the subsequent step of human intervention is completely lost, and any “meaningful” decision-making process is centralised within machinic agency. In drone warfare, the information received by the human operator when steering the drone is based on sensor data input that translates elements of a digital, numeric reality into a form of visualisation. Increasing machinic autonomy further increases the extent to which humans “in” or “on the loop” depend on the machinic interpretation of reality. This process of normalisation creates specific understandings of what an accurate representation of normality is, regardless of the level of technological sophistication.

A drone as a part of an apparatus of security can contribute to the surveillance of human movements, to the visual mapping of data, or to the identification of targets. Examples of these complex data processing and interpretation activities are found in “The Drone Papers”, published on the website “The Intercept”.[3] Practices such as the “Geolocational Watchlist” or “Small Footprint Operations” give evidence of how apparatuses of security unfold. The detailed insights into the “pattern of life” assessment and analysis of SKYNET in Pakistan, deployed with the aim of courier detection via machine learning (The Intercept 2015), give an impression of how programmes powered by learning algorithms scan the environment for patterns to establish an understanding of the normal and the abnormal. In this case, Skynet is a programme that analyses data from 55 million mobile phone users to approach a state of predictability and filter the most likely couriers of terrorist networks (Robbins 2016). This is a good example of how normalisation works: the assessment constructs a map of a specific normality. It provides data on normal behaviour in terms of movement patterns of mobile phone users in Pakistan. It initially detects patterns of anomaly, but the technologically calculated normality in combination with other data serves as the norm and deviation from

[3] https://theintercept.com/

this norm is considered as abnormal behaviour that is to be eliminated - individuals displaying this abnormal behaviour are potential targets of drone strikes. In this regard, how and when the use of force is expected to produce the most favourable outcome and is therefore appropriate does not depend on deliberative decision-making but is conveniently offered by the apparatus of security. In other words, a norm is established based on specific and varying calculations of what normality is. These procedural norms establish expectations for appropriate actions in a given operational scenario. The norm is no longer a fixed, a-priori standard of appropriate action, meeting specific ethical standards, but becomes a flexible technical benchmark.

Our conventional academic and political understandings of norms suggest that the use of force is governed by deliberative decisions of when and how violence is appropriate. Studies on the practice of use of force show that such understandings of appropriateness are often shaped in the various ways of patterned actions. It is pointed out that "[t]he premise is not that violence produces its targets (…) It is rather to focus on the dynamics through which systematic violence effectively creates worlds in which operations of tracking and targeting, done in the name of security, work as sociotechnologies of reciprocal (if also asymmetric) enmity and ongoing insecurity" (Suchman, Follis, and Weber 2017, 986). This suggests that security "solutions" produce the problems they are putatively addressing. In my view, we should, however, also consider how targeting produces violence as a process that is now mainly based on machinic agency. It produces specific subjectivities (the targets), which are already inherently linked to the "legitimate" use of force justified by referring to the alleged objectivity of the apparatus of security.

In this context, further research into the social qualities of algorithmic translations or interpretations in terms of trust and doubt becomes crucial (Amoore, 2018). Algorithmic processing means that doubt is reduced to probability and the multifaceted character of doubt in the social sense is condensed to a binary, numeric output grounded in the technological necessity to reach conclusive decisions. Furthermore, the technological element of "ground truth" is important. It describes the accuracy of training data, for example in the aforementioned version of supervised learning. The fact that the output of an algorithm might be false on the basis of a human assessment, "but its degree of truth will always remain intact in its relations to data" (Amoore, 2018: 6) further obfuscates the extent to which doubt is eliminated.

The question of trust in machines is certainly not novel and the well-known case of Stanislav Petrov, who correctly doubted the report of a U.S. nuclear missile attack produced by the Soviet early-warning satellite network in 1983, underlines this vividly. It shows that human agency is influenced by an electronically mediated reality but also that doubt is often only based on an intuitive understanding of social impossibility ("this feels wrong"), rather than on an informed, "superior" form of knowledge. Friendly-fire incidents, for example involving aircrafts engaging ground targets or anti-aircraft missile systems such as the Patriot, repeatedly prove to what extent human doubt is weakened by the electronically constructed reality (see Suchman and Weber 2016, 102): "[a]ccording to a summary of a report issued by a Pentagon advisory panel, Patriot missile systems used during battle in Iraq were given too much autonomy, which likely played a role in the accidental downings of friendly aircraft" (Singer 2005).

Increasing technological autonomy and related machinic agency seriously questions the viability of the concept of "meaningful human control" currently important in the debate on AWS at the UN -CCW. This concept, which is linked to the aforementioned differentiation of varieties of human intervention, is underdeveloped and does not take account of the fading capacity to articulate human doubt in the human-machine relationship: "[t]his human in the loop is, though, an impossible figure who can never meaningfully engage the plurality of posthuman doubts lodged within the calculus" (Amoore, 2018: 9).

In losing human control, the political-deliberative norm in the conventional sense of normativity is replaced by a norm derived from a technologically mediated distinction between normal and abnormal. The autonomy of weapons systems also promises a fusion of discipline, legal, and security mechanisms in the sense of how Foucault described their parallel existence. The SGR-A1, a stationary sentry robot developed in the 2000s by Samsung in South Korea and partly deployed in the Korean Demilitarized Zone is an example of a security and disciplinary machine that monitors, assesses and executes. The weapon is equipped with heat and motion sensors and an auto-firing system. While it is operated as a human-controlled system, it is noted that the system has the capacity to detect, select, and engage a human target without human intervention and can be hence switched into an autonomous mode (Goose and Wareham 2016).

The massive deployment of such systems is, however, still a future scenario. The information processed within an apparatus of security stems from different sources, such as conventional intelligence reports, and analytically identifying the "best" targets based on movement

patterns does not automatically translate into a kill list that is just implemented. The ability to process and deliberate data on-board is still very limited or non-existent in present systems (Boulanin and Verbruggen 2017, 28). But considering the implications for more advanced systems steered autonomously by algorithms and machine learning solutions, this would not only mean that the conventional, deliberative, separate and separable sequence of decision and implementation collapses because the system combines all steps in one instance. It also means that norms as standards of appropriateness are increasingly the product of machinic agency with serious consequences inter alia for ethical decision-making and political accountability. The effects of human-machine interactions on meaningful human control are yet largely unaddressed, not least in IR theory that is still strongly influenced by a conventional understanding of the meaning, relation and role of agency and structure as fundamental to social relations.

Conclusion

This article addressed the implications of AWS for the normative dimension of international relations and security policy. The normative dimension as an analytical object is understood as the entanglement of normality and normativity fundamental for perspectives on norms, which are however rarely considered comprehensively. In the absence of legal regulations of AWS, norms as standards of appropriate use of force are particularly important in this context (Bode and Huelss 2018). The primary objective and contribution of this article lies in proposing a theoretical approach to study how AWS with varying degrees of autonomy (and in this sense of machinic agency) can have an impact on norms. The article made two initial suggestions: first, to refrain from considering the role of sophisticated, AI-type solutions or algorithms as an issue separate from the function, material/social environment, and objective of a specific weapon system. Second, to consider AWS as an apparatus of security, which is a complex ensemble of interplaying material and non-material elements, including human and non-human elements of agency.

The conventional perspective on norms primarily considers how and under what conditions norms govern and influence actions, presupposing the existence of a pre-defined and established normativity translated by these norms. In this regard, technologies and techniques are expected to simply implement norms in practice. Viewpoints on public policy instruments and the materiality of things strive to accommodate the non-human aspects of impact and therefore offer a diverting perspective. In line with this, the emergence of autonomous qualities of machines such as AWS adds an important, potentially game-changing, aspect. We

have not yet reached a stage where machines replace the decision-making and implementation role of humans. Instead of overemphasising the relevance of self-standing autonomy in form of machinic agency and of AI in general for making technologies important, this article shows how AWS as a part of an apparatus of security can influence perceptions of normality as the basis of procedural norms. The decisive argument pertains to a closer investigation of human-machine interaction, particularly of how machinic agency influences and weakens human agency, primarily understood as "control" in the debate on AWS.

The Foucauldian process of normalisation outlined in this article is useful for understanding how norms are created, produced and reproduced by AWS. This process is at the heart of the mechanism of security, which is not based on a pre-defined, fixed norm like the mechanism of disciplines. It assesses populations to calculate the average normal, to define a norm on this basis, and to gain an understanding of normal and abnormal behaviour. The current practice of drone surveillance and the preparation of drone strikes exemplifies collecting and processing vast amounts of quantitative data to single out the cases most deviant from the normal. The outcome is a procedural norm that provides a benchmark to determine when the use of force is most appropriate. In that, it shapes understandings of normality when it comes to the use of force. The point is that these procedures and technological capacities replace the fundamental question of when the termination of human life is justified. The question of life and death turns from a normative deliberation exercised by humans into a calculative practice of machinic agency, where numeric probabilities and the elimination of human doubt play a major role.

At this point, it is uncertain what kind of technologies will be deployed in the future and whether a comprehensive regulation or ban of AWS will materialise. However, weapons with autonomous capacities, governed by algorithms, are currently developed, built, and tested. The comforting assumption that AI informed weapons in terms of "killer robots" are unrealistic and might never have any relevance is misleading and underestimates to what extent technological autonomy can be influential. It also underestimates the implications of an apparatus of security, comprising a complex arrangement of human and machinic agency and interaction that render restrictions, regulations and the vision of complete, meaningful human control of weapons systems, difficult. This is precisely what the apparatus of security tries to comprehend. Although we cannot predict with certainty what kind of weapons technologies will emerge in the future, human agency will be increasingly compromised by technological solutions such as learning algorithms. Therefore, contributing to an informed, theoretically

sophisticated debate on this theme is an urgent matter for the political and academic discourse.

References

Acharya, Amitav. 2004. "How Ideas Spread: Whose Norms Matter? Norm Localization and Institutional Change in Asian Regionalism." *International Organization* 58 (02): 239–75.

Adams, Thomas K. 2001. "Future Warfare and the Decline of Human Decisionmaking." *Parameters* 31 (4): 1–15.

Altmann, Jürgen. 2013. "Arms Control for Armed Uninhabited Vehicles: An Ethical Issue." *Ethics and Information Technology* 15 (2): 137–52.

Amoore, Louise. 2011. "Data Derivatives: On the Emergence of a Security Risk Calculus for Our Times." *Theory, Culture & Society* 28 (6): 24–43.

———. 2018. "Doubtful Algorithms: Of Machine Learning Truths and Partial Accounts." *Theory, Culture & Society* accepted manuscript.

Aradau, Claudia, and Tobias Blanke. 2018. "Governing Others: Anomaly and the Algorithmic Subject of Security." *European Journal of International Security* 3 (01): 1–21.

Arkin, Ronald. 2015. "The Case for Banning Killer Robots." *Communications of the ACM* 58 (12): 46–47.

Asaro, Peter. 2012. "On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making." *International Review of the Red Cross* 94 (886): 687–709.

Babuta, Alexander, Marion Oswald, and Christine Rinik. 2018. "Machine Learning Algorithms and Police Decision-Making: Legal, Ethical and Regulatory Challenges." Whitehall Report 3-18. London: Royal United Services Institute. https://rusi.org/publication/whitehall-reports/machine-learning-algorithms-and-police-decision-making-legal-ethical.

Barry, Andrew. 2001. *Political Machines: Governing a Technological Society*. London/New York: The Athlone Press.

———. 2013. *Material Politics*. Chichester, UK: John Wiley & Sons, Ltd.

Bode, Ingvild. 2016. "How the World's Interventions in Syria Have Normalised the Use of Force." The Conversation. February 2016. https://theconversation.com/how-the-worlds-interventions-in-syria-have-normalised-the-use-of-force-54505.

———. 2017. "'Manifestly Failing' and 'Unable or Unwilling' as Intervention Formulas: A Critical Analysis." In *Rethinking Humanitarian Intervention in the 21st Century*, edited by Aiden Warren and Damian Grenfell, 164–91. Edinburgh: Edinburgh University Press.

Bode, Ingvild, and Hendrik Huelss. 2017. "Why 'Stupid' Machines Matter: Autonomous Weapons and Shifting Norms." Bulletin of the Atomic Scientists. 2017. https://thebulletin.org/why-"stupid"-machines-matter-autonomous-weapons-and-shifting-norms11189.

———. 2018. "Autonomous Weapons Systems and Changing Norms in International Relations." *Review of International Studies* 44 (03): 393–413.

Bode, Ingvild, and John Karlsrud. 2018. "Implementation in Practice: The Use of Force to Protect Civilians in United Nations Peacekeeping." *European Journal of International Relations* OnlineFirst (October): https://doi.org/10.1177/1354066118796540.

Boulanin, Vincent, and Maaike Verbruggen. 2017. "Mapping the Development of Autonomy in Weapons Systems." Stockholm.

Brownlee, Jason. 2016. “Supervised and Unsupervised Machine Learning Algorithms.” *Machine Learning Mastery* (blog). March 15, 2016. https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/.
Campaign to Stop Killer Robots. 2016. “Country Views on Killer Robots.” 2016. www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC_CountryViews_13Dec2016.pdf.
———. 2017. “Diplomatic Efforts Falter.” 2017. https://www.stopkillerrobots.org/2017/05/diplomatsfalter/.
———. 2018. “Country Views on Killer Robots. 22 November 2018.” https://www.stopkillerrobots.org/wp-content/uploads/2018/11/KRC_CountryViews22Nov2018.pdf.
DARPA. 2018. “AI Next Campaign.” 2018. https://www.darpa.mil/work-with-us/ai-next-campaign.
DeJonge Shulman, Loren. 2018. “Precision and Civilian Casualties: Policymakers Believe Drones Can Be Precise. That May Not Be Enough.” Just Security. August 2, 2018. https://www.justsecurity.org/59909/precision-civilian-casualties-policymakers-drones-precise-enough/.
Dourish, Paul. 2016. “Algorithms and Their Others: Algorithmic Culture in Context.” *Big Data & Society* 3 (2): 1–11.
Ekelhof, Merel A.C. 2017. “Complications of a Common Language: Why It Is so Hard to Talk about Autonomous Weapons.” *Journal of Conflict and Security Law* 22 (2): 311–31.
Encyclopaedia Britannica. 2006. “Algorithm.” 2006. https://www.britannica.com/science/algorithm.
Foucault, Michel. 1991. “Question of Method.” In *The Foucault Effect. Studies in Governmentality*, edited by Graham Burchell, Colin Gordon, and Peter Miller, 73–86. Chicago: University of Chicago Press.
———. 2007. *Security, Territory, Population. Lectures at the Collège de France 1977-1978*. Basingstoke: Palgrave Macmillan.
Fowler, Chris, and Oliver JT Harris. 2015. “Enduring Relations: Exploring a Paradox of New Materialism.” *Journal of Material Culture* 20 (2): 127–48.
Garcia, Denise. 2014. “The Case Against Killer Robots.” *Foreign Affairs*, 2014.
———. 2015. “Killer Robots: Why the US Should Lead the Ban.” *Global Policy* 6 (1): 57–63.
———. 2016. “Future Arms, Technologies, and International Law: Preventive Security Governance.” *European Journal of International Security* 1 (01): 94–111.
Goose, Stephen, and Mary Wareham. 2016. “The Growing International Movement Against Killer Robots.” *Harvard International Review* 37 (3).
Grut, Chantal. 2013. “The Challenge of Autonomous Lethal Robotics to International Humanitarian Law.” *Journal of Conflict and Security Law* 18 (1): 5–23.
Gubrud, Mark. 2015. “Semi-Autonomous and on Their Own: Killer Robots in Plato’s Cave.” Bulletin of the Atomic Scientists. April 2015. https://thebulletin.org/semi-autonomous-and-their-own-killer-robots-plato’s-cave8199.
Gurney, Kevin. 2004. *Introduction to Neural Networks*. Taylor & Francis.
Hammond, Daniel N. 2014. “Autonomous Weapons and the Problem of State Accountability.” *Chicago Journal of International Law* 15: 652–87.
Heyns, Christof. 2016. “Autonomous Weapons Systems: Living a Dignified Life and Dying a Dignified Death.” In *Autonomous Weapons Systems. Law, Ethics, Policy*, edited by Nehal Bhuta, Susanne Beck, Robin Geiss, Hin-Yan Liu, and Claus Kress, 3–20. Cambridge: Cambridge University Press.

Huelss, Hendrik. 2017. “After Decision-Making: The Operationalisation of Norms in International Relations.” *International Theory* 9 (3): 381–409.
Human Rights Watch. 2012. *Losing Humanity: The Case against Killer Robots*.
Jefferson, Catherine. 2014. “Origins of the Norm Against Chemical Weapons.” *International Affairs* 90 (3): 647–61.
Kastan, Benjamin. 2013. “Autonomous Weapons Systems: A Coming Legal Singularity.” *Journal of Law, Technology & Policy* 45: 45–82.
Kaufmann, Mareile, Simon Egbert, and Matthias Leese. 2019. “Predictive Policing and the Politics of Patterns.” *The British Journal of Criminology* 59 (3): 674–92.
Knight, Will. 2017. “The Dark Secret at the Heart of AI.” MIT Technology Review. 2017. https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/.
Koskenniemi, Martti. 2011. *The Politics of International Law*. Oxford: Hart.
Leander, Anna. 2008. “Thinking Tools.” In *Qualitative Methods in International Relations: A Pluralist Guide*, edited by Audie Klotz and Deepa Prakesh, 11–27. Basingstoke: Palgrave Macmillan.
———. 2013. “Technological Agency in the Co-Constitution of Legal Expertise and the US Drone Program.” *Leiden Journal of International Law* 26 (04): 811–31.
Lemke, Thomas. 2015. “New Materialisms: Foucault and the ‘Government of Things.’” *Culture & Society* 32 (4): 3–25.
Leveringhaus, Alex. 2016. *Ethics and Autonomous Weapons*. London: Palgrave Macmillan.
Lin, Patrick. 2010. “Ethical Blowback from Emerging Technologies.” *Journal of Military Ethics* 9 (4): 313–31.
Lokhorst, Gert-Jan, and Jeroen Van den Hoven. 2012. “Responsibility for Military Robots.” In *Robot Ethics: The Ethical and Social Implications of Robotics*, edited by Patrick Lin, Keith Abney, and George A Bekey, 145–56. Cambridge, Mass.: MIT Press.
Lundborg, Tom, and Nick Vaughan-Williams. 2015. “New Materialisms, Discourse Analysis, and International Relations: A Radical Intertextual Approach.” *Review of International Studies* 41 (01): 3–25.
March, James G, and Johan P Olsen. 1989. *Rediscovering Institutions: The Organizational Basis of Politics*. New York: Free Press.
Purves, Duncan, Ryan Jenkins, and Bradley J. Strawser. 2015. “Autonomous Machines, Moral Judgment, and Acting for the Right Reasons.” *Ethical Theory and Moral Practice* 18 (4): 851–72.
Robbins, Martin. 2016. “Has a Rampaging AI Algorithm Really Killed Thousands in Pakistan?” *The Guardian*, 2016.
Roff, Heather M. 2015. “Lethal Autonomous Weapons and Jus Ad Bellum Proportionality.” *Case Western Reserve Journal of International Law*, no. 47: 37–52.
Roff, Heather M., and Richard Moyes. 2016. “Meaningful Human Control, Artificial Intelligence and Autonomous Weapons. Briefing Paper Prepared for the Informal Meeting of Experts on Lethal Autonomous Weapons Systems. UN Convention on Certain Conventional Weapons.” Geneva.
Sartor, Giovanni, and Andrea Omicini. 2016. “The Autonomy of Technological Systems and Responsibilities for Their Use.” In *Autonomous Weapons Systems: Law, Ethics, Policy*, edited by Nehal Bhuta, Susanne Beck, Robin Geiss, Hin-Yan Liu, and Claus Kress, 39–74. Cambridge: Cambridge University Press.
Schwarz, Elke. 2018. *Death Machines: The Ethics of Violent Technologies*. Manchester: Manchester University Press.
Sharkey, Noel 2008. “The Ethical Frontiers of Robotics.” *Science* 322 (5909): 1800–1801.
———. 2010. “Saying ‘No!’ To Lethal Autonomous Targeting.” *Journal of Military Ethics* 9 (4): 369–83.

———. 2016a. “Staying in the Loop: Human Supervisory Control of Weapons.” In *Autonomous Weapons Systems*, edited by Nehal Bhuta, Susanne Beck, Robin Geis, Hin-Yan Liu, and Claus Kres, 23–38. Cambridge: Cambridge University Press.
———. 2016b. “Staying in the Loop: Human Supervisory Control of Weapons.” In *Autonomous Weapons Systems*, edited by Nehal Bhuta, Susanne Beck, Robin Geis, Hin-Yan Liu, and Claus Kres, 23–38. Cambridge: Cambridge University Press.
Singer, Jeremy. 2005. “Report Cites Patriot Autonomy as a Factor in Friendly Fire Incidents.” SpaceNews.Com. March 14, 2005. https://spacenews.com/report-cites-patriot-autonomy-factor-friendly-fire-incidents/.
Sparrow, Robert. 2007. “Killer Robots.” *Journal of Applied Philosophy* 24 (1): 62–77.
Suchman, Lucy, Karolina Follis, and Jutta Weber. 2017. “Tracking and Targeting: Sociotechnologies of (In)Security.” *Science, Technology, & Human Values* 42 (6): 983–1002.
Suchman, Lucy, and Jutta Weber. 2016. “Human-Machine Autonomies.” In *Autonomous Weapons Systems: Law, Ethics, Policy*, edited by Nehal Bhuta, Susanne Beck, Robin Geiß, Hin-Yan Liu, and Claus Kreß, 75–102. Cambridge: Cambridge University Press.
Tannenwald, Nina. 1999. “The Nuclear Taboo : The United States Basis of and the Normative Nuclear Non-Use.” *International Organization* 53 (3): 433–68.
The Intercept. 2015. “SKYNET: Courier Detection via Machine Learning - The Intercept.” 2015. https://theintercept.com/document/2015/05/08/skynet-courier/.
U.S. Department of Defense. 2017. “Deputy Secretary of Defense. Memorandum. Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven).” 2017. https://www.govexec.com/media/gbc/docs/pdfs_edit/establishment_of_the_awcft_project_maven.pdf.
Walsh, James Igoe. 2015. “Precision Weapons, Civilian Casualties, and Support for the Use of Force: Precision Weapons.” *Political Psychology* 36 (5): 507–23.
Walters, William. 2014. “Drone Strikes, Dingpolitik and Beyond: Furthering the Debate on Materiality and Security.” *Security Dialogue* 45 (2): 101–18.
Wareham, Mary. 2017. “Banning Killer Robots in 2017.” The Cipher Brief. 2017. https://www.hrw.org/news/2017/01/15/banning-killer-robots-2017.
Warren, Aiden, and Ingvild Bode. 2014. *Governing the Use-of-Force in International Relations. The Post-9/11 US Challenge on International Law*. Basingstoke: Palgrave Macmillan.
———. 2015. “Altering the Playing Field: The U.S. Redefinition of the Use-of-Force.” *Contemporary Security Policy* 36 (2): 174–99.
Wiener, Antje. 2007. “The Dual Quality of Norms and Governance Beyond the State: Sociological and Normative Approaches to ‘Interaction.’” *Critical Review of International Social and Political Philosophy* 10 (1): 47–69.
———. 2014. *A Theory of Contestation*. Berlin: Springer.
———. 2016. “Contested Norms in Inter-National Encounters: The ‘Turbot War’ as a Prelude to Fairer Fisheries Governance.” *Politics and Governance* 4 (3): 20-36.
———. 2017. “A Theory of Contestation —A Concise Summary of Its Argument and Concepts.” *Polity* 49 (1): 109–25.
Williams, John. 2015. “Democracy and Regulating Autonomous Weapons: Biting the Bullet While Missing the Point?” *Global Policy* 6 (3): 179–89.
Winfield, A. F. T. 2012. *Robotics: A Very Short Introduction*. Very Short Introductions. Oxford: Oxford University Press.
Zarsky, Tal 2016. “The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making.” *Science, Technology & Human Values* 41 (1): 118–32.