Check for updates

ESSAY

REVISED Artificial Intelligence Technologies and Practical Normativity/Normality: Investigating Practices beyond the Public Space

[version 2; peer review: 4 approved, 1 approved with reservations]

Ingvild Bode, Hendrik Huelss

Department of Political Science and Public Administration, Syddansk Universitet, Odense, Region Syddanmark, Denmark

v2 **First published:** 22 Sep 2023, **3**:160
https://doi.org/10.12688/openreseurope.16536.1
Latest published: 11 Mar 2024, **3**:160
https://doi.org/10.12688/openreseurope.16536.2

Abstract

This essay examines how artificial intelligence (AI) technologies may shape international norms. Following a brief discussion of the ways in which AI technologies pose new governance questions, we reflect on the extent to which norm research in the discipline of International Relations (IR) is equipped to understand how AI technologies shape normative substance. Norm research has typically focused on the impact and failure of norms, offering increasingly diversified models of norm contestation, for instance. But present research has two shortcomings: a near-exclusive focus on modes and contexts of norm emergence and constitution that happen in the public space; and a focus on the workings of a pre-set normativity (ideas of oughtness and justice) that stands in an unclear relationship with normality (ideas of the standard, the average) emerging from practices. Responding to this, we put forward a research programme on AI and practical normativity/normality based on two pillars: first, we argue that operational practices of designing and using AI technologies typically performed outside of the public eye make norms; and second, we emphasise the interplay of normality and normativity as analytically influential in this process. With this, we also reflect on how increasingly relying on AI technologies across diverse policy domains has an under-examined effect on the exercise of human agency. This is important because the normality shaped by AI technologies can lead to forms of non-human generated normativity that risks replacing conventional models about how norms matter in AI-affected policy domains. We close with sketching three future research streams. We conclude that AI technologies are a major, yet still under-researched, challenge for understanding and studying norms. We

Open Peer Review

Approval Status ✓ ✓ ✓ ✓ ?

	1	2	3	4	5
version 2 (revision) 11 Mar 2024	✓ view		✓ view	✓ view	? view
version 1 22 Sep 2023	? view	✓ view			

1. **Marijn Hoijtink**, University of Antwerp, Antwerp, Belgium
2. **Marion Laurence**, Royal Military College of Canada, Toronto, Canada
3. **Damián Tuset Varela**, University of Jaén, Jaén, Spain
4. **Haekal Al Asyari**, Universitas Gadjah Mada, Yogyakarta, Indonesia
   University of Debrecen, Debrecen, Hungary
5. **Max Lesch**, Peace Research Institute Frankfurt, Frankfurt, Germany

Any reports and responses or comments on the article can be found at the end of the article.

should therefore reflect on new theoretical perspectives leading to insights that are also relevant for the struggle about top-down forms of AI regulation.

Plain language summary
This essay critically discusses the interplay between artificial intelligence (AI) technologies and norms, focusing on the discipline of International Relations (IR). Here, norms are usually conceptualised as the products of reflection processes where actors debate and decide on "agreeable" standards. IR research on AI governance therefore turns towards international public and deliberative processes and their outcomes. But this risks missing two crucial pathways of how AI technologies may shape norms: first, how practices of designing and using AI technologies performed by actors at sites outside of the public eye make norms; and second, how design and use AI technologies may contribute to not only shaping normativity, in the sense of ideas of oughtness and justice, but also normality, in the sense of the average and the normal. We conclude that AI technologies pose an important challenge for research on norms and consequently require new conceptual and regulative perspectives.

Keywords
artificial intelligence, governance, norms, practices, technologies, International Relations, military, weapon systems

This article is included in the Horizon 2020 gateway.

This article is included in the European Research Council (ERC) gateway.

This article is included in the Artificial Intelligence and the Social Sciences and Humanities collection.

Corresponding author: Ingvild Bode (bode@sam.sdu.dk)

Author roles: Bode I: Conceptualization, Formal Analysis, Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **Huelss H**: Conceptualization, Formal Analysis, Investigation, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 852123).

Copyright: © 2024 Bode I and Huelss H. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Bode I and Huelss H. **Artificial Intelligence Technologies and Practical Normativity/Normality: Investigating Practices beyond the Public Space [version 2; peer review: 4 approved, 1 approved with reservations]** Open Research Europe 2024, **3**:160 https://doi.org/10.12688/openreseurope.16536.2

First published: 22 Sep 2023, **3**:160 https://doi.org/10.12688/openreseurope.16536.1

REVISED **Amendments from Version 1**

Through engaging in detail with the review comments provided, we have changed the essay in the following ways:

First, we now reflect much more explicitly on the dynamic interplay between normativity and normality. We completely agree with the reviewers' suggestion to characterize these dimensions firmly as inseparable in practice and as two sides of the same coin.

Second, we have separated our thinking on post-human agency by putting this into a dedicated section and reflecting more explicitly on where "machine agency" comes from and how it remains tied to human agency. We also added an argument about "distributed normativity" here.

Third, in response to comments about the precise relationship between public-deliberative processes and operational practices in shaping normativity/normality we have added a new section at the end of the essay entitled "future theorising on AI and norms". This covers both this point, as well as the important questions the reviewers raised about actors beyond states and where our political engagement should lie.

Finally, we have added further clarification about the methodological problems that studying processes of designing AI technologies raises.

Any further responses from the reviewers can be found at the end of the article

Introduction

Artificial intelligence technologies (AITs) have gained increasing media attention, in particular since the roll-out of so-called generative forms of AI systems, such as ChatGPT that produce text based on probabilistic prediction of "characters, words, or sentences in a sequence" (Tarnoff, 2023). In response, the political and popular mood appears to fluctuate between a techno-optimist infatuation with the purported dramatic advances of 'AI' and techno-sceptical concerns with its societal risks, including worrying failures. At the moment of writing (February 2024), AI technologies remain lightly regulated but the growing number of applications based on AI technologies result in major governance demands for national, regional, and international policymakers to address (Paul, 2022; Tallberg *et al.*, 2023). The emerging governance and regulation regime of AI technologies takes various forms, with both general and policy-field specific approaches being sought by public and private actors – or a combination thereof (Veale *et al.*, 2023). AI governance outcomes sought after may be of a formal type, such as legislation or the creation of dedicated international institutions (Maas & Villalobos, 2023), but also include informal principles and suggestions. Internationally and regionally, governance processes associated with, chiefly, public actors have already produced outputs, such as *UNESCO's Recommendation on the Ethics of AI* (2019) and the *OECD's Recommendation on AI* (2019). Another noteworthy example is the Group of Governmental Experts (GGE) on emerging technologies in the area of lethal autonomous weapon systems (LAWS). Working in the context of the United Nations Convention in Certain Conventional Weapon Systems (CCW), the GGE on LAWS has provided a forum for states to discuss challenges associated with integrating AI technologies into weapon systems since 2017. On a regional scale, the European Union (EU) has agreed on its *Artificial Intelligence Act* in 2023, which looks set to become the first set of regional, legally enforceable rules *specifically pertaining* to AI technologies. Further, the Council of Europe's Committee on AI is working to develop a legal instrument, which is publicly available in draft form (December 2023, Draft Framework Convention on AI, Human Rights, Democracy and the Rule of Law).

These diverse governance efforts offer scholars of norm research in International Relations (IR) manifold opportunities to study forums and processes where actors *publicly* express what kind of (governance) challenges they associate with AI technologies and begin to frame new norms, for example the EU's vision of human-centric AI that is central to the *AI Act* (e.g. Amariles & Baquero, 2023; Carmel & Paul, 2022; Lukowicz, 2019). Norm research in IR has traditionally conceptualised norms as closely linked to international law. Early constructivist scholars, who introduced the empirical study of norms to the discipline, saw the institutionalisation of norms into soft or hard law as simultaneously the goal and the endpoint of norm emergence. Building on that, such scholarship focused on the specifics of state compliance with legally institutionalized norms via processes of diffusion, internalisation, and socialisation (Finnemore & Sikkink, 1998; Keck & Sikkink, 1998; Risse *et al.*, 1999).

Later work in critical norm research has accommodated the inherent flexibility of norms by emphasising enacted 'meaning-in-use' (Wiener, 2009; Wiener, 2017) across regions and contexts, focusing therefore on the much more dynamic and interactive processes of localisation, implementation, and contestation (Acharya, 2009; Acharya, 2004; Betts & Orchard, 2014; Deitelhoff & Zimmermann, 2018; Drubel & Mende, 2023; Wiener, 2014; Zimmermann, 2017). Importantly, this line of research challenges a supposed linearity of norm development, as well as the questions of who has agency therein and whose agency counts (Bucher, 2014; True & Wiener, 2019; Wiener, 2018). This scholarship also explores the very production of normativity in contestation processes (Hofferberth & Weber, 2015; True & Wiener, 2019; Wiener, 2018), understanding contestation as a pivotal, norm-generative societal practice (Orchard & Wiener, 2022, 54). These studies often remain connected/attached to soft or hard legally institutionalised norms or the public-deliberative process of drafting such norms and have produced many valuable insights.

But what such scholarship arguably does not account for is how the slow-moving nature of public-deliberative governance processes interacts with the, at times, fast pace of technological development, especially in (but not limited to) the field of 'AI'. Even though several AI governance initiatives have yielded preliminary results in the shape of soft international law (e.g. sets of principles and non-binding resolutions), many of them are only moving forward slowly and often need re-adjusting to keep in step with technological change. While the EU's AI Act has benefited from regulatory

momentum, it still moves slowly through the EU's institutions with more than two years passing between the presentation of its initial draft by the European Commission in April 2021 to its adoption by the European Parliament in June 2023. Further deliberations in EU member states create additional friction. The GGE on LAWS remains at a discussion rather than at a negotiation stage and insiders have invariably described its pace as glacial or even tectonic (Régimbal, 2023). At the same time, applications utilising AI technologies are spreading across various aspects social life and public policy. Such technologies, for example, play a role in the provision of medical care such as radiography (Obermeyer *et al.*, 2019) hiring decisions (Ajunwa *et al.*, 2016) to predictive policing (Lum & Isaac, 2016) and the approval of welfare beneficiaries (Eubanks, 2019). This development raises crucial regulatory questions because of the societal harms that such automated and algorithmic 'solutions' can produce, often reinforcing existing power asymmetries and hierarchies (Birhane, 2021b). Partly due to a perceived lack of expertise, partly due to concerns about stifling innovation (see UK Government, 2023), the governance of AI is cumbersome and dominated by the input provided by a narrow set of dominant stakeholders, including major tech corporations (Holland Michel, 2023, 4; Veale *et al.,* 2023).

This essay intervenes in the debate about governing AI by reflecting on governance through international norms, which we define as understandings of appropriateness (Bode & Huelss, 2022, 12).[1] We put forward two analytical additions to norm research in the discipline of IR, and, in particular, to how norms emerge in IR, with the ambition of shaping a research programme on *practical normativity/normality*. (1) We focus on how, often operational, practices of designing and using AI technologies that are performed by actors *at sites outside of the public eye* become sources of normative substance on 'AI'. This departs from a present focus in IR norm research on practices that are *observably* performed in public spaces, either in the form of "deliberation",[2] i.e. how actors publicly discuss, consider, and verbalise norms, or via behavioural contestation (Stimmer & Wisken, 2019; True & Wiener, 2019). But norms are not only constituted via practices associated with the public and/or deliberative space where they may transition into law, but also unfold in operational practices of design and use that are not subject to public debate, discursive codification, or public performance. This argument builds on scholarship at the theoretical interface of IR norm research and practice theories (Bode & Karlsrud, 2019; Bode, 2023b; Gadinger, 2022; Laurence, 2018; Pratt, 2022; Ralph & Gifkins, 2017; Wiener, 2022), but also draws from critical security studies, as well as constructivist strands of science and technology studies (STS). Approaches that merge practice theories and norm research typically emphasize the relational, unstable dimension of norms, evidenced by their preference for concepts such as "normative configuration" or "normative substance" as alternatives to the seemingly well-bounded 'norm' (Bode, 2024; Pratt, 2020; Pratt, 2022; Qiao-Franco & Bode, 2023).

(2) We consider how practices of designing and using AI technologies shape normative substance in an interplay between normativity, i.e. notions of moral duty and justice, and normality, i.e. notions of the typical and average (Huelss, 2017). IR norm research that focuses on publicly performed practices often, implicitly rather than explicitly, emphasises the dimension of normativity. By contrast, the dimension of normality and the interactive relationship between the two concepts have only received little dedicated attention (Gholiagha *et al.*, 2021; Huelss, 2020). However, both aspects are important to perceive and understand "previously obscured patterns and processes" (Bernstein & Laurence, 2022, 77) in the relationship between norms and practices, especially when it comes to AI technologies. Our distinction between normativity and normality is ideal-typical/analytical in nature – separating between these two dimensions in empirical practice is typically infeasible because they can be likened to two sides of the same coin. At the same time, this also implies that it is crucial to conceptualise the relationship not least for analytical reasons. We will use the descriptor normativity/normality and interchange which of two notions we start off with. Notwithstanding this methodological challenge, we consider paying attention to both normality/normativity as vital and missing component of understanding the significance and reach of norms.

Our arguments are abductively informed by our empirical study and analysis of military applications of AI, in particular, of how AI and autonomous technologies have become incorporated into weapon systems. This development is typically connected to the catch-all term autonomous weapon systems (AWS) as systems that can 'make' targeting 'decisions' without immediate human intervention.[3] Targeting algorithms, for example, introduce forms of non-human centred 'decision-making' into the conduct of warfare.

Before we proceed, let us define what we mean by 'AI' and associated terminology. We use the term 'AI' to refer to a broad area of scientific study that encompasses techniques like machine learning as well as research areas like computer vision or natural language processing. AI is the effort "to

[1] Our definition builds on the widely reproduced definition by Katzenstein: "The authors use the concept of norm to describe collective expectations for the proper behavior of actors with a given identity (…) Norms thus either define (or constitute) identities or prescribe (or regulate) behavior, or they do both" (1996, 5). However, we emphasise that norms are not necessarily given or ultimately stable factors shaping actions. They can emerge in practices in unstructured but sedimented or increasingly standardised procedures.

[2] When speaking of "deliberation", we do not refer to how the term has been shaped by the influential theoretical agenda of Jürgen Habermas and John Rawls. Rather, we use the term in its dictionary meaning as the "process of carefully considering or discussing something" (Oxford English Dictionary, 2020) in a public forum. Deliberation as public-discursive discussion, negotiation, criticism, or contestation therefore leaves "trails of communication that can be studied" (Björkdahl, 2002, 13).

[3] AWS may but do not necessarily integrate AI technologies (Garcia, 2023). Throughout the essay, we use inverted commas for action verbs that refer to AITs as agents as a form of distancing AIT processing from human decision-making. This is part of an effort to counter-balance the human tendency towards transference, i.e. "attributing understanding, empathy and other human characteristics to software" (Tarnoff, 2023).

create machines or things that can do more than what is programmed into them" (Gebru, 2023). We supplement the term 'AI' regularly with that of technologies, thus leading to the acronym AITs. In this, we are inspired by the critical public policy scholarship of Emma Carmel and Regine Paul who use AITs to highlight the complex, contingent, and variable insertion of 'AI' into the societal domain (Carmel & Paul, 2022; Paul, 2022). As a result of its frequent, ambiguous usage, 'AI' is a highly divisive term that is prone to numerous uncertainties (Afina, 2022; Holland Michel, 2023; Tucker, 2022). Problematically, the term 'AI' may, for example, lead us to consider machines with intelligence as equal to humans, thereby potentially triggering irrelevant sets of questions (Gunkel, 2023). As Whittaker critically contends, the publicly reported 'advances' in AI research are not so much the result of technological breakthroughs but "the product of significantly concentrated data and compute resources that reside in the hands of a few large tech corporations" (2021, 51). Further, we speak of both AI and autonomous technologies as these are closely related and arguably function along a trajectory of increasing technological complexity. Autonomy can be defined as the "programming machines to perform some tasks or functions that would ordinarily be performed by humans" (Scharre & Horowitz, 2015, 5). As such, a weapon system may, for example, integrate autonomous or AI technologies to perform different kinds of functions ranging from intelligence analysis to mobility and targeting, with targeting functions having, by far, garnered most attention (Boulanin & Verbruggen, 2017, 19–35). Finally, we follow Virginia Dignum's definition of an algorithm as "a set of instructions, such as computer code, that carries out some commands" (2019, 3). Algorithms are used to 'find' patterns in datasets and "specify how to transform a given set of inputs into a desired output" (Mackenzie, 2015, 43).

The remainder of this essay reviews our two analytical arguments outlined above and reflects on how they can advance a new research agenda of *practical normativity/normality* for the specific purpose of studying the impact of AI technologies on normative substance in international relations. We structure the two subsequent sections to include both a review of how far literature on practical normativity/normality has come, followed by an expanded discussion of our arguments on how operational design and use at sites beyond the public eye shape normative substance, as well as normativity/normality dynamics. In a third section, we briefly reflect on what AI technologies along with forms of 'machine agency' signify for established distinctions between agency and structure in norm research. In a fourth and final substantive part, we sketch three future research streams that draw on the insights presented but also take these a step further towards more experimental arguments that could allow us to critically study the emerging normativity/normality of AI technologies.

Practical normativity/normality I: practices of developing and using AI technologies as sources of norms

Sociology and political theory have long captured practices, broadly understood as patterned interactions in social context (Leander, 2008, 18), as the micro-level building blocks of the social world. As "nexuses of doings and sayings" (Schatzki, 2012, 14), practices make norms. On the ontological level, practices and norms are therefore inseparable, since action-relevant normative substance only manifests in practices (Bode & Huelss, 2018; Gadinger, 2022; Neumann, 2002, 200; Wiener, 2009; Wiener, 2018). This insight also applies to cases where a legal-institutionalized norm already exists that supposedly guides practices through its constitutive quality. The practical implementation of a norm always contains an element of improvisation and with it the potential to change its meaning (Bode & Karlsrud, 2019; Hofferberth & Weber, 2015; Huelss, 2017; Pratt, 2022; Wiener, 2018). In principle, dynamic practices can thus create new norms, maintain, and stabilize existing norms, and change the substance of existing norms. Indeed, this process may increase a norm's legitimacy (Wiener, 2018).

How practices maintain, stabilise, or change the substance of existing norms has already been well-covered in IR norm research as part of different sub-agendas on the resilience, robustness, and erosion of norms, focusing on deliberation, communication, and public-discursive exchange (Barbé & Badell, 2023; Deitelhoff & Zimmermann, 2020; Lantis & Wunderlich, 2018). From the start, IR norm research has studied how norms emerge in deliberative, international forums and therefore, conceptualized normative content as produced in public debate and international negotiation. What matters in this process are language-based strategies of presenting an argument such as framing, salience, and grafting (Acharya, 2004; Keck & Sikkink, 1998; Rosert, 2019). Indeed, "states are very verbal entities" (Hansen, 2006, 21). As Bower argues, "multilateral negotiations serve as focal points in the crystallization of emergent norms" (2015, 355).

Further, both localisation and contestation scholarship has invested much attention into offering deeply theorised thinking ways of thinking about how practices generate and shape normative substance. Contestation research, in particular, departs significantly from models considering norm emergence as sequentially distinct from the diffusion of norms, acknowledging instead the "norm-generative power that materialises through contestation" (Wiener, 2018, 11). That power is vested in the public space and is often, although not necessarily, connected to public-deliberative processes. It is thus crucial whether and to what extent actors have *access* to the forums where the constitutive debate about norms happens, thereby being able to ensure that their voices are *heard* (see Wiener, 2018). In broader IR norm research, this public-communicative access has also manifested in agency-oriented concepts such as norm entrepreneurs who "attempt to convince a critical mass of states […] to embrace a new norm" (Finnemore & Sikkink, 1998, 895). The reverse concept of "norm antipreneurs" (Bloomfield, 2016) likewise describes agents working through public-deliberative means.

However, Wiener's theoretical understanding of contestation explicitly goes beyond deliberation: "while mostly expressed through language not all modes of contestation involve

discourse *expressis verbis*" (2014, 1). She includes explicit forms of contestation, such as contention, but also implicit forms, such as the absence of practices (neglect) and practices that contradict a certain reading of a norm (negation, disregard) (Wiener, 2014, 7). In later conceptual and empirical work, contestation scholarship has begun to explore this behavioural dimension of contestation as "contestation by means of action that affects implementation" (Stimmer & Wisken, 2019). Yet, interestingly, even the types of practices that animate behavioural contestation are performed in and attached to some form of public space (e.g. True & Wiener, 2019, 560, 563, 570). Therefore, they become directly observable.

Despite recognising the norm-generative power of practices, norm contestation scholarship arguably does not yet engage comprehensively with the theorisations of practice offered by practice theories. From a norm contestation point of view, there might be pragmatic methodological reasons for this. Studying how norms emerge in operational practices means attempting to access processes that are not directly observable nor immediately accessible to researchers because their performance does not leave discursive trails. At the same time, Wiener's methodological toolkit has also included exercises such as the multilogue that work via reconstructing dialogues between different norm actors, thereby creatively handling an absence of readily available material (2018, 63–64). However, methodologies long used by practice scholars, such as praxiographies, feature many useful insights for rendering practices that are performed outside of the public space visible via for example, participant observation or interviews, that can be usefully applied to the study of emerging normative substance in operational practices interviews (Adler-Nissen, 2016; Bueger, 2014; Nicolini, 2013; Pouliot, 2016). Still, in the context of studying technological practices of design and development, such methodological problems are likely amplified by norm and governance scholars limited technical familiarity with AI. Yet, over the past decade, we have seen an increasing number of publications and courses that are dedicated to bridging such knowledge gaps by providing more common-sensical understandings of how AI technologies work (e.g. Crawford, 2021; Goldstaub, 2021; Russell, 2019; Walsh, 2018; Walsh, 2022).

Our arguments add to this existing body of knowledge in two ways: we broaden analytical attention (1) by *expanding the types of practices* whose performance generates/shapes/sustains normative substance beyond practices that are, in some shape or form, publicly performed, voiced, framed, or deliberated; and (2) by *expanding the social sites* where such practices are performed to sites outside of or hidden from the public eye. This *practical normativity/normality* is particularly well-suited to understanding the impact on normative substance that AI technologies in and beyond the military domain can unfold (Bode, 2023b).

A closer look at technologies in the military domain offers useful analytical inspiration for this line of thinking. Military technologies and weapon systems are often developed and deployed over many years before they even become the subject of public debate, leading to the public-deliberative emergence of new norms that can/should regulate existing practices (Bode, 2023b). Anti-personnel mines, for example, have been used since World War II on a wide scale, but public discussion of the practices associated with their use only began in the 1970s under the CCW, culminating in the 1997 Ottawa Treaty that banned anti-personnel landmines. A time lag between the development and deployment of military technologies and a public debate about their use is typical, as the further examples of submarines, chemical weapons, nuclear weapons, or cluster munitions show.

But even without practices performed in a public or deliberative space, other types of practices create, shape, sustain, and change normative substance. In the cases described above, these are practices associated with designing, training personnel for, and using military technologies. Practices of design, for example, anchor normative substance in AI technologies. Here, we follow constructivist STS insights, in understanding technology as fundamentally social and thus as a reservoir of social practices (Bijker & Law, 2010). These assumptions contrast with a positivist conceptualization of technology development as part of objective-neutral processes of experimentation and problem-solving. By contrast, STS thinking allows us to conceptualise the insertion of the material in the social. This perspective re-thinks existing dualisms, such as discourse vs. materiality, in favour of a "process ontology" that understands practices as relationally co-constituted by the social and the material (Schouten & Mayer, 2016, 311). It also reverses a conventional separation between social actors, on the one hand, and technologies as materiality, on the other hand, in favour of an interactive perspective on how these interact in performing technological practices (see also section 2). STS thinking therefore allow us to question the accuracy of the distinction between technological factuality and value-based, normative decisions.

In the design, development, and use processes of AI technologies and algorithmic systems, practices can constitute/shape normative substance in at least three ways: (1) All AI technologies are based on the generation of computer models that 'translate' complex social activities into a problem to be solved and set a planned goal for this purpose (Gillespie, 2014). This process is vaguely understood to be objectivist, but instead actively encodes values and therefore types of normative substance into AI technologies. Louise Amoore has therefore critically theorised algorithms as "ethicopolitical arrangement[s] of values, assumptions, and propositions about the world" (2020, 6). The ways in which algorithms find patterns in the data processed is inextricably linked to affording certain features and assumptions greater weight and therefore recognition than others (Amoore, 2020, 8). As Abeba Birhane and colleagues have found in their survey of 100 highly cited machine learning papers, values such as 'performance' and 'efficiency' significantly outrank considerations of societal need or negative potential (2021). In AWS, civil society organisations have investigated how the generation of computer models entails the construction of target profiles, "the set of conditions under

which such a system will apply force" (Moyes, 2019, 1). Reviewing such target profiles demonstrates how their design encodes normative choices in the AI or autonomous technologies that animate them. Current target profiles are basically particular patterns of sensor data – in the form of acoustic signatures, radar signatures, heat-shapes, or visual information – that are interpreted as a target. As such, target profiles will contain both intended and unintended targets, because a broad target profile reduces the number of false negative results that the military actors using the AWS want to avoid (Moyes, 2019, 4–5). In this way, target profiles shape normative substance at the stage of design practices in choosing instrumental parameters such as acceptable error rates for determining targets (Crawford, 2021, 96–97).

As the technology used in AWS grows in complexity, targeting profiles could become adaptive rather than based on fixed, encoded goals set by human designers (Moyes, 2019, 4). Using AI technologies, the system may suggest novel targets gleaned, for example, based on correlating multiple current forms of sensor data it has access to as part of a networked system with previous sensor data. This process describes a transition between different algorithms that sustain target profiles from rules-based types to generative learning types: the former are based on a designer-inserted hypothesis that the algorithm statistically tests, while the latter type experimentally constructs rules and concludes which of them were noteworthy (Amoore, 2020, 12). When it comes to military decision support, systems that enable precisely these practices have already been used on the battlefield. As public sources reported, Ukrainian military commanders use the "AI Platform for Defense" (AIP), an AI-assisted software produced by Palantir, to provide them with "likely scenarios for the course of the war and […] spit out automated field moves" (Fischer, 2023; see also Hudson & Khudov, 2023; Ignatius, 2022). Thus, such design practices encode value-based decisions that are not disclosed into the targeting process and its programming.

(2) The composition of the data sets used to program the algorithms that sustain AI technologies may contain data that does not cover the problem to be solved in a representative way, i.e. it may contain bias (Sharkey, 2018). It is a well-known adage that an AI system can only be as good as the data it is trained on. Critical studies or audits of prominent image datasets, such as Image Net, have therefore demonstrated that AI technologies replicate and augment racial and gender stereotypes and prejudices (Bolukbasi *et al.*, 2016; Chandler, 2021; Crawford, 2021, 96–97; Hooker, 2021). Well-known examples include systematically excluding women in results for occupations or associating women with particular types of occupations only, and systems performing best in classifying light-skinned males as human (Benjamin, 2019; Buolamwini & Gebru, 2018; Birhane, 2021a). Such datasets therefore include assumed patterns of normality that become the basis for and amplified by the AI technologies that use them as training material. Many AI researchers argue that it is possible to circumvent such bias in the training data, inspiring an increasing research stream on bias mitigation strategies and their potential trade-offs (Ferrara, 2023). To this, Amoore posits an interesting, more fundamental, two-fold critique. She holds that first, bias is an extricable part of algorithms as such systems need to contain forms of weightings to work and second, that learning algorithms are characterised by recursive, nested relations to input data rather than going through a linear sequence of steps thereby making it near impossible for humans to intervene in this non-linear process to remove bias (Amoore, 2019, 11).

(3) Practices of using systems integrating AI technologies, including weapon systems, are typically based on certain, not openly stated assumptions, including about human-machine interactions. The literature on autonomous systems refers to such assumptions as myths (Bradshaw *et al.*, 2013; Johnson *et al.*, 2014), while policy literature on AI talks about "technological myths" (Natale & Ballatore, 2020) or simply common policy assumptions that need recalibrating (Holland Michel, 2023). A persistent myth of autonomous systems is the "erroneous idea that once achieved, full autonomy obviates the need for human-machine collaboration" (Johnson *et al.*, 2014, 58). In the context of warfare, we find this thinking present in how autonomous and AI technologies may replace human soldiers. However, the use of weapon systems integrating autonomous technologies, such as the Patriot missile defence systems, shows the opposite: introducing autonomous and AI technologies increases the complexity of tasks to be performed by human operators and demands more, not less, "human expertise and adaptative capacity" (Johnson *et al.*, 2014, 84). Yet, such assumptions continue to shape use practices of AWS, such as the level of trust placed in system outputs or the involvement of the human operator in the decision-making loop of weapon systems (Bode & Watts, 2021; Bode & Watts, 2023), with clear effects on normative substance.

In summary, a focus on the operational practices of designing, training for, and using AI technologies enables an analytical view on modalities and sites where norms can emerge outside of the public space and a form of public debate. This perspective can diversify how and where we can localize norms on AI. Conceiving of practical normativity/normality beyond the public space thus offers scholars of norms and governance a potentially far-reaching review of how normative space is constituted. Although we draw these insights inductively from analysing autonomous and AI technologies, they point to a broader empirical phenomenon: the initial emergence of norms in operational practices, which become objects of public deliberation or observable in the public space only at a later point in time or do not become a part of public debate at all (Bode, 2023b).

Of course, we should not fail to recognise that such practices are still performed within a potentially dense normative structure. State actors who develop and use weapon system integrating AI and autonomous technologies, for example, do not operate in a normative vacuum but within the principles and prohibitions of international law, for example international humanitarian law (IHL) and international human rights law (IHRL). But international law provides practices with a

thoroughly ambiguous, sometimes contradictory, and potentially enabling normative basis (Hurd, 2017; Koskenniemi, 2011; Wiener, 2009). Based on existing international law, there can be very different judgments as to which practices are 'appropriate', and states therefore can have considerable practical and interpretative leeway (Bode, 2023a).

Practical normativity/normality II: the dynamic interplay between what is normative and what is normal

Through an emphasis on how practices of design, of training personnel for, and of using AI technologies performed at sites out of the public eye can become sources of normative substance, the interplay of normativity *and* normality also comes into focus. Practices can constitute ideas about which ideas of oughtness, justice, and associated moral duties apply (normativity) - but they can also make something appear normal or average behaviour (normality). So far, norm research in IR has only theorised the normality dimension of norms in a limited way and has, as a result, not examined how this dimension interacts with normativity. To do this, we can draw inspiration from IR scholarship beyond norm research. Critical security studies, such as scholarship associated with the so-called Paris School, for example, consider the emergence of *normality* within practices of exclusion and regulation (C.A.S.E Collective, 2006). Here, understandings of the 'normal' arise in a process that compares, weighs, calculates and executes resulting in and based on identifying the 'average'.

This work builds on Michel Foucault who offered an insightful theorisation of how social norms emerge beyond the law in his lectures on the *History of Governmentality* between 1977 and 1978. Foucault draws attention to the fact that the average is often considered the norm, but that this 'norm' is not identical with legal normativity (Huelss, 2017). To distinguish between the two processes and deviating from everyday language terms, he referred to the effect of legal norms as a process of *normation*, whereas a process of *normalization* refers to how norms emerge in the derivation from the average such as the statical normal distribution, i.e. the Bell curve.

Foucault's thinking offers interesting links to practice theories as the process of entrenching and spreading certain practices over time can lead to them becoming *routine* and therefore normalise practices outside of formal codification or acts of norm creation. Making something normal or normalising practices, is related to notions of "the common, the ordinary, the standard, the conventional, the regular" (Cryle & Stephens, 2017, 1).

Normativity describes understandings of oughtness and justice, of patterns of behaviour that are morally right and just. Often, scholars argue that this normativity must be discursively expressed and established – and even publicly voiced. But this public understanding of normativity is too limiting. If we take the notion of 'appropriateness' that is a common feature of how norms are defined and take this to represent normativity, then what counts as "appropriate" behaviour is therefore not only defined in public deliberation, but also generated by practices in various forms. This is particularly relevant when understandings of the substance of deliberative norms collide with the normativity of practices. The normativity constructed in practices can be more impactful than what is discursively established.

The main point we want to make here is that normativity is typically conceptualised as an outcome of deliberative norm-setting. The resultant norms are thought to shaping the normality of practices. But this sequence could also be conceptualised in the reverse order where the normality of practices creates aspects of normativity. There are conceptual challenges to be tackled here: first, to develop an understanding of normativity in practices that might be different from the deliberative normativity that is central for conventional norm studies. Second, it should be further considered whether normativity is rather broadly or narrowly understood. What are the limits of ethical standards of 'right and wrong'? Is normativity only relevant for certain types of practices? This is related to how broad or narrow we should understand norms to be and whether all norms also entail normativity. This question can be addressed in the context of a theoretical-philosophical discussion that, however, goes beyond the scope of this article. At this point, we just want to note that we find analytically beneficial to use both a broad understandings of norms (understandings of appropriateness) and a broad and relational concept of normativity. In other words, our position is that it is analytically too narrow to only focus on the most "fundamental norms" (Wiener, 2009) such as human rights' right to life, while leaving so-called procedural standards or similar aside and associating them with 'lower' overall rates of normativity. At the same time, broad norms should still be understood as entailing and informing generally elements of normativity.

As this conceptual discussion shows, understandings of normal behaviour can overlap with understandings of normative behaviour that is considered morally just or 'right'. In fact, we consider normativity and normality to be deeply connected and inextricably linked. What is normative and what is normal is often not easily separable, but we should account for both of these dimensions in theorising how normative substance emerges – in relation to AI technologies and beyond. It then becomes an empirical question whether norms normalize practices by shaping what is appropriate in a given situation or whether practices rather normalize norms by shaping their substance based on what is established as normal. In what follows below, we concentrate on highlighting the normality dimension as this has received comparatively less analytical attention in norm research and we want to draw attention to some interesting analytical tendencies that this focus reveals. But we are conscious of the fact that this is an ideal-typical conceptualisation rather than a clear-cut empirical one.

In the context of AWS, practices associated with the design and use of AI technologies, can work to establish patterns of normality in at least two ways: (1) the algorithmic definition of parameters and target profiles at the point of design determines what counts as 'normal' and what counts as 'abnormal' or anomaly. As Amoore argues, "contemporary algorithms

are not so much transgressing settled societal norms as establishing […] new thresholds of normality and abnormality" (2020a, 6). This encoded definition represents a clear limitation of what is conceivable as normal and abnormal in the future. AI technologies based on the technique of machine learning are particularly important to consider here, for example object recognition as part of the image-processing of computer vision. In recent years, the US military launched projects with significant financial budgets to conduct relevant research in this direction through collaborating with tech companies. Project "Maven", initially carried out in cooperation with Google, should be mentioned here as a prime example (see Suchman, 2020a).

When examining patterns of normalisation, the use of 'unsupervised learning' algorithms in the context of machine learning is noteworthy. Here, algorithms aim to identify patterns of interest in the data that represent an anomaly rather than searching for patterns that it has been trained on (Aradau & Blanke, 2018). As outlined above, "the algorithms […] generate all rules and the debate which of them was interesting" (Amoore, 2020, 12). A widely discussed empirical example of unsupervised learning in the military context was the US National Security Agency's (NSA) Skynet programme, which endeavoured to associate movement patterns of mobile phone users based on metadata with the users' potential terrorist activity. Controversially, this data could also be used in the wider targeting process to support target identification for drone strikes. However, public reporting showed that the system mis-identified prominent Al-Jazeera journalist Ahmad Muaffaq Zaidan in Pakistan as a member of terrorist organizations such as Al Qaeda and therefore as a potential target (Currier *et al.*, 2015; Robbins, 2016). The travel and movement patterns associated with his investigative work apparently generated an *anomalous* or suspicious movement profile (Aradau & Blanke, 2018; Bode & Huelss, 2021; Huelss, 2020). The processing and use of various form of digital data is part of common practices performed by the US military, as confirmed by the National Security Agency (NSA). The former director of the NSA and CIA, retired US Air Force General Michael Hayden, explicitly underlined: "We kill people based on metadata" (cited in Scahill & Greenwald, 2014). This is an example where digital normality detection may prevail over deliberative normativity considerations. In other words, what machine-learning algorithms identify as an anomalous and therefore interesting data pattern can have very immediate consequences for life and death decisions with high normative salience. Identifying a data anomaly as something abnormal that stands outside of normativity and needs to be normalized, brought in line with the normal, or rather eliminated, is central to this process that animates how technological, operational practices may shape norms. While the kill decision is not automated in scenarios such as the Skynet programme tests, the insertion of algorithms into targeting decision-making changes the quality of control that humans exercise therein. This development is therefore a decisive part of potentially normalising machine induced norm emergence that can limit human control over the use of force (Bode & Watts, 2021; Bode & Watts, 2023; Bode, 2023b).

(2) Practices associated with the design and use of AI technologies can also 'normalize' by informing and influencing human actors as part of human-machine interaction. Regardless of the extent to which the "machine" makes autonomous decisions, the repeated, collective execution of a similar range of practices based on human-machine interaction over a period of time can come to be considered "normal", increasingly "acceptable", perhaps even desirable. In these situations, we may find an interactive exchange between normality and normativity, in which incrementally changing, collectively held understandings of appropriateness develop.

As these illustrations show, the relationship between normality and normativity remains a research-analytical challenge and requires further, detailed studies to show the possible trajectories of transition from the normal to the normative or the interaction dynamics between these two dimensions. These trajectories do not only concern how the normality of practices can influence legally codified norms as a postulated expression of normativity and make inroads into deliberation. They also raise more fundamental questions about the extent to which long-established, legally codified norms are informed and shaped by normality. "Fundamental norms" (Wiener, 2007), such as human rights, for example, often align with dominant practices in certain cultures. Such practices typically precede a formal, legal codification. Given such possible dynamics, we could reconsider whether the effect of codified norms and the importance of deliberation are overemphasized in conventional, constructivist norm research – if such norms may correspond to normal, average behaviour anyway.

AI Technologies, Normality/Normativity, and a Post-Human Agency

Finally, returning to the empirical field of AI technologies, we also hold that the emergence of a new, post-human actor quality in practices is one of the great challenges that contemporary norm research encounters. Distributed forms of agency between humans and 'decision-making' technology that arise here contrast with the established analytical categories of human actor and normative structure. Studies at the intersection of STS studies and the "new materialism" have begun to take up this question of agency (Coole & Frost, 2010; Coole, 2013). A detailed discussion of this themes is beyond the scope of this essay, but we want to note how 'distributed agency' emerges in the interaction of people and the material. This has conceptual similarities to Latour's "actant" (2005), but differs in defining agency capacities more narrowly (Coole, 2013, 456–61). This form of agency does not replace human actors but is vested in the emergence of new systems where people and technologies have agency capacities in a material and a relational sense.

We should note that any conceptual notion of 'machine agency' cannot be clear-cut/discrete since much of it is the result of human action that has become invisible in the design process. This includes programmers who create the basic computational models and decide upon the algorithm's parameters, 'ghost workers' who prepare the data required for

training machine learning algorithms through a series of micro-tasks often referred to as "labelling data", and people whose data is used to train machine learning algorithms (Crawford, 2021; Jaton & Bowker, 2020; Penn, 2021). A distributed concept of agency should consequently affect our understanding about what AI technologies are in the first place, as well as the type of data work they require. For prominent STS scholar Lucy Suchman, AI is therefore "a cover term for a range of technologies of data processing and techniques of data analysis based on the iterative adjustment of relevant parameters, according to some combination of internally and externally generated feedback" (2020b).

Moreover, the idea of distributed agency can also lead us to consider a concept of distributed normativity/normality in the context of AI and human interaction. Given that the productive substance of AI, whether it is about generative task in terms of text production or similar applications, is essentially based on a recompositing or parroting of human creative output (fed-in as learning data), it would be misleading to speak of post-human normality/normativity in the strict sense.[9] Nevertheless, the normative outcome of AI decision-making would still be outside of strictly human controlled normativity/normality.

Notwithstanding these qualifications, the conventional view of agency and structure in norm research might no longer do justice to such technological developments. The influential premises of structuration theory, for example, assume a logical and temporal separation of both dimensions. But what if practices performed by using AI technologies or as part of human-machine interaction override public, human agency-based deliberations of normative substance? As discussed above, machine learning may lead to the production of normative substance based on data patterns, even if such algorithmic 'decision-making' is no longer comprehensible and verifiable by human actors. The introduction of algorithmic 'decision-making' is also an onslaught on the human ability to doubt (Amoore, 2019). The neat, seemingly objective outputs of AI that are based on a vast number of complex calculations are ultimately impossible to question in the logic of technology. Especially in the military context, but also beyond it, factors such as speed, lack of time, and high data load mean that it may no longer be possible for human operators who are in-the-loop of weapon systems to meaningfully control/understand the prompts offered by a system integrating autonomous or AI technologies (Bode & Watts, 2021; Bode & Watts, 2023).

What is more, how machines process information, identity anomalies, produce assessment about what constitutes an 'appropriate' action, and execute these actions collapses an ideal-type image of control and deliberation. In other words, systems integrating AI or autonomous technologies – regardless of whether a human is monitoring the system (in-the-loop or on-the-loop) – can create normative substance through analysing sensor input at negligible time intervals, detecting deviations from "the normal", and reacting to this deviation within the given parameters, for example by destroying the target. Therefore, machines are becoming *de facto* agents who create, maintain, or change forms of normative substance through performing practices. The resulting blend between practical normality/ normativity is a new dimension that should receive more attention in IR norm research. It may also result in fundamentally reconsidering how stable norms are over time: practical normality/normativity emerges in highly dynamic processes that no longer correspond to the idea of a changeable but largely stable normative structure.

Future theorising on AI technologies and norms

The conceptual concerns raised, and arguments presented above lead us towards sketching several additional research streams for theorising how AI technologies may shape norms. These avenues are preliminary but are intended to give a sense of further/future direction for issues that need to be further addressed if we are grasp the potential ramifications of practical normality/normativity related to AI technologies in an IR context.

Relationship between operational and deliberative processes of norm emergence

One of the fundamental tenets of our argument is to point IR norm research in the direction of more closely considering, conceptualising, and analysing operational practices performed outside of the public space as sources of normativity/ normality. With this theoretical move, however, we do not want to dismiss or side-line public-deliberative processes of norm emergence that happen in manifold forms in the context of AI technologies. In fact, such public-deliberative processes continue to diversify across state, regional, and global levels and may thereby result in an increasingly "thick" network of written principles, guidelines, standards, or regulations interacting with each other (Maas & Villalobos, 2023; Tallberg *et al.*, 2023). As Table 1 illustrates, in 2023 alone, stakeholders started at three new, longer-term cross-regional and international

Table 1. ***Governance initiatives on AI technologies in the military domain starting in 2023.***

Responsible AI in the Military Domain (REAIM) Summits, co-led by the Netherlands and the Republic of Korea, first multi-stakeholder summit held in February 2023
Political Declaration on the Responsible Use of AI in the Military Domain, led by the United States with currently 51 states participating
General Assembly First Committee (Disarmament and International Security) Resolution on Lethal Autonomous Weapons Systems (LAWS) adopted 164/8/5 in November 2023 (UN Document No. A/C.1/78/L.56), starts a process to report challenges and concerns raised by LAWS from humanitarian, legal security, technological, and ethical perspectives

public-deliberative processes and associated governance initiatives on AI technologies in the military domain. While such public-deliberative processes are moving slower in comparison to the pace and track record associated with actors designing, developing, and employing AI technologies, we can therefore expect them to produce a larger number of and more substantive outputs in the coming years.

It is important, then, to think about the public-deliberative and the operational processes not in isolation of each other but consider how the precise dynamics of their interaction can shape that emergent norms therein take via their respective constitutive practices. We started to sketch the interaction of these processes in our previous research on emergent normativity in the context of weapon systems integrating autonomous and AI technologies. Here, we distinguish between four potential interaction dynamics that are drawn inductively from arms control cases: negative acknowledgement, positive acknowledgement, collision, and (wilful) ignorance (Bode, 2023b, 995–96).

(1) Negative acknowledgement describes a process where actors involved in a public-deliberative process draw attention to the adverse impacts of normativity/normality originating in operational practices. Here, we would expect potential change in the normality/normativity that is emerging from practices as practices become the subject of public scrutiny, practitioners may rethink their performance of practices either as a result of awareness-raising or of new regulation (in whatever form) that specifically addresses the adverse impact of current practices. The scenario of negative acknowledgement therefore describes an interaction pattern where what is produced in the public-deliberative process has the highest potential shaping effect on the kinds of norms that emerge in the context of AI technologies. At the same time, the extent to which even specific public-deliberative outcomes in the form of guidelines etc change practices performed in operational context depends on a number of key factors, chief among them whether such agreements are legally binding nature, as well as their verification and implementation strategies.

(2) As part of positive acknowledgement, by contrast, actors involved in a public-deliberative process, consider and confirm existing practices and the normality/normativity that they constitute as "appropriate" and in line with (current) demands of the international normative framework. In this case, we would expect their public discussion not to change what kind of operational practices are performed and the associated normativity/normality. By contrast, their positive acknowledgement could instead be expected to lead to practitioners doubling down on such practices as they have now been publicly sanctioned as normal/normative.

(3) Collision describes a complex dynamic where contradictory practices constituting diverging normalities/normativities co-exist. Public deliberations point to the adverse consequences of (currently performed) practices in terms of constituting normativity/normality, attempt to address this situation through setting up a new normative framework and thereby to reconstitute normality/normativity. But this novel framework does not fully re-constitute normality/normativity as states continue to perform divergent practices that can be associated with the previously established normativity/normality. As such, a scenario of collision is a likely result when you have different international normative frameworks, including in the form of legal regulations, that different, potentially intersecting but also exclusive groups of states have signed and ratified. An example is the international normative framework on nuclear weapon systems that features two main sets of legal documents constituted by and producing colliding sets of practices: the 1970 Treaty on the Non-Proliferation of Nuclear Weapons and the 2021 Treaty on the Prohibition of Nuclear Weapons.

(4) A final interaction dynamic consists of actors in the public-deliberative process simply not speaking to or raising the emergent normativity/normality that is constituted by operational practices. Here, while the public-deliberative and the operational processes of norm emergence continue to run in parallel to each other, but do not appear to influence one another, at least not in direct ways. In such a set-up, the kind of operational practices that practitioners perform are not likely to change – at least not in the dramatic way that may result from negative acknowledgement.

These four interaction dynamics are not mutually exclusive. This means that we could see several of these dynamics at play in relation to operational and public-deliberative processes in the same policy field, e.g. the governance of AI technologies. Likewise, interaction dynamics in one policy field can change over time.

Landscape of actors whose practices map shape normativity/normality on AI technologies

Norm research has prominently focused on a comparatively limited group of actors as major agents shaping normality/normativity. This group chiefly includes civil society actors at international and national levels, international organisations, and state representatives, often acting in diplomatic settings. Our line of thinking about how operational practices of designing AI technologies can shape normativity/normality puts the focus on another set of actors: tech companies. Scholars of private authority have long acknowledged the influence of such private, corporate actors on global governance (Cutler *et al.*, 1999; Hall & Biersteker, 2002; Leander, 2010). Like the public-deliberative focus of norm research, corporate actors are described as having regulatory agenda-setting power through their participation in various, more or less formalised, international fora. But the role that tech companies play in shaping normativity/normality around AI technologies is, at the same time, more direct and less visible. Representatives from these tech companies are invited as recognised experts to discuss potential, public-deliberative governance solutions for AI technologies (Bode & Huelss, 2023). But these tech

companies also perform the operational practices of development and design that shape normality/normativity inherent to AI technologies in the first place.

This insight calls for reconsidering the landscape of relevant actors to study when norm scholarship wants to trace emerging normativity/normality around AI technologies. There is an emerging stream of studies who have started this endeavour, including a larger-scale research project theorising the significance of the platform nature of these companies. So far, the focus has mostly been on tracking the significant influence of "Big Tech", such as Google (Hoijtink & Planqué-van Hardeveld, 2022; Planqué-van Hardeveld, 2023). This denominator describes the five major tech companies that dominate the overall AI landscape in terms of the infrastructure needed to train learning algorithms in terms of data sets and data centres, the personnel with the requisite professional backgrounds needed to conduct such operations, and the output produced in the form of AI technologies. But the rise of tech companies such as OpenAI demonstrates that scholars need to cast their analytical net wider to account for the precise role of tech companies. Further, the example of AI technologies in the military domain demonstrates a changing ecosystem of companies: while governmental defence actors have initially primarily sought cooperation with the Big Tech and such defence contracts continue, over the past 2–3 years, we have seen a growing number of smaller, specialised, and often new tech companies who focus on producing AI technologies for the military domain. Such technologies are especially usable in the area of decision support. There is a lot to monitor here in terms of the roles of such tech companies versus the roles of political actors in shaping what governs AI technologies and what kind of normativity/normality these technologies produce, and how the dynamics of the evolving AI ecosystem with its growing diversity of tech companies producing AI technologies comes into play/this.

Direction of political engagement for critical International Relations scholarship

Finally, we close with reflecting on where these analytical observations and research streams about normativity/normality in the context of AI technologies lead us as critical theorists of International Relations. Our understanding of critical IR scholarship extends beyond the Critical Theory associated with the Frankfurt School. We share Robert Cox' basic distinction between problem-solving and critical theories famously summarised in the insight that "theory is always for someone, and for some purpose" (Cox, 1981, 128). For us, this manifests in the desire to sustain real-world transformative impacts through our theorising based on a broad recognition of dynamics of redistribution and recognition (Fraser, 1985). Based on this attitude, we want to add two comments about our key argument that operational practices related to designing, but also training people for using, and employing AI technologies shape normativity/normality – potentially at the expense of how normality/normativity is shaped as part of public-deliberative processes.

First, this argument points to how processes that typically run under the analytical radar should receive much closer attention and scrutiny. This should also make such operational practices a site of political engagement for both social scientists monitoring such processes closely and for the natural scientists involved in design and development. Among parts of the scientific community commonly involved in design processes of AI technologies there is a growing recognition of their responsibility manifesting in the notion of value-sensitive design (Friedman *et al.*, 2013). Specifically, this covers not only to the design phase but also extends to the ethical consequences of using technologies in particular contexts as well as their potential long-term impacts (Cawthorne, 2022; Cawthorne, 2023). Beyond recognising the potential normative/normal impacts of operational design and use practices, responsible research and innovation models focus on following the four principles of anticipation, inclusion, reflexivity, and responsiveness (Stilgoe *et al.*, 2013). This thinking around value-sensitive design that rejects neutral understandings of technology and recognises its normal/normative dimension also lends itself to interdisciplinary cooperation between the social and the natural sciences.

Second, the very insight that operational practices may contravene public-deliberative processes does not mean that we should abandon public-deliberative processes. Instead, such processes remain vital for drawing public attention to the potential adverse forms of normativity/normality that take shape in the context of operational practices. As we described about, negative acknowledgement of the normative/normal consequences of such operational practices can start of process of changing such practices and thereby the normativity/normality they shape. For that to be an option, any negotiation or drafting processes of soft or hard law on AI governance need to be either already accompanied by or followed by thorough thinking around their implementation. As we noted before, the indeterminacy of international law typically defers resolution elsewhere. This "elsewhere" is particularly important in the case of AI technologies because of what we noted about practices of design, development, training, and usage. Yet, the outcomes of some of the emerging governance initiatives on AI technologies remain at the level of abstract, ambiguous principles. Just to offer one example, the Bletchley Declaration by countries attending the AI Safety Summit in November 2023 calls for the development of AI "in such a way as to be human-centric, trustworthy and responsible" (Government of the UK, 2023). But there is no concrete pathway that would take such broad principles to the level of operational practices and thereby actually making a difference in terms of how AI technologies are developed, designed, and used (Nadibaidze, 2023). This requires thinking in very concrete ways about translating such principles in operational terms. Such a process could be accompanied by international standards organisations such as the IEEE SA or the ISO and will require a lot of additional attention to detail.

Conclusions

In this essay, we reflected on practices related to AI technologies that are performed beyond public deliberation and at sites beyond the public eye as sources of normativity and normality. Empirically inspired by our empirical research agenda investigating the integration of AI and autonomous technologies in weapon systems, we argue that these dynamics pose

several conceptual challenges, but also opportunities, for IR norm research. These include the relationship between the emergence of norms in operational practices and in a public space, the interplay between normativity and normality, and ultimately, how the increasing spread of AI and autonomous technologies in various social settings leads to the emergence of new, post-human agency qualities with a potential effect on normative substance. Accepting these challenges offers IR norm research a broader analytical view of how and by what means international normative space is constituted. Ultimately, these challenges also offer opportunities for constructing new analytical arguments that can contribute to a better understanding of human-machine interaction and its societal consequences, thus addressing questions that are socially and politically relevant far beyond the disciplinary boundaries of IR.

Data availability

There is no data associated with this article.

Acknowledgements

We are grateful for feedback received by the AutoNorms team – Cecilia Ducci, Anna Nadibaidze, Guangyu Qiao-Franco, and Tom Watts.

References

Acharya A: **How Ideas Spread: Whose Norms Matter? Norm Localization and Institutional Change in Asian Regionalism.** *INT ORGAN.* 2004; **58**(2): 239–75.
Publisher Full Text

Acharya A: **Whose Ideas Matter? Agency and Power in Asian Regionalism.** Ithaca, NY: Cornell University Press, 2009.
Reference Source

Adler-Nissen R: **Towards a Practice Turn in EU Studies: The Everyday of European Integration.** *JCMS-J COMMON MARK S.* 2016; **54**(1): 87–103.
Publisher Full Text

Afina Y: **Intelligence Is Dead: Long Live Artificial Intelligence.** Chatham House. 14 July 2022.
Reference Source

Ajunwa I, Freidler S, Scheidegger C, *et al.*: **Hiring by Algorithm: Predicting and Preventing Disparate Impact.** In: New Haven, CT, 2016.
Reference Source

Amariles DR, Baquero PM: **Promises and Limits of Law for a Human-Centric Artificial Intelligence.** *COMPUT LAW SECUR REV.* 2023; **48**: 105795.
Publisher Full Text

Amoore L: **Doubt and the Algorithm: On the Partial Accounts of Machine Learning.** *Theory Cult Soc.* 2019; **36**(6): 147–69.
Publisher Full Text

Amoore L: **Cloud Ethics: Algorithms and the Attributes of Ourselves and Others.** Durham and London: Duke University Press, 2020; 232.
Publisher Full Text

Aradau C, Blanke T: **Governing Others: Anomaly and the Algorithmic Subject of Security.** *Eur J Int Secur.* 2018; **3**(1): 1–21.
Publisher Full Text

Barbé E, Badell D: **Chasing Gender Equality Norms: The Robustness of Sexual and Reproductive Health and Rights.** *INT RELAT.* 2023; **37**(2): 274–97.
Publisher Full Text

Benjamin R: **Race after Technology: Abolitionist Tools for the New Jim Code.** Medford, MA: Polity, 2019.
Reference Source

Bernstein S, Laurence M: **Practices and Norms: Relationships, Disjunctures, and Change.** In: Drieschova, A; Bueger, C & Hopf, T. (eds) *Conceptualising International Practices.* Cambridge: Cambridge University Press. 2022; 77–99.
Publisher Full Text

Betts A, Orchard P: **Implementation and World Politics: How International Norms Change Practice.** Oxford: Oxford University Press, 2014.
Publisher Full Text

Bijker WE, Law J: **Shaping Technology/Building Society: Studies in Sociotechnical Change.** Cambridge, MA: MIT Press, 2010.
Reference Source

Birhane A: **Automating Ambiguity: Challenges and Pitfalls of Artificial Intelligence.** Dublin: University College Dublin, 2021a.

Birhane A: **Algorithmic Injustice: A Relational Ethics Approach.** *Patterns (N Y).* 2021b; **2**(2): 100205.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Birhane A, Kalluri P, Card D, *et al.*: **The Values Encoded in Machine Learning Research**. 2021.
Publisher Full Text

Björkdahl A: **Norms in International Relations: Some Conceptual and Methodological Reflections.** *CAMB REV INT AFF.* 2002; **15**(1): 9–23.
Publisher Full Text

Bloomfield A: **Norm Antipreneurs and Theorising Resistance to Normative Change.** *Rev Int Stud.* 2016; **42**(2): 310–33.
Publisher Full Text

Bode I: **Contesting Use of Force Norms Through Technological Practices.** *Heidelb J Int Law.* 2023a; **83**(1): 39–64.
Publisher Full Text

Bode I: **Practice-Based and Public-Deliberative Normativity: Retaining Human Control over the Use of Force.** *Eur J Int Relat.* 2023b; **29**(4): 990–1016.
Publisher Full Text

Bode I: **Emergent Normativity: Communities of Practice, Technology, and Lethal Autonomous Weapon Systems.** *Global Studies Quarterly.* 2024; **4**(1): 1–11.
Publisher Full Text

Bode I, Huelss H: **Autonomous Weapons Systems and Changing Norms in International Relations.** *Rev Int Stud.* 2018; **44**(3): 393–413.
Publisher Full Text

Bode I, Huelss H: **The Future of Remote Warfare? Artificial Intelligence, Weapons Systems, and Human Control.** In: *Remote Warfare. Interdisciplinary Perspectives.* edited by Alasdair McKay, Abigail Watson, and Megan Karlshøj-Pedersen, Bristol: E-International Relations, 2021; 218–33.
Reference Source

Bode I, Huelss H: **Autonomous Weapon Systems and International Norms.** Montreal: McGill-Queens University Press, 2022.
Reference Source

Bode I, Huelss H: **Constructing Expertise: The Front- and Back-Door Regulation of AI's Military Applications in the European Union.** *J Eur Public Policy.* 2023; **30**(7): 1230–1254.
Publisher Full Text

Bode I, Karlsrud J: **Implementation in Practice: The Use of Force to Protect Civilians in United Nations Peacekeeping.** *Eur J Int Relat.* 2019; **25**(2): 458–85.
Publisher Full Text

Bode I, Watts T: **Meaning-Less Human Control. The Consequences of Automation and Autonomy in Air Defence Systems.** Oxford and Odense: Drone Wars UK & Center for War Studies, 2021.

Bode I, Watts T: **Loitering Munitions and Unpredictability: Autonomy in Weapon Systems and Challenges to Human Control.** Odense, London: SDU Center for War Studies, Royal Holloway Centre for International Security, 2023.
Reference Source

Bolukbasi T, Chang KW, Zou C, *et al.*: **Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings**. 2016.
Reference Source

Boulanin V, Verbruggen M: **Mapping the Development of Autonomy in Weapon Systems.** Stockholm: SIPRI, 2017.
Publisher Full Text

Bower A: **Norms without the great power: International law, nested social structures, and the ban on antipersonnel landmines.** *International Studies Review.* 2015; **17**(3): 347–373.
Publisher Full Text

Bradshaw JM, Hoffman RR, Woods DD, *et al.*: **The Seven Deadly Myths of "Autonomous Systems".** *IEEE Intell Syst.* 2013; **28**(3): 54–61.
Publisher Full Text

Bucher B: **Acting Abstractions: Metaphors, Narrative Structures, and the Eclipse of Agency.** *Eur J Int Relat.* 2014; **20**(3): 742–65.
Publisher Full Text

Bueger C: **Pathways to Practice: Praxiography and International Politics.** *Eur Polit Sci Rev.* 2014; **6**(03): 383–406.
Publisher Full Text

Buolamwini J, Gebru T: **Gender Shades: Intersectional Accuracy Disparaties in Commercial Gender Classification.** *Proc Mach Learn Res.* 2018; **81**: 1–15.
Reference Source

Carmel E, Paul R: **Peace and Prosperity for the Digital Age? The Colonial Political Economy of European AI Governance.** *IEEE Technol Soc Mag.* 2022; **41**(2): 94–104.
Publisher Full Text

Cawthorne D: **Robot Ethics: Ethical Design Considerations.** In: *Foundations of Robotics.* edited by D Herath and D St-Onge, Singapore: Springer, 2022; 473–91.
Publisher Full Text

Cawthorne D: **The Ethics of Drone Design: How Value-Sensitive Design Can Create Better Technologies.** 1st ed. New York: Routledge, 2023.
Publisher Full Text

C.A.S.E Collective: **Critical Approaches to Security in Europe: A Networked Manifesto.** *Secur Dialogue.* 2006; **37**(4): 443–87.
Reference Source

Chandler K: **Does Military AI Have Gender? Understanding Bias and Promoting Ethical Approaches in Military Applications of AI.** Geneva: UNIDIR, 2021.
Reference Source

Coole D: **Agentic Capacities and Capacious Historical Materialism: Thinking with New Materialisms in the Political Sciences.** *Millennium-J Int St.* 2013; **41**(3): 451–69.
Publisher Full Text

Coole D, Frost S: **New Materialisms: Ontology, Agency, and Politics.** Durham: Duke University Press, 2010.
Reference Source

Cox RW: **Social Forces, States and World Orders: Beyond International Relations Theory.** *Millenn J Int Stud.* 1981; **10**(2): 126–55.
Publisher Full Text

Crawford K: **Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence.** New Haven: Yale University Press, 2021.
Publisher Full Text

Cryle P, Stephens E: **Normality: A Critical Genealogy.** Chicago, IL: University of Chicago Press, 2017.
Publisher Full Text

Currier C, Greenwald G, Fishman A: **US Government Designated Prominent Al Jazeera Journalist as "Member of Al Qaeda.** The Intercept, 8 May 2015.
Reference Source

Cutler AC, Haufler V, Porter T: **Private Authority and International Affairs.** Albany, NY: State University of New York Press, 1999.
Reference Source

Deitelhoff N, Zimmermann L: **Things We Lost in the Fire: How Different Types of Contestation Affect the Robustness of International Norms.** *Int Stud Rev.* 2018.

Deitelhoff N, Zimmermann L: **Things We Lost in the Fire: How Different Types of Contestation Affect the Robustness of International Norms.** *Int Stud Rev.* 2020; **22**(1): 51–76.
Publisher Full Text

Dignum V: **Responsible Artificial Intelligence. How to Develop and Use AI in a Responsible Way.** Springer, 2019.
Publisher Full Text

Drubel J, Mende J: **The Hidden Contestation of Norms: Decent Work in the International Labour Organization and the United Nations.** *Global Constitutionalism.* 2023; **12**(2): 246–68.
Publisher Full Text

Eubanks V: **Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.** First Picador edition. New York: Picador St. Martin's Press, 2019.
Reference Source

Ferrara E: **Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies.** arXiv, 2023.
Reference Source

Finnemore M, Sikkink K: **International Norm Dynamics and Political Change.** *Int Organ.* 1998; **52**(4): 887–917.
Reference Source

Fischer B: **Waging War Via Chatbot.** Handelszeitung, May 2023.
Reference Source

Fraser N: **What's Critical about Critical Theory? The Case of Habermas and Gender.** *New German Critique.* 1985; **35**: 97–131.
Publisher Full Text

Friedman B, Kahn PH, Borning A, *et al.*: **Value Sensitive Design and Information Systems.** In: *Early Engagement and New Technologies: Opening up the Laboratory.* edited by Neelke Doorn, Daan Schuurbiers, Ibo Van De Poel, and Michael E. Gorman, Philosophy of Engineering and Technology. Dordrecht: Springer Netherlands. 2013; **16**: 55–95.
Publisher Full Text

Gadinger F: **The Normativity of International Practices.** In: *Conceptualizing International Practices. Directions for the Practice Turn in International Relations.* edited by Alena Drieschova, Christian Bueger, and Ted Hopf, Cambridge: Cambridge University Press, 2022; 100–21.
Reference Source

Garcia D: **The AI Military Race: Common Good Governance in the Age of Military Artificial Intelligence.** Oxford: Oxford University Press, 2023.

Gebru T: **Don't Fall for the AI Hype.** Tech Won't Save Us Podcast with Paris Marx, 19 January 2023.
Reference Source

Gholiagha S, Hansen-Magnusson H, Hofius M: **Zum Verhältnis von Normativität Und Normalität in Der Normenforschung.** In: *Kritische Normenforschung in Den International Beziehungen.* edited by Stephan Engelkamp, Katharina Glaab, and Antonia Graf. Baden-Baden: Nomos, 2021.

Gillespie T: **Algorithm.** Culture Digitally, June, 2014.

Goldstaub T: **How to Talk to Robots: A Girls' Guide to a Future Dominated by AI.** London: 4th Estate, 2021.
Reference Source

Government of the UK: **The Bletcheley Declaration.** UK Government. 2023.
Reference Source

Gunkel DJ: **AI Is Not Intelligent.** IAI News, 22 June, 2023.
Reference Source

Hall RB, Biersteker TJ: **The Emergence of Private Authority in the International System.** In: *The Emergence of Private Authority in Global Governance.* Cambridge: Cambridge University Press, 2002; 3–22.
Publisher Full Text

Hansen L: **Security as Practice: Discourse Analysis and the Bosnian War**. London: Routledge, 2006.
Reference Source

Hofferberth M, Weber C: ***Lost in Translation*: A Critique of Constructivist Norm Research.** *J Int Relat Dev.* 2015; **18**(1): 75–103.
Reference Source

Hoijtink M, Planqué-van Hardeveld A: **Machine Learning and the Platformization of the Military: A Study of Google's Machine Learning Platform TensorFlow.** *Int Political Sociol.* 2022; **16**(2): olab036.
Publisher Full Text

Holland Michel A: **Recalibrating Assumptions on AI.** Chatham House, 12 April, 2023.
Reference Source

Hooker S: **Moving beyond "Algorithmic Bias Is a Data Problem".** *Patterns (N Y).* 2021; **2**(4): 100241.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Hudson J, Khudov K: **The War in Ukraine Is Spurring a Revolution in Drone Warfare Using AI.** The Washington Post, 26 July, 2023.
Reference Source

Huelss H: **After Decision-Making: The Operationalisation of Norms in International Relations.** *International Theory.* 2017; **9**(3): 381–409.
Publisher Full Text

Huelss H: **Norms Are What Machines Make of Them: Autonomous Weapons Systems and the Normative Implications of Human-Machine Relations.** *International Political Sociology.* 2020; **14**(2): 111–28.
Publisher Full Text

Hurd I: **How to Do Things with International Law.** Princeton, NJ: Princeton University Press, 2017.

Ignatius D: **How the Algorithm Tipped the Balance in Ukraine.** The Washington Post, 20 December, 2022.
Reference Source

Jaton F, Bowker G: **The Constitution of Algorithms: Ground-Truthing, Programming, Formulating.** Cambridge, MA: MIT Press. 2020.
Reference Source

Johnson M, Hawley JK, Bradshaw JM: **Myths of Automation Part 2: Some Very Human Consequences.** *IEEE Intell Syst.* 2014; **29**(2): 82–85.

Katzenstein PJ: **The Culture of National Security: Norms and Identity in World Politics.** New York: Columbia University Press, 1996.
Reference Source

Keck ME, Sikkink K: **Activists Beyond Borders: Advocacy Networks in International Politics.** Ithaca, NY: Cornell University Press, 1998.
Reference Source

Koskenniemi M: **The Politics of International Law.** Oxford: Hart, 2011.
Reference Source

Lantis JS, Wunderlich C: **Resiliency Dynamics of Norm Clusters: Norm Contestation and International Cooperation.** *Rev Int Stud.* 2018; **44**(3): 570–93.
Publisher Full Text

Latour B: **Reassembling the Social: An Introduction to Actor-Network Theory.** Oxford: Oxford University Press, 2005.
Reference Source

Laurence M: **An 'Impartial' Force? Normative Ambiguity and Practice Change in UN Peace Operations.** *Int Peacekeeping.* 2019; **26**(3): 256–80.
Publisher Full Text

Leander A: **Practices (Re)Producing Orders: Understanding the Role of Business in Global Security Governance.** In: *Business and Global Governance.* edited by Anna Leander and Morten Ougaard, London: Routledge, 2010; 57–77.
Reference Source

Leander A: **Thinking Tools.** In: *Qualitative Methods in International Relations: A Pluralist Guide.* edited by Audie Klotz and Deepa Prakesh, Basingstoke: Palgrave

Macmillan, 2008; 11–27.
Publisher Full Text

Lukowicz P: **The Challenge of Human Centric AI.** *Digitale Welt.* 2019; **3**(4): 9–10.
Publisher Full Text

Lum K, Isaac W: **To Predict and Serve?** *Significance.* 2016; **13**: 14–19.
Publisher Full Text

Maas M, Villalobos JJ: **International AI Institutions. A Literature Review of Models, Examples, and Proposals.** 1. AI Foundations Report, Legal Priorities Project, 2023.
Publisher Full Text

Mackenzie A: **Cutting Code. Software and Sociality.** New York: Peter Lang, 2015.

Moyes R: **Target Profiles.** 2019; 36.
Reference Source

Nadibaidze A: **AI Summits and Declarations: Symbolism or Substance.** *AutoNorms Blog.* 2023.
Reference Source

Natale S, Ballatore A: **Imagining the Thinking Machine: Technological Myths and the Rise of Artificial Intelligence.** *Convergence: The International Journal of Research into New Media Technologies.* 2020; **26**(1): 3–18.
Publisher Full Text

Neumann IB: **Returning Practice to the Linguistic Turn: The Case of Diplomacy.** *Millennium.* 2002; **31**(3): 627–51.
Publisher Full Text

Nicolini D: **Practice Theory, Work, and Organization: An Introduction.** Oxford: Oxford University Press, 2013.
Reference Source

Obermeyer Z, Powers B, Vogeli C, *et al.*: **Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations.** *Science.* 2019; **366**(6464): 447–53.
PubMed Abstract | **Publisher Full Text**

Orchard P, Wiener A: **Norms and Norm Contestation.** In: *Routledge Handbook of Foreign Policy Analysis Methods.* edited by Patrick A Mello and Falk Ostermann, 1st ed., London: Routledge, 2022; 51–66.
Publisher Full Text

Oxford English Dictionary: **Deliberation, Noun.** 2020.
Reference Source

Paul R: **Can *Critical Policy Studies* Outsmart AI? Research Agenda on Artificial Intelligence Technologies and Public Policy.** *Crit Policy Stud.* 2022; **16**(4): 497–509.
Publisher Full Text

Penn J: **Algorithmic silence: a call to decomputerise.** *J Social Comput*. 2021; 2(4): 335–356.
Publisher Full Text

Planqué-van Hardeveld A: **Securing the Platform: How Google Appropriates Security.** *Crit Stud Secur.* 2023; **11**(3): 161–75.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Pouliot V: **International Pecking Orders: The Politics and Practice of Multilateral Diplomacy.** New York: Cambridge University Press, 2016.
Publisher Full Text

Pratt SF: **From Norms to Normative Configurations: A Pragmatist and Relational Approach to Theorizing Normativity in IR.** *Int Theor.* 2020; **12**(1): 59–82.
Publisher Full Text

Pratt SF: **Normative Transformation and the War on Terrorism.** Cambridge: Cambridge University Press, 2022.
Publisher Full Text

Qiao-Franco G, Bode I: **Weaponised Artificial Intelligence and Chinese Practices of Human-Machine Interaction.** *Chin J Int Polit.* 2023; **16**(1): 106–28.
Publisher Full Text

Ralph J, Gifkins J: **The Purpose of United Nations Security Council Practice: Contesting Competence Claims in the Normative Context Created by the Responsibility to Protect.** *Eur J Int Relat.* 2017; **23**(3): 630–53.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Régimbal CM: **Opening Panel - Political View.** Presented at the Luxemburg Autonomous Weapon Systems Conference, Esch-sur-Alzette, Luxemburg, April 26, 2023.

Risse T, Ropp SC, Sikkink K: **The Power of Human Rights: International Norms and Domestic Change.** Cambridge: Cambridge University Press, 1999.
Publisher Full Text

Robbins M: **Has a Rampaging AI Algorithm Really Killed Thousands in Pakistan?** *The Guardian*. 18 February, 2016.
Reference Source

Rosert E: **Norm Emergence as Agenda Diffusion: Failure and Success in the Regulation of Cluster Munitions.** *Eur J Int Relat.* 2019; **25**(4): 1103–31.
Publisher Full Text

Russell SJ: **Human Compatible: Artificial Intelligence and the Problem of Control.** New York: Viking, 2019.
Reference Source

Scahill J, Greenwald G: **The NSA's Secret Role in the U.S. Assassination Program.** The Intercept. (blog). 2 October, 2014.
Reference Source

Scharre P, Horowitz MC: **An Introduction to Autonomy in Weapon Systems.** Center for New American Security, 2015.
Reference Source

Schatzki TR: **A Primer on Practices.** In: *Practice-Based Education.* edited by Joe Higgs, Rotterdam: Sense Publishers, 2012; 13–26.
Publisher Full Text

Schouten P, Mayer M: **Materiality.** In: *Routledge Handbook of International Political Sociology.* edited by Xavier Guillaume and Pinar Bilgin. Abingdon: Routledge, 2016.

Sharkey N: **The Impact of Gender and Race Bias in AI.** ICRC Blog. 28 August, 2018.
Reference Source

Stilgoe J, Owen R, Macnaghten P: **Developing a Framework for Responsible Innovation.** *Res Policy.* 2013; **42**(9): 1568–80.
Publisher Full Text

Stimmer A, Wisken L: **The Dynamics of Dissent: When Actions Are Louder Than Words.** *International Affairs (London).* 2019; **95**(3): 515–33.
Publisher Full Text

Suchman L: **Algorithmic Warfare and the Reinvention of Accuracy.** *Crit Stud Secur.* 2020a; **8**(2): 175–87.
Publisher Full Text

Suchman L: **AI at the Edgelands: Data Analytics in States of In/Security. Lecture at the Qualcomm Institute, UC San Diego.** You Tube, 2020b.
Reference Source

Tallberg J, Erman E, Furendal M, *et al.*: **The Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative Research.** *Int Stud Rev.* 2023; **25**(3): eloc.
Publisher Full Text

Tarnoff B: **A Certain Danger Lurks There: How the Inventor of the First Chatbot Turned Against AI.** The Guardian. 25 July, 2023.
Reference Source

True J, Wiener A: **Everyone Wants (a) Peace: The Dynamics of Rhetoric and Practice on "Women, Peace and Security".** *Int Aff.* 2019; **95**(3): 553–74.
Publisher Full Text

Tucker E: **Artifice and Intelligence.** Tech Policy Press. 17 March, 2022.
Reference Source

UK Government: **A Pro-Innovation Approach to AI Regulation.** GOV.UK. 29 March, 2023.
Reference Source

Veale M, Matus K, Gorwa R: **AI and Global Governance: Modalities, Rationales, Tensions.** *Annu Rev Law Soc Sci.* 2023; **19**: 1–30.
Publisher Full Text

Walsh T: **Machines That Think. The Future of Artificial Intelligence.** Amherst, NY: Promotheus Books, 2018.
Reference Source

Walsh T: **Machines Behaving Badly: The Morality of AI.** Cheltenham, Gloucesterhire: Flint, 2022.
Reference Source

Whittaker M: **The Steep Cost of Capture.** *Interactions.* 2021; **28**(6): 50–55.
Publisher Full Text

Wiener A: **The Dual Quality of Norms and Governance beyond the State: Sociological and Normative Approaches to 'Interaction'.** *Crit Rev Int Soc Political Philos.* 2007; **10**(1): 47–69.
Publisher Full Text

Wiener A: **Enacting Meaning-in-Use: Qualitative Research on Norms and International Relations.** *REV INT STUD.* 2009; **35**(1): 175–93.
Publisher Full Text

Wiener A: **A Theory of Contestation.** Heidelberg: Springer, 2014.
Publisher Full Text

Wiener A: ***A Theory of Contestation* - A Concise Summary of Its Arguments and Concepts.** *Polity.* 2017; **49**(1): 109–25.
Publisher Full Text

Wiener A: **Contestation and Constitution of Norms in Global International Relations.** Cambridge, 2018.
Publisher Full Text

Wiener A: **Societal Multiplicity for International Relations: Engaging Societal Interaction in Building Global Governance from Below.** *Coop Confl.* 2022; **57**(3): 348–66.
Publisher Full Text

Zimmermann L: **Global Norms with a Local Face: Rule-of-Law Promotion and Norm-Translation.** Cambridge: Cambridge University Press, 2017.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔ ✔ ✔ ?

Version 2

Reviewer Report 28 May 2024

https://doi.org/10.21956/openreseurope.18650.r39587

© **2024 Lesch M.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Max Lesch
Peace Research Institute Frankfurt, Frankfurt, Germany

Thank you for giving me the opportunity to review the essay for *Open Research Europe*, entitled "Artificial Intelligence Technologies and Practical Normativity/Normality: Investigating Practices beyond the Public Space" by Ingvild Bode and Hendrik Huelss. The article explores the normative dimension of governing artificial intelligence (AI) technologies. The authors contribute to this vivid field in international relations by providing a practice-based approach to international norm dynamics surrounding the rapid development of AI technologies in various areas, including autonomous weapons systems. Based on their concept of "practice normativity/normality", they provide an analytical framework for studying how the development and use of AI technologies shape international norms. This framework is based on the distinction (and interlinkage) between normality and normativity. Whereas normality connotes "the typical and average" (in practice), normativity refers to "principles of oughtness and justice". Rather than taking international negotiations as the starting point for analysing the development of international norms, they propose to begin with the practices and agency of tech firms that develop and deploy AI technologies. The essay provides an excellent contribution and comprehensive overview of the challenge that AI technologies pose for international relations, both politically and theoretically, and promises an exciting future research agenda on AI. In the remainder of this review, I will focus on the link between normality, deviance, and normativity, and, closely related, the concept of practice, where I believe conceptual clarification could further strengthen the central argument –particularly with regard to the norm-generative effects of practice.

How does practice become normatively meaningful? After all that is what the authors set out to study: "how artificial intelligence may shape international norms" (p. 1), which they define in line with conventional IR norms research as "understandings of appropriateness" (p. 4). Their central argument is that international norms on AI are currently mainly driven by AI practices themselves, which have been understudied in IR norms research due to a bias to focus on public deliberations. Therefore, as I understand the authors, the normality created in practice by AI technologies is the more salient source for shaping international norms and/or normativity than the public deliberations. In my view, it would help to strengthen this argument if it was more clearly distinguished from – ultimately also practice-based – approaches in other theoretical traditions

focussing on the norm-generative effects of for example norm violations, deviance and precedents. Here, the absence of normality or its rupture is the primary norm-generative factor.

This also points to a more fundamental conceptual issue: When we take norms as "understandings of appropriateness" and, very similarly, normativity as "principles of oughtness or justice", then I think it becomes difficult to see normativity (or norms) in normality. Normativity is only induced when these practices, this normality, are actualised as appropriate – or inappropriate. As Christoph Möllers (2020, 66) argues norms are more than representations of normality or regularity. From this perspective, regularity may even be devoid of normativity because we usually recognise and negotiate what is appropriate in the face of what is not normal, abnormal – or deviant. This point is of course recognized by Bode and Huelss, when they discuss "negative acknowledgements" or "collision" in their analytical framework (p. 11), but it could be made more explicit with regard to the underlying concept of norms and normativity as well as regarding the sites and actors that engage in negative acknowledgements beyond formal deliberative forums. Linking this argument to the sociology of deviance in IR (Adler-Nissen 2014), would suggest to recalibrate the metaphor of taking normality and normativity as "two sides of the same coin" (p. 4). Perhaps, normality and deviance are the two sides of that coin – the norm – and normativity is what distinguishes them. In my view, the abnormal or deviant is conspicuously absent from the notion of 'practical normativity/normality', at least at first sight, but is quite important for the making, maintaining and unmaking of international norms (Lesch 2023, 3). For example, we might only see some (informal or social) norms already at play/emerging in the governance of AI at the moment they are violated and labelled as such. In short, my suggestion would be to include the "non-normal" in the concept of norms and normativity, which would also better ground the framework developed on p. 11.

That being said, a focus on norm violations does of course not help empirically if we cannot observe negative reactions to norm violations or the labelling of deviance and stigmatization of norm breakers. While I wonder to what extent the discourse includes such responses – and the authors refer to some of them – the author's argument would probably be that the more decisive practice-driven factor (in this case) is normality and its productivity for international norms. Writing as someone very sympathetic with practice-based approaches to the study of norms in IR, I would like to invite the authors to invest a little more in explaining how practices really normalise? Where can we see the performative effects of practices on norms (or norm change) and how can we analytically capture this normativity? This also speaks to the agent-structure debate that the authors touch on briefly. Perhaps it would help to clarify the authors' interpretation of Giddens' structuration theory, which I would read slightly differently in terms of his suggestion that agency and structure are intimately intertwined through practice – in the duality of structure (Gidden 1979, 69). This is of course echoed by Antje Wiener's (2007) notion of "dual quality" of norms in IR. In other words, when do practices become structuring and when do they not? In my view, STS approaches could help to strengthen the argument about normalisation, the role of algorithms as "actants" etc. as creating normativity by virtue of being performed – also in view to other norm development processes that evolve below the radar of formal treaty making, for instance, through expertise. But there are also cases where normality is not norm-generative and my suggestion would be to use this distinction (to non-cases) to argue when and how these practices have performative effects and shape international norms.

Finally, and closely related, I wondered what kind of norms the authors are referring to that are shaped by international practice. The distinction between deliberative forums would suggest that

they have in mind some sort of formalised or codified international treaty norms – which do not keep pace with developments on the ground. However, the definition of norms given would also refer to social or informal norms that can emerge without, before or alongside legal norms. In international law, the shift to informality is also discussed as an alternative strategy to advance law-making processes where formal treaty-making is blocked (Mantilla 2024; Reiners 2024). While this can be productive for the development of international norms, the present case would probably point in the opposite direction, where practice-based and informal norm dynamics – from a normative point of view – hinder or undermine the regulation of AI technologies. A discussion of this aspect could also provide an opportunity (if space allows) to spend a few sentences on how we can generalize from the case of AI governance to other issue areas.

References

1. Adler-Nissen R: Stigma Management in International Relations: Transgressive Identities, Norms, and Order in International Society. *International Organization*. 2014; **68** (1): 143-176 Publisher Full Text
2. Lesch M: From Norm Violations to Norm Development: Deviance, International Institutions, and the Torture Prohibition. *International Studies Quarterly*. 2023; **67** (3). Publisher Full Text
3. Mantilla G: From treaty to custom: Shifting paths in the recent development of international humanitarian law. *Leiden Journal of International Law*. 2024. 1-20 Publisher Full Text
4. Schneider A: The possibility of norms: social practice beyond morals and causes. *Jurisprudence*. 2023; **14** (1): 118-126 Publisher Full Text
5. Reiners N: States as bystanders of legal change: Alternative paths for the human rights to water and sanitation in international law. *Leiden Journal of International Law*. 2023. 1-20 Publisher Full Text
6. Wiener A: The Dual Quality of Norms and Governance beyond the State: Sociological and Normative Approaches to 'Interaction'. *Critical Review of International Social and Political Philosophy*. 2007; **10** (1): 47-69 Publisher Full Text
7. anthony giddens: Central Problems in Social Theory: Action, Structure, and Contradiction in Social Analysis. *University of California Press*.

Is the topic of the essay discussed accurately in the context of the current literature?

Yes

Is the work clearly and cogently presented?

Yes

Is the argument persuasive and supported by appropriate evidence?

Yes

Does the essay contribute to the cultural, historical, social understanding of the field?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: international norms, international practices, international law, human rights, peace and security law

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 30 April 2024

https://doi.org/10.21956/openreseurope.18650.r39585

© 2024 Al Asyari H. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Haekal Al Asyari
[1] Universitas Gadjah Mada, Yogyakarta, Special Region of Yogyakarta, Indonesia
[2] University of Debrecen, Debrecen, Hajdú-Bihar, Hungary

The topic of the essay is accurate with the context of the current literature. The research proposed a rather obvious but often overlooked concept on how operational practices of technology companies contribute towards AI norm-making process. The take on interplay between normality and normativity as the basis of analysis adds value to the discourse of AI governance.

While it seems that the object of the research is regarding autonomous weapons systems, the authors does not seem to incorporate the discussion of AWS throughout the introduction consistently. Particularly throughout the first half of the essay, the narrative on the concept of AI implies a general example of AI technologies in the everyday social life. Furthermore, the statement of '*AI technologies that are performed by actors at* ***site outside of the public eye***' does not provide a direct correlation to the subject of AWS. Perhaps it would be better to further elaborate on the meaning of 'site outside of the public eye'.

The authors have provided an extensive explanation on the definition of AI. An eye-catching statement in that section states '*...both AI and autonomous technologies as these are closely related and arguably function along a trajectory of increasing technological complexity*'. It is arguable whether all AI technologies are autonomous. Perhaps providing a reference would add credibility to the statement.

The essay provides thorough explanation on how the practices of designing and using AI technologies that are performed by public actors in the military field, have become a phenomenon of normality, prior to the normative progress that shapes the existing legal framework on AI governance. The authors have done a great job at explaining the concepts of design and development proves of AI technologies and algorithmic systems. It is made clear from such examples, where the basis of 'normality' comes from.

It is very interesting to understand the proposed 'reverse order' conceptualization of norm setting which comes of normality of practices. The authors have tackled this challenge by addressing two main points; developing an understanding of normativity in practices that might be different from the deliberative normativity, and by considering whether normativity is broadly or narrowly

understood.

All in all, the research is conceptually thought provoking and opens a broad horizon for future research; as it is intended.

Is the topic of the essay discussed accurately in the context of the current literature?
Yes

Is the work clearly and cogently presented?
Yes

Is the argument persuasive and supported by appropriate evidence?
Yes

Does the essay contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Cyberspace Law, Cyberspace and AI Governance

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 15 April 2024

https://doi.org/10.21956/openreseurope.18650.r38848

© 2024 Hoijtink M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Marijn Hoijtink
Department of Political Science, University of Antwerp, Antwerp, Belgium

I have now had the opportunity to review the revised article. I am pleased with the authors' detailed response to my comments and the revisions made.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Critical security studies, Science & Technology Studies, military AI and modern warfare

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 12 April 2024

https://doi.org/10.21956/openreseurope.18650.r39578

© 2024 Tuset Varela D. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Damián Tuset Varela
University of Jaén, Jaén, Spain

The article under review presents an insightful analysis into the evolving landscape of artificial intelligence (AI) governance, particularly emphasizing the critical role operational practices play in shaping norms related to AI technologies. Through an interdisciplinary approach that melds international relations (IR), norm research, and science and technology studies (STS), the authors argue compellingly that the rapid advancements in AI have outpaced traditional public-deliberative governance processes. They propose that operational practices, especially those carried out behind closed doors by private tech companies in the design, development, and deployment of AI technologies, are pivotal in the formation of AI norms, including those concerning military applications.

Expanding on this, the paper delineates four distinct dynamics that characterize the interaction between public-deliberative processes and operational practices: negative acknowledgement, positive acknowledgement, collision, and ignorance. This nuanced analysis challenges the existing norm research narrative that predominantly focuses on state and civil society actors, calling instead for a broader lens that includes the influential role of tech companies in shaping normativity around AI technologies.

However, to ensure the article's scientific soundness and enhance its impact within the scholarly community and beyond, several areas require further attention and development. Firstly, the conceptual framework, particularly the distinction between normativity and normality, needs clearer definition and illustration. This will not only strengthen the paper's foundational arguments but also make the complex dynamics of AI governance more accessible to readers. The inclusion of concrete empirical examples demonstrating how operational practices in AI design and development influence norm formation would significantly bolster the paper's arguments, providing tangible evidence of the theoretical propositions put forth.

Moreover, a more balanced analysis that engages with counterarguments regarding the efficacy of public-deliberative processes in governing AI would enrich the discourse. Exploring perspectives that underscore successful instances where these processes have effectively shaped norms and regulated AI technologies could offer a counterbalance to the paper's emphasis on operational practices. Additionally, the article would benefit from an expanded discussion on policy implications. Articulating detailed recommendations for policymakers on developing adaptive and responsive governance frameworks capable of addressing the ethical, legal, and societal challenges posed by AI is crucial.

By addressing these points, the article can significantly enhance its contribution to the field,

offering a more comprehensive and nuanced understanding of AI governance and the interplay between operational practices and norm formation. This would not only solidify the paper's scientific soundness but also underscore its relevance and applicability to a broad spectrum of stakeholders engaged in navigating the complexities of AI technologies and their governance.

Is the topic of the essay discussed accurately in the context of the current literature?
Yes

Is the work clearly and cogently presented?
Yes

Is the argument persuasive and supported by appropriate evidence?
Yes

Does the essay contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Artificial Intelligence ; International Relations; Diplomacy ; Cybersecurity, UE Law

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 08 November 2023

https://doi.org/10.21956/openreseurope.17850.r35278

© 2023 Laurence M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Marion Laurence
[1] Department of Defence Studies, Royal Military College of Canada, Toronto, Ontario, Canada
[2] Department of Defence Studies, Royal Military College of Canada, Toronto, Ontario, Canada

This article uses international practice theory and evidence from the field of artificial intelligence technologies (AIT) to analyze shortcomings in the literature on norm emergence and norm contestation. The authors distinguish between the concepts of *normativity*, defined in terms of justice and 'oughtness,' and *normality*, understood in terms of ideas about what is 'standard' or 'average.' They build on this distinction to put forward a research programme on AI and practical normativity/normality, which has two main pillars. First, they argue that operational practices of

designing and using AI can make norms, even though they are typically performed outside the public domain. This focus on private norm-making serves as a corrective to much of the norms literature, which focuses disproportionately on public discourse and deliberation. Second, they argue that the interplay between normality and normativity is analytically influential in the making of norms, and that this has under-examined effects on human agency. If practices can transform and constitute norms, then normality shaped by AI technologies can lead to "non-human generated normativity," (p. 1) a possibility not currently accounted for in norms research and one with important implications for efforts to regulate AIT.

Report

The article provides a succinct overview of trends in the research on norm life cycles and norm contestation, including an accurate discussion of current debates in IR theory and gaps in the scholarly literature. The authors rightly point out that existing norms research has focused primarily on public deliberative processes (often in multilateral settings) while neglecting norm emergence and contestation that may occur outside the public eye (see also Stimmer & Wisken 2019). There is one additional point from the literature that might be worth noting: some practice theorists subscribe to a normative conception of practice that defines normativity very broadly (e.g. Rouse 2001, 2007; cf. Bernstein and Laurence 2022). In my view, the authors' distinction between normativity and normality is theoretically sound and analytically helpful, but others might disagree because they effectively collapse normativity and normality into one category.

The article's central argument is clear and well presented. It also provides an accessible introduction to debates about artificial intelligence technology and autonomous weapons systems for non-specialists. Indeed, the concept of non-human generated normativity is compelling, if somewhat alarming, for those not familiar with the field of AI. It raises important question about evolving dynamics of norm contestation and human agency in global governance.

The argument is original, persuasive, and well-supported by evidence, mostly drawn from the area of AI and autonomous technologies that are used in weapons systems. The focus on norm emergence and contestation outside the public eye is timely. The authors observe, correctly, that normativity is produced in and through practice (see also Wiener 2004; Krook and True 2010). This means that the concept of "normality," especially as used by AITs can easily have normative implications, with targeting algorithms for autonomous weapons systems being just one example (see p. 6-7). The authors also note that AIT can replicate and amplify existing forms of normativity, which are embedded within them by programmers (e.g. reinforcement of racial and gender stereotypes, see p. 6-7). This makes me wonder about the degree to which 'new' types of AI-generated normativity are derived from pre-existing norms. It seems like they are, at least initially, products of interactions between AIT and human inputs, but products that can take of a life of their own, so to speak. The authors seem aware of this, but it may be worth underlining the human roots of "post-human agency" and AI-generated normativity, if only to emphasize the importance of thinking carefully about what norms and assumptions humans choose to embed within AIT to begin with (p. 9).

The article makes important contributions to the fields of IR theory, global governance, and the literature on norm contestation. It also improves IR scholars' understanding of artificial intelligence technologies and global efforts to regulate them. In terms of broader relevance, I wonder about the extent to which norm scholars' tendency to focus on public deliberation instead

of private settings is an artefact of methodological decisions related to access (or lack thereof) to research sites and what can be easily observed or measured (noted on p. 5). In the field of AI, this is likely compounded by a lack of technical expertise among norm scholars. It might be interesting to expand on this methodological motivation in future work.

References

1. Rouse J: Two Concepts of Practices. In: Schatzki TR, KK Cetina and E von Savigny (eds) The Practice Turn in Contemporary Theory. *Routledge*. 2001. 198-208
2. Rouse J: Social Practices and Normativity. *Philosophy of the Social Sciences*. 2007; **37** (1): 46-56 Publisher Full Text
3. Bernstein S, Laurence M: "Practices and Norms: Relationships, Disjunctures, and Change," in Conceptualizing International Practices: Directions for the Practice Turn in International Relations, ed. Bueger, C., A. Drieschova and T. Hopf. *Cambridge: Cambridge University Press*. 2022.
4. Stimmer A, Wisken L: The dynamics of dissent: when actions are louder than words. *International Affairs*. 2019; **95** (3): 515-533 Publisher Full Text
5. Krook M, True J: Rethinking the life cycles of international norms: The United Nations and the global promotion of gender equality. *European Journal of International Relations*. 2012; **18** (1): 103-127 Publisher Full Text
6. Wiener A: Contested Compliance: Interventions on the Normative Structure of World Politics. *European Journal of International Relations*. 2004; **10** (2): 189-234 Publisher Full Text

Is the topic of the essay discussed accurately in the context of the current literature?

Yes

Is the work clearly and cogently presented?

Yes

Is the argument persuasive and supported by appropriate evidence?

Yes

Does the essay contribute to the cultural, historical, social understanding of the field?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: International relations theory; global security governance; political sociology of international organizations.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 01 Feb 2024

Ingvild BODE

We want to start by expressing our appreciation for this constructive and insightful review. We have enjoyed reviewing our article again in light of your comments and engaging with

them has considerably strengthened this revised version of our article. We have taken all comments on board which has added further depth to our argument. In particular, we now reflect much more explicitly on the dynamic interplay between normativity and normality. Further, we have separated our thinking on post-human agency by putting this into a dedicated section and reflecting more explicitly on where ”machine agency” comes from and how it still remains tied to human agency. We also added an argument about “distributed normativity” here. Finally, we have added further clarification about the methodological problems that studying processes of designing AI technologies raises.

Competing Interests: No competing interests were disclosed.

Reviewer Report 02 November 2023

https://doi.org/10.21956/openreseurope.17850.r35276

© 2023 Hoijtink M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Marijn Hoijtink
[1] Department of Political Science, University of Antwerp, Antwerp, Belgium
[2] Department of Political Science, University of Antwerp, Antwerp, Belgium

This essay brings an original and important perspective to the study of the emerging governance of AI technology, specifically within the field of defense. Based on an extensive discussion of existing scholarship on norms in International Relations, the essay argues that even if formal deliberations of norms related to military AI appear to be highly contested and slow, this does not mean that military AI is unregulated or that there is an absence of norms in this field. Rather, as the authors argue, in this context, norms are still produced in the operational practices of designing and using AI technologies. This is an important argument analytically, politically and (potentially) normatively. Yet, I do also think that the argument can be pushed in a number of ways, specifically when it comes to (1) the analytical distinction between normativity and normality; (2) the interaction between formal and deliberative processes of norm development (the ‘front stage’) and the practical development of norms (the ‘back stage’ of norm development); (3) the politics involved in the production of norms and normality; and (4) the normative implications of the practical development of normativity/normality.

First, related to the distinction between normativity and normality that the authors draw, I wonder if that distinction can and should be drawn in support of the authors argument. Looking at the examples that were given under sections I and II, it was not fully clear to me why these were examples that supported only normativity *or* normality. More fundamentally, it seems to me that normality necessarily comes into play when normativity is defined in the way the authors do, e.g. as emerging in and through operational practices of design and development. Perhaps this is what the authors have in mind when they refer to the “interaction” between normativity and normality, but this should have been clarified more: are they two sides of the same coin? Do they

have a different temporal register? In this context, it is not so helpful that the authors fall back on a more formal, legal or deliberative understanding of normativity when defining normality (in opposition to normativity it seems). Again, I assume that this is not the point they would like to make, but the fact that they distinguish between normativity and normality in different sections and provide different examples to support each aspect of norm development raises these questions.

Second, in addition to the interaction between normativity and normality, I would like to know more about the interaction between the formal and deliberative processes of norm development and the practical and operational context. How do the authors assess the relevance of those formal deliberations that are still happening (they even seem to proliferate), albeit in a very slow manner? Can some of the 'back stage' discussions and deliberations at the formal international governance stage not be part of the production of normativity/normality as the authors describe it? Where to draw the line? Or is there something very specific about the technical/operational context of design and development?

Third, and related, I would like to push the authors to be explicit about the political stakes involved in the production of normativity/normality within an operational context of design and development. The authors mention the role of new tech companies and the ways in which through their platforms and models, specific values and value-based decisions become inscribed into the technology, but I would like to know how to square this against other (public) actors involved in the development and deployment of AI technology. *Whose* values become inscribed, especially within a military context?

Fourth, and finally, I would like to encourage the authors to be more outspoken about their own into their own normative position or preferred ways forward. Where and how should norm development in relation to military AI take place? Are deliberative (democratic) processes leading to more formal rules and principles still the most ideal outcome, even if these are inevitably slow and potentially outpaced by the development of the technology? Or does the essay in the end call for more political engagement with designers and the design process. Where should our political engagement lie?

Is the topic of the essay discussed accurately in the context of the current literature?
Yes

Is the work clearly and cogently presented?
Partly

Is the argument persuasive and supported by appropriate evidence?
Partly

Does the essay contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Critical security studies, Science & Technology Studies, military AI and modern

warfare

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 01 Feb 2024

Ingvild BODE

We want to start by expressing our appreciation for this constructive and insightful review. We have enjoyed reviewing our article again in light of your comments and engaging with them has considerably strengthened this revised version of our article. We have taken all comments on board which has added further depth to our argument. In particular, we now reflect much more explicitly on the dynamic interplay between normativity and normality. We completely agree with your suggestion to characterise these dimensions as inseparable in practice and as two sides of the same coin. In response to your comments about the precise relationship between public-deliberative processes and operational practices in shaping normativity/normality we have added an additional section at the end of the essay entitled ”future theorising on AI and norms”. This covers both this point, as well as the important questions you raised about actors beyond states and where our political engagement should lie.

Competing Interests: No competing interests were disclosed.