# JF21 – NC – Contracts (2:33)

#### Permissibility negates:

#### [1] Semantics – Ought is defined as expressing obligation[[1]](#footnote-1) which means absent a proactive obligation you vote neg since there’s a trichotomy between prohibition, obligation, and permissibility and proving one disproves the other two. Semantics outweighs – A. it’s key to predictability since we prep based on the wording of the res B. It’s constitutive to the rules of debate since the judge is obligated to vote on the resolutional text.

#### [2] Safety – It’s ethically safer to presume the squo since we know what the squo is but we can’t know whether the aff will be good or not if ethics are incoherent.

#### [3] Logic – Propositions require positive justification before being accepted, otherwise one would be forced to accept the validity of logically contradictory propositions regarding subjects one knows nothing about, i.e if one knew nothing about P one would have to presume that both the “P” and “~P” are true.

#### Moral internalism is true:

#### [1] Disagreement – Externalist theories fail to explain why some agents have the differing motivation for actions – internalism solves by showing how agents’ motivations are dictated by internal desires. Markovitz

[Markovits 14, Markovits, Julia. Moral reason. https://philpapers.org/rec/ROCJMM Oxford University Press, 2014.//Scopa] SHS ZS

Relatedly, internalism about reasons seems less presumptive than externalism. **We should not assume** that **some of us have** special **epistemic access to what matters**, **especially in the absence of any criterion for making such a judgment**. **It’s better to start from the assumption**, as internalism does, **that everyone’s ends are equally worthy of pursuit** – **and correct this assumption** only **by appealing to standards that are** as **uncontroversial** as possible. **According to externalism** about reasons, **what matters normatively** – that is, what we have reason to do or pursue or protect or respect or promote – **does not depend in** any fundamental way on **what** in fact **matters to us** – that is, what we do do and pursue and protect and respect and promote. **Some of us happen to be motivated by what actually matters**, **and some** of us **are “wrongly” motivated**. **But externalists** can **offer no explanation for this supposed difference** in how well we respond to reasons – **no explanation of why some of us have the right motivations and some of us the wrong ones** – **that doesn’t** itself **appeal to the views about what matters** that they’re trying to justify. (They can explain why some people have the right motivations by saying, e.g., that they’re good people, but that assumes the truth of the normative views that are at issue.22) **A comparison to the epistemic case** helps **bring out what is unsatisfactory** in the externalist position. **We sometimes attribute greater epistemic powers to some people than** to **others** **despite not being able to explain why they’re more likely to be right** in their beliefs about a certain topic. **Chicken-sexing is a popular example** of this among philosophers. **We think some people are more likely to form true beliefs about the sex of chickens than others even though we can’t explain why they are better at judging the sex of chickens.** But in the case of chicken-sexing, **we have independent means of determining the truth, and so we have independent verification that chicken-sexers usually get things right**. **Externalism seems to tell[s] us that some of us are better reasons- sensors than others**, but **without providing the independent means of determining** which of us are in fact more reliably motivated by genuine normative reasons (or even that some of us are).

#### [2] Regress – a priori knowledge is merely an acceptance of an individual’s conception of rationality. Macintyre 81.

[Macintyre 81, Alasdair Macintyre, https://undpress.nd.edu/9780268035044/after-virtue/ After Virtue, 1981] SHS ZS

The most influential account of moral reasoning that emerged in response to this critique of emotivism was one according to which an agent can only justify a particular judgment by referring to some universal rule from which it may be logically derived, and can only justify that rule in turn by deriving it from some more general rule or principle; but on this view [**S]ince every chain of reasoning must be finite**, such **a process of justificatory reasoning must always terminate with the assertion of some rule or principle for which no further reason can be given.** ‘Thus a complete justification of a decision would consist of a complete account of its effects together with a complete account of the principles which it observed, and the effect of observing those principles. **If** [I] **the enquirer still goes on ask ing** ‘But why should I live like that?’ then **there is no further answer to give** him, because we have already, ex hypothesi, [we have already] said everything that could be included in the further answer.’ (Hare 1952, p. 69). **The terminus of justification is thus always**, on this view, a not further to be justified choice, **a choice unguided by criteria.** **Each individual implicitly or explicitly has to adopt his or her own first principles on the basis of such a choice.** The utterance of any universal principle is in the end an expression of the preferences of an individual will and for that will its principles have and can have only such authority as it chooses to confer upon them by adopting them.

#### [3] Empirically proven – the competition between competing externalists modes of ethics has been going for centuries. Leiter

[Leiter, Brian. “Moral Psychology with Nietzsche.” Oxford University Press. Published 2019] SHS ZS

With respect to very particularized moral disagreements — e.g., about questions of economic or social policy — which often trade on obvious factual ignorance or disagreement about complicated empirical questions, this seems a plausible retort. But **for over two hundred years**, **Kantians and utilitarians have** [developed] **been developing** increasingly systematic **versions of their respective positions**. The Aristotelian tradition in moral philosophy has an even longer history. **Utilitarians** [They] **have become** particularly **adept at explaining how they can accommodate** [**others**] Kantian and Aristotelian intuitions about particular cases and issues, **though** in ways that are usually found to be systematically unpersuasive to the competing traditions and which, in any case, **do nothing to dissolve the disagreement** about the underlying moral criteria and categories. Philosophers in each tradition increasingly talk only to each other, without even trying to convince those in the other traditions. And **while there may well be ‘progress’ within traditions** — e.g., most utilitarians regard Mill as an improvement on Bentham—**there does not appear to be any progress** [towards] **in moral theory**, in the sense of a consensus that particular fundamental theories of right action and the good life are deemed better than their predecessors. What we find now are simply the competing traditions — Kantian, Humean, Millian, Aristotelian, Thomist, perhaps now even Nietzschean — who often view their competitors as unintelligible or morally obtuse, but don’t have any actual arguments against the foundational principles of their competitors. **There is**, in short, **no sign** — I can think of none — **that we are heading towards any epistemic rapprochement** between these competing moral traditions. Are we really to believe that hyper-rational and reflective moral philosophers, whose lives, in most cases, are devoted to systematic reflection on philosophical questions, many of whom (historically) were independently wealthy (or indifferent to material success) and so immune to crass considerations of livelihood and material self-interest, and most of whom, in the modern era, spend professional careers refining their positions, and have been doing so as a professional class in university settings for well over a century — are we really supposed to believe that they have reached no substantial agreement on any foundational moral principle because of ignorance, irrationality, or partiality

#### [4] Motivation – A. Externalist ethics collapse to internalism because agents will only follow external demands if they are consistent with their internal account of the good. For instance, citizens only follow the law insofar as its consistent with their internal beliefs, even when external value structures are being placed upon them. B. Empirics – there is no factual account of the good since each agent has unique motivation and there is no way to combine these beliefs into a unified ethic.

#### Next, agents justify their actions based on individual moral preferences and deal with ethical dilemmas by prioritizing their own beliefs. Gauthier.

[David Gauthier, Canadian-American philosopher best known for his neo-Hobbesian social contract theory of morality, Why Contractarianism?, 1998, ///AHS PB] SHS ZS

Fortunately, I do not have to defend normative foundationalism. One problem with accepting moral justification as part of our ongoing practice is that, as I have suggested, **we no longer accept the world view on which it depends**. But perhaps a more immediately pressing problem is that we have, ready to hand, an alternative mode for justifying our choices and actions. In its more austere and, in my view, more defensible form, this is to show that **choices and actions maximize** **the agent ’s expected utility**, **where utility is a measure of considered preference.** In its less austere version, this is to show that **choices and actions satisfy, not a subjectively defined requirement such as utility, but meet the agent ’ s objective interests.** Since I do not believe that we have objective interests, I shall ignore this latter. But it will not matter. For the idea is clear; **we have a mode of justification that does not require the introduction of moral considerations.** 11 Let me call this alternative nonmoral mode of justification, neutrally, **deliberative justification**. Now moral and deliberative justification are directed at the same objects – **our choices and actions. What if they conflict?** And what do we say to the person who offers a deliberative justification of his choices and actions and refuses to offer any other? We can say, of course, that his behavior lacks moral justification, but this seems to lack any hold, unless he chooses to enter the moral framework. And such entry, he may insist, lacks any deliberative justification, at least for him. **If morality perishes, the justificatory enterprise**, in relation to choice and action, **does not perish with it**. Rather, one mode of justification perishes, a mode that, it may seem, now hangs unsupported. But not only unsupported, for it is difficult to deny that **deliberative justification** is more clearly basic, that it **cannot be avoided insofar as we are rational agents**, so that if moral justification conflicts with it, morality seems not only unsupported but opposed by what is rationally more fundamental. **Deliberative justification relates to our deep sense of self**. What distinguishes human beings from other animals, and **provides the basis for rationality**, is the capacity **for semantic representation**. You can, as your dog on the whole cannot, represent a state of affairs to yourself, and consider in particular whether or not it is the case, and whether or not you would want it to be the case. **You can represent to yourself the contents of your beliefs, and your desires or preferences**. But in representing them, you bring them into relation with one another. You represent to yourself that the Blue Jays will win the World Series, and that a National League team will win the World Series, and that the Blue Jays are not a National League team. And in recognizing a conflict among those beliefs, you find  rationality thrust upon you. Note that the first two beliefs could be replaced by preferences, with the same effect. Since in representing our preferences we become aware of conflict among them, the step from representation to choice becomes complicated. **We must, somehow, bring our conflicting desires and preferences into some sort of coherence. And there is only one plausible candidate for a principle of coherence – a maximizing principle**. We order our preferences, in relation to decision and action, so that we may choose in a way that maximizes our expectation of preference fulfillment. **And in so doing, we show ourselves to be rational agents, engaged in deliberation and deliberative justification**. There is simply nothing else for practical rationality to be. The foundational crisis of morality thus cannot be avoided by pointing to the existence of a practice of justification within the moral framework, and denying that any extramoral foundation is relevant. For an extramoral mode of justification is already present, existing not side by side with moral justification, but in a manner tied to the way in which we unify our beliefs and preferences and so acquire our deep sense of self. **We need not suppose that this deliberative justification is itself to be understood foundationally. All that we need suppose is that moral justification does not plausibly survive conflict with it.**

#### Thus, the standard is consistency with contractarianism. Agents must engage in the project of mutual self-restraint as to not impede upon the moral authority of others. Stanford.

[Stanford Encyclopedia of Philosophy. “Contractarianism.” <https://plato.stanford.edu/entries/contractarianism/> Published 18 June 2000] SHS ZS

A brief sketch of the most complete and influential contemporary contractarian theory, David Gauthier’s, is in order. **Gauthier’s project** in Morals By Agreement **is to employ a contractarian approach to grounding morality in rationality** in order **to defeat the moral skeptic.** (However, Anita Superson (2009) points out that Gauthier attempts to answer only the skeptic who asks “why should I be moral?” but leaves both the motive skeptic, who argues that it is enough to act morally but need not be motivated by morality, and the amoralist, who denies that there is any such thing as morality, that is, that there are true moral statements.) **It is** generally **assumed that humans can have no perfect natural harmony of interests** (otherwise morality would be largely superfluous), and that there is much for each individual to gain through cooperation. However, **moral constraint on the pursuit of individual self-interest is required because cooperative activities almost inevitably lead to a prisoner’s dilemma**: a situation in which the best individual outcomes can be had by those who cheat on the agreement while the others keep their part of the bargain. This leads to the socially and individually sub-optimal outcome wherein each can expect to be cheated by the other. But by disposing themselves to act according to the requirements of morality whenever others are also so disposed, they can gain each others’ trust and cooperate successfully. **The contractarian element of the theory comes in the derivation of the moral norms. The compliance problem—the problem of justifying rational compliance with the norms that have been accepted—must drive the justification of the initial situation and the conduct of the contracting situation**. **It is helpful to think of the contract situation as a bargain, in which each party is trying to negotiate the moral rules that will allow them to realize optimal utility**, and this has led philosophers to apply a number of bargaining solutions to the initial contract situation. Gauthier’s solution is the “minimax relative concession” (1986, ch. V). **The idea of minimax relative concession is that each bargainer will be most concerned with the concessions that she makes from her ideal outcome relative to the concessions that others make**. If she sees her concessions as reasonable relative to the others, considering that she wants to ensure as much for herself as she can while securing agreement (and thereby avoiding the zero-point: no share of the cooperative surplus) and subsequent compliance from the others, then she will agree to it. What would then be the reasonable outcome**? The reasonable outcome, according to this view, is the outcome that minimizes the maximum relative concessions of each party to the bargain** (Gauthier 1986, ch. V). Equally important to the solution as the procedure is the starting point from which the parties begin. For some contractarians (like Gauthier) there is no veil of ignorance—each party to the contract is fully informed of their personal attributes and holdings. However, without the veil of ignorance, contractors will be aware of the differences in bargaining power that could potentially affect the outcome of the bargain. **It is important, then, that the initial position must have been arrived at non-coercively if compliance to the agreement is to be secured.** A form of the “Lockean proviso” (modeled after Locke’s description of the initial situation of his social contract): that one cannot have bettered himself by worsening others, may turn out to be beneficial in cases without a veil of ignorance. In sum, **the moral norms that rational contractors will adopt** (and comply with) **are those norms that would be reached by the contractors beginning from a position each has attained through her own actions which have not worsened anyone else,** and adopting as their principle for agreement the rule of minimax relative concession (Gauthier 1986, ch. VII). On one line of thought, contractarianism produces liberal individuals who seem well suited to join the kind of society that Rawls envisioned (Gauthier 1986, ch. XI). On another line, the Hobbesian contractarian argument leads towards the sparse government of libertarianism (Narveson 1988). The controversy here turns on the primary motivation for individuals to make agreements and cooperate. As we said before, there are two such motivations for the Hobbesian contractarian: fear of the depredations of others and benefits from cooperation with others. Libertarianism results when the first of these is primary, whereas when the second is primary, the kind of reciprocity and supportive government that will be discussed in the final section becomes possible.

#### Prefer additionally:

#### [1] Actor specificity – states are not moral entities but derive authority from the contracts that allows them to constrain action. This outweighs on empiricism; states aren’t bound by moral obligations, but they are by their contracts to other entities.

#### [2] Collapses – Contracts takes into account all other ethical theories and allows agents to engage under the index of their own good so long as they don’t violate the constraints of their other. The NC functions as a meta constraint – meaning indicts don’t take it out but they rather prove the truth of a theory under a particular index.

#### [3] Culpability – Only contracts ensure agents are held to their agreements since there is a verifiable basis for judging their actions as wrong as well as a pre-established punishment for breaking it.

#### Negate:

#### [1] LAWs are an intrinsic good as they cannot deviate from their agreed upon programming, which means they are always consistent with their contracts from the programming and international law. Khurana 18, Ryan. Founder and Executive Director of the Institute for Advancing Prosperity, In Defense Of Autonomous Weapons, October 14, 2018, <https://nationalinterest.org/feature/defense-autonomous-weapons-33201//BA> PB

LAWS are artificially intelligent military technologies that are able to execute a decision to kill a combatant without human input. These weapons systems do not exist as of yet, though artificial intelligence and advanced robotics technologies have already proliferated on the battlefield. The most aggressive push in the use of unmanned technologies in warfare has been by the Russian military, which has been using minesweeping drones—such as the [Uran-6](https://sputniknews.com/russia/201601171033284874-russia-robots-systems-demining/) in Syria—to limit casualties since 2016. Earlier this year it was confirmed that the Russians had deployed a fully unmanned ground vehicle, the [Uran-9](https://taskandpurpose.com/russia-uran-9-robot-tank-syria/), into the conflict zone, equipped with anti-tank missiles, an automatic cannon, and a machine gun turret. While the unmanned vehicle is remote controlled and therefore does not have lethal autonomous capabilities, Russian arms manufacturer Kalashnikov announced their intent to release **“**[**autonomous combat drones**](https://www.defenseone.com/ideas/2018/04/russia-races-forward-ai-development/147178/)**” that can make lethal decisions without human input**. As these technologies continue to develop, a commitment by nations more willing to enforce ethical standards in war to not develop LAWS is an abdication of responsibility because it allows more aggressive nations to set the standards of their development and deployment. Artificial intelligence researchers in western liberal democracies have an obligation to aid in the development of lethal autonomous weapons as a means to ensure that they are tools that are designed well. **AI, in its current form, is fundamentally an** [**optimization technology**](https://towardsdatascience.com/artificial-intelligence-is-for-optimization-human-intelligence-is-for-innovation-f0bddce2ed79)**, meaning that the protocols it internalizes and the goals set for it reflect the values of its operation**. If those who are committed to higher ethical standards decide not to contribute to the development of LAWS in a way that reflects their values, the technology will be developed anyway but by less scrupulous members of the research community. Therefore, **it is vital to emphasize the means by which LAWS can be used to make warfare more humane**, and focus on regulating its use rather than calling for a blanket ban. In designing autonomous weapons to reflect humane considerations, nobody has contributed more than roboticist Ronald Arkin. Arkin has been working with the U.S. Department of Defense since 2006 to develop ethically-guided military robotics and has [written widely](https://www.cc.gatech.edu/ai/robot-lab/online-publications/aisbq-137.pdf) on the means by which **they could** [**improve battlefield conduct**](https://www.cc.gatech.edu/ai/robot-lab/online-publications/308_V5812_VP_PCPREVArkinREV.pdf). Much of his argument relies on the simple fact that **humans are not good rule followers, while robots are. In combat, humans tend to make many mistakes, be motivated by anxiety or anger, and as a result of the strong bonds between military personnel, tend to cover for each other in the case of illegal activity. LAWS suffer from none of these issues**. Those who critique the possibility of ethical robots in war tend to focus on technology’s lack of empathy that allows for immoral conduct, but neglect the ways **human emotion amplifies, rather than minimizes, error in conflict**. Advances in artificial intelligence have improved [object detection](https://aiindex.org/) and [navigation](https://www.fastcompany.com/40572522/googles-ai-is-learning-to-navigate-like-humans), allowing for lethal autonomous weapons to limit collateral damage. For example, **they can be programmed to** ensure that there are no women or children in an area before carrying out an operation, and will **cancel their course of action if civilian casualties are likely. LAWS can better internalize the rules of war, and they can be placed alongside human combatants to limit ethical abuses or more accurately report them.** Restricting the autonomy of advanced weapons systems would allow them to be subject more to the unethical conduct of individuals, rather than the ideal behavior a robotic agent can be developed to reflect.

[2] States would not consent to a ban on autonomous weapons since they are consistent with their rational self-interest. **Gayle 19**

But the UK is among a group of **states** – including Australia, Israel, Russia and the US – **speak**ing forcefully **against legal regulation**. As discussions operate on a consensus basis, their objections are preventing any progress on regulation.  The talks come as the UK military is ploughing tens **of** millions of pounds into **autonomous weapons**, most recently announcing on Thursday a £2.5m project for “drone swarms” controlled with the help of next-generation autonomy, machine learning, and AI.  The talks in Geneva are taking place under the convention on certain conventional weapons. First enacted in 1983, the convention is intended to restrict the use of weapons “that are considered to cause unnecessary or unjustifiable suffering to combatants or to affect civilians indiscriminately”. It already covers landmines, booby traps, incendiary weapons, blinding laser weapons and clearance of explosive remnants of war.  “We urgently need a ban on killer robots,” said Ben Donaldson, head of campaigns at the United Nations Association – UK. “The majority of states get it. A rapidly growing proportion of the tech community get it. Civil society gets it. But a handful of countries including the UK are blocking progress at the UN. The UK needs to listen to this growing coalition and join calls for a preemptive ban.”  Responding to the criticism, a Ministry of Defence spokesperson said: “The United Kingdom does not possess fully autonomous weapon systems and has no intention of developing them. We believe **a preemptive ban is premature** as there is still no international agreement on the characteristics of lethal autonomous weapons systems.”  The issue of human control is at the heart of discussions about killer robots, according to the British military, and its negotiators have sought to focus debates at the UN on building consensus on what that means. Britain’s negotiating team says that no UK offensive weapons systems will be capable of attacking targets without human control and input.  They are arguing against **a** preemptive **ban** on the basis that it **could jeopardise their ability to exploit** any potential **military advantages** they could gain by imbuing weapons **with AI.**

Gayle, Damien. “UK, US and Russia among Those Opposing Killer Robot Ban.” *The Guardian*, Guardian News and Media, 29 Mar. 2019, [www.theguardian.com/science/2019/mar/29/uk-us-russia-opposing-killer-robot-ban-un-ai](http://www.theguardian.com/science/2019/mar/29/uk-us-russia-opposing-killer-robot-ban-un-ai). //CF

#### [3] Moral obligations of states towards other states or citizens cannot exist because international politics is the state of war. Singh, [N. N. Singh (Faculty of Law, National University of Singapore). “THE ABSENCE OF A SOVEREIGN LEGISLATURE AND ITS CONSEQUENCES FOR INTERNATIONAL LAW.” December 1970.] MKThere being no international legislature for international society some writers have refused to concede the nature of true laws to the rules of international law.1 The 'Austinian' objection that law cannot emanate from a non-sovereign body clearly illustrates this view.2 John Austin had defined international law as a system consisting of rules of 'positive morality'.3 In his opinion : If the same system of international law were adopted and fairly enforced by every nation, the system would answer the end of law, but, for want of a common superior, could not be called so with propriety. If courts common to all nations administered a common system of international law, this system, though eminently effective would still, for the same reason, be a moral system.4 Similarly W.E. Heavn had declared that 'Law cannot be predicated of mere customs which are not even true commands much less the commands of any competent State'.5 Following John Austin's thesis, Holland had called international law a 'law by courtesy' or 'law by analogy'.6 These jurists and their followers clearly refused to accept anything except the acts of a sovereign legislature body or the commands of a sovereign as true laws.7 They had, thus, not only set up a very narrow definition of law, but thereby also expressed reluctance to recognise the historical evolution, or the changing character, of human institutions. Further, such an attitude had not only obscured the organic relationship of law to society8 but also the vital part which custom, as opposed to legislation, has always played in the legal systems of national states.9 However, the Austinian concept of law so far as it inde with commands of a sovereign seems to have been complet by modern writers; even from among the few modern writers who have denied the character of 'law' to international not one seems to have subscribed to it, Thus, the denial of the legal character of the rules of international law simply on the ground that they do not represent 'commands' of a sovereign or because there is no sovereign international legislature in existence is clearly a thing of the past. The few modern writers who have denied the legal character of the rules of international law10 have done so mainly by taking refuge in the smug opinion that the rules of international law, even though substantially observed, are not 'legal' because they cannot be effectively enforced against States if the States themselves do not submit to them. This, it must be submitted, brings us to a completely different problem, i.e., the problem concerning the 'enforcement' of the rules of interna tional law11 which is quite different from the question of its 'legal character '. The legal validity of a rule, based on the conviction that it is binding, is quite distinct from its enforceability: law is not law because it is enforced, on the contra 2. The question of the 'source' of obligation in international law Doubts about the character of international law as true law can, however, only be dispelled by showing that obligations upon States exist in international relations which are very similar to the normative obligations that exist in any system of law. To show this, however, involves an examination of the source of such obligations, for it is not possible, given the absence of an international parliament, to rely on the formal source of a sovereign legislative body. It is to an examination of the theories of the source of obligation in international law that we must, therefore, now turn. But as the term 'source' has been used by different writers implying different meanings it is perhaps necessary to indicate here the meaning attributed to this term; it has been used in the sense of 'originating Cause' and not in that of 'evidence'. Also no attempt has been made to distinguish sharply between the terms 'cause' and 'basis'. The question of the rise of obligations is certainly not the same as the question concerning the validity or the binding character of obligations after they hâve been established. It is only in connection with the latter question that we can use the term 'basis of obligation' and to that extent we can distinguish between 'basis' and the 'originating cause or causes'. But this must not lead us to believe that there is no correlation between the two. On the other hand, there is a vital connection between the originating causes, which explain the rise of obligations, and the basis of those obligations, in the sense that che 'originating causes' to a very large extent also provide the very basi for the continuing validity, application and the binding character of the obligations which they are instrumental in creating. Much depends also on how we formulate our question. The question that needs to be answered is in Ehrlich's words: 'Whence comes the rule of law, and who [in essence what] breathes life and efficacy into it (a) The 'will' of the State — the Continental approach A strong reaction against the Austinian concept of law was inevitable. On the Continent, Bergbohm had been one of the earliest writers io suggest that the absence of an international legislature should only lead us to the conclusion that a particular source of law does not exist within the society of States and, further, that it should not lead us to deny the 'legal character' of the rules of international law.13 To Bergbohm, the expression of State will, as evidenced in interna agreements, constituted a proper source of law.14 Similarly Jellinek traced the source of obligation in international law to the will of the State. First, he pointed out that States are not above law and can be bound by their own will — through a process of auto-limitation.15 Secondly, he emphasized that the international and the municipal systems of law are different systems having different objectives, that is, whereas the the municipal system of law envisages the 'subordination' of the members of a community, the international society is run on the principle of 'co-ordination'. Thus, international law was according to him a law between co-ordinate entities and different from the law of the States which emphasized the element of command. However, by asserting that the States are legally free to disengage themselves from any such obligation which runs counter to their interests,16 he called into question the very 'legal' character of international obligations.17 The difficulty with Jellinek was that he could not deny the States, especially within the framework of his theory of 'auto-limitation', the right to disengage themselves from those obligations which they had themselves created. The answer to this difficulty was provided by Triepel's Vereinbarung stheome.18 Triepel pointed out that although the will of the State is a necessary element, the will of any single State alone is not sufficient for the creation of international obligatio that purpose he envisaged that a fusion of several wills leading to the creation of a 'common will' is necessary. He called this 'common will', as expressed in treaties and agreements, by the name of Vereinbarung. By making the Vereinbarung the source of international obligations Triepel had also succeeded in creating 'a legal power over States' : only Vereinbarung could undo what it had created. But one of the serious charges levelled against Triepel's theory is that it does not explain the existence of customary or general international law and reduces the rules of international law to a conglomeration of particular law.19 Cavaglieri, Anzilotti and Strupp attempted to answer this question — raised by the criticism of Triepel's theory — by asserting that those States which do not participate in the formulation of a particular law later on become bound by it through certain processes.20 Claiming that rules of international law are expressions of the will of the State as evidenced in agreements, these writers also attempted to establish a basis for the binding force of international agreements in the rule of pacta sunt servanda, which in Cavaglieri's opinion is a rule of customary international law whereas Anzilotti describes it as an original hypothesis and a postulate incapable of proof.21 But, despite these variations in the approach of writers to explain the genesis of the rule or the fact that recourse to the principle of pacta sunt servanda leads them to a tautology (international agreements are binding because they are binding)22, such doctrinal assertions at least indicated a movement towards interpreting law on a much more practical basis than hitherto employed.23

# 2NR

## Extensions

### Normative

#### Extend the standard of consistency with contractarianism. Moral internalism is true – this means that ethical theories that are reliant upon external standards of measuring goodness will always fail – this takes out the aff framework.

#### [Disagreement] Extend Markovitz – externalist theories fail to explain why agents have different motivations for action. The fact that we’re not all perfect reasoners in the squo proves this argument – we obviously have different motivations and desires from each other which means there cannot be a universal truth, otherwise all agents would be bound to it.

#### [Regress] Extend Macintrye – Every externalist theory fails the problem of regress as we can always question why we ought to follow that theory. That takes out a priori knowledge and deontic theories – as it will never terminate in a chain of questioning. Contractarianism solves – it isn’t a normative theory in the sense that it says there is one ethical good, but rather, it says we should try and maximize people’s own interpretations of the good. Questioning Contractarianism is a fallacy because the only thing we presuppose is that we should attempt to be ethical – which is unquestionable because that is the intrinsic desire. Independently, don’t evaluate new 2AR argument why I bite into this if they weren’t made in the 1AR.

#### [Emperics] Extend Leiter – philosophers have been attempting to figure out the meaning of the world for centuries, and the fact that we don’t have moral agreement and we aren’t moving closer to moral agreement proves that there cannot be one, true, ethical theory. Therefore, we must let agents engage under their own index of the good and not impede them from restricting others.

#### [Motivation] Extend the motivation analytic – it makes two claims. First, externalist ethics necessarily collapse to internalize because we only follow an ethical demand if they are already consistent with our prior beliefs. This is why people break the law — although there are ethical external principles being enforced down upon them, if they just aren’t motivated to follow the law because it is not consistent with their internal desires they will not follow it but instead engage under their own index of the good. Secondly, if all agents are engaging under their own index it’s impossible to build a unified ethic because people don’t understand each other and their respective positions absent a unified starting point.

#### Next, extend the Gauthier card – agents justify their actions based on their own individual moral preferences. In our day to day lives, every single person has their own different ethical calculus that justifies them to take certain actions and have different thresholds for proving an action moral or immoral. That proves there isn’t a unified understanding of an ethic and the NC is key.

#### That justifies following the standard of consistency with contractarianism. Extend the Stanford Card – since all agents engage under their own unique understanding of the good it’s worthless to try and find a unified ethic that everybody can follow. Instead, we must allow people to engage under their own indexes and simply prevent coercive relationships. We ought to allow each party to negate the moral rules that allow them to achieve their best utility through contracts, but once those rules have been established we cannot deviate from them.

#### Extend the offense – LAWs are an intrinsic good because they are preprogrammed to follow the contracts contained within their initial programming. That means they make the perfectly justified decisions since they cannot make any mistakes and deviate from the programming. Banning of LAWs would be a reduction of the contractual relations within society which would be an intrinsic wrong.

#### This outweighs – Extend the Walsh evidence which indicates LAWs are on track to take over human soldiers. That means that the future of warfare is going to be even better from a perspective of contractarianism and anything else would be bad under the NC framework.

### Skep

#### I’ll concede that contractarianism fails, but all of the reasons why moral internalism is true are INDEPENDENT PERMISSIBILITY TRIGGERS.

#### Extend that permissibility negates –

#### Under a truth testing paradigm, I just have to prove a 1% chance that morality doesn’t exist and that negates – as the affirmative you have the burden of proving the resolution with 100% strength of link.

#### ALSO, do not allow them new 2ar answers to these cards. They would be new and should have been made in the 1AR. None of the implications of my arguments are changing, because I am just conceding that contracts fail but moral internalism is true – but their theory is externalist. That means you have to reject any skep bad arguments in the 2ar because they would be new – you had the burden of responding in the 1ar and anything else would constitutive infinite violations of fairness against the 2NR.

#### To clarify, I AM NOT CHANGING THE IMPLICATIONS OF ANY OF MY ARGUMENTS. I am simply conceding that contracts fails and winning that moral internalism is true. This means the only legitimate 2AR on substance would have to beat back moral internalism proper – which is impossible because they handled it horribly in the 1AR.

#### [Disagreement] Extend Markovitz – externalist theories fail to explain why agents have different motivations for action. The fact that we’re not all perfect reasoners in the squo proves this argument – we obviously have different motivations and desires from each other which means there cannot be a universal truth, otherwise all agents would be bound to it. Since the only other ethical theory in this round is a universal truth that triggers skepticism – since externalist theories fail, and I’ve conceded that my interanlist theory also fails.

#### [Regress] Extend Macintrye – Every externalist theory fails the problem of regress as we can always question why we ought to follow that theory. That takes out a priori knowledge and deontic theories – as it will never terminate in a chain of questioning. That triggers skepticism – the affirmative is an attempt at applying a universal rule through their deontically justified framework but I’ve won the fact that they necessarily fail.

#### [Emperics] Extend Leiter – philosophers have been attempting to figure out the meaning of the world for centuries, and the fact that we don’t have moral agreement and we aren’t moving closer to moral agreement proves that there cannot be one, true, ethical theory. That means moral theories cannot exist – if they would we woulld have made more substantial moral progress but insofar as we haven’t come closer to discovering the truth of morality it means it can’t exist. It doesn’t matter if they say we should keep on trying – if anything, moral theories have DIVERGED as time has gone on which proves that the longer we try to achieve morality the more violence and disagreement it will lead to. IF you think it’s more probable, at this time, that there doesn’t exist a moral truth then vote negative irrespective of future potentialities.

## Framework Frontlines

### AT Util

#### [1] Utilitarianism collapses into contractarianism.

John J. **Thrasher**, Assistant Professor in the Philosophy Department and the Smith Institute for Political Economy and Philosophy at Chapman University, Reconciling Justice and Pleasure in Epicurean Contractarianism, Ethical Theory and Moral Practice, Vol. 16, No. 2 (April **2013**), pp. 423-436 ///AHS PB

**If** you do not, on every occasion, refer each of your actions to the goal of nature, but instead turn prematurely to some other [criterion] in avoiding or pursuing [things], your actions will not be consistent with your reasoning (KD 25). **This goal of reasoning and action is the absence of pain** and the tranquility that comes from living without fear (KD 3).4 This kind of pleasure, ataraxia, is unhindered tranquility, rather than a sensation of active pleasure.5 It is a psychological fact, according to Epicurus, that we do actually seek ataraxia and that our lives go best, from a subjective point of view, when we pursue ataraxia. It is the natural goal of beings like us. If fear of the gods, death, and pain constitute sickness of the soul, removing those ailments constitutes its health. This psycho logical hedonism creates the justification for the normative hedonism that practical reason ing should aim at ataraxia.6 The normative ideal of Epicurean practical rationality is a hedonistic form of instrumental rationality with the final end of ataraxia. In the parlance of modern decision theory, it is a maximizing theory of rationality. Given a set of ordered preferences, individuals chose rationally when they choose to act on their highest valued goals. To choose less pleasure rather than more pleasure when given the choice is paradig matically irrational and contrary to nature. Given this conception of practical rationality and virtue, it is hard to see how one can single-mindedly pursue pleasure and accept the constraints of justice. Traditionally, virtue ethical theories solve this problem by making the virtue of justice constitutive of happiness with deontic restraints built into the formal conditions of happiness.7 To use the Rawlsian terminology, the right flows naturally out of the good.8 This solution, however, will not work for the Epicurean. Unlike in Aristotelian or Stoic virtue theory, the standard of Epicurean happiness is not an objective, formal standard, but rather the subjective, psychological state of ataraxia. The Epicurean has a reason to (j> only if he or she believes that (J)-ing will reliably lead to the final end of ataraxia. If all reasons are instrumental in this sense, how is it possible for the Epicurean to have reason to constrain his or her pursuit of the goal of nature by the deontic demands of justice? To give a plausible account of justice, the Epicurean needs to explain how to justify the demands of justice as a means to the final end of ataraxia. One version of this problem arises in the context of friendship. Epicurus claims . .every friendship is worth choosing for its own sake, though it takes its origin from the benefits it confers on us" (VS 23). Given this statement about the value of friendship and KD 25, how can friendship be non-instrumentally valuable while also being beneficial because of the benefit it confers? Some have argued that genuine friendship is impossible unless we amend the basic egoistic element of Epicurean practical rationality.9 In contrast, Matt Evans argues that there are two basic approaches to understanding friendship in a consistently egoistic way (Evans 2004, 413). Friendship as "indirect egoism" involves incorporating the good of a friend or of friendship generally into one's own good. This is the interpretation that Timothy O'Keefe favors (O'Keefe 2001a). The alternative is Evans's preferred view, "direct egoism," that one's own good "stands or falls" with the good of one's friend (Evans 2004, 413). Indirect egoism is, for O'Keefe, a two-level hedonistic theoiy. Choice of desires is governed directly by hedonic concerns and those desires then pick out particular actions, which are only indirectly related to the original hedonic calculus (O'Keefe 2001a, 300-302). In contrast, Evans's direct egoism applies the hedonic calculus to action selection. Evans maintains that Epicureans can "reason their way to friendship" through direct egoistic means (Evans 2004, 423). What is true of friendship will likely be true of justice so it is imperative to determine whether the Epicurean hedonic calculus is meant to apply to actions (direct egoism), desires (indirect egoism), or something else entirely. The direct egoist interpretation has the benefit of being the easiest to reconcile with KD 25. The indirect egoist interpretation makes it easier to understand how the Epicurean can incorporate friendship and justice into hedonism. Another possibility, between direct and indirect egoism, is what Gregory Kavka calls "rule egoism" (Kavka 1986, chap. 9). Although Kavka developed his version of rule egoism in the context of understanding Hobbes's ethical theory, there are enough similarities between the two accounts for a plausible Epicurean version as well. The hedonic calculus applies directly to rules rather than to desires or action. Furthermore, rules can be generalizations over desires or actions, e. g. "don't cultivate a desire for riches" or "seek out friends." The first is a rule that indicates what desires will lead to pleasure whereas the second is a rule that indicates a particular set of actions that will likely lead to pleasure, namely having friends. **Rule egoism has several benefits over direct and indirect egoism. First, it is more general. Both actions and desires are mentioned throughout KD and VS as the possible object of choice. Rule egoism recognizes the importance of both actions and desires to the end of ataraxia and accounts for both in terms of rules. Second, rule egoism is simpler and likely more reliable than direct or indirect egoism. It is reasonable to expect that the typical Epicurean would be bewildered in the face of the multiplicity and complexity of choices that would face him or her on any given day. The stress of deliberating over actions on the direct egoist interpretation of KD 25 would often create anxiety rather than tranquility. Similarly, it is not clear that, given the complexity of the world, the direct approach would reliably lead to ataraxia. The indirect approach is not better on this count partly because desires do not necessarily pick out unique action in decision situations, partly because the indirect egoist faces the same problem as the direct egoist at the level of desires. By using rules, however, the Epicurean can rely on the knowledge embodied in the rules without having to deliberate in each case.** This explains the reason that Epicurus spends so much time in his writing listing rules and maxims. He gives rules about how to reduce sexual passion (VS 18), the irrationality of suicide (VS 38), the danger of envy (KS' 53), and the dangers of great wealth (VS 67). In all of these cases, and many more, Epicurus is passing on wisdom about how to reliably achieve ataraxia. He is playing the part, of a guide who has walked down life's tangled road and is reporting to those who have yet to see everything he has seen. These maxims or rules are the embodiment of the successful use of practical rationality in the past. Following these types of rules is, therefore, an application of direct egoism in an indirect way. Given the limited cognitive capacity and time of the Epicurean rational agent, relying on rules as a guide can be, following Gigerenzer and Goldstein, a "fast and frugal" way of reasoning based on heuristics communicated as rules or maxims (Gigerenzer and Goldstein 1996). **Instead of choosing over the expected outcome of individual acts, the rule egoist chooses sets of rules to follow based on the expected outcome of following that rule or set of rules** (Kavka 1986, 358-359). In the next section we will see how understanding Epicurean practical rationality as "rule-hedonism" makes it possible to reconcile Epicurean practical rationality with justice. 3 The Possibility of the Contract Once we understand Epicurean practical rationality as applying to rules rather than to particular actions or desires, we can see how the Epicurean can reconcile the imperatives of practical rationality with the demands of justice. **A particular social contract is a set of rules that regulates behavior in certain public settings.** The Epicurean agrees to a particular set of rules in order to more reliably achieve and maintain personal ataraxia. We might wonder, however, why the Epicurean would need a contract at all. Why wouldn't the first personal application of practical rationality be sufficient for ataraxia? Why is the social **contract** necessary? In a world of practically rational Epicureans, the social contract seems either otiose or harmful. Either the contract recommends what practical rationality would recommend or it conflicts with practical rationality. On its face, Epicurean contractarianism looks either unnecessary or impossible. I will argue here that the Epicurean social contract is both necessary and possible. **The social contract is necessary, as I will argue in the next section, for its coordinating, assuring, and specifying functions**. The social contract is possible because of the role that rules can play in Epicurean practical rationality. In this section I will argue that the Epicurean social contract is consistent with Epicurean practical rationality and, hence, possible, while fulfilling an important social role. The Epicurean social contract is fundamentally instrumental; **it is a "pledge of reciprocal usefulness neither to harm one another nor be harmed**" (KD 35). To be consistent with Epicurean practical rationality, then, the contract must secure benefits that would not be possible without the contract. If, however, one only has reason to enter into a contract because of the benefits, what reason does one have to follow the contract when there are no benefits and only costs? This is the heart of the concern that the Epicurean cannot be a good citizen. If citizenship involves the possibility of sacrifice, why should we expect the Epicurean to comply? Here again, we see the same kind of problem that we saw in §2 concerning friendship; the solution is also similar.

#### [2] Pleasure and pain are only motivational to the individual who senses them, which means only a system of mutual self-restraint can enter agents into binding agreements to respect each-others’ pleasure and pain.

#### [3] Nagel proves individuals experience pain and pleasure and strives towards those things, but what counts as pain and pleasure are subject to disagreement between different agents which requires a contract that binds their conceptions of the good together.

#### [4] Cross apply how externalist theories fail – util is definitionally externalist as it applies some external metric, i.e. maximizing pleasure, and asserts that is the unconditional good for all of humanity – but different experiences cause different agents pleasure

## Offense Frontlines

### AT LAWs are Coercive

#### [1] The turn is consequentialist – it assumes that LAWs will be used in a coercive manner, but it’s not about the intrinsic nature of a LAW itself. States can use weapons in any way they choose – and the existence of a State owning a certain type of weapon says nothing about how it will be perceived by the rest of the world – if a benevolent State never imposes demands upon others, LAWs would not be coercive since the State itself isn’t coercive, which proves that this coercion isn’t intrinsic to the nature of LAWs but only a consequence of their existence.

#### [2] No link – Contracts need to be formed on a level playing field, which means even if LAWs were used to coerce, the contracts created wouldn’t be valid since the two states weren’t on an equal bargaining level – which proves the turn not relevant under the framework.

#### [3] The following of contracts outweighs their formation. A. Constitutivism – the purpose of following contracts is to put impositions on the moral behaviors of other citizens, so they do not violate them B. Bindingness – it’s the most verifiable metric to see if somebody has violated a contract or not, and it’s impossible to verify whether the starting positions were made in a coercive manner or not

#### [4] Turn – LAWs provide weaker states with far more bargaining power and recourse. Wilner

[Wilner, Alex. “Artificial Intelligence and Deterrence: Science, Theory and Practice.” NATO Otan] SHS ZS

The academic literature on AI and deterrence is exceptionally slim. With few exceptions, no study has yet to unpack the various ways in which the technology might intersect with deterrence logic, theory, and practice writ large. There is, however, a budding and related grey literature, mostly published by think tanks over the last 18 months, that explores AI and strategic studies, international security, global competition, and warfare. Lessons derived from this literature is useful for thinking about AI deterrence more specifically.32 What follows is a summary of several key studies. In September 2018, the Washington-based Brookings Institution began publishing a series, A Blueprint for the Future of AI, that explores the effects AI might have on society, including in health care provision, education, governance, business, and security. The series builds off an earlier, April 2018 piece co-published by Brookings’ President, John Allen, in which he and Darrel West explore the many ways in which AI and society will intersect. In security, West and Allen argue that AI will shorten the distance from intelligence gathering to decision-making. Under conditions of “hyperwar”, they argue, massive amounts of data will be “sifted in near real time – if not eventually in real time,” providing decision-makers with greater intelligence awareness and more options far more quickly. Under some conditions, they continue, perhaps especially in cybersecurity, AI might be able to make better sense of malicious code more quickly, alerting human handlers of incoming threats, and if empowered to, responding autonomously. And on the question of keeping the kill chain within human hands, West and Allen warn that **US adversaries**, like Russia, China, and North Korea, **are “not** nearly so **mired in this debate**.”33 A moral divide exists. As for the Brookings’ series itself, three papers focus on AI and security.34 Mara Karlin’s contribution explores how **AI might alter** the “**strategic level of national security**,” **and influence** “**the** management, employment, and **development of military force**.” She argues that AI will augment the influence the American private sector will have on national security decision-making, given their current control over AI research, development, and use. Like West and Allan, Karlin hypothesizes that **by providing decision-makers with new** and alternative **options** based on a wide-ranging assessment of an unimaginably large trove of data, that **AI may** eventually convince decision-makers to delegate some tasks (including targeting) to machines under specific time-sensitive conditions, and to **re-evaluate existing military** narratives, plans, and **paradigms**. In another piece, Alina Polyakova focuses on AI as a tool of asymmetric war; she calls it “AI- driven asymmetric warfare,” or ADAW. With Russia in mind, Polyakova illustrates how **weaker adversaries might “co-opt existing commercially available” AI technology to challenge stronger states with AI-enhanced cyberattacks** and AI-generated disinformation or political influence campaigns. For instance, she suggests that AI-driven “deep fake” technology – which allows a user to swap one person’s face for another in video content – can produce fake but highly realistic and customized content that can be used to strategically shift narratives and perceptions, and ultimately, behavior. Finally, Michael O’Hanlon provides the third Brookings’ article, in which he links advancements in AI and robotics together to illustrate how **new** **tactics on the battlefield might require a rethink at the strategic level.** For illustration, he argues that in a future battlefield, thousands of miniature autonomous drones deployed both at sea and in the air might lead to new swarm or “saturation” tactics, a topic Scharre first broached in 2014.35 O’Hanlon concludes that **these tactics might put an end to “the kind of impunity that US forces have enjoyed for decades**” in Europe, Asia, and elsewhere. Much like Brookings, the Center for a New American Security (CNAS), a Washington-based, non-partisan organization founded in 2007, developed its AI and Global Security Initiative to explore the different ways in which AI and security intersect. The Initiative has produced a series of publications, most of them linked to Paul Scharre and Michael Horowitz, who help direct CNAS’s Task Force on Artificial Intelligence and National Security.36 None of the publications explore deterrence, but they do provide a detailed assessment of the various ways in which AI might affect security and military affairs. Several highlights are worth unpacking. In a CNAS reprint from Foreign Policy, Horowitz suggests that as a general-use technology largely based on software developments, **competition for AI will be broad**, uniquely combining the efforts of countries and corporations alike. In another reprint from Foreign Policy, Scharre explores the ramifications of letting machines dictate the speed of warfare, including the associated risk of inadvertently augmenting the result of minor algorithmic accidents. In other pieces, Scharre, Horowitz, and CNAS colleagues outline the state of the art, and explore narrow AI’s applicability across the various security, intelligence, and defence disciplines. They argue that AI will help classify data; detect anomalies in patterns of behavior; predict future behavior; improve low-skilled human performance; facilitate labor-intensive activities; automate tasks in both physical and cyberspace; uncover new cyber vulnerabilities, threats, and solutions; develop, distribute, and counter “targeted propaganda”; and provide super-human qualities in speed, precision, reliability, patience, and vigilance. Elsewhere, Horowitz, writing in the Texas National Security Review – a new, policy-oriented publication linked to War on the Rocks – categorizes AI as the “ultimate enabler”, an all-purpose “technology with a multitude of applications” rather than as a weapon in any traditional sense.37 It allows developers, innovators, and adopters to apply it across the security spectrum in unique ways, unlike other, less flexible technological developments, like the ballistic missile or machine gun, which proved useful and revolutionary but in a much more limited context. Like Scharre and O’Hanlan, Horowitz suggests that **low-cost**, autonomous **drones**, coordinating their actions at machine speed, **might undermine high-cost, high-quality legacy weapon systems**. Horowitz argues further that the way AI develops in the future will help dictate the utility and advantage it might lend to its developers and early adopters. If advancements are led by the private sector, for instance, **AI might** more **quickly “diffuse” to militaries around the world, who purchase and adopt it for their own use.** That would reduce the original developer’s “first-mover advantage,” **and could narrow the balance of power between innovators and adopters**. But if AI – or certain types of AI particularly useful to defence and security – is developed primarily by states, government labs, and their militaries, the technology will be slower to spread between countries, and innovators may retain a technological edge that might translate into a longer-lasted strategic advantage. These assertions are explored further by M. L. Cummings in a 2017 brief published with Chatham House, a British think tank. Cummings suggests that private sector innovation in AI currently has the advantage because top engineering talent find more lucrative careers in the commercial applications of AI than they do in the more narrowly-focused aerospace and defence industry. This is especially true in the US, Canada, and Europe. “The global defence industry,” she warns, “is falling behind its commercial counterparts in terms of technology innovation.”38 Bridging the gab may be difficult. The RAND Corporation, another leading US think tank with a long track record of publishing work on strategic studies, held several workshops in 2017 that explored future US security challenges, circa 2040, and published a number of subsequent reports in 2018. One of these short reports explores the way **AI might interfere with nuclear strategy** and strategic **deterrence** in particular.39 Several useful findings stand out. First, if AI creates the perception among nuclear states that one country has the ability to detect, locate, and target all of another state’s nuclear weapon launchers – an infeasibility today – then vulnerable states may be especially inclined to use these weapons more quickly at the risk of losing them altogether. Second, hacking an AI’s system by, for instance, “poisoning its training data” or degrading hardware, might render AI- generated strategic advice suspect, fallible, and maliciously manipulated by third parties. Third, autonomous nuclear weapons akin to aerial or underwater drones – a “doomsday drone” – could be used to further dissuade a challenger from launching a debilitating nuclear first strike by augmenting the credibility and survivability of new-age second strike weapons. And fourth, AI might be used in arms control, augmenting the trust, control, and transparency that underpins the counter-proliferation verification process. Beyond the limited academic and think tank literature, AI and national security has been extensively explored by the US government. Leading the charge, in October 2016, the Executive Office of the President of the United States published a broad strategy document on AI, Preparing for the Future of Artificial Intelligence.40 Of interest to us are its brief discussions dedicated to exploring how AI might be used in cybersecurity and military affairs. On the former, the document suggests **AI might create more robust, agile, and cheaper forms of cybersecurity**, able to more quickly detect and respond to malicious events. Citing DARPA’s 2016 Cyber Grand Challenge, which sought to develop and test autonomous systems for exploiting and patching digital flaws within an all-machine tournament, these systems could eventually “perform predictive analytics to anticipate cyberattacks.” On the latter, the document suggests that offensive AI weapons systems may have “greater precision”, and could introduce “safer, more humane military operations” that rest on limiting the risk to military personnel, diminishing the number and type of munitions used, and producing less collateral damage. And defensive **AI** weapons **could be used for** “protecting people and high-value fixed assets,” and possibly, “**deterring attacks** through non-lethal means.” Finally, beyond the battlefield, the document argues that AI could “provide significant benefits,” in logistics, maintenance, base operations, veterans and personnel affairs, emergency and medical response, navigation, communications, and intelligence, all of which would make “American forces safer and more effective.” Some of these themes are picked up by Daniel Hoadley and Nathan Lucas in their 2018 brief to the US Congressional Research Service, in which they provide a detailed overview of the challenges and opportunities AI introduces to combat environments.41 While their work does not mention deterrence at all, they do provide a useful classification or clustering of the military effects of AI that lends itself to an exploration of coercion. First, AI will provide military systems with autonomy, replacing humans in dangerous, risky, complex, and labor-intensive jobs. Second, AI greatly augments the speed with which actions can be taken: it will have the “ability to react at gigahertz speed,” and accomplish “long-duration tasks that exceed human endurance.” Relatedly, AI will make sense of a huge quantity of data from disparate sources. The result is that AI might provide a decision-maker with advice, suggested actions, or solutions that allow it to outpace an adversary’s own assessment of and ability to strategically react to a situation if left to human analysis alone. Third, AI might boost the productivity and capability of human soldiers and of less sophisticated military systems. Like various other authors, Hoadley and Lucas focus on drones. Alone, a single drone is no match for a fighter jet, but algorithmically lassoed together, a fleet of thousands might well overwhelm it. And fourth, AI might provide out-of-the-box and unpredictable strategic advice that stretches the human imagination. 5.0 AI DETERRENCE: A SPECULATIVE ACCOUNTING Deterrence has been around a long time.42 It has, often and repeatedly, proven its theoretical flexibility in responding to shifting international dynamics and to emerging (and disruptive) technologies. The literature on deterrence is, as a result, robust, cumulative, and expansive. This evolution has occurred within the context of “four waves” of scholarship.43 Briefly, the first wave began following the Second World War, and focused on the development of nuclear weapons within a US-dominated strategic backdrop. The second wave corresponds to the great power dynamics of the Cold War period itself, and to the proliferation of nuclear weapons. Game theory and behavioral models were used to better understand alliances and escalation, with a focus on strategic and conventional deterrence by punishment, and on mutually rational and unitary state actors. The third wave followed in the 1970s, with a focus on empirically testing the frameworks and theories previously proposed. Concepts like communication, commitment, resolve, and expected utility, along with central versus extended deterrence, were refined. And the psychology of coercion – cognition, bias, fatigue, misperception – along with notions of strategic culture, bureaucracy, and leadership also emerged. The fourth wave began at the end of the Cold War, with a shift in focus from bipolarity, strategic weapons, and great power rivalry to an eclectic (and even chaotic) diffusion of interests on “rogue” states, violent non- state actors – from pirates to terrorists – and to processes like radicalization, rebellion, and insurgency. While a resurgent Russia and a rising China have garnered renewed interest in traditional coercive dynamics, the scope of deterrence nonetheless broadened during the fourth wave, with reinvigorated theoretical research on the logic of denial, delegitimization, and conventional deterrence. And altogether novel concepts for deterrence in both space and cyberspace, like cross-domain deterrence, were also developed. This article, and the larger AI deterrence project upon which it is based, reflects perfectly the speculative nature of fourth wave scholarship. At its logical core, **deterrence entails using threats to manipulate an adversary’s behavior**. More concretely, it rests on convincing an adversary to forgo an unwanted action. Compellence – a related concept – flips the process around, prompting an adversary to pursue an action it might otherwise have avoided. Deterrence restrains behavior while compellence encourages it. Both are subsumed within the logic of coercion more broadly. In all cases of deterrence, compellence, and coercion, a minimum of two actors are involved. Defenders issue threats to alter an opponent’s behavior; challengers consider these threats and decide whether or not to acquiesce. In some cases, a third actor is involved. With extended deterrence, for instance, a proxy, partner, or ally is also included in the process, either on the side of the defender or the challenger. **Threats of** punishment – usually in the form of **retaliation** – and promises of denial – usually in the form of **depriving an adversary the expected benefits** of a particular behavior – form the basis for most deterrent and compellent engagements, though ideological considerations that inform behavior have also been recently developed within the context of deterrence by delegitimization. Deterrence weighs on a challenger’s cost-benefit calculus, on the strategic choice it has available to it. Militarily crushing an adversary, Thomas Schelling writes, such that it has no alternative but to accept demands made of it is not deterrence, but rather military defeat.44 The same logic holds for denial: Developing a foolproof defence, however rare, does not deter by denial but rather deters by defeat. In this case, adversaries do not penetrate a target, for instance, because they are prevented from doing so, not because they decide against doing so. In deterrence, changes in behavior are an option, not a necessity. Finally, though deterrence is usually anchored to International Relations theory, it also functions below the level of the state among non-state actors, individuals, groups, and other organizations. Criminological deterrence, for instance, and scholarship on deterring terrorism both provide insights on deterring violence within communities and emanating from violent non-state actors. In practice, deterrence rests on a number of pre-requisites. First, actors must retain a degree of rationality, such that a combination of threats will suffice to shift their behavior. Second, challengers and defenders must agree, to some point, that non-violence is preferable to violence: if conflict is the only preference both sides share, then deterrence will not work. Third, defenders must not only communicate (or otherwise signal) their threats and expected changes in behavior, but must also reassure challengers that if they acquiesce to a threat, punishments will not be meted out. Fourth, defenders must have the perceived capability, and the resolve, to punish or deny as threatened. And finally, deterrence works best again a known or suspected adversary; anonymity robs deterrence of a coercive target. Taken together, how might AI and deterrence intersect? A speculative accounting – a laundry list of thoughts – follows. First, better defence equals better denial. **AI, by improving the speed and accuracy** of some defensive weapons, and by subsequently improving the reliability of defending infrastructure, weapons platforms, and territory against certain kinetic attacks (e.g. missile or rocket attack), **might deter some types of behavior by altogether denying their utility**. The same logic holds in defending cyber platforms: by denying aggressors access to information more persistently, a defender’s AI might compel a challenger not to bother attacking in the first place. Second, and on the flip side, under other conditions AI may augment the feasibility of certain types of attack, favouring punishment strategies over denial strategies. As noted, autonomous swarming drones have garnered the greatest attention. If and when these capabilities are developed and refined, swarming bots may provide challengers with a unique coercive tool not easily deflected or defeated. Saturation tactics that rely on thousands of disposable and easily replaced robotic platforms may tip the balance towards offensive measures and the promise of punishment strategies. Conversely but relatedly, swarms might likewise be used by a defender to fend off an attack against it employing conventional weapons – think of a defensive robotic swarm launched to counter an incoming fighter jet – providing it with a tactical-level threat of denial. But then again, and historically speaking, offensive developments usually spur defensive developments in kind: Just as AI feeds offensive swarming tactics, so, too, might it eventually result in defensive swarming responses. The resulting robotic dog fight might recalibrate coercion towards the middle. Third, and moving beyond kinetics alone, AI might improve a state’s ability to plan and carryout both offensive and defensive coercive threats, by improving logistics, navigation, communications, recruitment, training, deployment, and so on. The back-office AI that coordinates the machinery of warfare may make some coercive threats more robust and persuasive as a result. Fourth, by rapidly providing unique and novel advice to decision-makers that supersedes human innovation and capability, **AI may provide situational awareness that dips into predictive analytics**. By melding an improved analysis of what adversaries have done in the past and are currently doing today – indeed this very minute – AI will provide users with the ability to anticipate an adversary’s next action. Defenders can preemptively respond accordingly.45 If, over time, a challenger believes that a defender can anticipate its next move, it may be deterred from trying, or might alternatively calculate that only brash, unexpected, and novel actions will help it accomplish what it wants (at least, until the machine learns from these episodes, too). Fifth, by manipulating public information through deep fakes and other related processes, AI might provide users with a new form of deterrence by delegitimization. The threat, in this case, is the ability to create, release, and disseminate fake video or audio material threatening or embarrassing to a target. Think of Russia – or of any other state or non-state actor, if only because the technology will be cheaply available – surreptitiously threatening a US Democratic Party presidential nominee with engineered content that could influence the candidate’s standing among the electorate in the leadup to a future US presidential election. Because determining the veracity of AI-manipulated content and attributing its source is difficult to do, countering these types of coercive misinformation campaigns may be challenging.46 Sixth, fighting at “machine speed” may change the calculus of taking action. If AI-based decision-making provides one side of a conflict an advantage in responding quickly and decisively, then others will eventually mimic and come to rely on these processes, too. But as both sides of a contest come to rely on machines for insights, the very rationale of these AI-generated insights may degrade more quickly over time, as one side’s AI responds and reacts to another’s. Put another way, in this scenario an AI-generated insight may have a short shelf life, and windows of opportunity may prove fleeting. If so, the logic and value of striking first, and fast, may prevail, upending long-standing coercive and escalatory calculations along the way. Seventh, **AI might provide traditionally weak actors with novel means to coerce traditionally strong actors**. The dual-use nature of AI along with private-sector developments in the technology, suggests that many states – including relatively **weak ones** – and non-state actors, organizations, and individuals as well, **may** eventually be able to **purchase the technology for their own use**. While weak actors may face other limitations, like acquiring access to useful training data, **AI might nonetheless help level the playing field with more powerful actors**. If so, diffusion of the technology may diminish how the strong deter or compel the weak, and might otherwise allow the weak with new avenues for coercing the strong. Eighth, ubiquitous real-time surveillance could deter criminal behavior. If a state where to establish AI- powered surveillance of urban centers, border crossings, and other sensitive locations to generate biometric identification and behavioural analytics, and if it were to publicly announce its use of these tools, it might convince criminals, terrorists, spies, and other nefarious actors that their plans are unlikely to succeed. China’s experiment in countering jaywalking at busy urban intersections is informative.47 Facial recognition cameras monitor roads, snapping pictures of jaywalkers, and matching the offender to photo IDs stored in a database. The photos, along with some of the individual’s personal information, can then be displayed on roadside screens, and fines can be issued automatically. In the city of Ji’Nan, police report that this technology has reduced jaywalking by 90 percent. Used at scale, the technology could curtail other unwanted behavior and activity. And finally, ethical and legal limitations on how AI is used in battle may dictate how some countries behave and others respond. While some states, notably the United States and several European allies, are openly against providing AI with the right or the means to kill individuals without human intervention – French President Emanuel Macron explained, for instance, while promoting his country’s new AI innovation strategy in 2018, that he was “dead against” the idea48 – other countries appear far less concerned. China, Russia, Israel, and others, for example, may be more willing to delegate decisions – including those that result in human death – to Artificial Intelligence. Under certain conditions, doing so may provide these countries with a tactical, strategic, or coercive advantage over those inclined to keep humans in or on the loop. It may likewise provide these countries with a means to counter-coerce, influence, or otherwise manipulate countries that are more constrained and refrained in the way they use their AI in battle.

#### [5] LAWs are key to check US coercion because they are behind in their development which diminishes their place as world enforcer. Barlett 20, Barlett, Matt. "The AI Arms Race In 2020". *Towards Data Science*, 2020, <https://towardsdatascience.com/the-ai-arms-race-in-2020-e7f049cb69ac>. //Scopa On the otherwise celebratory occasion of the United Nation’s 75th anniversary in January this year, Antonio Guterres — the UN’s Secretary-General — gave a grim address best summarised by his description of the world as “off-track”. On Guterres’ list of existential threats were the climate crisis, geopolitical tensions and the abuse of new technologies — naming one in particular: “Lethal autonomous weapons — machines with the power to kill on their own, without human judgment and accountability — are bringing us into unacceptable moral and political territory.” While states might debate whether lethal autonomous weapon systems (or ‘killer robots’ in the popular imagination) are “unacceptably immoral”, there can be no doubt that Guterres is right on the urgency of the risk: development and use of autonomous weapons are both accelerating, and the stakes — ethical and political — are high. The world’s military powers have been competing to dominate this new class of intelligent weapons for years, with this AI arms race occurring against a contentious global landscape where an advantage in military AI could make a real difference to the balance of power. This geopolitical game theory driving such advancement in the sophistication of war machines has an unwanted blind spot — historically, human rights factor little into strategic calculations. With Covid-19, the acceleration of automation has taken on greater speed across a variety of different fronts. Military operations have had to be completely re-thought — physical distancing on a submarine is much harder than physical distancing in a supermarket. Lethal AI already had some mounting advantages over human equivalents, and can now add ‘immunity from catastrophic viruses’ to that list. For all of these reasons, keeping track of the AI arms race is more vital than ever. If It’s A Race, Who’s Winning? Almost every month, another innovation in autonomous weapons leaps off the headlines in military news — the autonomous Chinese Blowfish A3 helicopter drone equipped with machine guns or the Russian army of unmanned ‘Marker’ ground vehicles armed with mortars and grenade launchers. There is no question that new inventions in the world of military AI abound, but it is far less clear which country boasts the strongest tech. Key figures in the United States military have been forthright in warning of China’s might in this area. The US Defense Department’s relatively new Joint Artificial Intelligence Center is building command-and-control AI capability for the first time, explicitly citing the Chinese threat as the reason for the department’s urgency. The Center’s director Lt. Gen. Jack Shanahan has been clear about his desire to automate as much of the American military machine as possible: “What I don’t want to see is a future where our potential adversaries have a fully AI-enabled force and we do not.” In the last year, officials as senior as the US Defense Secretary have warned that Chinese technology may, in fact, already be more advanced than America’s. Secretary Mark Esper predicted that China might have “leapfrogged” existing American technology. With the military establishment suitably concerned, spending on lethal autonomous weapons in all branches of the American military seems set to go to another level in 2020 after already increasing in 2019. For China’s part, mounting investment in autonomous weapon development is a key plank in its ongoing effort to usurp American military dominance. Almost all large-scale AI programs in China benefit from massive governmental support and a huge trove of data, and its autonomous weapons program is the jewel in Beijing’s AI crown. China’s huge investment in lethal autonomous weapons predates other militaries, and its military theorists are ahead of the rest of the world in building futuristic “intelligentized” models of human-machine operations. A further dimension to China’s AI strategy is economic, with Beijing seemingly interested in profiting from its autonomous weapons program as a new export product. Already, China appears to be exporting many of its most high-tech aerial drones to wealthy buyers in the Middle East, explicitly marketing them as capable of advanced autonomous operations like assassinations. Last year, Zeng Yi, a senior executive at Norinco, China’s third-largest defense company, predicted that as early as 2025, “there will be no people fighting in battlegrounds”. As an arms race, autonomous weapons and AI more broadly is much harder to track than something like nuclear weapons with countable objects (eg stockpiles of warheads and launchers). How should we compare China’s development of autonomous aerial drones and hypersonic ‘smart’ missiles with America’s recent $2.7bn unveiling of huge unmanned ships? What about Russia’s autonomous tanks, deployed for the first time into combat in Syria? The short answer is that short of more extensive deployment of autonomous weapons, we cannot know for certain. Elsa B. Kania, an Adjunct Senior Fellow with the Technology and National Security Program at the Centre for a New American Security, agrees that the question of who leads an AI arms race between the great powers is not an easy one: There may be no single answer. For instance, potentially, the U.S. military may become more capable in leveraging AI in cyber operations, but the Chinese military could achieve greater advances in hypersonic weapons systems that can operate autonomously, and the Russian military may possess more experience in integrating unmanned systems in urban warfare. Given the most advanced autonomous weapons are developed behind closed doors in the world’s most secretive military labs, it is difficult to gauge the objective progress of military AI, let alone which country is closest to deploying sophisticated AI in all aspects of its military. However, there can be no doubt at this stage, with public declarations from American and Chinese officials alike, that this is the finish line: the automation of an entire military. ”Unacceptable” Is The New Acceptable For all Guterres’ stirring words about lethal autonomous weapons being an unacceptable global threat, the United Nations have been painfully slow in their response. There are no restrictions on nation-states or companies building, developing and integrating lethal autonomous weapons; and no sign that this will change in 2020. To be clear: there is an increasing groundswell of public pressure for autonomous weapons (or ‘killer robots’) to be regulated. Almost three in every four Europeans would like an international treaty prohibiting lethal autonomous weapons, according to a poll conducted in November by Human Rights Watch. 28 countries are advocating for the United Nations to adopt a similar measure. However, a small number of powerful countries are doing an effective job of stalling (if not derailing) progress at the international level. It will be no surprise that these countries — Russia, China and the United States — have by far the most advanced military AI. TIME has detailed how Russia has “steamrolled” the process by holding up discussion documents and postponing meetings. China and the United States have been more subtle, but no less obstructionist, in their own efforts to delay even the discussion of a substantive ban. Despite this, the UN’s disarmament chief Izumi Nakamitsu is optimistic that “within two years” the UN Convention on Certain Conventional Weapons (CCW) forum could release a definitive proposal for how lethal autonomous weapon systems could be regulated internationally. The CCW regime has previously been successful in limiting or regulating weapons like landmines or laser weapons that have “potentially indiscriminate” effects. With the potential for lethal autonomous weapons to potentially escape their programmed limitations either accidentally (through a software bug) or on purpose (the software is hacked), it is easy to see how their effects could be indiscriminate, falling within the criteria for CCW restrictions. But a hypothetical fit is only the qualifying mark in a long marathon to regulation. The influential Campaign to Stop Killer Robots is notably guarded about the UN process, warning that “incremental gains achieved to date are not impressive”. The lack of regulation or applicable international law leads to the grim conclusion that the AI arms race can only end with a global power developing and deploying ‘general’ AI into its military, with consequences we can only really speculate at this stage. If this truly is an “unacceptable” outcome, a dizzying amount of progress will have to be made on drawing up an agreement to ban or limit lethal autonomous weapons. In the meantime, war machines become ever smarter, and all the more alluring in a post-Covid world.

### AT I-LAW

#### [1] There is no warrant in this card – international law hasn’t been established that requires the banning of autonomous weapons – it is simply a proposed treaty but hasn’t been affected.

#### [2] Cross apply the Singh evidence which indicates contracts across States cannot be enforced – that means an external imposition on the States power through international LAW doesn’t turn the NC – because a contract would be meaningless if States didn’t follow them

#### [3] LAWs aren’t prohibited by the Martens clause. Hughes 18.

[Hughes, Joshua. “No, autonomous weapon systems are not unlawful under the Martens Clause.” Joshua Hueghes. [https://medium.com/@jghughes1991/no-autonomous-weapon-systems-are-not-unlawful-under-the-martens-clause-2653d18790e9 Published 21 August 2018](https://medium.com/@jghughes1991/no-autonomous-weapon-systems-are-not-unlawful-under-the-martens-clause-2653d18790e9%20Published%2021%20August%202018)] SHS ZS

First, let us consider customary law. For any non-international lawyers reading this, customary law is basically the rules which aren’t in treaties and states make up as they go along. It is made up of 2 parts: State practice, and opinio juris. **For a rule** of custom **to form**, ‘[n]ot only must **the acts** concerned amount to a settled practice, but they **must** also be such, or be carried out in such a way, as to **be evidence of a belief that this practice is** **rendered obligatory by the existence of a rule of law requiring it**.’([North Sea Continental Shelf Case, 1969](https://www.icj-cij.org/files/case-related/52/052-19690220-JUD-01-00-EN.pdf), para.77). **As technology has only recently reached a level where** some **systems could be considered fully autonomous** (e.g. [the Israeli Harpy/Harop](http://www.iai.co.il/2013/36694-16153-en/Business_Areas_Land.aspx)), **there has been no chance for settled state practice**, nor opinio juris **to form**. **Thus, there cannot be any customary rules preventing the existence or use of autonomous weapons**. Furthermore, many understandings of what an AWS actually is would not include the Harpy drone, and so in these conceptions customary rules are even further away. This leaves us with the principles of humanity, and the dictates of public conscience. To most people, these sound like moral or ethical concepts, and indeed this is how they are portrayed in HRW’s Heed the Call (pgs.19–22, 28–43). The problem is that the **Martens Clause comes from legal documents**, is a legal concept, **and in order to be applied it needs to be defined legally**. Indeed, Martens himself attempted to legally define ‘humanity’ ([Cassesse](http://www.ejil.org/article.php?article=511&issue=39), pg.188–189), and the concept has previously been referred to as the ‘laws of humanity’([Ticehurst](https://www.cambridge.org/core/journals/international-review-of-the-red-cross-1961-1997/article/martens-clause-and-the-laws-of-armed-conflict/19E402694542E42DD1EDA333027E490B), 129). **HRW** suggest that ‘The first principle of humanity requires the humane treatment of others’(pg.19) and that in order to do this, ‘one must exercise compassion and make legal and ethical judgments.’(pg.20) The effect of this is that ‘[w]hile compassion provides a motivation to act humanely, legal and ethical judgment provides a means to do so.’(pg.20) This all sounds well and good, and indeed morally right. But, this **is not a legal definition of a legal concept**. **In terms of legally defining ‘humanity’** we might first look to the Law of Armed Conflict (LoAC) principle of humanity, which suggests that one should not use any more force than necessary for ‘legitimate military purposes.’ ([UK Manual](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/27874/JSP3832004Edition.pdf), para.2.4) Whilst this conception of humanity exists as a principle in LoAC, **it is not** necessarily **the same thing as ‘the principles of humanity’ within the Martens Clause**. Still, Ticehurst accepts this conception as applying to the Martens Clause ([Ticehurst](https://www.cambridge.org/core/journals/international-review-of-the-red-cross-1961-1997/article/martens-clause-and-the-laws-of-armed-conflict/19E402694542E42DD1EDA333027E490B), pg.129). Would AWS breach this rule and use more force than necessary? I doubt it. For example, **if we imagine an autonomous uninhabited aircraft** (drone) **is deployed to destroy enemy tanks** using anti-tank missiles. **There is no reason to think that just because it is autonomous, the drone would fire many missiles when one would do the job**. So that isn’t a ground for AWS breaching the Martens Clause. Meron and Cassesse have similar conceptions of ‘humanity’ as it relates to the Martens Clause, which are different to Ticehurst. Both cite a number of cases and other influential instruments which link to the Martens Clause ([Meron](https://www.cambridge.org/core/journals/american-journal-of-international-law/article/div-classtitlethe-martens-clause-principles-of-humanity-and-dictates-of-public-consciencediv/F55EECE5BED3DDB9D78162DA4509A03A), pg.82–83; [Cassesse](http://www.ejil.org/article.php?article=511&issue=39), pg.202–207). For brevity, I’ll put their conclusions together, and not discuss the cases individually. Their analyses suggest that ‘**the principles of humanity’** as they relate to the Martens Clause **are: the rules of warfare in Common Article 3;** the distinction principle; prohibition on attacking civilians; sparing civilians as much as possible; limitations of means and methods of warfare; abiding by the notion of chivalry; prohibition on torture; prohibition on collective punishment. In terms of Common Article 3, there are two aspects which are problematic for AWS. First, is the prohibition on ‘violence to life and person’ of those hors de combat (‘out of the fight’, those who surrender, are prisoners, or who cannot fight due to incapacitation or sickness). Those who surrender could be protected if a system is programmed to recognise ‘hands up’ gestures (e.g. [Samsung SGR-A1](https://www.lawfareblog.com/foreign-policy-essay-south-korean-sentry%E2%80%94-killer-robot-prevent-war)) or white flags. Those prisoner in PoW camps, or in military hospitals can be protected through no-strike lists. This leaves us with those who are incapacitated through injury or sickness on the battlefield. One day, AWS may be able to recognise such protected individuals (although this is unlikely). However, **attackers need only take such individual into account where it would be reasonable**. For example, those prosecuting an aerial or artillery attack would find it difficult, or impossible, to evaluate injuries sustained by the enemy from afar. Thus, if it would be reasonable to use an AWS which cannot evaluate injuries or sickness on the battlefield, then this would be lawful (see [AMW Manual](https://usnwc.libguides.com/ld.php?content_id=2998228) Rule 15(b) commentary para.7). Therefore, it is possible that **AWS attacks could be lawful under Common Article 3.** Of course, ideally, a human operator would take control of an AWS where there is a risk of harming those hors de combat even if this means the system does not perform the operation wholly autonomou **The second aspect of common article 3 that could be problematic is the notion of ‘outrages upon personal dignity’.** According to HRW, human dignity is a second principle of humanity (pg.23). One conception that HRW mention (pgs.23–27) is Heyns, who has written on this subject as it relates to [armed conflict](https://www.tandfonline.com/doi/abs/10.1080/02587203.2017.1303903) and [law enforcement](https://muse.jhu.edu/article/617743/pdf) from a human rights perspective. As Heyns himself notes, **the notion of the ‘right to dignity’ is unclear** ([pg.49](https://www.tandfonline.com/doi/abs/10.1080/02587203.2017.1303903)). Therefore, despite being provided for in the African Charter on Human and Peoples Rights [Art.5](http://www.humanrights.se/wp-content/uploads/2012/01/African-Charter-on-Human-and-Peoples-Rights.pdf) (which HRW mistake as Art.7, see fn.76), due to the lack of clarity as to how this right might apply it is difficult to say for certain that it applies to AWS. Even if it does, it only applies to AWS to the extent that the African charter applies, as neither the charter nor the African Courts have expanded on extraterritorial application we should assume that its applicability extends to states parties within their national borders (for an overview of extraterritoriality as it applies to the African Charter, see [here](https://www.academia.edu/31026174/Extraterritorial_Applicability_of_the_African_Charter_on_Human_and_Peoples_Rights)). Regardless of extraterritoriality, **due to the unclear application of the right to dignity,** and what it entails as a matter of law, it would seem that **the notion as it applies to AWS is a suggestion of lex ferenda (what the law should be) rather than lex lata (what the law is).** Thus, **as it is not fully formed it cannot be applied as a legal concept and AWS cannot breach it**, nor the Martens Clause on this ground. Still, I would agree with this notion in a moral sense. To paraphrase Heyns words, AWS may be ‘lawful, but awful’. We shall now return to the other aspects which Meron and Cassesse found applied to ‘humanity’ in realtion to the Martens Clause. In terms of the distinction principle and the prohibition against attacking civilians, **an AWS programmed only to recognise and attack adverse military targets would abide by these rules**. Regarding sparing civilians as much as possible and limitations on means and methods, these duties would normally fall to humans programming the mission parameters into an autonomous system and so is not a ground for AWS breaching the Clause. In terms of chivalry, this has been confusingly defined in several works. However, Solis considers the concept really well, and essentially concludes that it means to act in good faith ([Solis](https://www.cambridge.org/core/books/the-law-of-armed-conflict/286ED6AD4C4164B473DF03E339F64A86), pg.5–6). **Unless AWS are programmed in bad faith, this is not a notion that could lead to AWS breaching the Martens Clause.** Finally, torture and collective punishments are things that would need to be decided upon by humans, and do not relate to the autonomous nature of an AWS. This, therefore would not be a ground for AWS breaching the humanity principle. **So, we can see that AWS do not breach the humanity principle according to Meron and Cassesse’s understandings of it,** although protecting those hors de combat can be problematic. So, we can now move onto the ‘dictates of public conscience.’ In terms of what this means, there is no agreed definition, but it is assumed to be synonymous with public opinion ([Meron](https://www.cambridge.org/core/journals/american-journal-of-international-law/article/div-classtitlethe-martens-clause-principles-of-humanity-and-dictates-of-public-consciencediv/F55EECE5BED3DDB9D78162DA4509A03A), pg.83). Indeed, HRW note that ‘“public” clarifies that these dictates reflect the concerns of a range of people and entities’ (pg.28) and I see no reason to disagree. HRW emphasise public opposition to AWS (pgs.30–31), and the groups which disagree with the production and usage of AWS, or find them morally reprehensible (pgs.32–43). However, as Greenwood notes, **the notion of ‘public conscience’ is so vague that, not only has it garnered little support, it is impractical to use** (pg.129, in [Fleck](https://books.google.co.uk/books?hl=en&lr=&id=ri-MB9EImPkC&oi=fnd&pg=PR15&dq=the+handbook+of+international+humanitarian+law+dieter+fleck+1st+edition&ots=wy65GAtKmc&sig=6fKElG_pa-ObWWVAP_fZTmQ0Gvg#v=onepage&q=the%20handbook%20of%20international%20humanitarian%20law%20dieter%20fleck%201st%20edition&f=false) 1995). Therefore, there is no threshold above which one could say there is enough public opinion opposing AWS that they should be deemed unlawful. Thus, despite the considerable opposition which HRW note, we cannot say that AWS go against the ‘dictates of public conscience’ because we simply cannot know what level of opposition is required, and how it should be measured. In conclusion, despite being morally problematic, **AWS do not breach the Martens Clause**. **Firstly, because there are no customary rules prohibiting the use of AWS.** **Second, because it by no means clear that AWS will breach the ‘principles of humanity’ as they are legally understood** — although safeguarding those hors de combat is problematic in some situations. **Finally, the ‘dictates of public conscience’ have no defined limit as to what level of public opposition is required to make a means or method of warfare unlawful, thus we cannot say that AWS go against these dictates**. Therefore, there is nothing inherent about the existence or use of autonomous weapon systems which makes them unlawful under the Martens Clause.

#### This outweighs:

#### [A] Methodology – It’s a meta-analysis on your applications as to why LAWs would violate Martens clause, so it logically comes first because it indicts the foundation of your study

#### [B] Specificity – It answers the internal links as to why LAWs would be morally problematic under Martens. Specifically, there are no customary rules prohibiting AWS since the technology is new and the original convention wasn’t made under the pretext of LAWs existing, LAWs themselves haven’t yet been meaningfully deployed in battle field so there isn’t a precedent for testing their legal status under international law, and there is 0 indication that LAWs will breach the principles of humanity because that presupposes bad programming, which isn’t constitutive to LAWs being but rather why the humans who operate the LAWs would be bad, which is NUQ under warfare.

#### [4] International bodies cannot form and enforce contracts among various States. Singh, [N. N. Singh (Faculty of Law, National University of Singapore). “THE ABSENCE OF A SOVEREIGN LEGISLATURE AND ITS CONSEQUENCES FOR INTERNATIONAL LAW.” December 1970.] MKThere being no international legislature for international society some writers have refused to concede the nature of true laws to the rules of international law.1 The 'Austinian' objection that law cannot emanate from a non-sovereign body clearly illustrates this view.2 John Austin had defined international law as a system consisting of rules of 'positive morality'.3 In his opinion : If the same system of international law were adopted and fairly enforced by every nation, the system would answer the end of law, but, for want of a common superior, could not be called so with propriety. If courts common to all nations administered a common system of international law, this system, though eminently effective would still, for the same reason, be a moral system.4 Similarly W.E. Heavn had declared that 'Law cannot be predicated of mere customs which are not even true commands much less the commands of any competent State'.5 Following John Austin's thesis, Holland had called international law a 'law by courtesy' or 'law by analogy'.6 These jurists and their followers clearly refused to accept anything except the acts of a sovereign legislature body or the commands of a sovereign as true laws.7 They had, thus, not only set up a very narrow definition of law, but thereby also expressed reluctance to recognise the historical evolution, or the changing character, of human institutions. Further, such an attitude had not only obscured the organic relationship of law to society8 but also the vital part which custom, as opposed to legislation, has always played in the legal systems of national states.9 However, the Austinian concept of law so far as it inde with commands of a sovereign seems to have been complet by modern writers; even from among the few modern writers who have denied the character of 'law' to international not one seems to have subscribed to it, Thus, the denial of the legal character of the rules of international law simply on the ground that they do not represent 'commands' of a sovereign or because there is no sovereign international legislature in existence is clearly a thing of the past. The few modern writers who have denied the legal character of the rules of international law10 have done so mainly by taking refuge in the smug opinion that the rules of international law, even though substantially observed, are not 'legal' because they cannot be effectively enforced against States if the States themselves do not submit to them. This, it must be submitted, brings us to a completely different problem, i.e., the problem concerning the 'enforcement' of the rules of interna tional law11 which is quite different from the question of its 'legal character '. The legal validity of a rule, based on the conviction that it is binding, is quite distinct from its enforceability: law is not law because it is enforced, on the contra 2. The question of the 'source' of obligation in international law Doubts about the character of international law as true law can, however, only be dispelled by showing that obligations upon States exist in international relations which are very similar to the normative obligations that exist in any system of law. To show this, however, involves an examination of the source of such obligations, for it is not possible, given the absence of an international parliament, to rely on the formal source of a sovereign legislative body. It is to an examination of the theories of the source of obligation in international law that we must, therefore, now turn. But as the term 'source' has been used by different writers implying different meanings it is perhaps necessary to indicate here the meaning attributed to this term; it has been used in the sense of 'originating Cause' and not in that of 'evidence'. Also no attempt has been made to distinguish sharply between the terms 'cause' and 'basis'. The question of the rise of obligations is certainly not the same as the question concerning the validity or the binding character of obligations after they hâve been established. It is only in connection with the latter question that we can use the term 'basis of obligation' and to that extent we can distinguish between 'basis' and the 'originating cause or causes'. But this must not lead us to believe that there is no correlation between the two. On the other hand, there is a vital connection between the originating causes, which explain the rise of obligations, and the basis of those obligations, in the sense that che 'originating causes' to a very large extent also provide the very basi for the continuing validity, application and the binding character of the obligations which they are instrumental in creating. Much depends also on how we formulate our question. The question that needs to be answered is in Ehrlich's words: 'Whence comes the rule of law, and who [in essence what] breathes life and efficacy into it (a) The 'will' of the State — the Continental approach A strong reaction against the Austinian concept of law was inevitable. On the Continent, Bergbohm had been one of the earliest writers io suggest that the absence of an international legislature should only lead us to the conclusion that a particular source of law does not exist within the society of States and, further, that it should not lead us to deny the 'legal character' of the rules of international law.13 To Bergbohm, the expression of State will, as evidenced in interna agreements, constituted a proper source of law.14 Similarly Jellinek traced the source of obligation in international law to the will of the State. First, he pointed out that States are not above law and can be bound by their own will — through a process of auto-limitation.15 Secondly, he emphasized that the international and the municipal systems of law are different systems having different objectives, that is, whereas the the municipal system of law envisages the 'subordination' of the members of a community, the international society is run on the principle of 'co-ordination'. Thus, international law was according to him a law between co-ordinate entities and different from the law of the States which emphasized the element of command. However, by asserting that the States are legally free to disengage themselves from any such obligation which runs counter to their interests,16 he called into question the very 'legal' character of international obligations.17 The difficulty with Jellinek was that he could not deny the States, especially within the framework of his theory of 'auto-limitation', the right to disengage themselves from those obligations which they had themselves created. The answer to this difficulty was provided by Triepel's Vereinbarung stheome.18 Triepel pointed out that although the will of the State is a necessary element, the will of any single State alone is not sufficient for the creation of international obligatio that purpose he envisaged that a fusion of several wills leading to the creation of a 'common will' is necessary. He called this 'common will', as expressed in treaties and agreements, by the name of Vereinbarung. By making the Vereinbarung the source of international obligations Triepel had also succeeded in creating 'a legal power over States' : only Vereinbarung could undo what it had created. But one of the serious charges levelled against Triepel's theory is that it does not explain the existence of customary or general international law and reduces the rules of international law to a conglomeration of particular law.19 Cavaglieri, Anzilotti and Strupp attempted to answer this question — raised by the criticism of Triepel's theory — by asserting that those States which do not participate in the formulation of a particular law later on become bound by it through certain processes.20 Claiming that rules of international law are expressions of the will of the State as evidenced in agreements, these writers also attempted to establish a basis for the binding force of international agreements in the rule of pacta sunt servanda, which in Cavaglieri's opinion is a rule of customary international law whereas Anzilotti describes it as an original hypothesis and a postulate incapable of proof.21 But, despite these variations in the approach of writers to explain the genesis of the rule or the fact that recourse to the principle of pacta sunt servanda leads them to a tautology (international agreements are binding because they are binding)22, such doctrinal assertions at least indicated a movement towards interpreting law on a much more practical basis than hitherto employed.23

### AT LAWs Not Human

#### [1] Actor specificity – the State has an obligation of fostering good conditions to their citizens, and by instituting LAWs that is an intrinsic good under the framework. For instance, just because nuclear weapons aren’t agents doesn’t mean the use of nuclear weapons are permissible, but ownership controls culpability

#### [2] LAWs are moral agents. Maiyane doesn’t agree with the conclusion of the NC.

[Maiyane, Karabo. “Ethics of artificial intelligence: virtue ethics as a solution to artificial moral reasoning in the context of lethal autonomous weapon systems.” Published 2019] SHS ZS

On the question of the moral status of AWS, I argue that **AWS**, **by virtue of their definition** and espoused functions, **can no longer be considered as mere weapons** and thus instruments of war. As such, in the context of warfare, **they should be considered as combatants**, meaning that their moral status should be **that of moral agents**. Ascribing AWS a moral status places us at a strategic position with regards to responding to the critique of whether AWS will be able to conform with IHL. Knowing that AWS are now considered as combatants means that **we are now aware of which parts of IHL they must specifically comply with**. For example, as combatants **they would have to comply with** jus in bello principles such as: **discrimination and non- combatant immunity** and proportionality, [see [4], [5], [6], [7]]. It also means that with regards to responsibility, **they can now be held responsible the same way other combatants are.** Knowing which parts of IHL AWS must comply with, what follows is to resolve the questions on how AWS must be programmed to ensure such compliance. **Regarding programming** morality for artificial moral agents (AMA’s), Wallach and Allen [8] published a book titled: Moral Machines: Teaching Robots Right from Wrong. In this book they explore different programming possibilities for AMA’s. They make a case for both top down and bottom up approaches, and argue that both such systems have shortcomings: top down in that explicit procedural systems don’t work in all situations, especially in warfare where some situations cannot be anticipated; and bottom up in that learning systems need the position from which to learn from. They argue that “If neither a pure top-down approach nor a bottom-up approach is fully adequate for the design of effective AMAs, then **some hybrid will be necessary**” [8, p. 117]. This is an approach that combines both top-down and bottom up approaches. **Many** [see [4], [5], [8], [9]] who argue in favour of hybrid approaches **consider virtue ethics as an important normative approach**. Specifically, **the** **focus is on Aristotle’s conception of virtue as twofold, moral and intellectual**. For Aristotle [10], **intellectual virtues can be taught, and moral virtues can be cultivated through habit over time.** It is because of this distinction he makes that researchers in the field find his conception of virtue attractive for hybrid approaches. Wallach and Allen [8] argue that **intellectual virtues can be programmed using a top-down programming, whilst moral virtues can come about as a result of a “bottom up process of learning and discovery by an individual through practice**”. I argue that a hybrid program, **using a virtue ethics normative approach, would be the solution for moral reasoning in the case of AWS.** I wish to clarify here that by a virtue approach I do not mean programming specific virtues into AWS; I mean **using the framework that Aristotle argues should be followed by human agents in terms of cultivating virtues**. I argue this because for Aristotle, virtues are those characteristics that makes one good at what they do. Thus, **their acquisition is through learning how to be good at that which one does, by doing it constantly – from being taught rules (intellectual virtues), to learning them by continuous practice over tim**e (moral virtues). I will illustrate that Aristotle’s conception of how virtues are acquired offers a good framework for building an architecture of an autonomous moral agent, especially one that operates under defined contexts such as AWS in warfare. In this sense the **top down approach would be used to program IHL while the bottom up approach would use machine learning as learning program to train AW**S. I am not arguing that a virtues-based architecture is the only one that can work in programming morality in all AMAs, only that in the case of AWS in warfare it is ideal. This is because such an architecture allows for universal programming of rules (IHL) on the one hand but accepts dynamic applicability (applicability based on context and agent) on the other. **Which is what is required of combatants in warfare.**

# Cards

## Valley Offense

First, moral obligations of states towards other states or citizens cannot exist because international politics is the state of war. **Forde:**

Hobbes does not deny that there are objective ethical principles, or a "law of nature." Rather, like most realists, he argues that **ethical principles** are simply suspended under the conditions of international politics. Hobbes constructs a lengthy list of laws of nature, which are duties that apply to human beings in and out of the state of nature. These laws differ fundamentally from more traditional systems of natural law in that their **foundation is self-interest**. In essence, accord- ing to these laws, people have a "duty" only to preserve themselves (De Cive Ch. 2, 123; Leviathan Ch. 14, 184). Seeking peace is a corollary of this, as are, consequently, "duties" to be fair and forgiving; but the primacy of self-preservation dictates that individuals are required to take no significant risks to comply with them. **Only a contract with a** sovereign **power to enforce** obedience **can remove just fear and give these** duties **real force. In the international realm, there is by definition no such power.** Such compacts as states have among themselves may suffice for limited purposes; but they are incapable of lifting the fundamental state of war in international politics. We might paradoxically say, therefore, that Hobbes accepts a universal code of ethics, but it is one that no state is morally bound to follow. And Hobbes sees no prospect of a compact to remedy this. In the original state of nature, conflict and insecurity were so immediately threatening that individuals were effectively driven into a social compact, but international conflict does not threaten individuals in the same way. And states themselves are not threatened as mortally by the state of war. Thus the formation of a supranational sovereignty never becomes imperative. In other respects, however, the inter- national state of war is worse than the original state of nature. The equality characteristic of the original state does not obtain among nations, making security wholly relative, depending upon thestrength of a particular nation's adversaries. Nations must therefore heed a fluctuating balance of forces, and arm in accordance with the circumstances. One of the few duties Hobbes is willing to give his sovereign is the duty to maintain adequate armaments, and he insists on the propriety of sovereigns doing whatever they can to undermine the power of other states, whether "by slight or force" (De Cive Ch. 13, 262). All of this can take place in what is ordinarily regarded as peacetime, but which Hobbes considers only "a breathing time" in the interminable state of war (260). Hobbes goes so far as to say a sovereign may, without injustice, inflict any harm whatsoever on "innocent" foreign persons in the pursuit of state interests, also in "peacetime" (Leviathan Ch. 28, 360). For only by compact can the state of war be terminated; and **a state can have no binding compacts**, in the Hobbesian sense of the word, **with people** who are **not its subjects. The conditions of** international **politics preclude any real justice**, or even simple humanity, among states.

Steven Forde (Professor of Political Science, University of North Texas), “Classical Realism,” in *Traditions of International Ethics*, Cambridge UP, 1992.

Thus LAWS are used as a deterrent which coheres the international arena. The nature of autonomous weapons increases a leader’s credibility with deterrent threats. **Wong et al 20:**

On the one hand, **autonomous systems** could **enhance the credibility of** U.S. **conventional** extended **deterrence because the risk to** U.S. **military personnel** of employing these systems **is** much **lower than** with traditional **kinetic military means.** Additionally, the **operational advantages of autonomous systems** relative to more traditional military means—such as faster decision cycles, the ability to stay ready to strike much longer, and greater precision than human personnel—**could lead adversaries** and allies **to conclude that** U.S. **leaders will** be more willing and likely to **employ autonomous systems** in situations in which allied interests are threatened. On the other hand, U.S. allies could interpret Washington’s increased reliance on autonomous systems as a reflection of a growing U.S. unwillingness to put American lives on the line in severe crises and confrontations with adversaries. Thus, although the United States may see fielding autonomous systems as a way to reduce risk to U.S. military personnel by substituting machines for humans, reducing risk to U.S. personnel in overseas commitments may paradoxically reduce Washington’s ability to assure U.S. allies. Autonomous systems may also affect the credibility of deterrent threats. **States with autonomous systems** might **appear more credible when making deterrent threats** than states without them. Nonetheless, as with other conventional weapons, opponents who do not possess autonomous systems will not simply accede to the deterrent or coercive threats of states that do have them. Instead, they will develop strategies, operational approaches, and capabilities designed to counter, avoid, or mitigate the advantages of autonomous systems. When confronting states that do possess autonomous systems of their own, using autonomous systems could come to be seen as low-risk and thus attractive means for mounting probing attacks against adversaries. This could result in “salami” tactics employed to slice away at the adversary’s interests without overtly crossing a threshold or red line that invites the opponent to strike back.

Wong, Yuna Huh, John Yurchak, Robert W. Button, Aaron Frank, Burgess Laird, Osonde A. Osoba, Randall Steeb, Benjamin N. Harris, and Sebastian Joon Bae, Deterrence in the Age of Thinking Machines, Santa Monica, Calif.: RAND Corporation, RR-2797-RC, 2020. As of June 22, 2020: <https://www.rand.org/pubs/research_reports/RR2797.html> //CF

Second, states would not consent to a ban on autonomous weapons since they are consistent with their rational self-interest. **Gayle 19**

But the UK is among a group of **states** – including Australia, Israel, Russia and the US – **speak**ing forcefully **against legal regulation**. As discussions operate on a consensus basis, their objections are preventing any progress on regulation.  The talks come as the UK military is ploughing tens **of** millions of pounds into **autonomous weapons**, most recently announcing on Thursday a £2.5m project for “drone swarms” controlled with the help of next-generation autonomy, machine learning, and AI.  The talks in Geneva are taking place under the convention on certain conventional weapons. First enacted in 1983, the convention is intended to restrict the use of weapons “that are considered to cause unnecessary or unjustifiable suffering to combatants or to affect civilians indiscriminately”. It already covers landmines, booby traps, incendiary weapons, blinding laser weapons and clearance of explosive remnants of war.  “We urgently need a ban on killer robots,” said Ben Donaldson, head of campaigns at the United Nations Association – UK. “The majority of states get it. A rapidly growing proportion of the tech community get it. Civil society gets it. But a handful of countries including the UK are blocking progress at the UN. The UK needs to listen to this growing coalition and join calls for a preemptive ban.”  Responding to the criticism, a Ministry of Defence spokesperson said: “The United Kingdom does not possess fully autonomous weapon systems and has no intention of developing them. We believe **a preemptive ban is premature** as there is still no international agreement on the characteristics of lethal autonomous weapons systems.”  The issue of human control is at the heart of discussions about killer robots, according to the British military, and its negotiators have sought to focus debates at the UN on building consensus on what that means. Britain’s negotiating team says that no UK offensive weapons systems will be capable of attacking targets without human control and input.  They are arguing against **a** preemptive **ban** on the basis that it **could jeopardise their ability to exploit** any potential **military advantages** they could gain by imbuing weapons **with AI.**

Gayle, Damien. “UK, US and Russia among Those Opposing Killer Robot Ban.” *The Guardian*, Guardian News and Media, 29 Mar. 2019, [www.theguardian.com/science/2019/mar/29/uk-us-russia-opposing-killer-robot-ban-un-ai](http://www.theguardian.com/science/2019/mar/29/uk-us-russia-opposing-killer-robot-ban-un-ai). //CF

Third, a civilian’s role in war is one which is forced upon them without their consent. A soldier’s killing of a civilian is a violation of mutual benefit as the civilian never consented to the role of a soldier. **Benbaji et al 19:**

We turn now to a different concern about the permission to inflict collateral damage on civilians. **From a contractarian perspective**, **there** seems to be **a** profound **difference between civilians and soldiers.** As argued above, **soldiers willingly** and knowingly join the army and **accept their role**." **Since soldiery is** part of a **mutually beneficial** and fair arrangement, it can be assumed that they freely accepted their role, together with the rules that define it. This acceptance involves waiving moral rights and undertaking moral duties. They knowingly come to occupy their role and (perhaps unknowingly) accept the package of norms that constitute it. Acceptance is the locus of the moral efficacy of the war agreement: **combatants have an equal right to kill** **each other** in war because, by joining the army, they accept a regime under which they are released from the horror, and it was different from reading about the crimes of another country. Familial and national relations provoke positive emotions as well: the pride felt for the achievements of a compatriot resembles the pride felt for the achievements of a brother or cousin. This identification of **civilians** with their countries undermines the potential claim that since they **never explicitly undertook the role** of a civilian**, it was forced upon them against their consent.**

Benbaji, Yitzhak, and Daniel Statman. War by Agreement: A Contractarian Ethics of War. First Edition, Oxford University Press, 2019. //CF

And, high civilian casualties due to collateral damage can be attributed to human error. AI solves. **Lewis 20:**

I should know. As a senior advisor for the State Department on civilian protection in the Obama administration, I was a member of the US delegation in the UN deliberations on lethal autonomous weapons systems.  As part of that delegation, I contributed to international debates on autonomous weapons issues in the context of the Convention on Certain Conventional Weapons, a UN forum that considers restrictions on the design and use of weapons in light of the requirements of international humanitarian law, i.e, the laws of war. Country representatives have met every year since 2014 to discuss the future possibility of autonomous systems that could use lethal force. And talk of killer robots aside, several nations have mentioned their interest in using **artificial intelligence in weapons** to better protect civilians. A so-called smart weapon—say a ground-launched, sensor-fused munition— could more **precisely and efficiently target enemy fighters and deactivate** itself **if it does not detect the** intended **target**, thereby **reducing** the **risks** inherent in more intensive attacks like a traditional air bombardment**.** I’ve worked for over a decade to help reduce civilian casualties in conflict, an effort sorely needed given the fact that **most** of those **killed in war are civilians**. I’ve looked, in great detail, at the possibility that automation in weapons systems could in fact protect civilians. Analyzing over 1,000 real-world incidents in which civilians were killed, I found that **humans make mistakes** (no surprise there) **and** that there are specific ways that **AI could** be used to **help avoid** them. There were two general kinds of mistakes: either military personnel **missed indicators that civilians were present**, **or civilians** were **mistaken as combatants** and attacked in that belief. Based on these patterns of harm from real world incidents, artificial intelligence could be used to help avert these mistakes. Though the debate often focuses on autonomous weapons, there are in fact three kinds of possible applications for artificial intelligence in the military: optimization of automated processing (e.g., improving signal to noise in detection), decision aids (e.g., helping humans to make sense of complex or vast sets of data), and autonomy (e.g., a system taking actions when certain conditions are met).

Lewis, Barry. “Killer Robots Reconsidered: Could AI Weapons Actually Cut Collateral Damage?” Bulletin of the Atomic Scientists, 10 Jan. 2020, thebulletin.org/2020/01/killer-robots-reconsidered-could-ai-weapons-actually-cut-collateral-damage/. //CF

## Misc.

#### And, this outweighs, LAWs are on track to replace humans in the battlefield. Walsh doesn’t agree with the NC advocacy.

[Walsh, Toby. “The one job that will disappear by 2062 – the job of fighting wars.” ThePrint. https://theprint.in/pageturner/excerpt/job-will-disappear-by-2062-fighting-wars/409314/ Published 26 April 2020] SHS ZS

**There is one job likely to disappear** through automation by 2062 which I and many others especially fear. This is **the job of fighting wars**. Indeed, **replacement has already started to happen**. **An arms race has begun in the development of robots that can replace humans in the battlefield**. The media like to call them ‘killer robots’, but the technical term is ‘lethal autonomous weapons’, or LAWs. The problem with calling them killer robots is that this conjures up a picture of the Terminator, and hence of technologies that are a long way off. But it is not Terminators that worry me or thousands of my colleagues working in AI. It is much simpler technologies that are, at best (or at worst), less than a decade away. It is not smart AI but stupid AI that I fear. We’ll be giving machines that are not sufficiently capable the right to make life-or-death decisions. Take a Predator drone. This is a semi-autonomous weapon, which can fly itself much of the time. However, there is still a soldier, typically in a container in Nevada, in overall control. And importantly, it is still a soldier who makes the decision to fire one of its missiles. But it is a small technical step to replace that soldier with a computer. Indeed, it is already technically possible to do so. **And once we build such simple autonomous weapons, there will be an arms race to develop more and more sophisticated versions.** The world will be a much worse place if, in twenty years’ time, lethal autonomous weapons are commonplace and there are no laws about LAWs. This will be a terrible development in warfare. But it is not inevitable. We get to choose whether we go down this road – and we’ll be choosing which road we go down in the next few years.

1. <https://www.merriam-webster.com/dictionary/ought> [↑](#footnote-ref-1)