

# How to Engage the Public on the Ethics and Governance of Autonomous Weapon Systems

Jason Millar<sup>1</sup> and AJung Moon<sup>2</sup>

## Abstract

There are active international discussions taking place on the ethics and governance of Autonomous Weapon Systems (AWS)—robots that can kill without direct human intervention or oversight. It is imperative that we critically examine the role and nature of public engagement intended to inform decision makers. The Martens Clause, included in the additional protocols of the Geneva Conventions, makes explicit room for the public to have a say on what is deemed permissible in matters of armed conflict, especially where new technologies are concerned. However, many measures of public opinion, using methods such as surveys and polls, have been designed in such a way that makes them subject to potential biasing effects. For example, some only consider specific applications instead of general aspects/features unique to the technology under consideration. In this paper, we survey various studies that have been conducted to gauge public opinion on the use of military drones (autonomous and remotely operated), including the recent international poll conducted by the Open Roboethics initiative (ORi). By drawing on evidence from moral psychology, we highlight some potential biasing effects that particular question framings could have on outcomes, and outline considerations that can be taken into account when designing and determining the applicability of public opinion measures to questions of the governance of AWS. Such considerations can help public engagement objectives live up to the spirit of the Martens Clause.

## Introduction

The list of militaries using robotic weapon systems to aid in carrying out combat operations is growing (Tucker 2014). Today's robotic weapon systems are remotely operated, and sometimes referred to as *Remotely Operated Weapon systems* (ROWS). With ROWS, the decision to use lethal force (i.e. to fire a weapon at a target) remains a *human* decision. However, systems are being developed that could allow military robots to autonomously make the decision to use lethal force. As defined by the US Department of Defense, these are systems that, "once activated, can select and engage targets without further intervention by a human operator"

---

<sup>1</sup> Jason Millar is a Postdoctoral Fellow at the University of Ottawa Faculty of Law, and teaches robot ethics and philosophy at Carleton University (Ottawa). He is the Chief Ethics Analyst at the Open Roboethics initiative ([openroboethics.org](http://openroboethics.org)), and the Working Groups Coordinator at the Foundation for Responsible Robotics ([responsiblerobotics.org](http://responsiblerobotics.org)).

<sup>2</sup> AJung Moon is a PhD candidate at the University of British Columbia (Vancouver), and co-founder of the Open Roboethics initiative ([openroboethics.org](http://openroboethics.org)). She has authored and co-authored several papers on Roboethics and human-robot interaction.

(Department of Defense 2012[in Horowitz 2016a]). In other words, in the not-too-distant future military robots could kill without human intervention.

The prospect of those new *Autonomous Weapon systems* (AWS) has sparked an ethical debate that has been taken up by the public media (Evelith 2014; Parkon 2015; Devlin 2016), governments (Lin et al. 2008), civil society (PAX 2014; Human Rights Watch 2015; Article 36 2015), and the United Nations (UNIDIR 2014). On one side of the debate we have various arguments to the effect that although ROWS may be ethically permissible, AWS are not, so AWS should be internationally banned (e.g. Human Rights Watch 2015; Article 36 2015). On the other side of the debate we have various arguments to the effect that both ROWS and AWS are ethically permissible technologies, and that a ban is unwarranted (e.g. Arkin 2015).

What might help inform the debate over the ethical permissibility of ROWS and AWS?

In this paper, we focus on public engagement as a means of informing the debate over the ethical permissibility of ROWS and AWS. We do not aim to directly address the central question in the debate—Are AWS ethically permissible? Rather, supplementing previous work in general survey design literature<sup>3</sup>, we aim to inform the way researchers go about engaging the public on that, and related, questions about new weapons technologies. We start by making the case that, on this particular issue, public engagement is an important and necessary policy input. We then identify a distinction between two general kinds of questions typically used in those surveys: questions that focus on the *application* of the technology, and questions that focus on the *nature* of the technology. If researchers are interested in gauging the ethical permissibility of particular applications of ROWS and AWS, asking *application-specific questions* will be useful. However, we argue that application-specific questions can introduce problematic biases in surveys meant to shed light on the debate over the ethical permissibility of AWS. We then examine two moral psychological biasing effects that are of particular concern when asking people about robots: moral emotional priming, and anthropomorphic framing. Finally, we survey existing public opinion polls and examine them in light of the considerations just outlined.

### **Why engage the public?**

Decisions whether or not to build or use new kinds of weapons are a matter of democratic and humanitarian concern. The 20<sup>th</sup> century was particularly illustrative on this matter. Over the course of mere decades, soldiers and civilians were introduced to new and unique forms of suffering resulting from military decisions to develop and use such technologies as chemical weapons, landmines, incendiary weapons, cluster munitions and nuclear weapons on the battlefield. Various international governance initiatives have arisen over the years in response to the nature of those weapons.

---

<sup>3</sup> For a brief, yet excellent, overview of questionnaire design see the section on Questionnaire Design in Pew Research Center (2016).

Most notable are the Geneva Conventions, drafted in 1949 following the Second World War, which set out the rules of armed conflict and “remain the cornerstone of contemporary international humanitarian law” (ICRC 2009).<sup>4</sup> The Geneva Conventions provide specific protections for prisoners of war, non-combatants, civilians, and victims of armed conflict, and have been ratified by 196 countries. They also make clear that those who are party to armed conflict have limited rights in choosing the methods and means of warfare (International Justice Resource Centre 2016).

More specifically, Article 36 of Additional Protocol I to the Geneva Conventions explicitly raises the introduction of new weapons technologies as a concern:

“In the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party.”

However, it is not always clear how well, or if, existing rules apply to new weapons. In cases of uncertainty, international humanitarian law (IHL) underscores the importance of public engagement as a means of informing determinations of the permissibility of new weapon systems. The Martens Clause, included in the Preamble of Additional Protocol II of the Geneva Conventions, makes explicit room for the public to have a say on what is, and is not, deemed permissible in matters of armed conflict, especially where new technologies are concerned. It reads:

“The High Contracting Parties, ... Recalling that, in cases not covered by the law in force, the human person remains under the protection of the principles of humanity and the dictates of the public conscience,...” (Additional Protocol II to the Geneva Conventions)

Though legal scholars often disagree on how best to interpret and implement the Martens Clause (Cassese 2000), it remains a fact that, from the perspective of the Clause, “the dictates of the public conscience” are meant, in certain cases, to help refine and extend the international laws of armed conflict. Indeed, at two separate meetings of the UN Convention on Certain Conventional Weapons (CCW), first in 2014 and again in 2015, experts and state representatives convened specifically to discuss the prospect of AWS and consider an international response. The Martens Clause was cited as a relevant guiding governance consideration in the meeting proceedings. Groups such as Human Rights Watch (2015) and the ICRC (2015) have questioned whether or not current laws are capable of addressing the unique issues raised by AWS.

These questions help to underscore a key distinction between AWS and ROWS: AWS transfer the power to make lethal decisions *from a human to a machine*. That distinction is at the heart

---

<sup>4</sup> The Geneva Conventions of 1949 amended three existing conventions dating from 1864, 1906 and 1929 and added a fourth. The Geneva Conventions have since been “supplemented and developed” by three Additional Protocols (ICRC 2009).

of the controversy surrounding AWS. Discussions at the UN CCW on AWS have focused on the importance of maintaining “meaningful human control” over decisions to use lethal force. But it is difficult to define meaningful human control (Article 36 2015; Horowitz & Scharre 2015). For example, meaningful human control could be interpreted as a design requirement just to make the weapon system meet the requirements of existing IHL, or it could form an additional requirement for a weapon system even if that weapon system could conform to IHL (Horowitz & Scharre 2015). Furthermore, the process of automating decisions to fire at a target involves progressively automating a number of different functions over time. It might not be entirely clear when a particular automation technology has become so automated that it has “crossed the line” from meaningful human control, into an ethically impermissible territory (Millar 2015). The Martens Clause seems particularly useful here, since, if meaningful human control is an important guiding concept in the ethics of AWS, and if the Martens Clause applies to the AWS debate (as we have argued it does), then meaningful human control is bound to the public conscience. Thus, we argue that the challenging task of defining meaningful human control can, and should, be informed by, and aligned with, the public conscience.

We assert that the prospect of AWS is a case that fits the purpose of the Martens Clause. Public engagement is one way to support the requirements set out in the Clause, and to gain traction on the permissibility of AWS.

### **Two kinds of questions about technology**

Public opinion polling can help achieve the task set out in the Martens Clause by helping us to better understand the public conscience on the permissibility of new weapons technologies. However, as with any survey, the questions must be designed to probe participants’ opinion on the heart of the matter. Here, rather than reiterating best practices from survey design literature, we highlight some of the pitfalls specific to the discussion of ROWS/AWS issues. In the ROWS/AWS debate, we assert that questions that contain reference to specific applications can mislead participants. Rather, questions must be designed to reveal changes to the public’s response with respect to the *nature* of the technologies in question.

Consider the following two questions:

- (A) Should your government develop and use autonomous weapon systems?
- (B) Should your government develop and use autonomous weapon systems to guard against terrorism?

Question (A) is a question about the *nature* of AWS; it is an *application-neutral* question. In order to answer it in a way that would satisfy the researcher’s intent, one would need to have specific knowledge about AWS, perhaps about specific features unique to the technology, or about how those unique features compare to, say, unique features of other types of weapons. As such, a survey posing only question (A) to participants might require a preamble containing a (somewhat) substantive description of AWS, or it would run the risk of confusing participants and generating relatively uninformative data. Without a preamble to inform participants about

the nature of AWS, questions like (A) could be said to occur “in a vacuum” (Horowitz 2015, 3), biasing results by causing participants to recall popular images of so-called *killer robots*, such as the *Terminator*, an effect we call Terminator Bias<sup>5</sup>, when answering survey questions.

Question (B) is an *application-specific* question: it can be meaningfully answered without any particular knowledge of AWS. It is a question about a particular *application* for which AWS could potentially be used. In contrast to question (A), a survey posing only question (B) could rely only on the assumption that participants have an opinion on guarding against terrorism. This is because the application could do the heavy lifting in cases where participants know nothing at all about nature of the underlying technology. To illustrate this point, consider the following question:

(C) Should your government develop and use phased cyclotrons to guard against terrorism?

Depending on how committed one is to guarding against terrorism, the use of phased cyclotrons can be rendered irrelevant in answering the question. Indeed, knowing nothing about phased cyclotrons might effectively increase a participant’s focus on the application, exacerbating the problem.

Asking questions (A) and (B) in a survey, in order to create context (Horowitz 2015) and avoid Terminator Bias, could backfire. This is so first, because it is not clear whether (B) creates meaningful context or a problematic, application-specific distraction in the participant and, second, because the use of (B) to create context could lead researchers away from providing a meaningful description of the nature of a technology.

Now consider a fourth question:

(D) Should your government develop and use autonomous weapon systems to kill known terrorists’ family members?

(D) focuses participants on an application that is both illegal and (we assume) broadly considered ethically impermissible, rendering it subject to social desirability bias<sup>6</sup>. A survey posing only questions (B) and (D) illustrates a further problem with application-specific questions. If the goal is to inform a policy debate about the permissibility of AWS, differences in answers to (B) and (D) will help to inform debates about particular AWS applications, but might reveal little about participants’ views on the permissibility of AWS.

---

<sup>5</sup> Media references to *The Terminator* are so prevalent in reporting on this issue that we feel it’s worth coining the phrase. See, for example, (see Gibbs 2015; Brant 2016; McDonald 2016, and countless more).

<sup>6</sup> Social desirability bias refers to people’s tendency to answer survey questions Paulhus, D.L. (1991).

Pairing application-specific questions featuring different technologies may seem like a useful strategy to isolate a variable, but does not offer a satisfactory solution. This point is illustrated by considering a survey containing only questions (B) and (C). As with stand-alone questions focused on a particular application, depending on a survey participant's reaction to the application—in this case guarding against terrorism—the survey results might not reveal anything informative about the development and use of phased cyclotrons or AWS. For example, participants might recognize that different technologies are featured in (B) and (C) yet answer both questions the same way because of the questions' heavy reliance on the application context, or they might feel the need to respond differently just because they are two different questions.

Not all application-specific questions distract from the nature of a technology. Applications that are unique, or paradigmatic, of a particular technology can help to distinguish those technologies from other technologies. For example, AWS can be used to find and kill targets without direct human intervention. Focusing a question on an application that is unique to a particular technology could help to uncover opinions on the distinguishing characteristics of that technology. As we have argued, focusing on applications that are common to a number of technologies or methods of waging war, for example, guarding against terrorism, or keeping troops out of harm's way, make drawing conclusions about the permissibility of AWS difficult. Thus, we should distinguish between *technology-specific* applications (i.e. those applications that are unique to a particular class of technologies), and *technology-neutral* applications (i.e. those applications that many different classes of technologies could be put to) when thinking about study design.

### **Moral psychological biasing effects**

Different study designs can trigger different, complex, psychological responses in participants. In this section, we focus on two possible *moral* psychological biasing effects that seem particularly relevant in the context of studies on ROWS and AWS: inducing moral emotional responses and anthropomorphic framings. We review some of the moral psychology literature to set up considerations that should be taken into account when polling publics on the development and use of new weapons.

Any bias introduced by moral emotional responses, as opposed to emotional responses generally, could be particularly troublesome in surveys intended to gauge the “public conscience”, precisely because the public conscience is a measure of moral judgment. Measuring the public conscience is all about eliciting moral judgments: are ROWS or AWS *morally permissible*?

However, the link between moral psychology and public polling in the context of weapon systems is relatively new and, understandably, underdeveloped. Our goal in this section is merely to describe some moral psychological effects that we have reason to believe could introduce bias in polling results with the understanding that measuring the effects, and

identifying workable solutions to counteract, or mitigate, such biases, would require more research.

### *Moral Emotional Priming*

Anger is a moral emotion. It is associated with, and often accompanies judgments of *right* and *wrong*. Other moral emotions include shame, disgust and guilt (Haidt 2013, Prinz & Nichols 2010). Perhaps not surprisingly, experiencing these moral emotions tends to affect the quality of our moral judgments.

When we experience anger, for example, the characteristic response is to seek revenge or retribution (Prinz & Nichols 2010). The psychological profile of anger is something like this: Amy witnesses an autonomy violation (i.e. a harm or injustice) that results in an attribution of blame. For example, when watching a news story about Bob harming Chris, Amy blames Bob for the harm. Amy then feels angry at Bob for harming Chris. The result is that Amy wants to see Bob punished for harming Chris (Prinz & Nichols 2010).

A study conducted by Lerner et al. (1998) illustrates this effect. Researchers induced anger in one group of subjects by showing them a video clip in which a bully “humiliates and beats up a teenager” (Lerner et al. 1998, 566). Another group of subjects was shown a video depicting abstract figures, known to elicit no emotional response. Subjects in both groups were then shown another video in which a person’s negligence led to an injury. Subjects were asked how harshly the negligent person should be punished. Subjects who were *primed* in the anger response group were more punitive in their judgments than those in the emotion-neutral group. Similar anger responses can be evoked merely by asking people to recall an event in which they were treated unjustly (Mikula 1986).

Moral emotional priming effects can result when moral emotions other than anger are elicited. In another study that used disgust as the moral emotional variable, two groups of participants were shown videos depicting moral transgressions (e.g. harms), and were then asked to judge the transgressions they had just witnessed. Participants sitting at a dirty table while watching the videotapes judged the transgressions more harshly than participants sitting at clean tables (Schnall et al. 2008).

Thus, eliciting moral emotional responses in a design setup can have a noticeable biasing effect on participants’ moral judgments. The effect is often subtle and intuitive, as demonstrated in the above studies, though statistically significant. That is, the participant need not be aware of the emotional response for it to have an effect.

Moral psychological research suggests, therefore, that we must be aware of moral emotional priming in study design. Application-specific questions that prime participants to think about things that make them angry or disgusted could bias answers by inducing the corresponding

emotional response in the participant: in the case of anger, the participant could form a desire for revenge or retribution (Prinz & Nichols 2010).

This relates to a common bias that may be at work in polling people about ROWS and AWS. For example, a question mentioning terrorism or specific people considered “enemies”, such as question (B) above, could induce an anger or disgust response. The effects could bias answers toward developing and using those weapons as a means of retribution, regardless of the kind of weapon in question. This could also be true of questions focused on saving soldiers’ lives, or protecting them from threats. Participants who experience an anger response to the thought of “friendly” soldiers dying at the hands of enemy soldiers or foreign nationals could be similarly biased towards developing and using whatever weapon might prevent what they might perceive as an “unjust” state of affairs.

### *Anthropomorphic Framings*

We have known for some time now that humans have a strong tendency to anthropomorphize robots (Johansson 1973; Weizenbaum 1976; Duffy 2003; Bartneck et al. 2007; Kahn et al. 2011; 2012; Darling 2015). That means we tend to attribute human-like properties to non-human entities in the world. We’ll say, for example, that when the little colour pinwheel on our computer screen is spinning, our computer is “thinking”, or that when two objects move close or follow each other’s motion that they are attracted to one another, or that when a computer-generated opponent in a video game shoots at us “he is trying to kill me”. Strictly speaking, the computer is not thinking, nor does the computer-generated opponent “want” anything at all. Yet, this psychological tendency is now well demonstrated.

Kahn et al. (2012) invited participants to interact with a *social robot* named Robovie, in order to determine to what extent people would anthropomorphize and attribute moral blame to Robovie. Participants’ interactions with Robovie started with a long conversation, during which Robovie asked questions, paid compliments and attempted jokes. This long interaction was intended to *frame* Robovie as a social entity—it was an anthropomorphic framing in which Robovie was presented as having human-like social abilities. Ultimately, each interaction led to a situation in which Robovie *mistakenly* denied each participant a \$20 “prize” that they were promised. Researchers debriefed participants in order to assess their anthropomorphic responses to Robovie. All participants interacted socially with Robovie (e.g. answered questions, responding to requests, etc.), and though they held Robovie less responsible than humans for being denied the prize, participants held Robovie more accountable than a vending machine (Kahn et al. 2012). When asked if Robovie was conscious, 50% answered “yes”, 73% indicated they believed Robovie could think, and 63% thought Robovie could be trusted (Kahn et al. 2012, 37).

In another experiment involving a social robot, Bartneck et al. (2007) teamed each participant up with a robot to play a game of Mastermind.<sup>7</sup> The researchers varied the robot's perceived level of intelligence by varying the quality of its gameplay suggestions: the "smart" version of the robot made good suggestions to its human partner (the participant), while the "stupid" version made unsuccessful suggestions (Bartneck et al. 2007, 218). The researchers also varied the robot's agreeableness, that is, the "agreeable" version of the robot was patient and polite to its human partner (the participant), while the "non agreeable" version insisted that it was its turn. After the game of Mastermind, the researchers asked the participant to turn the robot off, at which point the robot began begging to be left on. Participants paired with the "agreeable" and "smart" robot hesitated significantly longer when asked to turn the robot off compared to those participants paired with the "stupid" and "non agreeable" robot. Bartneck et al. (2007, 221) conclude that intelligent robots are "perceived to be more alive".

Darling (2015) conducted a similar study. The study involved three separate groups of participants. For each group Darling asked participants to observe a small bug-like robot moving about, then crush it with a mallet, and she measured their hesitation prior to doing so. Each group was introduced to the little robots with a different framing narrative. One group of participants, the non-framing condition, received no framing. However, the other two groups were both subject to anthropomorphic framing during their introduction to the robots, that is, they were told a story about the little robots. For one of those groups of participants Darling framed the introduction using a "personified backstory" in which the robot has a name, "lives in the lab", "likes to play", and so on. With the other group Darling framed the introduction to the robot using a "non-personified" framing, but one that "lent itself to anthropomorphic projection", in which the robot "gets around but doesn't go too far. Last week, though, it got out of the building". Participants in the anthropomorphic framing condition hesitated prior to striking the robots significantly longer than those in the non-framing condition. According to Darling, the longer the hesitation, the more emotionally engaged the participant. Darling (2015) concludes that subjecting participants to anthropomorphic framing conditions increases their empathetic response to robots.

These studies have direct implications for the design of studies involving ROWS and AWS. Questions that cast robots in social roles, such as *guardians*, *protectors*, *ethical agents* or *decision-makers*, could inadvertently induce empathetic, or other anthropomorphic, responses in participants. For example, (B), which mentions *guarding* against terrorism, could have an anthropomorphic framing effect by casting the robot as guardian, and could therefore cause participants to feel sympathy toward the robot, thus biasing their responses. This effect could also be triggered in a survey containing questions both about deploying soldiers to accomplish a task, and deploying robots to accomplish similar tasks. Questions about soldiers could inadvertently have anthropomorphic framing effects that spill over into participants' reactions to questions about robots—participants might imagine robots as soldiers in such cases.

---

<sup>7</sup> Mastermind is a game that involves one player trying to guess the correct sequence of coloured "pegs" that has been arranged by the opponent (see [https://en.wikipedia.org/wiki/Mastermind\\_\(board\\_game\)](https://en.wikipedia.org/wiki/Mastermind_(board_game))).

Empirical evidence from moral psychology and human-robot interaction, like that mentioned here, is relevant to study design involving military robots. However, more work is required to confirm whether or not biasing effects are significant or problematic in practice, how biases are introduced into poll responses, and how we might control for problematic biasing effects.

### **Existing surveys on ROWS and AWS<sup>8</sup>**

Overall, few public opinion polls have been conducted to gauge support for ROWS or AWS. In this section, we examine existing surveys through the lens of the above discussion on application-specific and application-neutral questions, and moral psychological biasing effects. Our goal here is to survey the kinds of questions that have been used in public engagement polls, and comment on potential biasing effects that might be present as a result of study design.

#### *Washington Post-ABC News Poll (2012):*

In a large survey conducted in 2012, the Washington Post and ABC News (WP-ABC) posed two questions related to the use of “drones” (in this case ROWS) for killing foreign terrorist suspects, and American terrorist suspect living abroad. They asked participants to indicate their level of approval of “The use of unmanned ‘drone’ aircraft against terrorist suspects overseas.” In a follow up question directed at participants who approved of the use of drones for such strikes, they asked “What if those suspected terrorists are American citizens living in other countries? In that case do you approve or disapprove of the use of drones?”

Both of these questions are application-specific and the applications mentioned in them are technology-neutral. As such, these questions might not provide any useful information about the use of drones specifically, since one could answer the question by relying only on an opinion about killing foreign terrorist suspects—the fact that the question mentions drones as the weapon of choice could fade into the background. Similarly, one could answer the follow-up question based only on an opinion about killing American citizens who are also terrorist suspects living overseas, regardless of the weapon of choice.

The Washington Post – ABC questions also seem to prime participants along moral emotional lines. Both questions mention “terrorists”, which seems to run the risk of inducing anger or disgust responses.

Whether or not these biasing effects actually played a significant role in the results is an open question. What is clear, however, is that the Washington Post interpreted the results as revealing something specifically about Americans’ opinion on the *nature* of drone technology:

---

<sup>8</sup> Note: When reviewing each survey we focus only on those questions and aspect of the surveys that contained reference to ROWS or AWS, or were directly related, and ignore demographic questions or other questions not directly addressing ROWS and AWS.

the title of the report was, “The American Public Loves Drones” (Cillizza 2013). We think their conclusion is a misinterpretation of the data. The article might more accurately have been titled, “The American Public Loves Eliminating Terrorist Suspects Living Abroad.”

### *PublicMind*

As part of a PublicMind (2013) survey, Woolley and Jenkins asked participants four questions intended to gauge their opinion on using “drones” (in this case ROWS) to attack foreign nationals and American citizens living abroad, who have been deemed a threat to the US. The survey is therefore application-specific, as it is about a very specific use of drones that could be accomplished by other means. Here are the questions:

1. You probably know that the United States uses remotely controlled aircraft called “drones” to spy on and sometimes attack people and other kinds of targets around the world. How much have you heard or read about these drones?
2. To the best of your knowledge, can the U.S. target U.S. citizens living in other countries with drones, or is that illegal?
3. In general, do you approve or disapprove of the U.S. Military using drones to carry out attacks abroad on people and other targets deemed a threat to the U.S.?
4. In general, do you approve or disapprove of the C.I.A. using drones to carry out attacks abroad on people and other targets deemed a threat to the U.S.?

Questions (2)-(4) provide the most information about participants’ opinions on the use of drones for attacking individuals. Though questions (3) and (4) are application-specific, they seem to be designed to minimize any moral emotional priming. “People and other targets deemed a threat to the US” is a relatively emotionally neutral language, compared to the Washington Post survey’s use of “terrorist suspects”. This could have the effect of controlling for bias induced by the nature of the application. It could also have the effect of underscoring the role that drones are playing in the survey, that is, it could help participants focus on the use of drones as the means to carry out the application.

At the same time, questions (2)-(4) could potentially be answered without much knowledge of drones, owing to their application-specific, and technology-neutral, nature.

Question (1) assesses participants’ knowledge of the technology, and could also have the effect of helping researchers understand how emotionally primed the participants are, thus controlling for bias. For example, if a participant has read or heard a lot about drones, and most popular media accounts of drones involve emotionally engaging language, researchers might assume something about that participant’s emotional reaction to drones in general. However, a better way of controlling for bias could be to measure their level of emotional engagement while thinking about drones. Of course, this would likely complicate the experimental set-up, making it difficult or costly to carry out.

### *Pew Research Center*

In 2013, the Pew Research Center posed the following question about “drones” (in this case ROWS) to participants in twenty countries:

1. Do you approve or disapprove of the United States conducting missile strikes from pilotless aircraft called drones to target extremists in countries such as Pakistan, Yemen and Somalia?

This question is highly application-specific and technology-neutral, and we can predict that it also induces moral emotional responses, and anthropomorphic framing effects, in participants because of the specific mention of the US as attacker, and of extremists as targets. In this question, the specific use of drones in carrying out the objective could quickly fade into the background. We suspect it reveals less about public opinion on ROWS, and much more about people’s sympathies towards US foreign policy and “extremists in countries such as Pakistan, Yemen and Somalia”.

#### *Huffington Post – YouGov Polls*

Two separate Huffington Post/YouGov Omnibus surveys were administered in early 2013 in order to gauge public opinion on the use of “drones” (in this case ROWS). The questions read as follows (Swanson 2013):

YouGov Omnibus Poll, Jan 10-11 2013:

1. Do you approve or disapprove of the Obama administration using drones to kill high-level terrorism suspects overseas?
2. Do you approve or disapprove of the Obama administration using drones to kill high-level terrorism suspects overseas, even if those suspects are American citizens?

YouGov Omnibus Poll, Feb 6-7 2013:

1. Do you approve or disapprove of the Obama administration using drones to kill high-level terrorism suspects overseas?
2. Do you approve or disapprove of the Obama administration using drones to kill high-level terrorism suspects overseas, even if those suspects are American citizens?
3. Do you approve or disapprove of the Obama administration using drones to kill high-level terrorism suspects overseas, even if innocent civilians may also be killed?
4. Which comes closest to your opinion about who should be permitted to order drone strikes against suspected terrorists who are US citizens?
5. Who do you think should be responsible for setting the rules for when the CIA or U.S. military may use drone strikes on American citizens suspected of being terrorists?

Both of these polls use application-specific, technology-neutral, questions with language that likely induces moral emotional responses in participants. We suspect these surveys contribute little to the debate on ROWS v. AWS (see the above analyses on the Washington Post, Public Mind and Pew Research Center surveys for more detail).

*Charli Carpenter*

As part of a different YouGov Omnibus survey, Charli Carpenter (2013, 2014) posed a series of questions to try to gauge differences in opinion between ROWS and AWS. Carpenter was explicitly interested in using her study to help satisfy the requirements set out in the Martens Clause. She posed the following two questions:

1. Drones are remotely piloted by a human controller, but several nations are developing robotic weapons that can independently make targeting and firing decisions without a human in the loop. How do you feel about the trend toward using completely autonomous [robotic weapons/lethal robots] in war?
2. Some non-profit citizen groups have launched [a campaign to ban the use of fully autonomous weapons / the Campaign to Stop Killer Robots], which seeks a global treaty requiring human involvement in all decisions to take human life. To what extent would you support or oppose such a campaign?

Carpenter randomly varied the language in both questions, to create two treatment groups. One group was presented with a version of the questions containing the “robotic weapons/fully autonomous weapons” wording pair, while the other group responded to a version of the questions containing the “lethal robots/killer robots” wording pair. These variations were meant to test if participants had a different response based on the nature of the robot descriptions being used. This study design choice is useful in its ability to expose biases that might be introduced due to anthropomorphic framing (casting the robot as “killer”), and moral emotional priming (a moral aversion to “killing”). However, the difference in wording had no effect on the results. We interpret this finding as an indication that there were likely no significant differences in moral emotional responses between treatment groups. Perhaps this finding is not too surprising, since the word “killer” is buried in the larger phrase “The Campaign to Stop Killer Robots,” and the word “lethal” can be used in morally neutral contexts (e.g. “the lethal dose of substance x is...”).

Question (1) is application-neutral and, as such, it contains a preamble to explain the technology to participants. As discussed above in relation to (A), questions that are application-neutral cannot be answered adequately unless the participant has an understanding of the technology in question, hence the need for a preamble, albeit a short one in this case.

Aside from the variations in wording between treatment groups, there are no obvious moral psychological biasing worries raised by the wording of (1) or (2). We consider this survey well designed and useful in the context of the debate over ROWS and AWS, at least insofar as it avoids application-specific, technology-specific language, and also avoids any obvious moral psychological biasing traps.

*Michael C Horowitz*

In two separate but related public opinion surveys conducted in 2015, Horowitz (2016a,b) claims to have improved on Carpenter’s study, which he criticizes for asking questions about

AWS “in a vacuum”. His argument is, first, that asking questions about ROWS or AWS directly “makes it hard to distinguish support or opposition to weapons in general from autonomous weapons in particular” (Horowitz 2016a, 4). Furthermore, asking questions about AWS without reference to “particular scenarios and contexts for usage and/or development”, he says, might induce Terminator Bias, since participants might imagine “*The Terminator*, *The Matrix*, or other portrayals in the media” when responding (Horowitz 2016a, 4). Questions in a vacuum, he concludes, are good for establishing a baseline, but you need to introduce particular use scenarios (applications) in order to overcome these weaknesses.

In his studies he asked a baseline question similar to Carpenter’s:

1. **Baseline Condition:** Drones are remotely piloted by a human controller, but autonomous weapon systems are robotic systems that, once activated, can independently make targeting and firing decisions without a human in the loop. Would you approve or disapprove of the United States developing autonomous weapon systems?

Question (1) is a slight variation on Carpenter’s question (see above), subject to a similar analysis.

However, he immediately follows with several questions, each one introducing particular scenarios. For example, he poses the following additional question, which he calls the “Protect + More Effective Condition” (Horowitz 2016a):

2. **Protect + More Effective Condition:** Drones are remotely piloted by a human controller, but autonomous weapon systems are robotic systems that, once activated, can independently make targeting and firing decisions without a human in the loop. Would you approve or disapprove of the United States developing autonomous weapon systems if:
  - a. Autonomous weapon systems would be used to protect US military personnel on naval ships and military bases from attacks by foreign militaries and militant groups
  - b. Autonomous weapon systems would be more effective than other options at protecting US military personnel on naval ships and military bases from attack by foreign militaries and militant groups

Question (2)(a) is an application-specific, technology neutral question, subject to the same criticisms previously mentioned. It also casts the technology as “protector”, making it susceptible to anthropomorphic framing effects, and subjecting it to criticisms already mentioned.

Question (2)(b) has the same features as (2)(a), with the addition of a technology-specific feature: the AWS being more effective than all other weapon options. Carpenter (2016) criticizes Horowitz for the addition of this condition claiming it biases the study in favor of killer robots, since all of the conditions align with pros of AWS. Horowitz’ (2016 b) first experiment of the two includes the following conditions: the one stated in (2)(b); an “Attack + More Effective Condition”; an “Attack + Not More Effective Condition”; and a “Protect + Not More Effective Condition”. As Carpenter (2016) argues, an “objective ‘contextual’” study would include many

more conditions, including scenarios that prime participants to think of the cons of the technology that have been widely discussed in the literature on banning AWS. Though a more balanced contextual study might have the effects Carpenter envisions, we feel that contextual studies are generally problematic from the start since they are susceptible to too many biasing effects, such as moral psychological biasing effects.

There is an important point worth mentioning here. Though we agree with Carpenter's assessment of the bias in Horowitz' choice of conditions, we do not think all of the features of his study conditions are as problematic as Carpenter claims. One technology-specific condition—AWS being more effective than all other military options—escapes some criticism. Recall that a technology-specific feature provides information that can focus the participant's attention on aspects of a technology not shared by other technologies. Stating that a technology is more effective than all other military options is a feature that distinguishes it from those other technologies, in a way that being “no more”, or “less effective than all other military options” do not. Furthermore, stating that a technology is either “no more” or “less effective than all other military options” could help to focus participants on a feature that is morally relevant in the context of the Martens Clause: the kind of feature that has consequentialist relevance. However, more research is necessary to understand if the “no more” or “less effective” conditions would have this effect.

One important difference between Carpenter's and Horowitz' studies is the general research methods applied: Carpenter's (2013, 2014) study included a qualitative component in which she asked participants why they liked or disliked the idea of autonomous weapons, while Horowitz' was strictly quantitative. This is a significant difference that Carpenter mentions in her critique of Horowitz' work. A qualitative approach using open-ended questions allows participants to define the scenarios and contexts that are relevant to the moral discussion (Carpenter 2016). As such, Carpenter argues that a qualitative study design provides a more direct route to the public conscience, which is the heart of the Martens Clause.

#### *Moon, Danielson, and Van der Loos*

In contrast to the abovementioned poll and survey questions that are designed to collect quantitative metrics on the public perception of ROWS and AWS (with the noted exception of Carpenter 2013, 2014), there has been an effort to get a hybrid of quantitative and qualitative responses using a simple, yet unique survey platform. Moon, Danielson, and Van der Loos (Moon et al., 2012) conducted a study where the participants were provided with a preamble about a particular technology in question, followed by a question they could answer with a “Yes”, “No”, or “Neutral” response as well as a qualitative, text-based reasons for their response. The participants—instead of contributing reasons for their response—also had the option to vote for an existing response that other participants have contributed to the discussion. In a series of nine questions pertaining to various roboethics issues, the first and second questions concerned ROWS and AWS technologies respectively.

The preamble given for both ROWS and AWS questions were from the same report (Lin et al. 2008). For the ROWS question, the preamble and the question read:

“The Predator is a remote controlled aerial robot. “These robots have dual applications: they can be used for reconnaissance without endangering human pilots, and they can carry missiles and other weapons. ... Predators are used extensively in Afghanistan. They can navigate autonomously toward targets specified by GPS coordinates, but a remote operator located in Nevada (or in Germany) makes the final decision to release the missiles.

Should remote controlled Predators be armed with lethal weapons in combat?”

Subsequently, the participants were given a quotation from a report on AWS:

“A proposed modification of the Predator remote controlled aircraft would allow it to select targets autonomously, without the intervention of human operators. “The technology to (responsibly) create fully autonomous robots is near but not quite in hand.”

Should fully autonomous lethally armed Predators be developed?”

While the results of this study is subject to order bias (the question order was not randomized), it is also subject to bias in the language used in the preamble—in particular, the mention of “without endangering human pilots” and its contrast with the description of the technology being used in Afghanistan. The introduction to AWS, on the other hand, remains neutral in its wording. Given the question’s reliance on the preamble of the previous question, it is likely that the participants assumed the same application context in responding to the AWS question.

Keeping the biasing effect that potentially applies to this study, though, the mapping of quantitative responses with qualitative reasons given by the participants provided a richer understanding of the moral psychology at work when participants were thinking through the two different technologies. For example, the authors found that the participants were giving much more rational and consequentialist reasons for supporting or rejecting ROWS. In contrast, the same participants who were in support of ROWS for functional or practical rationales (e.g., effectiveness of the technology, consideration for physical safety of soldiers) shifted to strongly normative reasoning when rejecting AWS: that humans should always make life/death decisions.

Such a shift in the nature of qualitative responses—though analyze qualitative surveys adds cost over quantitative studies—can be powerful in revealing nuanced, yet significant differences nature of weapons technologies can have on participants’ moral psychology. The particular unique survey platform itself also has limitations, though a discussion of that is beyond the scope of this paper. Nonetheless, the qualitative set of reasons collected from this study helped inform the design of a larger study conducted by the Open Roboethics initiative.

*Open Roboethics initiative (ORi)*

ORi (2015) conducted an international online public opinion poll in 2015 to gauge public opinion on the use and development of ROWS and AWS. The survey contained six questions about ROWS and AWS, and each question was accompanied by a brief preamble reminding participants of the terminology.

The preamble read (ORi 2015):

**Terminology:**

**ROWS (Remotely Operated Weapon systems):** weaponized systems in which *a person in a remote location* makes the decision to use lethal force.

**LAWS (Lethal Autonomous Weapon systems):** weaponized systems in which *the system, without requiring human intervention,* makes the decision to use lethal force.

The questions read as follows:

1. Consider **Lethal Autonomous Weapon systems (LAWS)** technology becoming available to your country in the near future. If your country goes to war against another country, would you support the use of LAWS over **Remotely Operated Weapon systems (ROWS)**? With LAWS, the system will make the decision to use lethal force without human input, whereas ROWS requires a person to make the decision for it.
2. Hypothetically, if **Lethal Autonomous Weapon systems (LAWS)** technology becomes available to a country that is attacking your country, would you rather be under attack by LAWS than **Remotely Operated Weapon systems (ROWS)**? With LAWS, the system will make the decision to use lethal force without human input, whereas ROWS requires a person to make the decision for it.
3. If certain types of LAWS are to be internationally banned from development and use, what types of LAWS do you feel most strongly should be banned? (Choose one)
  - a. air (e.g., drones, fighter pilots)
  - b. sea (e.g., ships, submarines)
  - c. land (e.g., infantry)
  - d. all of the above should be banned
  - e. none of the above should be banned
4. What do you think is the main reason for supporting the development and use of LAWS in battlefields? (Choose one)
  - a. Cost of war will be cheaper to use LAWS than ROWS
  - b. Autonomous machines will make more ethical life/death decisions than humans
  - c. LAWS will save human military personnel from psychological harm of war, such as post-traumatic stress disorder (PTSD)
  - d. LAWS will save human military personnel from physical harm of war
  - e. Development of LAWS will lead to the development of useful, non-military technologies
  - f. There are no valid reasons for developing and using LAWS over ROWS
  - g. Other \_\_\_\_\_
5. What do you think is the main reason for rejecting the development and use of LAWS in battlefields? (Choose one)
  - a. The risk of the technology falling into the wrong hands is too big
  - b. Humans should always be the one to make life/death decisions (i.e., it is wrong for machines to make the decision)

- c. It is uncertain who will be responsible when things go wrong
  - d. LAWS will kill more lives than it will save
  - e. It is doubtful that LAWS technology of the near future will be technically robust and reliable enough to be trusted
  - f. There are no valid reasons for rejecting the development and use of LAWS over ROWS
  - g. Other \_\_\_\_\_
6. In general, should LAWS that make life/death decisions without human operator be developed or used?
- a. Yes, LAWS should be developed and used for both defense and offense purposes
  - b. Yes, but LAWS should be developed and used for defense purposes *only*
  - c. LAWS should be developed, but never used
  - d. No, LAWS should **not** be developed and used
  - e. Other: \_\_\_\_\_

Questions (1) and (2) are application-neutral questions, containing no obvious moral psychological biasing traps. Those questions were asked prior to the others specifically to avoid any biasing effects that could spillover from the contextual options contained in (4)-(5).

Questions (4)-(5) contain specific mention of value-laden aspects associated with the development and use of AWS, which could result in moral psychological biasing effects. These questions were also always asked prior to (6), which could result in spillover effects from any bias introduced in (4)-(5). However, the value-laden aspects mentioned in (4)-(5) are intended to elicit rational justifications for using/banning the technology. Though a qualitative design approach would have been more effective, as Carpenter (2016) points out, this poll was intended for an international audience and, as such, practical translation limitations eliminated the possibility of open-ended questions. A quantitative approach requires some mention of the possible answers, and these choices represent arguments featured in the AWS literature (e.g. PAX 2014). Ideally, these questions should be asked last to avoid biasing effects.

### *Moshkina and Arkin*

As a descriptive-explanatory study, the scope of Moshkina and Arkin's (2007) survey was extensive, and widest across the similar surveys on the topic of AWS. For example, Moshkina and Arkin examined the specific roles weaponized robotic systems could take on, such as crowd/mob control, sentry, prison guard, hostage rescue, reconnaissance, direct combat. Demographic groups also included robotics researchers, policymakers, the military, and the general public. They also explored the stakeholder perception across different situations: open warfare with war on foreign territory, war on home territory, covert operations on foreign territory, and covert operations on home territory. However, one aspect of their study deserves mention because it is unique to their study and applies generally to their preamble, and to many of the questions they asked.

As a general approach, for each point of interest Moshkina and Arkin (2007) ask three consecutive questions: the first deals with human soldiers performing task X, the second deals with humans and robots as extensions of humans performing task X, the third deals with

autonomous robots performing task X. Here is the preamble followed by an example of a set of three such questions:

Preamble:

For the rest of the survey, we will be using some terms that should be defined the same for everyone. Please look through them and try to think of them whenever we use them in the questions. They will be repeated at the top of each page for your convenience. These terms are:

- Robot: as defined for this survey, an automated machine or vehicle, capable of independent perception, reasoning, and action
- Robot acting as an extension of a human soldier: a robot under the direct authority of a human, including authority over the use of lethal force
- Autonomous robot: a robot that does not require direct human involvement, except for high-level mission tasking; such a robot can make its own decisions consistent with its mission without requiring direct human authorization, including decisions regarding the use of lethal force

Set of three questions:

1. To what extent do you agree or disagree that it is acceptable for a human soldier to take a human life, in a manner consistent with the existing laws of war, in Each of the following situations: [list of cases]
2. To what extent do you agree or disagree that it is acceptable for a robot acting as an extension of a human soldier to take a human life, in a manner consistent with the existing laws of war, in Each of the following situations: [list of cases]
3. To what extent do you agree or disagree that it is acceptable for an autonomous robot to take a human life, in a manner consistent with the existing laws of war, in Each of the following situations: [list of cases]

In all such cases we feel that the first question could bias participants by framing the issues in the context of human soldiers, thus inducing an anthropomorphic response. Indeed, the three questions step participants through a gradual distancing, keeping the human soldier in view until the third question. By the time participants encounter the autonomous robot they could very well be thinking of it more like a human soldier than they would otherwise have. Of course, more work would need to be done to test these concerns.

### Concluding remarks

It will take some time to settle the debate over the ethics and governance of autonomous weapon systems. Attempts to gauge the public conscience, in the spirit of the Martens Clause, will play a role in settling that debate. Though there are numerous public polls and surveys that have been conducted to measure the public beliefs about ROWS and AWS, many of the polls are subject to biases either well known in the survey design literature (Carpenter 2016), or discussed in this paper, namely, application-specific and technology-neutral questioning, moral emotional priming, and anthropomorphic framing. It is important to engage the public, but we must be aware of the pitfalls highly applicable to poll/survey questions on this topic. Only then,

we will be able to have a meaningful discussion with the public and effectively inform governance decisions on AWS.

It is also worth noting that all but two of the surveys examined here were conducted in the U.S. alone.<sup>9</sup> AWS are an international issue. Their effects will spill over borders, affecting different publics to different degrees. It is therefore crucial that we look internationally when conducting public polling on this, and other, arms control issues. The data we use to frame conclusions about which weapons are acceptable, and which are not, ought to be representative of those who stand to confront those weapons in conflict. AWS promises to leave very few people untouched. Therefore, the public conscience, as mentioned in the Martens Clause, cannot be interpreted as applying only to single nation states. Researchers should, whenever possible, work to design their studies well, and gather data from beyond their own borders, to help paint as accurate a picture of the global moral landscape as is possible.

---

<sup>9</sup> The exceptions are Pew Research Center (2013), and the Open Roboethics initiative (2015).

## References

- Article 36. (2015). *Killing By Machine: Key Issues for Understanding Meaningful Human Control*.  
Online: <http://www.article36.org/autonomous-weapons/killing-by-machine-key-issues-for-understanding-meaningful-human-control/>
- Arkin, R.C. (2015). "Warfighting Robots Could Reduce Civilian Casualties, so Calling for a Ban Now Is Premature," *IEEE Spectrum* (5 Aug). Online:  
<http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/autonomous-robotic-weapons-could-reduce-civilian-casualties>
- Bartneck, C., van der Hoek, M., Mubin, O., Al Mahmud, A. (2007). "'Daisy, Daisy, Give Me Your Answer Do!': Switching Off a Robot." *HRI'07 Proceedings*:217-222.
- Brant, T. (2016). "U.S. Military Eyes Autonomous Weapons." *PCMAG.com* (4 Mar.). Online:  
<http://www.pcmag.com/article2/0,2817,2500351,00.asp>
- Carpenter, C. (2013). "How Do Americans Feel About Fully Autonomous Weapons?"  
*DuckofMinerva.com* (19 Jun.). Online: <http://duckofminerva.com/2013/06/how-do-americans-feel-about-fully-autonomous-weapons.html>
- . (2014). "Who's afraid of killer robots? (and why)." *The Washington Post* (30 May). Online:  
<https://www.washingtonpost.com/news/monkey-cage/wp/2014/05/30/whos-afraid-of-killer-robots-and-why/>
- . (2016). "How (Not) to Measure the 'Public Conscience'." *DuckofMinerva.com* (7 Mar.).  
Online: <http://duckofminerva.com/2016/03/how-not-to-measure-the-public-conscience.html>
- Cassese, A. (2000). "The Martens Clause: Half a Loaf or Simply Pie in the Sky?" *European Journal of International Law* 11(1):187-216.
- Cillizza, C. (2013). "The American Public Loves Drones." *Washingtonpost.com*. Online:  
<https://www.washingtonpost.com/news/the-fix/wp/2013/02/06/the-american-public-loves-drones/>
- Darling, K. (2015). "Who's Johnny?"
- Devlin, H. (2016). "Killer Robots and Digital Doctors: How Can We Protect Society from AI?" *The Guardian* (16 Mar.). Online:  
<http://www.theguardian.com/technology/2016/mar/15/killer-robots-driverless-cars-alpha-and-the-social-impact-artificial-intelligence>.
- Duffy, B.R. (2003). "Anthropomorphism and the Social Robot." *Robotics and Autonomous Systems* 42:177-190.

- Evelith, R. (2014). "So What Exactly Is A Killer Robot?" *The Atlantic* (20 Aug.). Online: <http://www.theatlantic.com/technology/archive/2014/08/calling-autonomous-weapons-killer-robots-is-genius/378799/>
- Gibbs, S. (2015). "Musk, Wozniak and Hawking Urge Ban On Warfare AI and Autonomous Weapons." *The Guardian* (27 July). Online: <http://www.theguardian.com/technology/2015/jul/27/musk-wozniak-hawking-ban-ai-autonomous-weapons>
- Haidt, J., (2013). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. (Vintage Books).
- Horowitz, M., Scharre, P. (2015). *Meaningful Human Control in Weapon Systems: A Primer*. Centre for a New American Security. Online: [http://www.cnas.org/sites/default/files/publications-pdf/Ethical\\_Autonomy\\_Working\\_Paper\\_031315.pdf](http://www.cnas.org/sites/default/files/publications-pdf/Ethical_Autonomy_Working_Paper_031315.pdf)
- Horowitz, M.C. (2016a). "Public Opinion and the Politics of the Killer Robots Debate." *Research & Politics*. Online: <http://rap.sagepub.com/content/3/1/2053168015627183>
- . (2016b). "Online Appendix for Public Opinion and the Politics of the Killer Robots Debate." *Research & Politics*. Online: [http://rap.sagepub.com/content/srap/suppl/2016/02/16/3.1.2053168015627183.DC1/Supplementary\\_material.pdf](http://rap.sagepub.com/content/srap/suppl/2016/02/16/3.1.2053168015627183.DC1/Supplementary_material.pdf)
- Human Rights Watch (2012) *Losing humanity: The case against killer robots*. Online: [http://www.hrw.org/sites/default/files/reports/arms1112\\_ForUpload.pdf](http://www.hrw.org/sites/default/files/reports/arms1112_ForUpload.pdf)
- . (2015). *Mind the Gap: The Lack of Accountability for Killer Robots*. Online: <https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots>.
- International Committee of the Red Cross (ICRC). (2009). "The Geneva Conventions of 1949: Origins and Current Significance." ICRC. Online: <https://www.icrc.org/eng/resources/documents/statement/geneva-conventions-statement-120809.htm>
- . (2015). *Autonomous Weapon Systems: Is It Morally Acceptable for a Machine to Make Life and Death Decisions?* (13 Apr.). Online: <https://www.icrc.org/en/document/lethal-autonomous-weapons-systems-LAWS>
- International Justice Resource Centre. (2016). "International Humanitarian Law." Online: <http://www.ijrcenter.org/international-humanitarian-law/>
- Johansson, G. (1973). "Visual perception of biological motion and a model for its analysis." *Perception & Psychophysics* 14(2):201–211.

- Kahn, P.H. et al. (2011). "The New Ontological Category Hypothesis in Human-Robot Interaction." *Proceedings of the 6<sup>th</sup> International Conference on Human-Robot Interaction*:159-160.
- . (2012). "Do People Hold a Humanoid Robot Morally Accountable for the Harm It Causes?" *HRI'12 Proceedings*:33-40.
- Kreps, S. (2014). "Flying Under the Radar: A Study of Public Attitudes Towards Unmanned Aerial Vehicles." *Research and Politics* (April-June):1-7.
- Lerner, J.S., Goldberg, J.H., Tetlock, P.E. (1998). "Sober Second Thought: The Effects of Accountability, Anger, and authoritarianism on attributions of responsibility." *Personality and Social Psychology Bulletin* 24:563-574.
- Lin, P., Bekey, G., Abney, K. (2008). *Autonomous Military Robots: Risk, Ethics, and Design. A report commissioned by U.S. Department of Navy/Office of Naval Research*. Online: [ethics.calpoly.edu/ONR\\_report.pdf](http://ethics.calpoly.edu/ONR_report.pdf).
- McDonald, G. (2016). "Armed Robots Could Decide to Pull Trigger." *Discovery News* (8 Mar.). Online: <http://news.discovery.com/tech/robotics/armed-robots-could-decide-to-pull-trigger-160307.htm>
- Mikula, G. (1986). "The Experience of Injustice: Toward a Better Understanding of Its Phenomenology." In H.W. Bierhoff, R.L. Cohen & J. Greenberg (Eds.) *Justice and Social Relations*. (New York: Plenum):103-124.
- Millar, J. (2015). "Meaningful Human Control and Dual-Use Technologies." UN Informal Meeting of Experts on Lethal Autonomous Weapons (LAWs) under the framework of the Convention on Certain Conventional Weapons (CCW), April 13-17, 2015, Geneva, Switzerland. Online: [http://www.unog.ch/80256EDD006B8954/\(httpAssets\)/F483D421E67D230FC1257E2F0033E690/\\$file/Jason+Millar+-+Meaningful+Human+Control+and+Dual-Use+Technology.pdf](http://www.unog.ch/80256EDD006B8954/(httpAssets)/F483D421E67D230FC1257E2F0033E690/$file/Jason+Millar+-+Meaningful+Human+Control+and+Dual-Use+Technology.pdf)
- Moon, A., Danielson, P., Van der Loos, H.F.M. (2012). "Survey-Based Discussions on Morally Contentious Applications of Interactive Robotics." *International Journal of Social Robotics* 4(1):77-96.
- Moshkina, L., Arkin, R.C. (2007). *Lethality and Autonomous Systems: Survey Design and Results*. Mobile Robot Laboratory, Georgia Institute of Technology. Online: <https://smartech.gatech.edu/handle/1853/20068>
- Open Roboethics initiative (ORi). (2015). "The Ethics and Governance of Lethal Autonomous Weapon systems: An International Poll." Open Roboethics initiative (9 Nov.). Online: [http://www.openroboethics.org/wp-content/uploads/2015/11/ORi\\_LAWS2015.pdf](http://www.openroboethics.org/wp-content/uploads/2015/11/ORi_LAWS2015.pdf)

- Paulhus, D.L. (1991). "Measurement and control of response biases." In J.P. Robinson et al. (Eds.) *Measures of personality and social psychological attitudes*. (San Diego: Academic Press).
- Parkon, S. (2015). "Killer Robots: The Soldiers that Never Sleep." *BBC* (16 Jul.). Online: <http://www.bbc.com/future/story/20150715-killer-robots-the-soldiers-that-never-sleep>.
- PAX. (2014). *Deadly Decisions: 8 Objections to Killer Robots*. Online: <http://www.paxforpeace.nl/stay-informed/news/stop-killer-robots-while-we-still-can>
- Pew Research Center. (2013). "U.S. Use of Drones, Under New Scrutiny, Has Been Widely Opposed Abroad." *Pewresearch.org*. Online: <http://www.pewresearch.org/2013/02/06/u-s-use-of-drones-under-new-scrutiny-has-been-widely-opposed-abroad/>
- . (2016). *U.S. Survey Research*. Online: <http://www.pewresearch.org/methodology/u-s-survey-research/>.
- Prinz, J.J., Nichols, S. (2010). "Moral Emotions." In J.M. Doris (Ed.) *The Moral Psychology Handbook*. (Oxford: Oxford University Press): 111-146.
- Schnall, S., Haidt, J., Clore, G., & Jordan, A. (2008). "Disgust as Embodied Moral Judgment." *Personality and Social Psychology Bulletin* 34:1096-1109.
- Swanson, E. (2013). "Drone Poll Finds Support For Strikes, With Limits." *HuffingtonPost.com* (8 Feb.). Online: [http://www.huffingtonpost.com/2013/02/08/drone-support-poll\\_n\\_2647051.html](http://www.huffingtonpost.com/2013/02/08/drone-support-poll_n_2647051.html)
- Tucker, P. (2014). "Every Country Will Have Armed Drones Within 10 Years." *Defense One* (6 May), Online: <http://www.defenseone.com/technology/2014/05/every-country-will-have-armed-drones-within-ten-years/83878/>
- UNIDIR. (2014). *Framing Discussions on the Weaponization of Increasingly Autonomous Technologies*. Online: <http://www.unidir.org/en/publications/framing-discussions-on-the-weaponization-of-increasingly-autonomous-technologies>
- University of Massachusetts Amherst. (2013). "New Survey Shows Widespread Opposition to 'Killer Robots,' Support for New Ban Campaign." *Newswise* (19 Jun.). Online: <http://www.newswise.com/articles/new-survey-shows-widespread-opposition-to-killer-robots-support-for-new-ban-campaign>
- Washington Post a. (2012). [http://www.washingtonpost.com/wp-srv/politics/polls/postabcpoll\\_020412.html](http://www.washingtonpost.com/wp-srv/politics/polls/postabcpoll_020412.html) (raw data)
- Weizenbaum, J. (1976). *Computer Power and Human Reason: From Judgment to Calculation*. (San Francisco: W.H. Freeman & Company).

Woolley, P., Jenkins, K. (2013). "Public Says It's Illegal to Target Americans Abroad As Some Question CIA Drone Strikes." PublicMind. Online:  
<http://www.publicmind.fdu.edu/2013/drone/>