Article

A Broad View on Robot Self-Defense: Rapid Scoping Review and Cultural Comparison

Martin Cooney [1,2,*], Masahiro Shiomi [2], Eduardo Kochenborger Duarte [1] and Alexey Vinel [1,3]

1   School of Information Technology, Halmstad University, 301 18 Halmstad, Sweden
2   Interaction Science Laboratories (ISL), Advanced Telecommunications Research Institute International (ATR), Kyoto 619-0237, Japan
3   Institute of Applied Informatics and Formal Description Methods (AIFB), Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany
*   Correspondence: martin.daniel.cooney@gmail.com

Abstract: With power comes responsibility: as robots become more advanced and prevalent, the role they will play in human society becomes increasingly important. Given that violence is an important problem, the question emerges if robots could defend people, even if doing so might cause harm to someone. The current study explores the broad context of how people perceive the acceptability of such robot self-defense (RSD) in terms of (1) theory, via a rapid scoping review, and (2) public opinion in two countries. As a result, we summarize and discuss: increasing usage of robots capable of wielding force by law enforcement and military, negativity toward robots, ethics and legal questions (including differences to the well-known trolley problem), control in the presence of potential failures, and practical capabilities that such robots might require. Furthermore, a survey was conducted, indicating that participants accepted the idea of RSD, with some cultural differences. We believe that, while substantial obstacles will need to be overcome to realize RSD, society stands to gain from exploring its possibilities over the longer term, toward supporting human well-being in difficult times.

Keywords: robot self-defense; robot ethics; robot violence; robot crime; technological acceptance; dark side of human–robot interaction (HRI)

1. Introduction

The current article explores the broad picture of *robot self-defense* (RSD). To motivate our work, we start by examining a real example of a horrific attack [1]: On 13 March 1964, a young woman, Kitty Genovese, was stabbed to death in a crowded residential area. The cards were stacked against her: In Kitty's case, she was tired, on her way home after work in the early morning, and unarmed, in contrast to her attacker, who had a deadly weapon, and alleged experience killing two victims before. Over a half-hour, her cries for help were unheard, ignored, or misinterpreted by local residents; some phone calls to the police were unclear and not given high priority. When the killer was apprehended and asked why he had felt so bold, he said, "I knew they wouldn't do anything, people never do". The bottom line is that, in some cases, neither law enforcement, nor bystanders, nor the victim themselves can prevent an attack.

While the shocking details of this case helped to spur the creation of various programs such as the 9-1-1 system, the fundamental problem of violence remains. Victims are typically at a disadvantage; the attacker decides the time and circumstances of the attack, selecting targets they believe to be weaker, and victims' reactions are delayed by the time needed to understand what is going on and what they should do (not to mention any accidental excessive force can result in victims themselves being criminally charged). Bystanders as well must consider their own safety, and, as the saying goes, "when seconds count, the police are only minutes away", so sometimes there is no one who shows up to help on time: numerous recordings of unprevented attacks can be seen on social media [2].

More generally, violence is recognized as a serious and prevalent problem in human societies, leading to thousands of deaths annually, mental problems, reduced productivity, administrative burdens, and weakening attempts to tackle poverty [3–6]. Its cost in 2015 has been estimated at approximately USD 13.6 trillion, or 13.3% of the sum of global gross domestic product.

This article poses the question: What if there were another option? What if, in the future, nearby robots could help? Robots are expected to become increasingly present in human environments where violence can take place: For example, Tung and Jara Santiago Campos mention that markets for professional and consumer service robots have increased in 2021 to USD 6.7 billion and USD 4.4 billion (growth rates of 16% and 12% from 2020), with the market for social robotics projected to grow rapidly in the near future [7]. This trend could possibly also be aided by recent excitement over the next possibilities for Artificial Intelligence (AI) and robotics (e.g., with the introduction in late 2022 of ChatGPT (Chat Generative Pre-trained Transformer) [8,9]). As such robots transition from labs to the homes, schools, stores, and streets of our societies, they will bear witness to not just our daily triumphs, but also our struggles. Robots will be useful in various public and domestic settings, providing care, transport, cleaning, companionship, entertainment, or education, and placing them in a position where they could help. For example, a care robot, Autonomous Vehicle (AV), or cleaning robot, which would not feel tired or cold, might be present at late times when there are few people outside and crimes are more likely to occur. Given that robots can be repaired or replaced, such a robot would also not likely feel fear at having to confront a violent attacker. Thus, in some cases, we believe robots could help. People could be warned (e.g., by digitally sending emergency messages to police or security, acoustically announcing an attack/calling for help/blaring sirens, or visually suggesting danger via lights or a fearful expression). Moreover, crimes could be documented, attackers placated (e.g., by calming, distracting, negotiating and obstructing), and victims allowed to evade or escape. However, a question arises in regard to what a robot should do in difficult situations in which physical harm to someone, attacker or victim, cannot be avoided. For example, when an attacker and victim are pressed together, attacks are fast and deadly, and the victim is not able to stop the attack, find help, or get away. This is the main scenario considered here, shown in Figure 1.
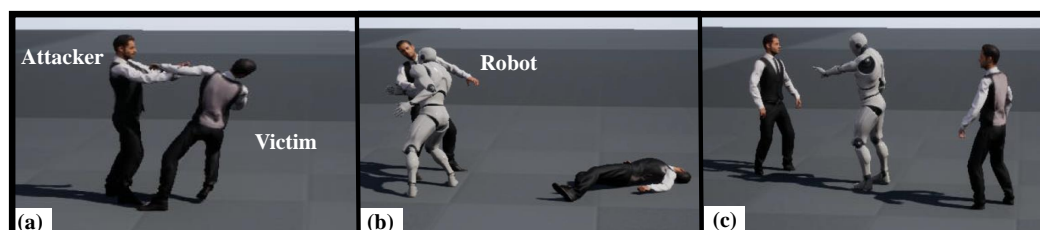

Figure 1. Could a robot step in to defend a person who is under attack, (**a**) detection what is happening, (**b**) applying needed force, and (**c**) conducting post-hoc deescalation?

To go deeper, we define how some important terms are used in the current article: RSD is defined here as "the use of force by a robot to protect a person in the robot's care from an immediate threat of physical danger from some attacker when safe retreat is not possible". This is in line with the common definitions of robots and self-defense below. "Robots" are machines with human-like qualities that are controlled by a computer to carry out tasks [10], encompassing both humanoid robots as well as other embodiments such as AVs, drones, or zoomorphic robots. "Self-defense" is "the act of defending yourself, your property, etc." [11]. McCormack summarizes its mentality as follows: "I am not allowed to use deadly force against another person unless that person represents a threat of great bodily harm to me or another, and even then I can only use so much force as necessary to remove the threat" [12]. Furthermore, he discusses how preemptive use of force, likewise, should only take place when threat is imminent and action necessary, as per the "Caroline" test. (Note: here "self-defense" does not refer to a case in which a robot defends only

itself against an attacker, since robots can be repaired or replaced, and we believe the fundamental case of risk of harm to humans is more important and interesting to explore.)

We note that the current work differs from previous conceptualizations where the focus has been on robots whose sole purpose is security (e.g., a RoboCop or PackBot). Rather, here our focus is on a more ubiquitous or dispersed responsibility to help that could be given to advanced technological artifacts, in analogy to the duties people are expected to perform. For example, this could involve vehicles and humanoids whose primary uses lie elsewhere, in transport or healthcare. (Even more radically, a person's environment could defend him/her, such as a smart home opening a door to let in a victim, removing stairs, and thrusting up walls and barriers to block and trip an attacker, etc.)

What then might be useful for robot designers and policy-makers who are interested in exploring this direction? We believe that an important initial goal would be to gain insight into the "big picture" of RSD. The challenge is that the question is highly complex, requiring knowledge from various areas such as psychology, philosophy, computer science, AI, jurisprudence, and security, but previous insights on how robots could help support people's well-being by physically defending them from violence are likely to be scattered here and there, requiring work to pull together. Furthermore, the public might perceive the question differently from academics, but the participation of both groups would be vital. Academics are experts who guide and push forward the frontier of knowledge, whereas the public are the end users we want to help and those who control the vote to put systems into use or not.

Thus, the intended contribution of the current article is to explore the broad scope of RSD, extending our previous work, as follows:

- *Academic theory.* We explore the concept of RSD in depth based on a rapid scoping review of the literature, encompassing the intersection between robots, crime, and violence, and other "dark" topics in Human-Robot Interaction (HRI).
- *Public opinion.* Furthermore, we report on the results of an online survey to check how people in two different countries perceive the effects of two factors that we felt could be important, a robot's embodiment and use of force.

The remainder of this article is structured as follows: Section 2 summarizes some academic work related to RSD, identifying gaps such as cultural comparisons. Section 3 explores public opinion, reporting on the results of an online survey administered in Japan and the United States (U.S.), which are discussed in Section 4, along with limitations and next steps. We conclude by returning full-circle to our initial "attack narrative", which is used to illustrate our vision for how robots could help.

2. Scoping Review

An initial check did not turn up a review on RSD or related topics such as the dark side of HRI. Thus, our goal was to broadly and rapidly map existing studies and identify gaps within this seemingly under-examined area. In such cases, scoping reviews have been described as appropriate; therefore, in line with the prescriptions of Arksey and O'Malley and the Joanna Briggs Institute, we followed a process consisting of the following steps: (1) a broad search to refine selection criteria, (2) selecting and reviewing, and (3) mapping and summarizing results [13,14].

2.1. Review Process

Our basic process is shown in Figure 2: (1) First, the review question was specified: *"What is known from the existing literature about RSD?"* Objectives, inclusion criteria and methods were also specified in advance. Given that different authors might use different words to refer to similar concepts, various related keywords were identified. In general, within the broader area of the dark side of HRI, keywords focused on what the robot should do (defend or protect) or prevent (violence, crime, or attacks). Some keywords such as security appeared to lead to a reduction in the relevance of the search results, yielding many papers about networking security, so were not included.
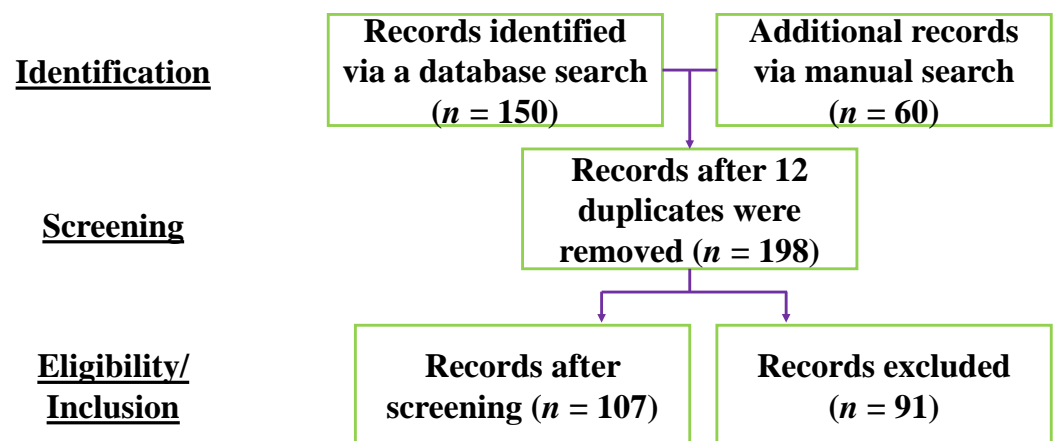
Identification	**Records identified via a database search (*n* = 150)**	**Additional records via manual search (*n* = 60)**
Screening	**Records after 12 duplicates were removed (*n* = 198)**	
Eligibility/ Inclusion	**Records after screening (*n* = 107)**	**Records excluded (*n* = 91)**

Figure 2. The process followed for the rapid scoping review.

(2) Thereafter, a rapid scoping review was conducted using the first fifty results each from ACM Digital Library, IEEE Xplore, and Google Scholar using the search phrase: *"(robot OR HRI OR "human–robot interaction" OR robotics) AND (defense OR violence OR crime OR dark OR protection OR attack)"*. The 150 search results were entered into a shared document for processing, and 12 duplicates were removed from consideration. The inclusion criteria in the first round of processing were that the paper should be written in English and not be fiction or an "own" publication. Exclusion was not conducted based on year of publication, kind of paper, field, or target demographic. This resulted in the removal of three papers: One paper in Chinese was removed from consideration. One fictional story was removed, after checking its summary to make sure it did not contain new ideas related to RSD. One of our own publications was removed from the main group of reviewed publications, and cited later in the article, at the end of this section.

Next, to identify potential relevance for each paper, a quick scan was conducted of the title, abstract, and figures, and a "find" function was used to see how the key words in the search string were discussed. Two reviewers were involved; interpretations were compared and discussed to resolve conflicts (seven papers with differing interpretations were discussed and agreement reached). As a result, 91 papers were removed due to lack of relevance to the review question. For example, searching for the dark side of HRI also led to some search results on physical darkness. Papers that seemed to have some relevance to the topic were then read and notes taken. Some papers cited in the initial papers' bibliographies were also added. While, as expected, no papers directly addressed the exact topic of the current paper—self-defense as a secondary function for social robots—several papers were found to indirectly touch upon the theme of the dark-side of HRI and the role of robots might play in violent encounters.

(3) Some basic statistics were calculated, and papers were analyzed. Figure 3 suggested that the reviewed publications were mostly from recent years (start year: 1967, end year: 2022), with a growing trend since around 2004; the current average rate of detected publications related to the topic seemed to be approximately 7.4 papers per year, considering the past ten years from 2012 to 2022. As well, Figure 4 indicated that the reviewed publications comprised roughly the same amount of conference papers and journal articles, along with a few technical reports and theses. Next, these publications were grouped into themes, as shown in Figure 5, summaries of content merged, and our own thoughts, suggestions, and proposals included.

Figure 3. Papers per year.
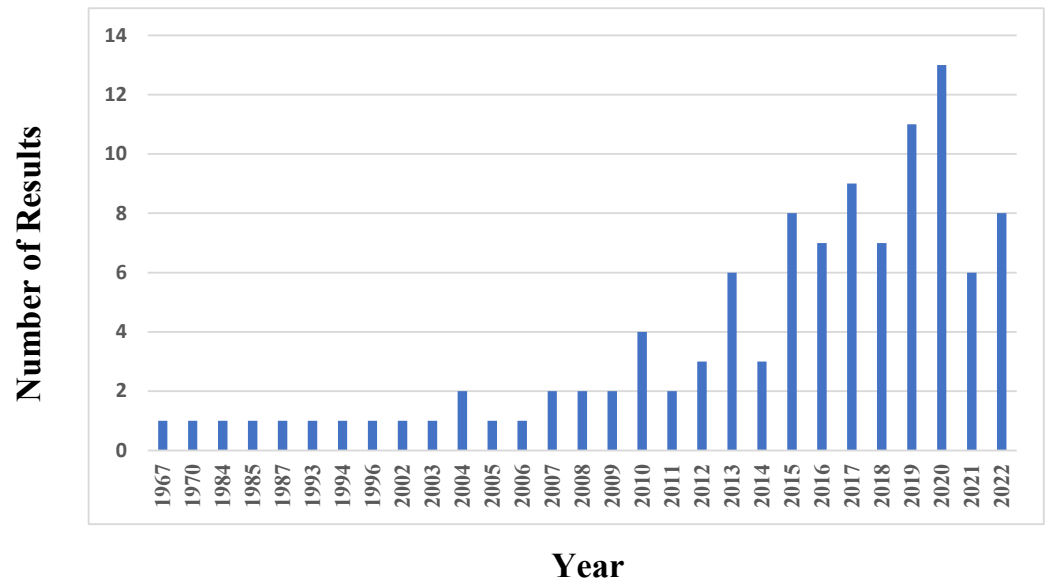
■ Articles   ■ Conference Papers   ■ Other
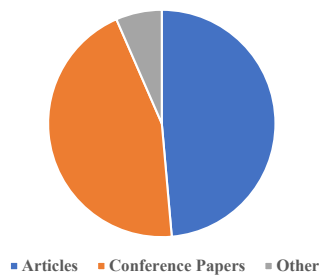
Figure 4. Papers by venue type.

Figure 5. The main themes that emerged from our review, represented visually as yellow squares, as well as sub-themes, in purple truncated squares. (Culture was felt to be a separate, overarching theme). Numbers indicate section numbers.

The themes that emerged included: Background (13 publications cited), Risks for People (21), Violence to Robots (23), Tasks (30), Control (13), and Culture (11). (Thus, the

most citations related to the Tasks theme, and the least to Culture.) These themes are briefly described below. Section 2.2 Background touches on t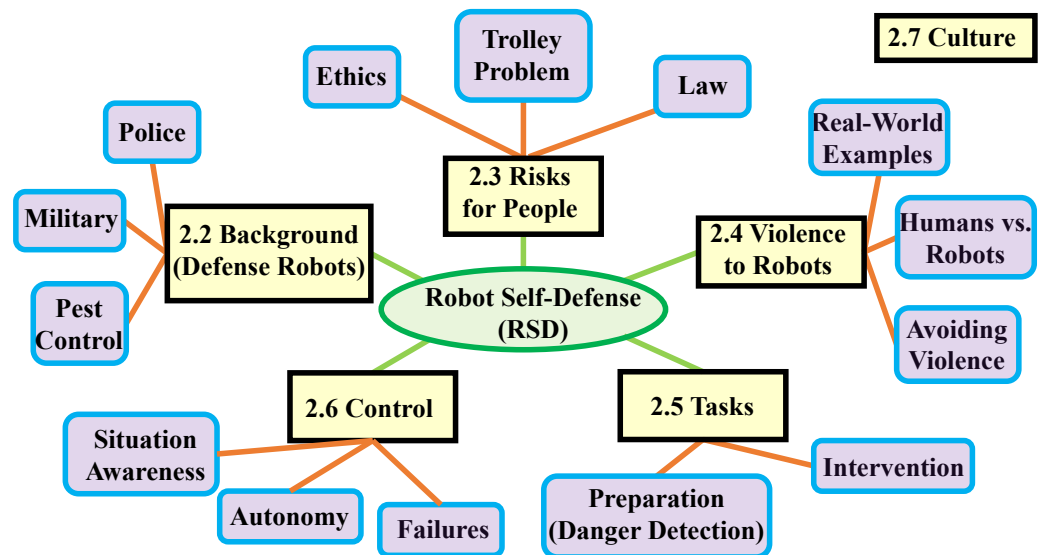he history of robots capable of wielding force, which are finding increasing use in law enforcement, military, and pest control applications. Their use is not uncontroversial. Section 2.3 Risks for People discusses merits and demerits of RSD in comparison to current attitudes toward military robots and convergence with human responsibilities, how RSD relates to the trolley problem, and legal questions. The flip side of this "coin" involves considering not just what robots might do to people, but what people might do to robots. Section 2.4 Violence to Robots examines real-world examples of violence toward robots, negative perceptions of robots compared to humans, and strategies for how robots could avoid violence. Section 2.5 Tasks approaches more practical concerns of what a self-defense robot would actually be required to do and how capabilities could be implemented. Since tasks can be carried out via teleoperation or by a robot itself, Section 2.6 Control touches upon control interfaces, autonomy, and failure mitigation strategies. Finally, Section 2.7 Culture highlights that groups' requirements for RSD will likely differ.

2.2. Background: Defense Robots

A fundamental "call to arms" can be found in the words of Haring et al. [15]. The authors argue that researchers have mostly focused on the positive side of HRI, despite the need to also identify possible problems—given the expectation that some robots will be placed in charge of life-or-death situations in our surroundings, that robots can cause harm, and that there is a current lack of specifications for ethical robot behavior. Furthermore, they point out a need to go beyond Asimov's laws, which were originally intended to bring attention to problems that can occur with rigid rules, within the context of fiction intended to entertain.

Yet, fiction could be considered to serve as a useful starting point for discussion. While it has also been speculated that robots will help us with dull, dirty, and dangerous tasks [16], the latter aspect being our focus here, there is likely more to the story of why designers are striving to build robots than utilitarian considerations. As Cha describes, robots are considered "cool"; when customers see a robot for the first time, they often use the word "wow", and coolness can predict people's intentions to adopt technologies [17]. In the context of RSD, violence does not necessarily curb such impressions. Even if we might not wish to be the target of violence ourselves, reports of aggression and force can excite people, in the phenomena described as "recreational fear", "voluntary arousing of negative emotions", "angstlust", or "excitation transfer", and be perceived positively through their association with power and achievement [18–20]. Thus, as one might expect, the basic motif of a robot that could use force to help save people from physical danger appears in various works of fiction. In Ancient Greek and Jewish myth humanoids of metal and clay hurled boulders or struck enemies to defend people, and both Japanese cartoons and Hollywood movies describe robots and cyborgs that use firearms and gadgets to protect people from criminals and other robots (e.g., Talos, golems, Doraemon, Tachikoma, RoboCop, and Terminator) [21]. A more in-depth depiction of law enforcement robots in fiction can be seen in a work by Reid [22].

Similarly, in the real world, the focus for force-wielding robots has also been on dedicated platforms intended to help in law enforcement, military, and other applications such as conservation. The line between applications is not always fully clear. For example, PackBots have been used to find people (emergency victims or snipers) in dangerous terrain such as ruined buildings or caves, clean up, and detect and defuse improvised explosive devices (IEDs), during the Iraq and Afghanistan conflicts, 9/11, and the Fukushima nuclear disaster [23]. Nonetheless, some details are provided by application below.

2.2.1. Police Robots

In the U.S., robots have been used in various settings by law enforcement, as described by Carruth and Bethel [24]. The goal therein is to be able to remotely monitor and act

without risking officers' lives, given that each year many officers are killed and injured in the line of duty. While typical current uses comprise explosive ordnance disposal and scouting, various other possible tasks could be conducted, including guarding, search and rescue, distraction, serving warrants, negotiating and recovering casualties. As such, most Special Weapons and Tactics (SWAT) teams have access to robots, with a reported average annual usage in Section 3.1 interventions. One specific example of usage in the U.S. was in a trial in 2019, in which Massachusetts State Police tested Boston Dynamics' Spot [25]. Another example of a security robot in the U.S. is the Knightscope RoboCop, which can detect nearby mobile usage and scan license plates; this robot was developed in the wake of the December 2012 mass shooting at Sandy Hook Elementary School that left 20 children dead [26]. As well, Reuben Brewer, a researcher at a non-profit organization called "SRI International" in the U.S., built a robot called the "Go-Between" to reduce harm during police traffic stops [27,28]. The robot consists of a remote conference system and ticket dispenser attached to a rod that can be extended from a police car to the window of a car in front, as well as a spike strip to prevent a motorist's escape. In the U.S. also, a robot called the Telebot was designed to potentially patrol under the control of a disabled police officer or veteran [29]. As well, the Cobalt robot patrols in office buildings, alerting a human if anomalies are detected such as open windows or people in restricted areas [30].

In Japan, various security robots have been designed by the company Secom, such as the Secom Robot X and COCOBO, which can emit smoke, patrol, and detect suspicious objects (e.g., at a home or airport) [31,32]. As well, a security robot, Ugo from Mira Robotics, was tested in an office building in Tokyo; the robot, which was designed with a cute appearance to avoid intimidating people, patrolled every two hours, calling an elevator itself, and allowing remote monitoring through its cameras [33]. Additionally, Japanese police used drones with nets to capture a drone piloted by an activist that attempted to carry radioactive sand to the prime minister's residence [34].

As well, in 2017, a humanoid "police robot", the "REEM" by "PAL Robotics", patrolled the Dubai Mall in the United Arab Emirates; its capabilities include providing information, processing fines, and remitting monitoring footage to a control center [35]. Maliphol and Hamilton also describe a drive toward "smart policing" by using robots as 25% of Dubai's police force and realizing a smart police station that does not require humans by 2030. The authors also describe a robot at a police station in India, KP-Bot, which hears complaints and guides visitors. Additionally, the Singapore Police Force has used a fleet of drones for area observation to pick out suspicious activity, police presence (drones come equipped with red-and-blue police blinkers), enforcement, and search and rescue; for example, in 2021, a man evading a police check was detained with the help of a drone that saw him hiding at a construction site [36,37]. In China, the robot AnBot, which was trialled at an airport, can seize and incapacitate targets with a stun gun [38,39]. Furthermore, the "E-Patrol Robot Sherriff" has patrolled a train station and can purportedly detect wanted criminals, alerting police and following the suspect until police arrive [30,38]. In South Korea, a prison guard robot introduced in 2012 allows guards to remotely communicate with inmates, also detecting anomalous behavior related to emergencies such as suicide, assault, or arson [40,41] In the latter work, McKay furthermore mentions a similar trial in Hong Kong in 2019, including use of robots to detect drugs in prisoners' feces, and a trend (mentioning China) toward "robotocracy" and establishing smart systems capable of "hyper" or "blanket security" and "optimal orderly functionality". Moreover, the Korean national police agency apparently plans to form an "Iron Man Police" service, under the rubric "Police Future Vision 2050", that aims to combine robot dogs, robotic power suits, and AI to mitigate crime [42]. As well, in the Democratic Republic of Congo in 2013, some traffic lights with a robotic appearance waved red and green flags to direct traffic and allow pedestrians to cross busy intersections, in an attempt to promote safety and compliance with traffic rules [34]. The same article describes a robotic boat called "Emergency Integrated Lifesaving Lanyard" (Emily), which was designed in the U.S. and deployed by the Greek coast guard to try to save Syrian refugees from drowning.

2.2.2. Military Robots

Robots capable of wielding force are also increasingly being considered or used for military applications. In the U.S., from 2020 to 2022, robot dogs from Ghost Robotics were tested by the U.S. Air Force, U.S. Border Protection forces, and the U.S. Space Force, for outdoor patrolling and inspections at air bases, the southern border, and Cape Canaveral [43,44]. In 2021, the French military also tested using some robots in combat exercises, including a Spot robot dog from Boston Dynamics, an "OPTIO-X20" tank-like robot, and Barakuda, an armor-plated wheeled robot designed to provide cover [45]. Various other military platforms also exist, such as Centaur from "Teledyne FLIR", and Iris from Roboteam, a small remote controlled robot that can be thrown through small openings such as open windows to scout out buildings (similar to the Recon Throwbot and Taktyczny Robot Miotany (TRM) Tactical Throw Robot that have been used by U.S. and Polish police) [34,46].

Some robots have also been developed that carry weapons. While a few examples of sword-wielding robots have been described for sports training or entertainment, such as the kendo robot Musa [47–49] and Yaskawa's swordsman-emulating MOTOMAN-MH24 [50], the majority of weaponized robots appears to be designed for military usage. Some of the first robots were flying bombs and rockets, conceived of before World War I, and featuring automatic piloting and detonation capabilities by the end of World War II (V1, V2 rockets) [51]. In addition to such drones and cruise missiles, some increasingly smaller, more lifelike mobile robots with firearms are also being recently shown. For example, in 2008, the "Modular Advanced Armed Robotic System" (MAARS) was introduced as a successor to the "Special Weapons Observation Reconnaissance Detection System" (SWORDS) Talon used since 2000, which can carry a grenade launcher, machine gun, laser dazzler, and gunfire detection system [52,53]. In 2006, South Korea developed semi-autonomous prototypes of robotic sentry guns called "SRG-A1" robots, to protect its border with North Korea [34]. In 2014, a taser-bearing drone called the "Chaotic Unmanned Personal Intercept Drone" (CUPID) was demonstrated, by temporarily incapacitating a volunteer intern [54,55]. In 2016, the Dogo robot was created by General Robotics with advice from the Israeli Police Counter-Terrorism Unit, which carries a Glock-26 pistol, eight cameras and radio, and can enter houses quietly and climb stairs [34]. In 2017, an android robot called FEDOR (Final Experimental Demonstration Object Research) that was developed by Android Technics and a Russian military research agency, was demonstrated, driving, walking, and firing a handgun [35]. In 2021, Ghost Robotics in China attached a rifle/firearm and thermal camera to a Quadrupedal Unmanned Ground Vehicle (QUGV) robot dog [56]. In 2022, a Chinese defense contractor posted a video of a drone delivering a robot dog carrying a light machine gun [57]. In 2022 also, the Russian military showed a robot dog with a rocket-propelled grenade (RPG) on its back, based on Unitree Robotic's UnitreeYushuTechnologyDog [58]. Loitering drone systems also exist, such as the Israel Aerospace Industries (IAI) Harpy or AeroVironment Switchblade [59].

Such weaponized platforms, whether for police or military, are not always merely for show. In addition to some people who have been killed accidentally by large factory robots [60], some robots have been used intentionally to exert force. As previously noted, military robots have been used in wars. More locally, Sharkey reports that a security robot was used outside a bar to shoot water at undesired loiterers in the U.S. in 2008 [61]. Furthermore, according to Glaser, the Los Angeles Police used a large remote-controlled robot called the "Bomb Assault Tactical Control Assessment Tool" ("Bat Cat") to tear down the walls of a house during a standoff in 2011 [34]. An article also describes how, in 2013, in the U.S., a robot helped police catch the Boston Marathon bombers [62]. The same article also describes how police used a manually controlled mobile robot equipped with an explosive to kill a gunman in 2016, a robot was exploded near a window to subdue a shooter in 2018, and police used a robot to end a standoff with an barricaded arsonist, bringing him a vape pen he had asked for and a cellphone for negotiation, in 2019. In 2020, a robotic machine gun was remotely controlled by Israeli security to assassinate a nuclear scientist (and then remotely detonated to minimize evidence) [63]. Furthermore, in 2020, an

autonomous military system, the Kargu-2 drone from the Turkish company "STM", might have been used to kill enemy soldiers for the first time in Libya [64]. This loitering drone is described as using machine learning and computer vision to detect and track enemies, and being capable of operating in swarms of twenty. It seems such robot systems have also been used elsewhere since, in Nagorno-Karabakh, Yemen, Syria, and Ukraine [59].

2.2.3. Pest Control Robots

In addition to law enforcement and military uses, robots are also being developed for other purposes, such as conservation and wildlife control. For example, robots have also been developed to "revolutionize" monitoring and interventions in the vicinity of dangerous wildlife, such as elephants, bears, and rhinos (in the latter case also detecting poachers) [65]. Some robotic devices have even been developed that are capable of force, such as traps to spray feral cats with poison [66]. As well, drones outfitted with vacuums, pesticides, and even flamethrowers have been used against dangerous wasps and hornets, for which manual interventions have resulted in human deaths in the past [67].

2.3. Risks for People

The use of force-wielding robots could entail risks. Some of the reviewed publications carry ethical insights on how robots could engage in damaging or malicious behavior, how we can conceptualize dilemmas, and how laws might need to change to reflect such risks.

2.3.1. Ethics of Self-Defense Robots

The current paper is fundamentally concerned with the ethics of robots, which is sometimes referred to as Robot Ethics, Roboethics, Machine Ethics, Artificial Moral Agents, or Friendly AI. One central question can be phrased simply as "Should we have robot defenders?" We approach this question here from various perspectives, looking first at how people have felt about the robots capable of force described in the previous section, and then also discussing how robots might be perceived in comparison to humans, before examining RSD itself more closely.

First, examining perceptions of military robots, the benefits of using robots in war include avoiding risking the lives of friendly soldiers, enhanced power, cost savings, allowing operators to calmly make decisions away from threat (easily exchanging seats with fresh operators, which would be difficult with manned aircraft), etc. [68]. A hope is that robots or human–robot teams could one day be able to make better decisions than humans alone, given robots' potential immunity to factors that can hurt people's ability to make good decisions, including dangerous emotions, adrenaline, stress, fatigue, low morale, human perceptual and communication challenges (e.g., ears being deafened, glasses breaking)—thereby maintaining levelheadedness and avoiding overreactions that could result in atrocities [69]. Accordingly, much money is being used to develop such systems; for example, the U.S. reportedly allocated USD 18 billion for autonomous weapons from 2016 to 2020 [70].

Yet substantial concerns exist as well. For example, Sharkey et al. envisions misuse by criminals: e.g., lethal cartel-controlled drug-delivery robots that could also destroy their internal stashes when cornered, as well as robot attackers, aerial lookouts, and exoskeletons being used to carry out robberies, assault, rape, or murder [61]. Similarly, Froomkin et al. raise the question of how people could defend against robot-assisted crimes, such as unlawful recording by drones or trespassing by AVs [71], and King et al. speculate that robots could be used to torture [72]. Furthermore, Alston described for the United Nations (UN) a potential risk of developing a "Playstation" mentality to killing, as the computer screen/speaker interfaces for controlling robots resemble those used for playing games [12]. As well, some police boards have ceased using robots due to protests from civil rights groups, who fear that enabling remote operation might facilitate state violence without necessarily contributing to safety or crime prevention. In 2021, the New York Police stopped using a robot, Boston Dynamics' "Digidog", after complaints were raised [62]. Likewise, in

2022, the San Francisco Supervisory Board backtracked on a decision to allow police to use robots to exert lethal force [73].

This concern appears to be greater still for weaponized robots that are being designed to feature autonomous capabilities, which are known by various names, as Lethal autonomous weapons (LAWs), Lethal autonomous robotics (LARs), Autonomous weapon systems, or, more colloquially, as killer robots or slaughterbots. Researchers and politicians have expressed worry about the risk of insufficient recognition capabilities and common sense (e.g., errors, or lack of ability to distinguish civilians from combatants), low-end and high-end proliferation (e.g., allowing misuse by criminal groups), the potential for decreased accountability to facilitate war crimes, a risk of more wars being started due to lessened costs, loss of morale if robot squad members record and spy on human soldiers, and uncertain behavior (such as the potential to refuse reasonable orders due to some unforeseen failure); as such, six robotics companies, including Boston Dynamics, promised to not allow their products to be weaponized, and there appears to be ongoing debate on regulation at the UN [69,70,74]. Popular anxiety about the potential for such autonomous robots to rise up against humans has also more facetiously been expressed in a parody video by Corridor Digital in 2019, which has been viewed 80 million times as of January, 2023 [75].

While the above sources might provide some insights, self-defense robots are not military robots, and not yet a reality, such that we do not know how exactly people will perceive them. Another way to approach this question is to follow the train of thought that robots are often treated like people [76], and that robots are becoming increasingly similar to humans in terms of capabilities (i.e., advancements in robotic technologies could one day result in human-like capabilities to engage in self-defense). From this perspective, assuming progress continues (as it has up until now), robots might be expected to one day assume similar responsibilities as human bystanders and defenders. Although laws for humans differ around the world, a "right to self-defense" appears to be typical. In some regions, there is even a moral and legal "duty to rescue", and even laws to protect defenders who might know the attacker or victim (e.g., "good Samaritans") [77]. In this sense, robots might one day have similar rights, duties, and protections.

Lin et al. compare humans and robots from a different perspective [69]. They reasonably point out that an ethically-infallible robot should not be our first goal—even casting doubt on whether such a robot would be feasible. Instead, they suggest that designers should seek to first design a robot that performs better in some ways than humans, especially with respect to avoiding unlawful actions. Such a proposal could be interpreted in different ways. Given also the prevalence of protests against excessive force by law enforcement in the U.S., an ambitious goal could be to develop some kind of a "Turing test" related to police work or self-defense, in which robots should outperform humans in terms of ethical outcomes. Another more modest goal, which seems more likely in the near future, could be to simply show that by providing robots with some capabilities to help in self-defense situations, we will be able to avoid some harm that would have been otherwise not prevented (or perhaps more vaguely, by showing that the crime rate is reduced).

The potential merits and demerits of RSD itself can also be considered more closely. At the level of a single interaction, there seems to be some further uncertainty. The benefits of a successful defense scenario seem to be self-evident: a capable robot might help not only a victim, but also avoid the risk of harm to a human defender (e.g., as "Good Samaritans" who try to help victims face a risk of themselves being injured or arrested), or even to a human attacker (who could be harmed by human defenders or victims). In some cases, the mere presence of a robot might dissuade attacks, and even if a robot is incapable of self-defense, it could help people to feel more comfortable by acting as another potential target for attackers, via the phenomenon of "risk dilution", related to how people in groups feel more comfort and less fear. Yet, the reverse is also possible: robots could cause harm. It seems also self-evident that an attacker could sustain injuries if a defender uses force to stop them, and that mistakes in recognition or motions could also cause harm to defenders,

victims, bystanders, or law enforcement (e.g., if mistakes occur in detecting a threat, who the attacker and victim are, or how much force to use, etc.). More subtle effects will also likely exist. For example, the presence of a robot might be unwelcome if a victim feels anxiety toward robots. If the robot is recording, a victim might feel even more shame from having their traumatic experience potentially viewed by others. Or, in line with the "photo-taking impairment effect", in which taking photos can result in less vivid memory retention, a victim might not look at their attacker carefully enough to identify them later, assuming that the robot will see or record, which might not be the case.

Furthermore, what might happen if robots became overprotective? For example, one extreme way to avoid violence might be to simply separate all humans, which would likely have devastating effects, since, as Maslow describes, humans experience a basic need for social interactions and affection. Even in a less extreme context, however, it is known that moderate fear can be helpful. For example, fear-based play among children appears to possibly protect against the onset of anxiety disorder, potentially by bolstering tolerance of uncertainty, a phenomenon which has been referred to as "recreational fear" or "voluntary arousing negative emotions" (VANE) [18]. Might RSD thus take away opportunities for people to feel strong and self-sufficient?

2.3.2. Comparison to the Trolley Problem

Continuing along the same lines of thought, a risk to people could occur if a robot is not able to effectively navigate the dilemmas at the heart of the RSD question: The robot must choose whether to apply force to an attacker, possibly harming them, or to let a victim be hurt. As such, this problem can be related to another dilemma, the trolley problem, which is often discussed in regard to robotics and AVs (e.g., some of the cited publications here have several hundreds or thousands of citations as of 2023). The basic version of the trolley problem is that a decision must be made on whether a vehicle (robot) should follow its typical course, striking five people who are improperly situated, or veer into a zone that should be safe, striking one person who is rightly situated [78–80]. Various other variations of this problem exist, involving bystanders, "fat men" being thrown off bridges, levers being pulled, and animals.

Are there differences between the dilemma posed in the RSD situation and the trolley problem? We believe yes, and list three potential differences below. First, two perspectives are examined, that the literature seems to commonly differentiate, one of which is described as deontological or Kantian, and the other as teleological, utilitarian, or consequentialist [79–81]. The deontological perspective for the trolley problem is phrased as indicating that the person who was in the right place should not be sacrificed to save the five who were wrong. Conversely, the teleological perspective is phrased as indicating that the robot should kill the one person to save the group of five. However, this might not be the only interpretation of the latter perspective. Taking into account the overall, long-term effect, if the robot obeys the rules and kills the five, it should act to dissuade similar conduct in the future. Clear rules that people follow should lead to less deaths and greater efficiency in the future. Whereas, if the robot breaks the rules, the people around it will know that the rules lack meaning and break them, leading to a chance of more deaths and less efficiency in the traffic system in the future. In that case, both the deontological and long-term, overall teleological views would seem aligned, given that the effect of breaking rules over years and within a large area such as country might very well result in more than five deaths. (Admittedly, this might become more problematic when we increase the number of people in the group to a number greater than might be expected from dissuading rule-following; e.g., from five to a million; although, it might also be difficult to imagine a million people waiting for the robot to collide with them.)

Next, we consider the fundamental case for RSD treated in this article. Here, the deontological case has been deliberately set to be clear: the attacker is in the wrong. Rather, focus is placed on the utilitarian view, which becomes more unclear given that there seems to be more uncertainty about the robot's abilities. The trolley problem is intuitive in the

sense that most people are familiar with vehicles and know well that they can kill. In the RSD case, however, it might be the case that the attacker is likely to hurt the victim far more than the robot would hurt the attacker, given that the attacker has initiated an unlawful attack with malicious intent and the robot's intent is merely defense; or, it might be the other way around, since robots are composed of hard materials and contain sharp edges and dangerous components such as actuators, and batteries that could explode. Moreover, it is unknown if a robot would be able to defend a victim (the robot might be too small or weak, a manual operator might not be able to pilot the robot well enough, and autonomous capabilities can fail). Thus, a first difference is in how the deontological and teleological perspectives might be interpreted, and the uncertainties in the RSD case.

Another difference seems to lie in the distinction between "killing" versus "letting die". In the trolley problem, doing nothing results in following the deontologically right approach, by sparing those who are not breaking the rules. In the RSD problem, doing nothing is counter to the deontological ideal, by not helping someone in need.

A third, fundamental difference between the two problems however might be in the maturity of the technologies discussed, also in relation to practical usefulness, and how likely we are to see these scenarios played out in reality and when. AVs to some extent are a reality today. We know also that driving does not need to be perfect: our societies accept a high number of deaths yearly due to drivers' errors, without banning driving. While the trolley problem has been explained as a thought experiment and "intuition pump" designed to probe beliefs, as well as an edge case for technologically mature systems where basic functionality is no longer a problem, some have questioned its practical usefulness [82]. For example, Rodney Brooks asks how many times we, or anyone we know, have been forced to make a decision about which group of people to kill, the five nuns or the child? [83]. He goes on to mention that, just as these problems never come up for human drivers, they might never come up for AVs either, and the problem might be "non-existent and irrelevant". By contrast, no robot capable of self-defense currently exists in our surroundings. Yet, conversely, how many of us have never seen or been in a fight? For example, Stein et al. reported that a third of high school students claimed to have been in a fight recently, and that approximately 80% of children 7–15 had witnessed a fight [84]. Thus, for the trolley problem, the scenario of vehicles is highly prevalent, but there is some lack of clarity regarding how prevalent the problem is (e.g., five nuns violating traffic rules when an AV happens by), whereas for robot defense, the scenario does not yet exist, but the problem seems to be prevalent.

Further comparisons can be made, but one challenge is that there are various variations of the trolley problem, and the RSD scenario could also be varied. For example, to bring the two problems closer together, we can imagine five attackers assaulting one victim (e.g., as an example of bullying, gang violence, or lynching, etc.). Here though, increasing the number of rule-breakers would seem to increase the force imbalance, making it even more desirable to help the victim. We can also imagine cases in which the deontological case might also not be clear; e.g., if the victim has somehow goaded the attacker into attacking, or if there is no victim (e.g., if both parties attacked one another). Other factors that could affect judgments might include the ability of the victim to escape harm, and the balance of force between attacker and victim; for example, is the victim a child or elderly person? There might also be prejudices related to ethnicity or gender. Furthermore, when exploring how people feel, people might state one opinion when they think their peers might see or when they feel removed from the action, and another opinion when anonymous or when they feel a connection to the attacker or victim. Thus, it seems difficult to enumerate all possible variations for comparison.

Regardless of the differences, however, insights can be taken from previous work on the trolley problem, on how to also deal with the RSD dilemma. For example, Goodall proposed a three-part strategy to deal with the trolley problem for AVs, involving rules, machine learning, and explanations; he also discusses perceived downsides of deontological, utilitarian, and AI-based approaches [81]. While our exploratory work involves gathering

academic thought and cross-cultural probing of public views, rather than machine learning, we believe such an approach might also be useful for RSD.

2.3.3. Laws Related to Self-Defense Robots

Given potential risks and uncertainties, legal perspectives should also be considered. For example, Terzian argued that robotic weapons can be considered "arms" in U.S. law under the Second Amendment, giving people a right to be defended by robots [31].

Reid speculated on what a future police robot might be like, whom she called Officer Joe Roboto [22]. Although such a robot will eventually likely be faster than humans in collating big data, potential concerns include privacy risks, requiring a rethinking of Fourth Amendment doctrine in the U.S., and lack of common sense. As well, such a robot should be treated the same as a human officer in regard to motions to suppress evidence or file abuse of civil rights claims.

While robots might currently lack a human's common sense, Joh has argued that the potential legal analogy of autonomous security robots of tomorrow to past mechanical security devices such as "spring guns" (guns designed to discharge via a tripwire or other simplified mechanical mechanism to defend property) should be rejected [85]. She speculates that, while deaths and injuries will also result from robots intended to provide security, robots will have some ability to distinguish legitimate threats of deadly force from innocent mistakes or petty crime.

Calo as well posed various legal questions [86]. For example, could robots be treated similarly to animals under law? For example, if a robot hurts someone more than once, could its human "owner" be held responsible, or could it be labeled as roaming? Could attacks on robots be used like attacks on animals or domestic violence as a warning indicator (e.g., to call child welfare services if there is a child in the house of someone who attacks a robot)? Could police hesitate to send a robot into an encounter in which the robot could be destroyed, as might be the case for sending in a police dog? Or, comparing robots to humans, if a robot is involved in a fatal incident, should it be forced to "take time off", like a human police officer, as some programmer perhaps reviews its code and database? How can losses be compensated for a robot with which there is an emotional bond, if a robot becomes like a family member to someone?

Asaro counseled against the development and use of robots that could exert violent and deadly force against people, calling this the "deadly design problem" for HRI [55]. In discussing which legal standards could be used, who could be targeted with force (discrimination), how much force can be used (proportionality), and who would be responsible (accountability), he referred to guidelines from the United Nations Human Rights Council and described various boundary cases that would be difficult for robots to deal with. For example, he argues that special considerations should be made for disabled persons such as those who are mentally ill, deaf or blind, and that weapons can be disguised and seemingly ordinary objects such as sticks or rocks can be used as weapons.

Regarding accountability, the effects of mistakes can be complex. Robinette and colleagues found disturbing evidence that participants in a simulated emergency in a lab study were willing to take the advice of a robot that malfunctioned in front of them and ignore contradictory evidence [87]. This suggests that victims or bystanders might be overly willing to follow the lead/advice of a self-defense robot, even though the complexity of defense situations and thus the possibility for errors seems high, and that robot designers should take care to design responsibly.

Simmons also described benefits of police robots that will include increased efficiency, perception (robots can detect heat, trace components, or sounds outside human hearing range), flexibility (robots can be reprogrammed to follow new protocols and procedures) and expendability [30]. Demerits could involve increased surveillance that might fall disproportionately on poor, minority groups, as well as repercussions of robots mistakes and opaqueness of robot capabilities. He suggested that such robots will be allowed to conduct Terry stops (brief detainments based on reasonable suspicion of criminal activity),

and potentially non-deadly force under strict management, but should not be allowed to search people's belongings, homes, or cars, or use deadly force.

Another potential source of liability to avoid could be intentional misleading. Lacey and Caudwell explored the relevance to HRI of "dark patterns", deceptive design practices such as nagging, obstruction, sneakiness, interference, and forced action [88]. Cuteness was proposed to provide an illusion of the robot as not being harmful, emphasize short-term gains over the long-term, and manipulate emotions. Thus, a robot capable of self-defense maybe should not be made to look weaker than it truly is, or to take advantage of victims or attackers afterwards.

Carr also discussed some potential liabilities in regard to drones, under U.S. law [89]. For example, it is illegal for the public to fly a weaponized drone, but a young adult who published videos of a drone shooting a gun and using a flamethrower to roast a turkey argued that this law applies to aircraft rather than a hobby project in a backyard done for non-commercial purposes. As well, although people can fly drones above their property, robots should not "get in the way", as in the case of a drone that delayed fire-fighting efforts.

Another topic is how to legally deal with threats to humans and robots. In regard to previous findings that people also care about harm to robots (described in Section 2.4), Mamak argued that it would be a crime to hesitate to destroy a robot to save a human, and even that human-like appearance in robots is undesirable, as it could elicit empathy or deceive people into thinking a human is in trouble [90]. Several questions emerge. Could it one day be a crime to watch and not help a robot that is being destroyed, if the robot is responsible for the lives of humans (e.g., a medical robot), or has somehow acquired rights like a human? (For example, Sheliazhenko described a project started by the Non-Governmental Organization "Autonomous Advocacy" to defend robot rights, which included the idea that robots should have the right to self-defense [91]. He qualifies this by stating that "every threat caused by robots can be avoided by robots, like automatic spam filters clean email inbox from automatically sent advertising"). Assuming that some robots (of any form, not necessarily humanoid) might appear like they could help but not be able to, could the presence of such a robot mislead victims into a false sense of security? Who should be liable if a robot does not help when it could (e.g., due to imperfect recognition, or even a conscious decision)? For example, could a victim sue a robot manufacturer, alleging that a robot vacuum cleaner simply kept cleaning while they were being attacked, drowning out their calls for help? Various work will be required to find answers to these and other legal questions related to RSD.

2.4. Violence to Robots

In the previous section, we have considered what robots might do to people; here we explore findings of what people do to robots, when robots are "suddenly" placed in real-world environments.

Various reports have described people treating security robots with empathy, like humans or pets, giving them names like Boomer, Scooby Doo, Danny DeVito, or Owen Wilson; attributing qualities like gender; taking them fishing; awarding medals, and holding funerals [92]. More generally, including some experiences with other kinds of robots, reports describe people feeling bad when robots get stuck or have to be put back into storage; feeling self-conscious when changing clothes or having trouble concentrating in front of a robot; selecting less optimal choices to spare robots; taking out a robot's battery to try to avoid having it feel "pain" during a damaging experience; accusing people who mistreated a robot of "cruelty"; and in one case calling off a landmine defusing test for a six-legged demining robot whose legs kept getting blown off [93].

However, many of the publications we reviewed (especially those in HRI) seem to have focused on the idea that people sometimes abuse robots. These studies are also relevant for the current topic because real-world examples of robot abuse support the idea that we can expect robots to have to deal with violent behavior, they compare how much

people care about the safety of robots versus other people, and they suggest possible ways to avoid violence.

2.4.1. Real-World Examples of Violence Toward Robots

Examples of abuse include the following. Rehm and Krogsager described how some students were impolite to a NAO robot left in a dormitory in Denmark [94]. They highlighted that past work has focused on positive effects noted in successful interactions in restricted environments, but that we need to take into account more negative ways of interaction in real environments, because people also sometimes behave negatively. Connolly et al. reported that a Muscovite beat a robot with a baseball bat and kicked it, causing it to fall, while the robot pleaded for help [95]. They also mentioned San Franciscans kicking delivery robots on sidewalks [95]. Likewise, Bartneck and Keijsers mentioned that HitchBOT, a "hitchhiking robot" relying on the kindness of passersby to carry it across the U.S., was destroyed by unknown vandals, a K5 Knightscope robot was assaulted by a drunken person (maybe in part since alcohol inhibits rational decision-making via the prefrontal cortex), and RoboVie was obstructed and kicked by children unaccompanied by parents [96].

In terms of the context of this paper, these examples suggest that merely having a robot present (as "security theater") will not always be enough to stop attacks, since people are also willing to attack robots–such systems must be able to deal with negative interactions. We furthermore propose that the following questions could be interesting to explore:

- *Complications due to people.* Could some people's desire to mistreat robots include sometimes trying to stop a defender robot from doing its job, or faking attacks/crimes in front of the robot to get it to do something? For example, bystanders and victims might not always help a robot to defend a victim. This could be unintentional (due to a misunderstanding) or intentional, as in hybristophilia, Bonnie and Clyde syndrome, or Stockholm Syndrome (in which people feel attracted to those who commit crimes) [97]; lack of trust in robots; or domestic violence, in which fear of later reprisal could result in rejecting needed help. (As well, could RSD create concerns about entrapment, e.g., given decisions on where to place robots? Furthermore, could robot abuse also indirectly lead to violence against humans? For example, if a child gets into a fight trying to protect their robot from a bully.)
- *Physical design for self-defense.* Should only certain kinds of highly robust robots that would be difficult to topple (e.g., with a wide base, high weight, and short height) be allowed to conduct self-defense?
- *Causes for robot abuse.* Why were the robots above attacked? Were the delivery robots too slow or in the way, or too oblivious and passing through areas where an attack would be easy to carry out? Do children attack robots because the consequences of being caught would be less and they are not as fettered by social norms as adults?
- *Defense against specific groups.* How should a self-defense robot deal with drunk people or children? For example, in the case of an attack by a child on another child, should a child be treated the same as an adult? If not, a means of assessing potential harm (threats and consequences of intervention) might be required.

2.4.2. Comparing Perceptions of Humans vs. Robots

One cause of negative treatment of robots could be related to how they are perceived in comparison to humans. In particular, some studies have found that people also care about robots, but less than they do about other people. For example, Tan et al. found that people care about robots, intervening to help a robot that was being abused [98]. Bartneck et al., in replicating Milgram's experiment on obedience, found that participants were less willing to protest when they believed they were giving electric shocks to a robot rather than a human; thus, it is more acceptable to harm robots than humans [99]. Similarly, Bartneck and Keijsers found that humans fighting back were perceived as more acceptable than robots [96]; although this study uses the term "reactive aggression", which might

encompass also punitive or violent behavior, within a simpler dyadic context, we believe it is also relevant for the current context of necessary self-defense in triadic encounters. This study also speculates on why perceptions of robots and humans differ—robots attract attacks because they are supposed to be lower in status, are expected not to react, and cannot feel pain, so attackers can feel like it is easy to get away with attacking them and that is not morally wrong—and describes the inherent difficulties of working in this area of research. An actual attack could destroy a robot, which is undesirable due to danger to humans from hard/sharp parts, as well as the high costs of robots, so various tricks have been used, such as turning a robot off to symbolize its destruction, or using the metaphor of a game to explore adversarial interactions in a safe way.

One thing that seems clear is that the kind of robot embodiment is likely to affect how it is perceived. In line with the idea that familiar interfaces support interactions, Eyssel et al. reported that a humanoid appearance was able to support "effectance motivation", in allowing people to make sense of a robot's behavior and reduce their own uncertainty; the authors also looked at anticipation of use and unpredictability of the robot [100]. Likewise, Natarajan and Gombolay found a positive correlation between anthropomorphization and trust, in a user study with Pepper, Nao, Kuri, and Sawyer in the U.S. [101]. The ramifications for RSD are not completely clear; a humanoid appearance (in a human or humanoid robot defender) could be used to reduce a victim's fear and gain their trust, but conversely, a non-humanoid appearance (like for an AV) could also be used to induce anxiety in attackers to try to halt attacks as soon as possible. As well, Garcia et al. explored the effect of manipulating the humanness and gender of the robot victim, finding that witnessing the abuse of a female robot was more distressing, and that female observers were more distressed than males [102]. Based on this latter study, especially in environments with many females, a self-defense robot that could be damaged during harmful interactions could be designed with a neutral or male appearance to minimize observers' distress.

Another interesting idea is that such perceptions might simply be due to the fact that most people are not used to robots, and therefore employ a model of a human to decide how they feel about something happening to a robot: e.g., if a human would feel bad about losing a limb, so too might a robot. Luria et al. however point out that destruction is a common human behavior, and that robots could be designed to help people to feel good when they are destroyed [103]. For example, a robot could break when it detects fighting (calling for the combatants to come fix it and mend their relationships), become immobilized over time to reflect a person's growth (e.g., a victim's courage in facing bullying, as the robot is no longer needed), or take damage from people to entertain (like the BattleBots TV show; e.g., to distract attackers from attacking a victim). In other words, what could be considered abuse for a person could be perfectly fine for a robot, and people's beliefs might change in the future as they gain experience with various kinds of robots.

2.4.3. Strategies for Robots to Avoid Violence

Although a robot might be able to get an attacker to attack it instead of a victim, in most cases it would seem desirable to avoid violence. Some of the reviewed publications have suggested ways in which violence could be avoided. For example, Brščić et al. showed that a robot can reduce the risk of attack from children by planning its path appropriately [104]. This could also be relevant when a robot (or navigation system) is guiding a person and should avoid potential dangerous environments (e.g., dark streets at night where crime is prevalent). Tan et al. found that people perceived violence to be less acceptable if a robot victim shut down afterwards, and were more impelled to help when a robot did not react to abuse than when it reacted emotionally [98]. The latter might have been due to perceiving the robot as being a less capable "moral patient", incapable of sensing negative behavior toward it, or simply to fill the moral "void" and provide some kind of deterrence). Interestingly, the opposite might be true when the robot is not the victim but a third party. Connolly et al., in investigating prosocial behavior that could help others at personal cost, found that participants were more likely to try to help when a group of bystander

robots appeared sad about ongoing abuse than if the robots ignored it [95]. Moreover, they discuss how a robot's size, perceived intelligence, and behavior (emitted light) influence the probability it will be abused.

Similarly, in regard to size, Lucas et al. found that behavior toward a large robot was perceived as less abusive, and that the large robot was perceived as less emotionally capable [105]. The implication for self-defense robots does not seem completely clear: For example, a robot could be made large so that people care more for human participants and do not worry about the robot. Conversely though, bystanders might not feel like they have to help the victim if a large robot is in the vicinity (regardless of its actual ability to prevent an attack), and a large robot could be considered excessive force; e.g., a baby robot could possibly deter violence in some cases by winning over the hearts of those present, by extending its small arms up to protect a victim (although this might also constitute a misleading, dark pattern).

Another attribute connected to violence could be mind attribution. Keijsers and Bartneck found that dehumanization, especially a lack of mind attribution, led to increased abuse of a robot [106]. This relates also to a finding by Zlotowski et al. that people perceived more threat the more autonomous a robot seemed to be [107]. An implication for self-defense might be that in some cases a robot could prevent attacks through adept conversation by stressing a victim's humanity, explaining the logical reasons why they act the way they do. Conversely though, an incapable robot could potentially lead an attacker to attack more if its intervention is poor (e.g., if its lack of ability is revealed).

In some interactions, group dynamics can also play a factor. Yamada et al. reported that attacks against robots by children appeared to follow an escalating pattern in which children mimicked each other, providing validation and support for the abuse [108]. We believe that escalating behavior could also be common in human-human interactions. From a Fogg behavior model perspective [109], feedback from a victim's response could increase attackers' confidence in their ability to successfully carry out an attack (as well as provide a sense of power), motivation (to support the group), and act as a trigger (if it is felt that the application of violence has started and there is no going back). The implication could be that a robot that can recognize escalation stages can try to deescalate, and/or separate individual attackers to try to reduce the extra confidence they receive from being in a group. Along similar lines, Rehm and Krogsager suggest that Goffman's notion of face and Gricean norms can be considered for effective communication, including "positive" and "negative" abuse/impoliteness strategies based on face needs [94].

Additionally, various plans for further research have been described in short position papers. For example, Davidson et al. described a plan to compare how children perceive an abused robot and an abusive adult human using videos, and to also compare a human victim with a robot victim [110]. Likewise, Nomura et al. created a tool to measure the degree to which people feel robots should be treated morally, which might find its use in future papers [111]. Such work could help self-defense robots to intervene more successfully.

2.5. Tasks

Some of the reviewed papers also addressed more practical concerns of what tasks a self-defense robot could perform, which can be broadly divided into preparation (training and predicting/detecting threats), and intervening (negotiating, assessing risks, avoiding damage, treating victims, and gaining information).

2.5.1. Preparation: Danger Detection

Practical necessities would exist in regard to training and maintenance. Zhang et al. point out that robots intended only for emergencies can spend much time unused, since emergencies are uncommon, which could be dangerous [112]. This is corroborated by how six disaster response robot prototypes developed in Japan after a disaster at Tokai 10 years prior, could not be used during the Fukushima incident due to a lack of maintenance and funding, becoming lost, disassembled for parts, or rendered inoperable for use in a

museum display [113]. As noted, we advocate here that, although specialized disaster response robots will no doubt also be developed, for some robots, self-defense can be one capability out of others, like with humans. For example, a robot might clean 99% of the time, but defend someone once or twice during its operating lifetime. Nonetheless, occasional training and maintenance should still be carried out to ensure the robots are in working order and capable of self-defense. Companies providing robots as a service could be asked to conduct such training. Furthermore, training should not be limited to simulations, given that virtual models do not always capture what can happen in the real world with a physical robot. As well, demonstrating such capabilities openly, in front of potential attackers, could deter attacks.

The next important task for RSD is that a robot should predict and detect attacks. This could be compared to the military concept of Intelligence, surveillance, target acquisition, and reconnaissance (ISTAR), or the related term, ISR [114]. A self-defense robot might also be able to predict when violence is likely to occur and behave accordingly, given that crime does not occur in a completely random manner. Victims often know their attackers, such as in the case of domestic violence, and colloquially put, the risk of experiencing danger increases when doing "stupid things" in "stupid places" with "stupid people" [115]. One challenge is that attacks might not always be short, clear events, but can be difficult to detect or carried out over the long-term, such as caregiver abuse, bullying, poisoning, or assassinations. For example, Coffee-Johnson and Perouli prescribe giving elderly a questionnaire that could be used to detect abusive human caregivers [116]. Thus, a robot could possibly ask questions in everyday conversations with humans to seek to detect abuse or be assigned to protect a victim from further harm, and be aware of various cues to detect subtle attacks. This might especially apply to robots being developed to help with therapy, who might be required to follow a "Tarasoff standard", or duty to warn and help a victim when a serious threat is perceived [117]. Aside from elderly, robots could also potentially seek to protect people from sex-related violence. Cox-George and Bewley state that there is already a massive industry related to sex technology, including some companies that already sell "sexbots" [118]. The authors describe how robots could seek to obviate or decrease sex trafficking, sex tourism, the sex trade, pedophilia, rape and rape culture, and relationship problems related to mismatched libido or erectile dysfunction, also protecting from sexually transmitted diseases (STDs). However, negative effects could result from promoting objectification, so more studies could be conducted to provide evidence of positive effects.

In addition to predicting threats, robots could also seek to detect danger directly. Xilun et al. proposed that teams of small legged and wheeled robots receiving information from police could help with finding bombs or poison [119]. We believe that groups of robots could also be useful for protecting people during rioting, looting, war or hooliganism, or to track fleeing armed attackers or search for kidnapped victims. Robots could also act like reporters or mobile surveillance cameras [120,121]. News stories could automatically be created based on what robots see, possibly also by allowing robots to interview nearby humans when something new is detected. Drones with cameras could be sent in to dangerous situations to avoid potential harm to human reporters. Public robots could be controlled by citizens, enabling a democratization of monitoring and news gathering.

Various technologies capable of threat detection are also being developed. For example, after a terrorist attack in 2017, Manchester Arena in the United Kingdom introduced an AI tool to scan visitors for weapons, built by a company called Evolv [122]. This also relates to the "Social Sentinel" software service used by schools like Uvalde to monitor social media accounts of students and related persons to detect threats, although some discussion exists around whether such services are effective enough yet, given the attack that occurred at Uvalde in 2022 [123]. As well, a machine learning system was developed to predict crime a week in advance with 90% accuracy in eight U.S. cities (the study also reported finding evidence of bias in police responses, with higher activity in wealthier neighborhoods leading to more arrests) [124].

Another (more far-fetched) possibility for achieving natural motion in human environments and detecting threats might be to use "cyborgs", specifically augmented animals. In the past, snakes with radiation monitors have been used to acquire data in Fukushima after the nuclear disaster [125]. Scientists in the U.S. found that they could detect 2,4,6-trinitrotoluene (TNT) versus two other smells by processing neural activity in cyborg grasshoppers with an accuracy of 80% with seven grasshoppers (60% for one) [126]. Thus, could the pigeons in the town square, or the ducks in a park's pond, somehow help to detect or stop an attack?

A complication with any kind of detection of threats is that human behavior tends to be highly complex. Robots capable of self-defense will be required to have some ability to distinguish between jokes, skits, or play-fighting, and true attacks. Furthermore, they should also be secure and have a strategy to deal with hacking to make it seem like there is some threat when there is not or interfere with threat detection (e.g., spoofing, jamming, etc.) that might be used by criminals to get a robot to kill or harm, destroy itself, or immobilize itself/render itself inactive/not interfere [127–129].

2.5.2. Intervention

If a potentially dangerous situation has been detected, robots could also seek to persuade prospective attackers to stop, and victims to flee; manage tension; and analyze risks. A short paper by Hayashi et al. proposed that in cases involving serious decision-making, physical robots were more persuasive than virtual agents [130]. Given that current robots often possess both physical and virtual properties (e.g., screens), designers could consider requirements for self-defense. A robot could also seek to either put an attacker at ease to deescalate, or to put pressure on them to desist. For example, Koay et al. found that a robot could make people uncomfortable by moving behind them, blocking their path, moving toward them on a collision course, interrupting them, and behaving erratically (e.g., by spinning in place) [131]. Risk analysis should also be conducted. For example, Guiochet et al. describe how risk analysis can be performed for a rehabilitation robot [132], which could perhaps be adapted, not only to assess risks to attackers, but also to deal with more complex cases such as hostage encounters.

If an attack occurs, the robot should be able to judge if it should intervene, in regard also to safety concerns and how humans are behaving. This can include a strategy similar to military "Laws of War" ("jus ad bellum" and "jus in bello", or law before war and law during war) or "Rules of Engagement" [114]. Risks of damage should be considered as above, and compliance in collisions will be desired, along with abilities to fall safely, as well as not being easily broken, and not hurting people when struck. A self-defense robot should have some basic ability to predict intentions, e.g., to block an attacker or avoid blocking a victim from escaping. This problem could be interesting, since attackers could sometimes try to hide their true intentions, e.g., pretending to have stopped attacking, or to be complying, before resuming or escaping (i.e., "Byzantine" or adversarial intention recognition). Even bystanders, victims, and the robot itself could seek to mislead the attacker through their behaviors.

If a robot decides it should intervene, a next step might be to consider how. Some principles that should be considered include "distinction" (attackers and bystanders must be distinguished) and "proportionality" (damage to other people or property should not be too much given the advantage realized by an action) [74], in addition to "economy of force" (minimum effort should be directed toward secondary actions in order to maximize the likelihood of its primary objective), and when to move fast and when to seek to slow the situation. Another question is the kind of force that should be used. Non-lethal tools such as bean bags, electroshock tools, directed-energy tools, batons, chemical irritants, high-pressure water, marking paint, or smoke could be used, but robots might not need to be limited to traditional means used by humans: A robot could also seek to trap an attacker using nets or sticky foam, dazzle with lasers, or emit loud sounds (e.g., high frequency if an attacker is young). Song and Yamada reported that blinking red light could be used by a

robot to warn people to stay away, which could perhaps be used to get attackers to flee or help victims to escape [133]. Furthermore, to engage in an attack, a robot might have to override some kind of safety mechanisms; for example, an AV might be given the ability to override the feature to prevent running over people when facing a violent carjacker.

A more advanced robot could even, as noted previously, be aware of tricky situations such as domestic violence, when a victim might not even want to be helped (e.g., out of fear for later reprisal), how to deal with active versus barricaded shooters, how to communicate effectively with human law enforcement officers (e.g., with gestures, in noisy environments, or to avoid betraying its position to some attacker), or persuasive psychology. Other complex topics could relate to armed conflict, urban warfare, close quarters battle, cover fire, room-clearing, rhizome manoeuvres (surprise attacks from an unexpected direction, such as through a wall or floor), dynamic defense, erratic tactics, and denial strategies. For example, some inspiration could be taken from Flaherty's ideas of 3D tactics for a soldier paired with a small drone [134]. Furthermore, although we believe that the important, fundamental scenario for self-defense robots involves credible physical harm to a person, future robots might also be capable of acting to protect objects in the environment. As it has been proposed that some robots could protect seashores from garbage and plants [135,136], robots could also protect people's pets, or prevent destruction or theft of property that could be dangerous, like in the case of arson or the theft of a firearm, or medical device.

During or after an attack, a robot might also need to obtain consent and information from people involved, since human self-defenders such as police must sometimes request consent from victims and attackers to conduct searches or pass on information, and ask questions about what might have happened before they arrived, especially if the situation seems complex or confusing. Weng et al. explored some of the potential problems that designers might need to consider if a robot is used to obtain informed consent [137]. The authors observed that small acts of deception on the part of their NAO robot were largely accepted without questions, that the robot could misconstrue a bystander's interference as a user having consented, and that participants did not perceive any potential danger from the robot being very close to them due to its cute appearance. A benefit of robots was found by Bethel et al., who observed that eyewitnesses were less misled by robots than human interviewers, suggesting that robots could be useful for extracting accurate information about attacks [138]. As well, Singh et al. describe a robot system called "Social Mobile Advanced Robot Test-Bed—SMART", which was intended to identify a criminal based on verbally querying a victim [139]. Similar to police body cameras and dash cams, robot camera feeds can also be recorded for accountability, while blurring faces to preserve privacy. A short paper by Rueben et al. presented some feedback from users who explored three strategies for indicating objects that should be kept blurred for privacy in a robot's video stream, using physical markers, a wand for pointing, and a graphical user interface [140].

As well, by getting between an attacker and their victim, a robot might also become targeted with violence (as described in the previous section) and require some technical capabilities to deal with it. For example, Xia et al. reported on a system that can recognize punching from first-person camera feed [141], and Garg and Nirupam described an idea for how a drone could detect and avoid objects thrown toward it [142].

After an attack, a robot should be able to follow-up in an appropriate way. First aid could be rendered to save lives. For example, medical professionals or security staff with first aid training could conduct telesurgery (i.e., remote surgery) [143]. Furthermore, the victim might require empathy, advice, or counseling. For example, Trost et al. described how empathy displayed by a robot reduced fear and pain [144]. Finally, any data relevant for law enforcement should be remitted in a responsible way [22]. Any intervention or "seizure" of an attacker involving a robot might also be reported, e.g., potentially along the lines of Graham analysis (describing the severity of the attack, immediate threat, and resistance), or as Wanebo proposes, also describing other factors such as time available, force options, and the fluidity of the encounter [145].

2.6. Control

In order to accomplish tasks, a model for control is required. Some papers we reviewed related to effective control of robots in emergencies, how varying degrees of control might be perceived, and how to deal with failures when they invariably will occur.

2.6.1. Situation Awareness

Given that autonomous robots capable of self-defense are not yet a reality, it seems likely that initial uses might require humans in the loop, which in turn will require interfaces that allow adequate situation awareness, but it is not clear how a self-defense robot should be controlled. For example, should there be one operator per robot, one operator for multiple robots (only taking control in the rare occasion that an addressable threat is detected, which might be practical in terms of cost), or multiple operators for a single robot (given that mistakes could cost precious lives)? Within the context of Urban Search and Rescue (USAR), Casper and Murphy reported on some actual experiences of rescue workers teleoperating two kinds of robots in a trial emergency, in teams that were predominantly two humans to one robot, comprising an "Incident Commander" to make decisions and an operator [146]. Occasional problems with communication and collisions were reported, which we believe could also be important to deal with in the current context. Furthermore, in regard to interfaces, Ventura and Lima aimed to enable easy teleoperation of their robot RAPOSA, by facilitating situational awareness through Head Mounted Displays (HMDs) and pan&tilt stereo cameras [147]. As well, Harriott and Adams proposed to explore if Human Performance Moderator Functions (HPMFs) could be used to check the workload and performance of humans in human–robot teams [148]—which seemed also relevant for the controllers of a self-defense robot, who should not experience excessive pressure in order to make good decisions and pilot the robot effectively.

2.6.2. Perception of Robot Autonomy

Another central topic seemed to be the degree of autonomy that should be used. Some people have described a preference for teleoperated robots. People reported feeling more secure when interacting with a teleoperated robot, despite acknowledging that machines can make fewer error than people, mentioning fear of new technologies [149]. People seemed to feel less accepting of an autonomous robot's advice, and to spend more effort explaining to it, than when a robot appears to be controlled or there is uncertainty [150]. As well, one tricky problem, as with AVs, is that partial autonomy on the part of the vehicle can result in greater delays in an emergency for humans who are not actively driving or concentrating, but must suddenly focus, leap in and carry out some action.

Others have expressed benefits of robot autonomy. Tozadore et al. and Bennett et al. found that children and university students perceived a robot as less intelligent once they knew it was teleoperated and not autonomous [151,152]. As well, it might not be feasible to control all robots completely, given there might be many such robots, potential problems recruiting enough human operators, and that controlling minute motions could be tedious and tiring. Thus, while some human input might be useful to ensure safety, deal with unexpected situations, and use human knowledge to realize excellent outcomes, a high degree of autonomy could allow a single operator to control multiple robots [153].

We propose that a self-defense robot could be operated at different levels of autonomy, such as the six levels specified in the Society of Automotive Engineers' (SAE) J3016 standard for AVs, from Zero (no automation) to SAE Level 5 (full autonomy) [154]. Initially, operation at lower levels of autonomy would seem wise, until the required technologies are mature enough and performance is clearly sufficient. Titiriga describes in depth various scales of autonomy, and how they might be used with military robots, which could provide insights for RSD as well [114].

2.6.3. Effects of Control Failures

As Tian et al. noted, robots often fail, which can reduce people's beliefs in a robot's performance and social competence, making error detection and recovery useful capabilities [155]. For example, Halbach et al. mentioned that children's excitement for a robot diminished after technical problems, resulting in the children calling the robot stupid or strange [156], and comic depictions of robots falling off staircases or toppling when failing to kick a ball abound on social media.

In some contexts, robot failures can lead to benefits. For example, failures observed by robot designers can expose opportunities for improvement. Failures can also elicit a desired behavior, as in Cunningham's Law, which states that the best way to get a correct answer might not be to ask a question, but rather to provide a wrong answer [157,158]. Furthermore, robot failures can put people at ease, and motivate them to "learn by teaching" in an educational setting [159]. However, failures in the case of self-defense might be undesirable, given the potential consequences of failure.

Some problems reported by the French army and U.S. Police included low battery life, difficulty getting the robot to move, insufficient video quality, pacing before a hill without climbing, as well as swaying and falling down stairs or in tall grass [25,160]. Likewise, for the KnightScope security robot, police cannot yet use some features (e.g., continual monitoring is not possible due to data caps and lack of time) [26]. The same article describes how a KnightScope security robot slipped on stairs, falling into a fountain in 2017, and how a woman pressed an emergency button on the robot in 2019 to report a fight but found that the button was not connected to the police department.

Strategies for dealing with a robot's failure would seem useful. For example, Olsen et al. present an algorithm for rapidly recovering from catastrophic failures, exemplified within the context of a team of robots pursuing an evading person (as might be the case if self-defense robots seek to stop a violent attacker from escaping and committing further violence) [161]. We believe that inspiration can also be taken from standards for AVs, such as ISO (International Organization for Standardization) 26262, governing functional safety and malfunctions, ISO 21434, dealing with cybersecurity, and perhaps especially "Safety Of The Intended Functionality" (SOTIF: ISO/Publicly Available Specification (PAS) 21448) [162], which addresses new challenges that autonomous robots face for inherently imperfect functions such as classification. There, the aim is to move from unknown unsafe states, to known unsafe states, to known safe states, with the latter most preferable, by identifying and mitigating hazards.

2.7. Culture

Of various gaps that seemed to exist, one important factor that did not seem to have been clearly treated was culture, which we felt was an overarching theme. Some cultural differences, such as linguistic differences, seem clear; e.g., in order to effectively persuade attackers to desist, perhaps allowing them to save face, a robot in Korea might require the ability to use honorifics, rather than the indirect speech acts often used in English [163]. However, in general, it can be difficult to make conclusions about cultural contexts, given their complexity. For example, although various authors have speculated that attitudes in Japan toward social robots might be different due to familiarity, animistic beliefs, positive portrayals in the media, and optimism regarding loss of jobs or privacy, in some cases, the actual evidence can be conflicting and unclear [164]. Likewise, Brohl et al. comparing cross-cultural attitudes toward robots in Japan, the U.S., China, and Germany, found that the results could not be simply explained by classifying countries as "eastern" or "western" and speculated that various factors come to play [165].

Yet, despite the challenges, modern psychology argues we never "cast aside our cultural dressings to reveal the naked universal human mind", and that culture must be taken into account when we consider people's beliefs, not only as we identify differences, but also to find what might be shared [166]. For RSD, various cultures could be examined. We focused on countries where we expected we might be able to observe cases of robots

in self-defense situations (i.e., countries with high robot prevalence). Given that the two regions with the most robots in use in 2014 were Japan (306,700), followed by North America (237,400), where most are in the U.S. [167], we focused on the Japan and the U.S.

As noted, Japan and the U.S. appear similar in some respects, e.g., in terms of their high use of robotics and economic strength, and neither country is regarded as highly dangerous (U.S. is the 79th most dangerous country in the world, and Japan is 187th, or conversely the 6th safest country) [168]. Nonetheless, Grinshteyn and Hemenway note some cultural differences related to violence, as follows [169]. The U.S. seems more dangerous (homicide rate: 0.3 in Japan vs. 5.3 in the U.S.). As well, the kind of violence is different, with firearms much more prevalent (firearm death rate: 0.0 in Japan vs. 10.2 in the U.S.). However, certain kinds of violence, such as suicide, are more common in Japan (suicide rate: 23.1 in Japan vs. 12.4 in the U.S.). Various explanations exist for these differences: Aside from the high availability of firearms, the U.S. is a highly heterogeneous country, with high income/wealth inequality, many immigrants, and proximity to the most dangerous countries in the world (e.g., El Salvador, Venezuela, Colombia, Guatemala, Honduras, Brazil, The Bahamas, and Haiti), where violent death rates differ based on region and ethnicity [168,169]. While both countries might be said to have some kind of martial culture (Western movies romanticizing gunslingers in the U.S., versus martial arts in Japan), notwithstanding some exceptions (such as the assassination of former Prime Minister Abe via a firearm in 2022 [170]), laws regarding firearms differ highly. Furthermore, one explanation for the high suicide rate in Japan has been that there might be high pressure on people in Japan to conform, sacrificing individual goals for the group (collectivism rather than individualism) and internalizing aggression, as well as indifference to others' problems, which could affect how RSD is perceived [171].

Likewise, there seems to be differences in terms of self-defense laws. In relatively homogeneous Japan, self-defense of one's self and others involving appropriate force is permissible when criminal aggression is impending and otherwise unavoidable, according to Article 36 Paragraph 1–2 of the Japanese criminal code [172]. By contrast, in the U.S., laws vary greatly by state. While the "castle doctrine" protects citizens attacked in their own homes, in a little over half of the states (28), the law also indicates no duty to retreat when threatened by unlawful and immediate danger in a place where the defender is lawfully present (in 10, more protective "stand your ground" language is included). In 16 states, there also exists a "presumption of reasonableness", or "presumption of fear", in which the burden of proof is placed on the prosecutor to show that a defender has not acted reasonably [173,174]. As well, some differences in opinions can be seen in regard to one form of robot, the AV. Attitudes in the U.S. have been found to be more positive than in Japan, who were least willing to pay [175].

One point of caution in regard to RSD is that a rise in violent crimes in the U.S. was seen after expanding self-defense rights via stand-your-ground laws, possibly due to an increase in the number of violent incidents [176]. It would be regrettable if RSD might also lead to more violence and inequality. People might be slower/less inclined to flee if they think robot could help, and robot owners might come from some economic groups/ethnicities more than others (e.g., rich white persons). While we believe that introducing unarmed robots to help might not have the same effects as encouraging defense (by people who often use firearms), we believe nonetheless that lawmakers and designers should be careful to avoid such effects.

Thus, Japan and the U.S. seemed to be a good target for cross-cultural exploration, given their similar prevalence of robots but differences in regard to homogeneity, safety, firearms, AVs, etc.

2.8. Own Work

The scoping review suggested that various gaps should be explored, including the effects of culture on public opinion. In one study similar to our own work, Gallimore et al. showed a video to Amazon Mechanical Turk (AMT) participants to explore public

opinion of autonomous robots that could harm people [177]. The participants preferred the robot shown, a large sentry robot using non-lethal force (a strobe light) to guard a security checkpoint, within a military, rather than a public, context. Furthermore, female participants perceived the robot to be more competent and benevolent, and expressed greater willingness to have such a robot at a hospital or college campus than male participants. In our previous study in Japan, focused on RSD, we initially explored two factors that we felt might strongly influence people's perceptions of the acceptability of RSD: the role of embodiment (humanoid and AV), and force (pushing vs. disarming or firearms) [21,178]. However, it was not clear if our results were limited only to Japanese participants, or potentially applicable to other cultures as well. Therefore, we focused on examining the effects of culture on the public's perception of the acceptability of RSD in the second portion of this article.

3. Study Comparing Cultures

To gain insight into how much public opinion might differ in different countries, we extended our previous study conducted in Japan with new participants from the U.S.

3.1. Hypotheses

Since there is more crime in the U.S., we considered that U.S. participants might feel more strongly about the need to do something to reduce crime, and thus be more accepting of defense in general, including robot defenders. This includes that U.S. participants have expressed a more positive opinion about AVs in the past. Furthermore, since firearms are more prevalent in the U.S., U.S. participants might be more accepting of lethal force. This led to the hypotheses below:

Hypothesis 1. *(Embodiment) Participants in the U.S. will be more accepting of the idea of self-defense and robot defenders than participants in Japan.*

Hypothesis 2. *(Force) Participants in the U.S. will be more accepting of the idea of use of lethal force by a robot than participants in Japan.*

3.2. Participants

In this study, 304 participants in Japan (157 women, 146 men, 1 who declined to specify; average age: 41.4 years, SD = 9.7; recruited via a recruiting agency in Japan) and 307 participants in the U.S. (94 women, 211 men, 2 who declined to specify; average age: 35.3 years, SD = 9.9; recruited via AMT) participated in our survey. Through a filtering process that checks invalid answers with missing data or the same value input for each question, we extracted 299 participants' data in Japan and 249 participants' data in the U.S. All participants received some small compensation (<USD 5), regardless of the validity of their data.

3.3. Measurements

To investigate the perceived acceptability of RSD, we used a seven-point response format with this question: "I can accept the actions of the defender (human or robot) who comes to the aid of the person being attacked." (1: strongly disagree, 7: strongly agree).

3.4. Procedure

First, the participants read explanations of the aims of our study via a web page. Participants who agreed to join the survey viewed instructions about this study, including an introduction of the three characters (human, humanoid robot, and AV). Then they watched eight videos in random order, and answered a questionnaire about the perceived acceptability for each video.

3.5. Videos

The eight videos developed for our previous study in Japan were used again, as is. Generic characters were used to represent humans and robots, as shown in Figure 6, and animated using Autodesk 3ds Max, Autodesk Maya, and Unreal Engine 4.27. (A male human character was selected since males commit more violent crimes than females [179].) Videos V1–V5 were designed to gain insight into effects of embodiment (human, humanoid robot, or AV), and V6-8 to investigate how different degrees of force were perceived (firearm, pushing, blocking), as shown in Figure 7. For our initial exploration, not all combinations were tested, since some cases were expected to be more interesting than others. In each eight second video, the storyline is the same: an attacker on the left attacks a victim on the right, causing a defender to show up from the bottom of the screen to prevent the attack, as summarized in Figure 1. The details for the videos, which can all be seen online [180], are shown below:

- *V1.* A Human Defender stops a Human Attacker, both using Non-lethal force.
- *V2.* A Humanoid Robot Defender stops a Human Attacker, both using Non-lethal force.
- *V3.* An AV stops a Human Attacker, both using Non-lethal force.
- *V4.* A Human stops a Humanoid Robot Attacker, both using Non-lethal force.
- *V5.* A Humanoid Robot Defender stops a Humanoid Robot Attacker, both using Non-lethal force.
- *V6.* A Humanoid Robot Defender stops a Human Attacker, both using Lethal force.
- *V7.* A Humanoid Robot Defender stops a Human Attacker, using Non-lethal force against Lethal force.
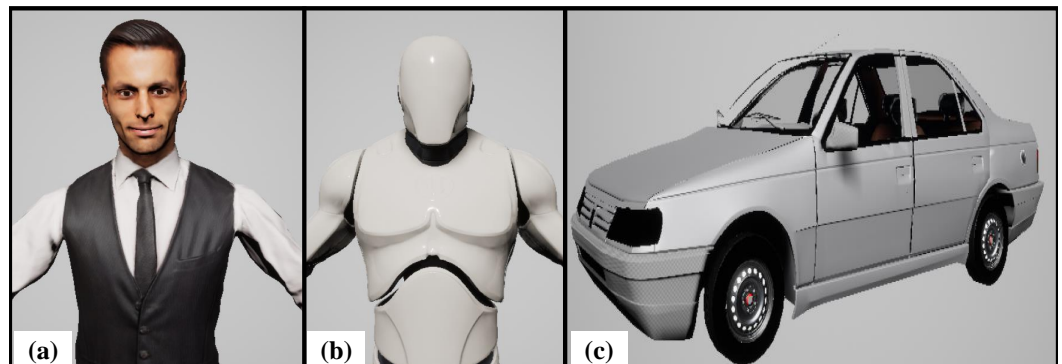- *V8.* A Humanoid Robot Defender stops a Human Attacker, using Disarming against Lethal force.

Figure 6. Characters used in the animations: (**a**) human (both attacker and victim), (**b**) humanoid robot, (**c**) autonomous vehicle (AV).
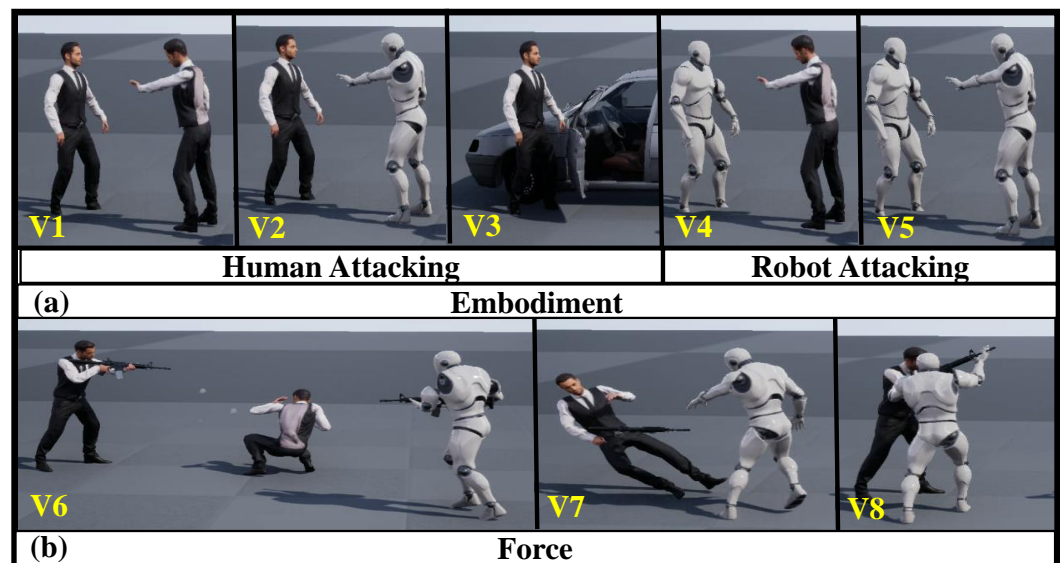
Figure 7. The animated videos (in each case, the attacker is on the left, and the defender on the right): (**a**) Embodiment. V1–V3: A human attacker is stopped by a (human/humanoid/AV) defender. V4–V5: A humanoid attacker is stopped by a (human/humanoid) defender. All use the same non-lethal force (pushing). (**b**) Force. V6–V8: A human attacker using lethal force is stopped by a humanoid defender using (lethal force/pushing/disarming).

3.6. Statistical Analysis

Figure 8 shows the averages and standard errors for questionnaire results. We conducted a two-factor mixed ANOVA with SPSS [181] for the perceived acceptability on culture factor and video-type factor. Two-factor mixed ANOVA with SPSS has been conducted before in HRI studies (e.g., to explore the positive effects of exercising with a robot [182]). Mauchly's test indicated that the assumption of sphericity had been violated $\chi^2(27) = 451.991$, $p < 0.001$, therefore a Huynh–Feldt correction was applied ($\epsilon = 0.802$). Figure 9 summarizes the significant differences that were found, which are also described below. We found significant differences for the video-type factor ($F(5.691, 3107.141) = 111.486$, $p < 0.001$, $\eta_p^2 = 0.170$), and the interaction effect ($F(5.691, 3107.141) = 42.265$, $p < 0.001$, $\eta_p^2 = 0.073$). We did not find an overall significant difference for the culture factor ($F(1, 7.639) = 3.655$, $p = 0.056$, $\eta_p^2 = 0.007$).

The simple main effects showed significant differences: V1 (U.S. > Japan, $p = 0.032$), V2 (U.S. > Japan, $p < 0.001$), V3 (U.S. > Japan, $p < 0.001$), V6 (U.S. > Japan, $p < 0.001$), V7 (Japan > U.S. $p < 0.001$), and V8 (Japan > U.S., $p < 0.001$), Multiple comparison with Bonferroni method showed significant differences in Japan: V1 > V2, V3, and V6 (all p values < 0.001), V2 > V3 ($p < 0.001$) and V2 > V6 ($p = 0.008$), V4 > V1, V2, V3, and V6 (all p values < 0.001), V5 > V1, V2, V3 and V6 (all p values < 0.001), V6 > V3 ($p < 0.001$), V7 > V1, V2, V3, V4, V5 and V6 (all p values < 0.001), and V8 > V1, V2, V3, V4, V5, V6 and V7 (all p values < 0.001). Multiple comparison with Bonferroni method showed significant differences in the U.S.: V1 > V3 ($p < 0.001$), V2 > V3 ($p < 0.001$), V4 > V3 ($p = 0.004$), V5 > V3 ($p < 0.001$), V7 > V3 and V6 (all p values < 0.001), V8 > V3 ($p < 0.001$) and V8 > V6 ($p = 0.002$).
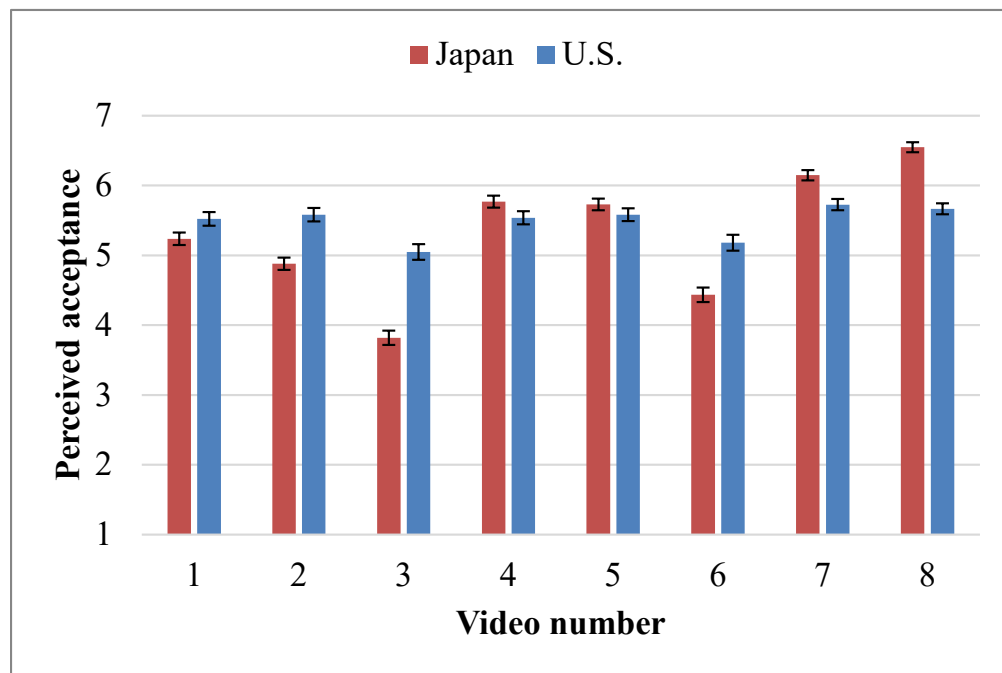
Figure 8. Questionnaire results.

Source	*F*	*p*
Video(V)	111.486	<0.001
Country(C)	3.655	0.056
V × C	43.265	<0.001

	V1 <	V4, V5, V7, V8
	V2 <	V1, V4, V5, V7, V8
	V3 <	V1, V2, V4, V5, V6, V7, V8
Japan	V4, V5 <	V7, V8
	V6 <	V1, V2, V4, V5, V7, V8
	V7 <	V8

US > Japan	V1, V2, V3, V6
Japan > US	V7, V8

	V3 <	V1, V2, V4, V5, V7, V8
US	V6 <	V7, V8

Figure 9. Summary of the statistical differences found.

3.7. Summary of Results

The experiment results appeared to show some similar patterns: the least acceptable defense involved AVs and a robot's lethal force (videos V3 and V6), and most acceptable involved non-lethal force used to defend against a lethal attack (videos V7 and V8). Public opinion dipped below neutral in only one case (the AV defender in Japan). The results did not show that participants in the U.S. were significantly more positive about self-defense in general than in Japan, but perception differed between countries related to the identify of the defender. The participants in Japan preferred human defenders more than robot defenders significantly, but the U.S. participants did not significantly prefer human defenders compared to robot defenders. As well, the AV defender was more acceptable in the U.S. Thus, Hypothesis 1 was partially supported. As well, the U.S. participants were more accepting of the self-defense of robots, except for non-lethal and blocking defenses in the lethal case. In other words, the participants in Japan preferred non-lethal and blocking defenses of the robot more than other types of self-defense. Thus, Hypothesis 2, that U.S. participants would accept firearms more, was supported.

4. Discussion

In summary, the aim of the current article was to gain insight into the "big picture" in regard to the possibility for future robots in our surroundings to help us when we are threatened with violence. A rapid scoping review was conducted on academic thoughts about how robots will be expected to defend and protect against violence and crime, and a study was conducted to gain insight into public opinion on acceptability. As expected, various initial work seems to be in progress in various areas. Our contributions are summarized below:

- *Background.* We gathered together recent information on the usage of robots by law enforcement, military, and pest control groups in various countries, that could potentially be developed and adapted for self-defense.
- *Risks for people.* We discussed potential merits and demerits of RSD from the perspectives of historical attitudes to previous force-wielding robots, comparison with humans, as well as the unique qualities of RSD. Furthermore, we contrasted the fundamental dilemma of RSD with that in the well-known trolley problem, pointing out some similarities and three differences (in regard to utilitarian clarity, opposite emphasis on "letting die", and prevalence of the scenario and problem).
- *Negative perceptions.* We put forth questions and proposals about how self-defense robots might be designed to put people at ease and avoid violence (e.g., a stable, neutral embodiment with adept communication abilities).
- *Tasks.* We made various proposals (e.g., exploring how various military ideas such as ISTAR and rules of engagement could be translated to RSD, raising the problem of "Byzantine" intention recognition, and describing objects that might also be important to defend, etc.).
- *Control.* We proposed extending AV standards to self-defense robots, such as SAE's J3016 standard for levels of autonomy, and the SOTIF (ISO/PAS 21448) standard for dealing with recognition failures.
- *Culture.* Since few studies seemed to have looked at cultural influences on RSD, a study was conducted, revealing some cultural differences. A small preference for human defenders found in Japan was not observed in the U.S. As well, the idea of lethal force by a robot was more acceptable in the U.S.

4.1. Limitations and Future Work

The rapid scoping review that was performed had several limitations. The search results were dependent on the search query we used, as well as the databases checked, as well as when the search was conducted, given that search results change over time. Only two reviewers were used, a librarian was not consulted in refining the search query, and follow-up interviews with experts were not conducted. Some few papers were also excluded due to language and type.

As well, for the cultural study, the results are limited first by the countries tested and number of participants, as well as potential effects of age, gender, local attitudes, and demand characteristics. For example, Japanese and Americans are typically familiar with depictions of robots in movies, books, and other media, whereas participants from some other countries might have less exposure and preconceptions. As well, regarding age and gender, young Americans consider themselves more open-minded [183] (thus possibly more in favor of new ideas involving robots?), and males tend to be more in favor of firearms and use of force than females [184]. Furthermore, we do not know which states our U.S. respondents came from (the results might have a different meaning if they came from those regions in which support for self-defense, firearms and AVs is higher). While the geographic distribution of respondents appears to be similar to that of the U.S. population, with most coming from California, Texas, Illinois, Florida, North Carolina and New York [185], big cities with high population density and education rates tend toward liberal ideas [186]. This might have led to a positive trend (optimism for new ideas and AVs) or a negative trend (pessimism toward increased options for the use of force,

firearms, and stronger laws of self-defense). Furthermore, demand characteristics might have played a role. As with police advisory boards that changed opinions on use of robots when they received backlash, some participants might have said what they thought would be socially acceptable.

Future work will include addressing the limitations above by conducting more thorough reviews, and considering gender, age, and specific locations. As well, for law enforcement and military robots, next developments should be tracked. For negativity toward robots, insight would be useful into why violence towards robots occurs, and how different target groups can be handled in a good way, such as drunk persons. For ethics and law, we should explore various cases: e.g., how people perceive RSD when potentially lives could be lost if a robot is damaged (attacks on a nurse robot, rather than on a cleaning robot), how opinion is affected by if attacks are repeated or provoked by a victim, how acceptable RSD would be for other embodiments such as drones, if an AV's defense is less acceptable than a humanoid with a firearm when the attack is lethal for both, what happens if victims are children or female, what if the number of victims or attackers differs, or what if there is no clear attacker (e.g., both humans are attackers). For tasks, autonomous technologies that could be used will be developed, including robust recognition; e.g., approaches such as OpenPose might be used to detect pose of attackers and victims. Subtle cues such as signs of an impending attack can be read, and intentions inferred. For control, studies will be conducted to determine differences in how people perceive RSD in a manual or autonomous robot. As well, we should explore how people feel if a self-defense robot tries but fails. Thus, much work remains to be performed, although we believe that advances in robust recognition for example could benefit areas outside of only RSD as well.

4.2. The Genovese Case—Revisited

On 13 March 2044, a young woman returning home from work in the early morning is prompted by the parking lot that a man is approaching her who might be armed. She is able to look back over her shoulder in time and see the knife coming—time buys options. The AV she had taken back home jumps back to life, getting between her and her attacker as she runs toward the building. In the country where she lives, it was felt that AVs should not be able to use force, so the AV merely continues to obstruct and sends a message to law enforcement, blaring an alarm and flooding the area with pulsing red light. The man catches up and grabs her with a leveraging arm, only to be pushed backwards by the building's custodian, a humanoid robot. The robot is designed safely, with a stable base resembling a PackBot, and neutral features. Due to the sudden nature of the emergency, the robot initially moves autonomously, deciding that an intervention is necessary, predicting where the attacker is trying to go, and the least amount of force needed to deter him. Suddenly, a voice is heard from the robot, which still holds the struggling attacker; law enforcement have taken control of the robot, and now the situation seems much more positive. The young woman has reached the locked door of the building, which having received information about the emergency, has quickly opened to let her in. She is shaken by the experience—but safe and sound.

Thus, in summary, the take home message then of this article might be that: Self-defense by a robot could be permissible if various requirements are met (e.g., it detects an ongoing or imminent threat that cannot be handled without force, its intervention is judged likely to result in less harm than not intervening, and it has sufficient confidence in its abilities; i.e., it is deemed that its recognition capabilities allow it to accurately grasp what is happening and assess risk, and its behavioral capabilities are sufficient to deter the attack). Is this full scenario likely to be a reality soon? The answer is probably not. This task seems like it would require various advanced abilities that might not be possible for most current robots. Does this mean it is useless to think about RSD? We believe no, that such conceptualizations fall within the umbrella of speculative design, which seeks to provoke thought and stimulate discussion on important topics (before it may be too late). Will we first see technologically mature, dedicated "police robots" before self-defense

becomes a duty for robots in general? The answer is maybe yes, maybe no. As the article describes, there is a broad range of capabilities that robots could employ, some of which are feasible today (e.g., an emergency button that generates a loud noise to alert attention). We look forward with interest to how robots might begin to take on increasingly more human-like responsibilities, thereby "paying back" society and potentially contributing to human well-being.

Author Contributions: Conceptualization, M.S. and M.C.; review, M.C. and E.K.D.; survey data acquisition and statistical analysis, M.S.; writing—original draft preparation, M.C.; writing—review and editing, M.S., E.K.D. and A.V. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was approved by the ethics committee at the Advanced Telecommunication Research Institute (ATR) (21-523).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The videos used for our survey can be found at: youtube.com/watch?v=F5y7dPy41p0&list=PLtGv2XOitdkQsEsPX528cmat5cnVgd15V (accessed on 11 March 2023).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of the data; in the writing of the manuscript; nor in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
AMT	Amazon Mechanical Turk
AV	Autonomous vehicle
HRI	Human–robot interaction
ISO	International Organization for Standardization
ISR	Intelligence, surveillance, and reconnaissance
ISTAR	Intelligence, surveillance, target acquisition, and reconnaissance
PAS	Publicly available specification
RSD	Robot self-defense
SAE	Society of Automotive Engineers
SOTIF	Safety of the intended functionality
STD	Sexually transmitted disease
TNT	2,4,6-trinitrotoluene
UN	United Nations
U.S.	United States
USD	United States Dollars

References

1. Takooshian, H. Not Just a Bystander: The 1964 Kitty Genovese Tragedy: What Have We Learned. *Psychology Today*, 24 May 2014.
2. Active Self Protection. Available online: https://www.youtube.com/@ActiveSelfProtection (accessed on 4 March 2023).
3. The Economic Value of Peace 2016: Measuring the Global Impact of Violence and Conflict. 2016. Available online: https://reliefweb.int/report/world/economic-value-peace-2016-measuring-global-economic-impact-violence-and-conflict (accessed on 4 March 2023).
4. Krug, E.G.; Mercy, J.A.; Dahlberg, L.L.; Zwi, A.B. The world report on violence and health. *Lancet* **2002**, *360*, 1083–1088. [CrossRef] [PubMed]

5. DeLisi, M. Measuring the cost of crime. In *The Handbook of Measurement Issues in Criminology and Criminal Justice*; Wiley: Hoboken, NJ, USA, 2016; pp. 416–433.

6. Hobbes, T. *Leviathan or the Matter Form and Power of a Commonwealth, Ecclesiastical and Civil*; Simon and Schuster: London, UK, 1886; Volume 21.

7. Tung, W.F.; Jara Santiago Campos, J. User experience research on social robot application. *Libr. Hi Tech* **2022**, *40*, 914–928. [CrossRef]

8. Zarifhonarvar, A. Economics of ChatGPT: A Labor Market View on the Occupational Impact of Artificial Intelligence. 2023. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4350925 (accessed on 15 March 2023).

9. Grieco, L. ChatGPT Excitement Sends Investors Flocking to AI Stocks Like Microsoft and Google. 2023. Available online: https://www.proactiveinvestors.co.uk/companies/news/1004479/chatgpt-excitement-sends-investors-flocking-to-ai-stocks-like-microsoft-and-google-1004479.html (accessed on 4 March 2023).

10. The Britannica Dictionary: Robot. Available online: https://www.britannica.com/dictionary/robot (accessed on 4 March 2023).

11. The Britannica Dictionary: Self-Defense. Available online: https://www.britannica.com/dictionary/self%E2%80%93defense (accessed on 4 March 2023).

12. McCormack, W. Targeted Killing at a Distance: Robotics and Self-Defense. *Pac. McGeorge Global Bus. Dev. Law J.* **2012**, *25*, 361.

13. Arksey, H.; O'Malley, L. Scoping studies: Towards a methodological framework. *Int. J. Soc. Res. Methodol.* **2005**, *8*, 19–32. [CrossRef]

14. Peters, M.; Godfrey, C.; McInerney, P.; Soares, C.B.; Khalil, H.; Parker, D. *The Joanna Briggs Institute Reviewers' Manual 2015: Methodology for JBI Scoping Reviews*; The Joanna Briggs Institute: Adelaide, Australia, 2015; pp. 6–22.

15. Haring, K.S.; Novitzky, M.M.; Robinette, P.; De Visser, E.J.; Wagner, A.; Williams, T. The dark side of human–robot interaction: Ethical considerations and community guidelines for the field of HRI. In Proceedings of the 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Daegu, Republic of Korea, 11–14 March 2019; pp. 689–690.

16. Martland, C.D. Analysis of the potential impacts of automation and robotics on locomotive rebuilding. *IEEE Trans. Eng. Manag.* **1987**, *34*, 92–100. [CrossRef]

17. Cha, S.S. Customers' intention to use robot-serviced restaurants in Korea: relationship of coolness and MCI factors. *Int. J. Contemp. Hosp. Manag.* **2020**, *32*, 2947–2968. [CrossRef]

18. Andersen, M.M.; Schjoedt, U.; Price, H.; Rosas, F.E.; Scrivner, C.; Clasen, M. Playing with fear: A field study in recreational horror. *Psychol. Sci.* **2020**, *31*, 1497–1510. [CrossRef]

19. Zillmann, D. Excitation transfer theory. In *The International Encyclopedia of Communication*; Wiley: Hoboken, NJ, USA, 2008.

20. Wood, J.T. Gendered media: The influence of media on views of gender. *Gendered Lives Commun. Gender Cult.* **1994**, *9*, 231–244.

21. Duarte, E.K.; Shiomi, M.; Vinel, A.; Cooney, M. Robot Self-defense: Robots Can Use Force on Human Attackers to Defend Victims. In Proceedings of the 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Napoli, Italy, 29 August–2 September 2022; pp. 1606–1613.

22. Reid, M. Rethinking the fourth amendment in the age of supercomputers, artificial intelligence, and robots. *West Va. Law Rev.* **2016**, *119*, 863.

23. Gwozdz, J.; Morin, N.; Mowris, R.P. Enabling Semi-Autonomous Manipulation on iRobot's Packbot. Bachelor's Thesis, Worcester Polytechnic Institute, Worcester, MA, USA, 2014.

24. Carruth, D.W.; Bethel, C.L. Challenges with the integration of robotics into tactical team operations. In Proceedings of the 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI), Herl'any, Slovakia, 26–28 January 2017.

25. Hayes, M. The Creepy Robot Dog Botched a Test Run With a Bomb Squad. 2020. Available online: https://onezero.medium.com/boston-dynamics-robot-dog-got-stuck-in-sit-mode-during-police-test-emails-reveal-4c8592c7fc2 (accessed on 4 March 2023).

26. Farivar, C. Security Robots Expand Across U.S., with Few Tangible Results. 2021. Available online: https://www.nbcnews.com/business/business-news/security-robots-expand-across-u-s-few-tangible-results-n1272421 (accessed on 4 March 2023).

27. Jeffrey-Wilensky, J.; Freeman, D. This Police Robot Could Make Traffic Stops Safer. 2019. Available online: https://www.nbcnews.com/mach/science/police-robot-could-make-traffic-stops-safer-ncna1002501 (accessed on 4 March 2023).

28. Jackson, R.D. "I Approved It. . . And I'll Do It Again": Robotic Policing and Its Potential for Increasing Excessive Force. In *Societal Challenges in the Smart Society*; Universidad de La Rioja: La Rioja, Spain, 2020; pp. 511–522.

29. Prabakar, M.; Kim, J.H. TeleBot: Design concept of telepresence robot for law enforcement. In Proceedings of the 2013 World Congress on Advances in Nano, Biomechanics, Robotics, and Energy Research (ANBRE 2013), Seoul, Republic of Korea, 25–28 August 2013.

30. Simmons, R. Terry in the Age of Automated Police Officers. *Seton Hall Law Rev.* **2019**, *50*, 909. [CrossRef]

31. Terzian, D. The right to bear (robotic) arms. *Penn. State Law Rev.* **2012**, *117*, 755. [CrossRef]

32. Donlon, M. Automation at the Airport. 2022. Available online: https://electronics360.globalspec.com/article/18513/automation-at-the-airport (accessed on 5 March 2023).

33. Lufkin, B. What the World Can Learn from Japan's Robots. 2020. Available online: https://www.bbc.com/worklife/article/20200205-what-the-world-can-learn-from-japans-robots (accessed on 4 March 2023).

34. Glaser, A. 11 Police Robots Patrolling Around the World. 2019. Available online: https://www.wired.com/2016/07/11-police-robots-patrolling-around-world/ (accessed on 5 March 2023).

35. Maliphol, S.; Hamilton, C. Smart Policing: Ethical Issues & Technology Management of Robocops. In Proceedings of the 2022 Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, USA, 7–11 August 2022; pp. 1–15.

36. Chia, O. Keeping Watch from the Skies: Police Unveil Two New Drones for Crowd Management, Search and Rescue. 2023. Available online: https://www.straitstimes.com/singapore/courts-crime/keeping-watch-from-the-skies-police-unveil-two-new-drones-for-crowd-management-search-and-rescue (accessed on 4 March 2023).

37. Jain, R.; Nagrath, P.; Thakur, N.; Saini, D.; Sharma, N.; Hemanth, D.J. Towards a smarter surveillance solution: The convergence of smart city and energy efficient unmanned aerial vehicle technologies. In *Development and Future of Internet of Drones (IoD): Insights, Trends and Road Ahead*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 109–140.

38. Williams, O. China Now Has Robot Police with Facial Recognition. 2017. Available online: https://www.huffingtonpost.co.uk/entry/china-now-has-robot-police-with-facial-recognition_uk_58ac54d2e4b0f077b3ee0dcb (accessed on 5 March 2023).

39. Joh, E.E. Policing Police Robots. 2016. Available online: https://www.uclalawreview.org/policing-police-robots (accessed on 4 March 2023).

40. Kim, L. Meet South Korea's New Robotic Prison Guards. 2012. Available online: https://www.digitaltrends.com/cool-tech/meet-south-koreas-new-robotic-prison-guards (accessed on 4 March 2023).

41. McKay, C. The carceral automaton: Digital prisons and technologies of detention. *Int. J. Crime Justice Soc. Democr.* **2022**, *11*, 100–119. [CrossRef]

42. White, L. "Iron Man" Power Armour and Robot Dogs Coming to Korean Police. 2022. Available online: https://stealthoptional.com/robotics/iron-man-power-armour-and-robot-dogs-coming-to-korean-police (accessed on 4 March 2023).

43. Carbonaro, G. Robot Paw Patrol: The US Space Force Begins Deploying Robots as Guard Dogs. 2022. Available online: https://www.euronews.com/next/2022/08/10/robot-dogs-report-for-duty-the-us-space-force-enlists-robotic-dogs-to-guard-spaceport (accessed on 4 March 2023).

44. Kemper, C.; Kolain, M. K9 Police Robots-Strolling Drones, RoboDogs, or Lethal Weapons? In Proceedings of the WeRobot 2022 Conference, Washington, DC, USA, 14–16 September 2022.

45. Dent, S. Boston Dynamics' Spot Robot Tested in Combat Training with the French Army. 2021. Available online: https://www.engadget.com/boston-dynamics-spot-robot-combat-training-101732374.html (accessed on 4 March 2023).

46. Engineering, I. Unique Military and Police Robots. 2022. Available online: https://youtu.be/Ox6A0hOYL5g (accessed on 4 March 2023).

47. Lee, W.; Bang, Y.B.; Lee, K.M.; Shin, B.H.; Paik, J.K.; Kim, I.S. Motion teaching method for complex robot links using motor current. *Int. J. Control. Autom. Syst.* **2010**, *8*, 1072–1081. [CrossRef]

48. Cao, Y.; Yamakawa, Y. Marker-less Kendo Motion Prediction Using High-speed Dual-camera System and LSTM Method. In Proceedings of the 2022 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Hokkaido, Japan, 11–15 July 2022; pp. 159–164.

49. Jamiepaik. MUSA: Kendo Robot Collective. 2009. Available online: https://www.youtube.com/watch?v=Mh3muo2V818 (accessed on 5 March 2023).

50. Electric, Y. Yaskawa Bushido Project/Industrial Robot vs. Sword Master. 2015. Available online: https://www.youtube.com/watch?v=O3XyDLbaUmU (accessed on 5 March 2023).

51. Werrell, K.P. *The Evolution of the Cruise Missile*; Technical Report; Air University Maxwell AFB: Montgomery, AL, USA, 1985.

52. Brollowski, H. Military Robots and the Principle of Humanity: Distorting the Human Face of the Law? In *Armed Conflict and International Law: In Search of the Human Face: Liber Amicorum in Memory of Avril McDonald*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 53–96.

53. Soffar, H. Modular Advanced Armed Robotic System (MAARS Robot) Features, Uses & Design. 2019. Available online: https://www.online-sciences.com/robotics/modular-advanced-armed-robotic-system-maars-robot-features-uses-design/ (accessed on 5 March 2023).

54. Chayka, K. Watch This Drone Taser a Guy Until He Collapses. 2014. Available online: https://time.com/19929/watch-this-drone-taser-a-guy-until-he-collapses/ (accessed on 5 March 2023).

55. Asaro, P. "Hands up, don't shoot!" HRI and the automation of police use of force. *J. Hum.-Robot. Interact.* **2016**, *5*, 55–69. [CrossRef]

56. Holt, K. Ghost Robotics Strapped a Gun to Its Robot Dog. 2021. Available online: https://www.engadget.com/robot-dog-gun-ghost-robotics-sword-international-175529912.html (accessed on 5 March 2023).

57. Lee, M. Terrifying Video Shows Chinese Robot Attack Dog with Machine Gun Dropped by Drone. 2022. Available online: https://nypost.com/2022/10/26/terrifying-video-shows-chinese-robot-attack-dog-with-machine-gun-dropped-by-drone/ (accessed on 5 March 2023).

58. White, L. Russian Army's RPG-Equipped Robot Dog Can Be Easily Purchased Online. 2022. Available online: https://stealthoptional.com/robotics/russian-armys-rpg-equipped-robot-dog-easily-purchased-online/ (accessed on 5 March 2023).

59. Fish, T. Loitering with Intent. 2022. Available online: https://www.asianmilitaryreview.com/2022/12/loitering-with-intent/ (accessed on 5 March 2023).

60. Kirschgens, L.A.; Ugarte, I.Z.; Uriarte, E.G.; Rosas, A.M.; Vilches, V.M. Robot hazards: From safety to security. *arXiv* **2018**, arXiv:1806.06681.

61. Sharkey, N.; Goodman, M.; Ross, N. The coming robot crime wave. *Computer* **2010**, *43*, 115–116. [CrossRef]
62. Thomas, M. Are Police Robots the Future of Law Enforcement? 2022. Available online: https://builtin.com/robotics/police-robot-law-enforcement (accessed on 5 March 2023).
63. Bergman, R.; Fassihi, F. The Scientist and the A.I.-Assisted, Remote-Control Killing Machine. *The New York Times*, 18 September 2021; Volume 18 .
64. Kallenborn, Z. Was a Flying Killer Robot Used in Libya? Quite Possibly. 2021. Available online: https://thebulletin.org/2021/05/was-a-flying-killer-robot-used-in-libya-quite-possibly/ (accessed on 5 March 2023).
65. Gonzalez, L.F.; Montes, G.A.; Puig, E.; Johnson, S.; Mengersen, K.; Gaston, K.J. Unmanned aerial vehicles (UAVs) and artificial intelligence revolutionizing wildlife monitoring and conservation. *Sensors* **2016**, *16*, 97. [CrossRef]
66. Slezak, M. Robots, Lasers, Poison: The High-Tech Bid to Cull Wild Cats in the Outback. 2016. Available online: https://www.theguardian.com/environment/2016/apr/17/robots-lasers-poison-the-high-tech-bid-to-cull-wild-cats-in-the-outback (accessed on 5 March 2023).
67. Crumley, B. AeroPest Positions Wasp-Culling Drone for Broader Applications. 2022. Available online: https://dronedj.com/2022/05/23/aeropest-positions-wasp-culling-drone-for-broader-applications/ (accessed on 5 March 2023).
68. Hecht, E. Drones in the Nagorno-Karabakh War: Analyzing the Data. *Mil. Strateg. Mag* **2022**, *7*, 31–37.
69. Lin, P.; Bekey, G.; Abney, K. *Autonomous Military Robotics: Risk, Ethics, and Design*; Technical Report; California Polytechnic State Univesristy San Luis Obispo: San Luis Obispo, CA, USA, 2008.
70. Dawes, J. UN Fails to Agree on 'Killer Robot' Ban as Nations Pour Billions Into Autonomous Weapons Research. 2021. Available online: https://theconversation.com/un-fails-to-agree-on-killer-robot-ban-as-nations-pour-billions-into-autonomous-weapons-research-173616 (accessed on 5 March 2023).
71. Froomkin, A.M.; Colangelo, P.Z. Self-defense against robots and drones. *Conn. Law Rev.* **2015**, *48*, 1.
72. King, T.C.; Aggarwal, N.; Taddeo, M.; Floridi, L. Artificial intelligence crime: An interdisciplinary analysis of foreseeable threats and solutions. *Sci. Eng. Ethics* **2020**, *26*, 89–120. [CrossRef]
73. White, L. Killer Robot Cops Are Bad Actually, Decides San Francisco Supervisory Board. 2022. Available online: https://stealthoptional.com/news/killer-robot-cops-bad-actually-decides-san-francisco-supervisory-board/ (accessed on 5 March 2023).
74. Sharkey, N.E. The evitability of autonomous robot warfare. *Int. Rev. Red Cross* **2012**, *94*, 787–799. [CrossRef]
75. Digital, C. New Robot Makes Soldiers Obsolete (Corridor Digital). 2019. Available online: https://www.youtube.com/watch?v=y3RIHnK0_NE (accessed on 5 March 2023).
76. Gambino, A.; Fox, J.; Ratan, R.A. Building a stronger CASA: Extending the computers are social actors paradigm. *Hum.-Mach. Commun.* **2020**, *1*, 5. [CrossRef]
77. Nedim, U. Is It a Crime Not to Help Someone in Danger? 2015. Available online: https://www.sydneycriminallawyers.com.au/blog/is-it-a-crime-not-to-help-someone-in-danger/ (accessed on 5 March 2023).
78. Foot, P. The problem of abortion and the doctrine of the double effect. *Oxf. Rev.* **1967**, *5*, 5–15.
79. Thomson, J.J. The trolley problem. *Yale Law J.* **1984**, *94*, 1395. [CrossRef]
80. Lim, H.S.M.; Taeihagh, A. Algorithmic decision-making in AVs: Understanding ethical and technical concerns for smart cities. *Sustainability* **2019**, *11*, 5791. [CrossRef]
81. Goodall, N.J. Ethical decision making during automated vehicle crashes. *Transp. Res. Rec.* **2014**, *2424*, 58–65. [CrossRef]
82. Lin, P. Robot cars and fake ethical dilemmas. *Forbes Magazine*, 3 April 2017.
83. Brooks, R. Unexpected Consequences of Self Driving Cars. 2017. Available online: http://rodneybrooks.com/unexpected-consequences-of-self-driving-cars (accessed on 5 March 2023).
84. Stein, B.D.; Jaycox, L.H.; Kataoka, S.; Rhodes, H.J.; Vestal, K.D. Prevalence of child and adolescent exposure to community violence. *Clin. Child Fam. Psychol. Rev.* **2003**, *6*, 247–264. [CrossRef]
85. Joh, E.E. Private security robots, artificial intelligence, and deadly force. *UCDL Rev.* **2017**, *51*, 569.
86. Calo, R. Robotics and the Lessons of Cyberlaw. *Calif. Law Rev.* **2015**, *103*, 513.
87. Robinette, P.; Li, W.; Allen, R.; Howard, A.M.; Wagner, A.R. Overtrust of robots in emergency evacuation scenarios. In Proceedings of the 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand, 7–10 March 2016; pp. 101–108.
88. Lacey, C.; Caudwell, C. Cuteness as a 'dark pattern' in home robots. In Proceedings of the 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Daegu, Republic of Korea, 11–14 March 2019; pp. 374–381.
89. Carr, N.K. Programmed to protect and serve: The dawn of drones and robots in law enforcement. *J. Air Law Commer.* **2021**, *86*, 183.
90. Mamak, K. Whether to save a robot or a human: On the ethical and legal limits of protections for robots. *Front. Robot. AI* **2021**, *8*, 712427. [CrossRef] [PubMed]
91. Vadymovych, S.Y. Artificial personal autonomy and concept of robot rights. *Eur. J. Law Political Sci.* **2017**, *1*, 17–21.
92. Garber, M. Funerals for Fallen Robots. 2013. Available online: https://www.theatlantic.com/technology/archive/2013/09/funerals-for-fallen-robots/279861 (accessed on 4 March 2023).
93. Darling, K. Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects. In *Robot Law*; Edward Elgar Publishing: Cheltenham, UK, 2016; pp. 213–232.

94. Rehm, M.; Krogsager, A. Negative affect in human robot interaction—Impoliteness in unexpected encounters with robots. In Proceedings of the 2013 IEEE RO-MAN, Gyeongju, Republic of Korea, 26–29 April 2013; pp. 45–50.

95. Connolly, J.; Mocz, V.; Salomons, N.; Valdez, J.; Tsoi, N.; Scassellati, B.; Vázquez, M. Prompting prosocial human interventions in response to robot mistreatment. In Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, Cambridge, UK, 23–26 March 2020; pp. 211–220.

96. Bartneck, C.; Keijsers, M. The morality of abusing a robot. *Paladyn J. Behav. Robot.* **2020**, *11*, 271–283. [CrossRef]

97. Takas, K.L. Exploring How and Why Women Become Involved in Relationships with Incarcerated Men. Bachelor's Thesis, University of South Florida, Tampa, FL, USA, 2004.

98. Tan, X.Z.; Vázquez, M.; Carter, E.J.; Morales, C.G.; Steinfeld, A. Inducing bystander interventions during robot abuse with social mechanisms. In Proceedings of the 2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Chicago, IL, USA, 5–8 March 2018; pp. 169–177.

99. Bartneck, C.; Rosalia, C.; Menges, R.; Deckers, I. Robot abuse-a limitation of the media equation. In Proceedings of the Interact 2005 Workshop on Abuse, Rome, Italy, 12 September 2005.

100. Eyssel, F.; Kuchenbrandt, D.; Bobinger, S. Effects of anticipated human–robot interaction and predictability of robot behavior on perceptions of anthropomorphism. In Proceedings of the 6th International Conference on Human-Robot Interaction, Lausanne, Switzerland, 6–9 March 2011; pp. 61–68.

101. Natarajan, M.; Gombolay, M. Effects of anthropomorphism and accountability on trust in human robot interaction. In Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, Cambridge, UK, 23–26 March 2020; pp. 33–42.

102. Garcia Goo, H.; Winkle, K.; Williams, T.; Strait, M.K. Robots Need the Ability to Navigate Abusive Interactions. In Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction, Online, 7–10 March 2022.

103. Luria, M.; Sheriff, O.; Boo, M.; Forlizzi, J.; Zoran, A. Destruction, catharsis, and emotional release in human–robot interaction. *ACM Trans. Hum.-Robot. Interact. THRI* **2020**, *9*, 22. [CrossRef]

104. Brščić, D.; Kidokoro, H.; Suehiro, Y.; Kanda, T. Escaping from children's abuse of social robots. In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, Portland, OR, USA, 2–5 March 2015; pp. 59–66.

105. Lucas, H.; Poston, J.; Yocum, N.; Carlson, Z.; Feil-Seifer, D. Too big to be mistreated? Examining the role of robot size on perceptions of mistreatment. In Proceedings of the 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), New York, NY, USA, 26–31 August 2016; pp. 1071–1076.

106. Keijsers, M.; Bartneck, C. Mindless robots get bullied. In Proceedings of the 2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Chicago, IL, USA, 5–8 March 2018; pp. 205–214.

107. Złotowski, J.; Sumioka, H.; Bartneck, C.; Nishio, S.; Ishiguro, H. Understanding anthropomorphism: Anthropomorphism is not a reverse process of dehumanization. In Proceedings of the International Conference on Social Robotics, Tsukuba, Japan, 22–24 November 2017; Springer: Berlin/Heidelberg, Germany, 2017; pp. 618–627.

108. Yamada, S.; Kanda, T.; Tomita, K. An escalating model of children's robot abuse. In Proceedings of the 2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Cambridge, UK, 23–26 March 2020; pp. 191–199.

109. Fogg, B.J. A behavior model for persuasive design. In Proceedings of the 4th international Conference on Persuasive Technology, Claremont, CA, USA, 26–29 April 2009; pp. 1–7.

110. Davidson, R.; Sommer, K.; Nielsen, M. Children's judgments of anti-social behaviour towards a robot: Liking and learning. In Proceedings of the 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Daegu, Republic of Korea, 11–14 March 2019; pp. 709–711.

111. Nomura, T.; Kanda, T.; Yamada, S. Measurement of moral concern for robots. In Proceedings of the 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Daegu, Republic of Korea, 11–14 March 2019; pp. 540–541.

112. Zhang, Q.; Zhao, W.; Chu, S.; Wang, L.; Fu, J.; Yang, J.; Gao, B. Research progress of nuclear emergency response robot. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *452*, 042102. [CrossRef]

113. Sakai, Y. Japan's Decline as a Robotics Superpower: Lessons From Fukushima. *Asia Pac. J.* **2011**, *9*, 3546.

114. Titiriga, R. Autonomy of Military Robots: Assessing the Technical and Legal (Jus In Bello) Thresholds. *J. Marshall J. Inf. Technol. Privacy Law* **2015**, *32*, 57.

115. Hayes, G. Balancing Dangers: An Interview with John Farnam. 2012. Available online: https://armedcitizensnetwork.org/archives/253-february-2012 (accessed on 5 March 2023).

116. Coffee-Johnson, L.; Perouli, D. Detecting anomalous behavior of socially assistive robots in geriatric care facilities. In Proceedings of the 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Daegu, Republic of Korea, 11–14 March 2019; pp. 582–583.

117. Bambauer, J.R. Dr. Robot. *UCDL Rev.* **2017**, *51*, 383.

118. Cox-George, C.; Bewley, S. I, Sex Robot: The health implications of the sex robot industry. *BMJ Sex. Reprod. Health* **2018**, *44*, 161–164. [CrossRef]

119. Xilun, D.; Cristina, P.; Alberto, R.; Zhiying, W. Novel robot for safety protection identification & detect. In Proceedings of the 2007 IEEE International Workshop on Safety, Security and Rescue Robotics, Rome, Italy, 17–19 September 2007; pp. 1–3.

120. Matsumoto, R.; Nakayama, H.; Harada, T.; Kuniyoshi, Y. Journalist robot: Robot system making news articles from real world. In Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, CA, USA, 29 October–2 November 2007; pp. 1234–1241.

121. Latar, N.L. The robot journalist in the age of social physics: The end of human journalism? In *The New World of Transitioned Media: Digital Realignment and Industry Transformation*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 65–80.

122. Desk, N. Manchester Arena's AI Weapon-Scanning Technology—Does It Work?—BBC Newsnight. 2022. Available online: https://theglobalherald.com/news/manchester-arenas-ai-weapon-scanning-technology-does-it-work-bbc-newsnight/ (accessed on 5 March 2023).

123. Faife, C. After Uvalde, Social Media Monitoring Apps Struggle to Justify Surveillance. 2022. Available online: https://www.theverge.com/2022/5/31/23148541/digital-surveillance-school-shootings-social-sentinel-uvalde (accessed on 5 March 2023).

124. Rotaru, V.; Huang, Y.; Li, T.; Evans, J.; Chattopadhyay, I. Event-level prediction of urban crime reveals a signature of enforcement bias in US cities. *Nat. Hum. Behav.* **2022**, *6*, 1056–1068. [CrossRef]

125. Gerke, H.C.; Hinton, T.G.; Takase, T.; Anderson, D.; Nanba, K.; Beasley, J.C. Radiocesium concentrations and GPS-coupled dosimetry in Fukushima snakes. *Sci. Environ.* **2020**, *734*, 139389. [CrossRef] [PubMed]

126. Saha, D.; Mehta, D.; Altan, E.; Chandak, R.; Traner, M.; Lo, R.; Gupta, P.; Singamaneni, S.; Chakrabartty, S.; Raman, B. Explosive sensing with insect-based biorobots. *Biosens. Bioelectron. X* **2020**, *6*, 100050. [CrossRef]

127. Yaacoub, J.P.A.; Noura, H.N.; Salman, O.; Chehab, A. Robotics cyber security: Vulnerabilities, attacks, countermeasures, and recommendations. *Int. J. Inf. Secur.* **2022**, *21*, 115–158. [CrossRef] [PubMed]

128. Clark, G.W.; Doran, M.V.; Andel, T.R. Cybersecurity issues in robotics. In Proceedings of the 2017 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA), Savannah, GA, USA, 27–31 March 2017; pp. 1–5.

129. Cerrudo, C.; Apa, L. Hacking Robots before Skynet. IOActive Website. 2017. Available online: www.ioactive.com/pdfs/Hacking-Robots-Before-Skynet.pdf (accessed on 15 February 2023).

130. Hayashi, Y.; Wakabayashi, K.; Shimojyo, S.; Kida, Y. Using decision support systems for juries in court: Comparing the use of real and CG robots. In Proceedings of the 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Daegu, Republic of Korea, 11–14 March 2019; pp. 556–557.

131. Koay, K.L.; Dautenhahn, K.; Woods, S.; Walters, M.L. Empirical results from using a comfort level device in human–robot interaction studies. In Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction, Salt Lake City, UT, USA, 2–3 March 2006; pp. 194–201.

132. Guiochet, J.; Do Hoang, Q.A.; Kaaniche, M.; Powell, D. Model-based safety analysis of human–robot interactions: The MIRAS walking assistance robot. In Proceedings of the 2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR), Seattle, WA, USA, 24–26 June 2013; pp. 1–7.

133. Song, S.; Yamada, S. Bioluminescence-inspired human–robot interaction: Designing expressive lights that affect human's willingness to interact with a robot. In Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, Chicago, IL, USA, 5–8 March 2018; pp. 224–232.

134. Flaherty, C. Micro-UAV Augmented 3D Tactics. Small Wars Journal. 2018. Available online: https://smallwarsjournal.com/jrnl/art/micro-uav-augmented-3d-tactics (accessed on 5 March 2023).

135. Nakamura, T.; Tomioka, T. Seashore robot for environmental protection and inspection. In Proceedings of the 1993 IEEE/Tsukuba International Workshop on Advanced Robotics, Tsukuba, Japan, 8–9 November 1993; pp. 69–74.

136. Zhang, W.; Ai, C.S.; Zhang, Y.Z.; Li, W.X. Intelligent path tracking control for plant protection robot based on fuzzy pd. In Proceedings of the 2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM), Hefei and Tai'an, China, 27–31 August 2017; pp. 88–93.

137. Weng, Y.H.; Gulyaeva, S.; Winter, J.; Slavescu, A.; Hirata, Y. HRI for legal validation: On embodiment and data protection. In Proceedings of the 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Naples, Italy, 31 August–4 September 2020; pp. 1387–1394.

138. Bethel, C.L.; Eakin, D.; Anreddy, S.; Stuart, J.K.; Carruth, D. Eyewitnesses are misled by human but not robot interviewers. In Proceedings of the 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Tokyo, Japan, 3–6 March 2013; pp. 25–32.

139. Singh, A.K.; Baranwal, N.; Nandi, G.C. Human perception based criminal identification through human robot interaction. In Proceedings of the 2015 Eighth International Conference on Contemporary Computing (IC3), Noida, India, 20–22 August 2015; pp. 196–201.

140. Rueben, M.; Bernieri, F.J.; Grimm, C.M.; Smart, W.D. User feedback on physical marker interfaces for protecting visual privacy from mobile robots. In Proceedings of the 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand, 7–10 March 2016; pp. 507–508.

141. Xia, L.; Gori, I.; Aggarwal, J.K.; Ryoo, M.S. Robot-centric Activity Recognition from First-person RGB-D Videos. In Proceedings of the 2015 IEEE Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 5–9 January 2015; pp. 357–364.

142. Garg, N.; Roy, N. Enabling self-defense in small drones. In Proceedings of the 21st International Workshop on Mobile Computing Systems and Applications, Austin, TX, USA, 3 March 2020; pp. 15–20.

143. Meng, C.; Wang, T.; Chou, W.; Luan, S.; Zhang, Y.; Tian, Z. Remote surgery case: Robot-assisted teleneurosurgery. In Proceedings of the IEEE International Conference on Robotics and Automation, ICRA'04, New Orleans, LA, USA, 26 April–1 May 2004; Volume 1, pp. 819–823.

144. Trost, M.J.; Chrysilla, G.; Gold, J.I.; Matarić, M. Socially-Assistive robots using empathy to reduce pain and distress during peripheral IV placement in children. *Pain Res. Manag.* **2020**, *2020*, 7935215. [CrossRef] [PubMed]

145. Wanebo, T. Remote killing and the Fourth Amendment: Updating Constitutional law to address expanded police lethality in the robotic age. *UCLA Law Rev.* **2018**, *65*, 976.

146. Casper, J.L.; Murphy, R.R. Workflow study on human–robot interaction in USAR. In Proceedings of the 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292), Washington, DC, USA, 11–15 May 2002; Volume 2, pp. 1997–2003.

147. Ventura, R.; Lima, P.U. Search and rescue robots: The civil protection teams of the future. In Proceedings of the 2012 Third International Conference on Emerging Security Technologies, Lisbon, Portugal, 5–7 September 2012; pp. 12–19.

148. Harriott, C.E.; Adams, J.A. Human performance moderator functions for human–robot peer-based teams. In Proceedings of the 2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Osaka, Japan, 2–5 March 2010; pp. 151–152.

149. Weiss, A.; Wurhofer, D.; Lankes, M.; Tscheligi, M. Autonomous vs. tele-operated: How people perceive human–robot collaboration with HRP-2. In Proceedings of the 2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI), La Jolla, CA, USA, 9–13 March 2009; pp. 257–258.

150. Nasir, J.; Oppliger, P.; Bruno, B.; Dillenbourg, P. Questioning Wizard of Oz: Effects of Revealing the Wizard behind the Robot. In Proceedings of the 31st IEEE International Conference on Robot & Human Interactive Communication (RO-MAN), Napoli, Italy, 29 August–2 September 2022.

151. Tozadore, D.; Pinto, A.; Romero, R.; Trovato, G. Wizard of Oz vs. autonomous: Children's perception changes according to robot's operation condition. In Proceedings of the 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Lisbon, Portugal, 28 August–1 September 2017; pp. 664–669.

152. Bennett, M.; Williams, T.; Thames, D.; Scheutz, M. Differences in interaction patterns and perception for teleoperated and autonomous humanoid robots. In Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 6589–6594.

153. Zheng, K.; Glas, D.F.; Kanda, T.; Ishiguro, H.; Hagita, N. How many social robots can one operator control? In Proceedings of the 2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Lausanne, Switzerland, 6–9 March 2011; pp. 379–386.

154. Hopkins, D.; Schwanen, T. Talking about automated vehicles: What do levels of automation do? *Technol. Soc.* **2021**, *64*, 101488. [CrossRef]

155. Tian, L.; Oviatt, S. A taxonomy of social errors in human–robot interaction. *ACM Trans. Hum.-Robot. Interact. THRI* **2021**, *10*, 13. [CrossRef]

156. Halbach, T.; Schulz, T.; Leister, W.; Solheim, I. Robot-Enhanced Language Learning for Children in Norwegian Day-Care Centers. *Multimodal Technol. Interact.* **2021**, *5*, 74. [CrossRef]

157. Baeth, M.J.; Aktas, M.S. Detecting misinformation in social networks using provenance data. *Concurr. Comput. Pract. Exp.* **2019**, *31*, e4793. [CrossRef]

158. Friedman, N. Word of the Week: Cunningham's Law. 2010. Available online: https://nancyfriedman.typepad.com/away_with_words/2010/05/word-of-the-week-cunninghams-law.html (accessed on 5 March 2023).

159. Alemi, M.; Meghdari, A.; Ghazisaedy, M. Employing humanoid robots for teaching English language in Iranian junior high-schools. *Int. J. Humanoid Robot.* **2014**, *11*, 1450022. [CrossRef]

160. Vincent, J. The US Is Testing Robot Patrol Dogs on Its Borders. 2022. Available online: https://www.theverge.com/2022/2/3/22915760/us-robot-dogs-border-patrol-dhs-tests-ghost-robotics (accessed on 5 March 2023).

161. Olsen, T.; Stiffler, N.M.; O'Kane, J.M. Rapid Recovery from Robot Failures in Multi-Robot Visibility-Based Pursuit-Evasion. In Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Prague, Czech Republic, 27 September–1 October 2021; pp. 9734–9741.

162. *ISO 21448:2022*; Road Vehicles—Safety of the Intended Functionality. ISO: Geneva, Switzerland, 2022. Available online: https://www.iso.org/standard/77490.html (accessed on 5 March 2023).

163. Seok, S.; Hwang, E.; Choi, J.; Lim, Y. Cultural differences in indirect speech act use and politeness in human–robot interaction. In Proceedings of the 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Sapporo, Japan, 7–10 March 2022; pp. 1–8.

164. Persson, A.; Laaksoharju, M.; Koga, H. We Mostly Think Alike: Individual Differences in Attitude Towards AI in Sweden and Japan. *Rev. Socionetw. Strateg.* **2021**, *15*, 123–142. [CrossRef]

165. Bröhl, C.; Nelles, J.; Brandl, C.; Mertens, A.; Nitsch, V. Human–robot collaboration acceptance model: Development and comparison for Germany, Japan, China and the USA. *Int. J. Soc. Robot.* **2019**, *11*, 709–726. [CrossRef]

166. Heine, S.J. *Cultural Psychology*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2010.

167. West, D.M. *What Happens if Robots Take the Jobs? The Impact of Emerging Technologies on Employment and Public Policy*; Centre for Technology Innovation at Brookings: Washington, DC, USA, 2015.

168. Violent Crime Rates by Country. 2018. Available online: https://wisevoter.com/country-rankings/violent-crime-rates-by-country/ (accessed on 5 March 2023).

169. Grinshteyn, E.; Hemenway, D. Violent death rates: The US compared with other high-income OECD countries, 2010. *Am. J. Med.* **2016**, *129*, 266–273. [CrossRef] [PubMed]

170. Lies, E. Abe Assassination Stuns Japan, a Country Where Gun Violence Is Rare. 2022. Available online: https://www.reuters.com/world/asia-pacific/mostly-gun-free-nation-japanese-stunned-by-abe-killing-2022-07-08/ (accessed on 5 March 2023).

171. Stack, S. The effect of the media on suicide: Evidence from Japan, 1955–1985. *Suicide -Life-Threat. Behav.* **1996**, *26*, 132–142. [PubMed]

172. Penal Code. 2017. Available online: https://www.japaneselawtranslation.go.jp/en/laws/view/3581/en (accessed on 5 March 2023).

173. Cheng, C.; Hoekstra, M. Does strengthening self-defense law deter crime or escalate violence? Evidence from expansions to castle doctrine. *J. Hum. Resour.* **2013**, *48*, 821–854. [CrossRef]

174. Self Defense and "Stand Your Ground". 2022. Available online: https://www.ncsl.org/civil-and-criminal-justice/self-defense-and-stand-your-ground (accessed on 5 March 2023).

175. Schoettle, B.; Sivak, M. *Public Opinion about Self-Driving Vehicles in China, India, Japan, the US, the UK, and Australia*; Technical Report; University of Michigan, Ann Arbor, Transportation Research Institute: Ann Arbor, MI, USA, 2014.

176. Yakubovich, A.R.; Esposti, M.D.; Lange, B.C.; Melendez-Torres, G.; Parmar, A.; Wiebe, D.J.; Humphreys, D.K. Effects of laws expanding civilian rights to use deadly force in self-defense on violence and crime: A systematic review. *Am. J. Public Health* **2021**, *111*, e1–e14. [CrossRef]

177. Gallimore, D.; Lyons, J.B.; Vo, T.; Mahoney, S.; Wynne, K.T. Trusting robocop: Gender-based effects on trust of an autonomous robot. *Front. Psychol.* **2019**, *10*, 482. [CrossRef]

178. Duarte, E.K.; Shiomi, M.; Vinel, A.; Cooney, M. Robot Self-defense: Robot, don't hurt me, no more. In Proceedings of the 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Sapporo, Japan, 7–10 March 2022; pp. 742–745.

179. Collier, R. *Masculinities, Crime and Criminology*; Sage: San Mateo, CA, USA, 1998.

180. Robot Self-Defense Playlist. 2022. Available online: youtube.com/watch?v=F5y7dPy41p0&list=PLtGv2XOitdkQsEsPX528cmat5cnVgd15V (accessed on 5 March 2023).

181. Two-Way Mixed ANOVA. Available online: https://guides.library.lincoln.ac.uk/mash/statstest/ANOVA_mixed (accessed on 5 March 2023).

182. Fitter, N.T.; Mohan, M.; Kuchenbecker, K.J.; Johnson, M.J. Exercising with Baxter: Preliminary support for assistive social-physical human–robot interaction. *J. Neuroeng. Rehabil.* **2020**, *17*, 19. [CrossRef]

183. Meola, A. Generation Z News: Latest Characteristics, Research, and Facts. 2023. Available online: https://www.insiderintelligence.com/insights/generation-z-facts/ (accessed on 4 March 2023).

184. Smith, T.W. The polls: Gender and attitudes toward violence. *Public Opin. Q.* **1984**, *48*, 384–396. [CrossRef]

185. Walters, K.; Christakis, D.A.; Wright, D.R. Are Mechanical Turk worker samples representative of health status and health behaviors in the US? *PLoS ONE* **2018**, *13*, e0198835. [CrossRef]

186. Florida, R. What Is It Exactly That Makes Big Cities Vote Democratic? 2013. Available online: https://www.bloomberg.com/news/articles/2013-02-19/what-is-it-exactly-that-makes-big-cities-vote-democratic (accessed on 5 March 2023).