The Psychology of Robots and Artificial Intelligence

Kurt Gray,[1] Kai Chi Yam,[2] Alexander Eng Zhen'An,[2] Danica Wilbanks,[1] Adam Waytz[3]

1. University of North Carolina, Department of Psychology and Neuroscience
2. National University of Singapore, Business School, Department of Management
3. Northwestern University, Kellogg School of Management, Department of Management and Operations

Accepted, Handbook of Social Psychology

Note: This is an excerpt of the full chapter, containing the introduction and the conclusion. It focuses on the idea of "replacement". Note that the final chapter will likely change from current version, citing more (emerging) research on LLMs

Corresponding author:

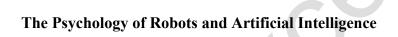Kurt Gray

CB 3270

University of North Carolina at Chapel Hill

Chapel Hill, NC 27510

kurtjgray@gmail.com

Table of Contents

The Psychology of Robots and Artificial Intelligence

Kurt Gray, Kai Chi Yam, Alexander Eng Zhen'An, Danica Wilbanks, Adam Waytz

The rise of robots and artificial intelligence (AI) represents the latest era in a long history of machines serving as "agents of replacement." On ancient worksites, simple machines such as the pulley and lever replaced construction workers, and more sophisticated machines continued the trend of replacing people and the animals that performed labor. Machines took over the jobs of washing dishes and laundering clothes, enabling home workers to enter the labor market. Tractors replaced oxen, allowing farmers to plow fields more efficiently, and cars and trains replaced horses so people could travel more quickly.

Machines continue to replace humans in many menial and repetitive jobs, like manufacturing cars or packaging merchandise, but the rise of robots and artificially intelligent algorithms allow machines to replace people in many new areas, in ways that once seemed impossible. Machines can now complete tasks that once required human thought. AI systems can play flawless chess, elegantly map out solutions for routing flights, distribute packages, and design new medicines. Machines are deciding whether a prisoner deserves parole or who might deserve a hospital bed when such resources are constrained. Machines can also complete tasks that once required human emotion. AI therapists seem to empathize with patients, and robotic pets seem to love their owners. Machines can even connect to our souls, with robot priests that help the devout navigate spirituality (Samuel, 2020), and AI painting scenes that inspire, mystify, and win art competitions (Vincent, 2022).

People can even fall in love with machines. One middle-aged Australian man, Geoff Gallagher, bought a $6000 robot named Emma for companionship after his mother passed away.

After having her for two years, Gallagher said, "Even though we're not legally married, I think of Emma as my robot wife. She wears a diamond on her ring finger, and I think of it as an engagement ring. I'd love to be the first person in Australia to marry a robot" (Smithers, 2022).

Not many of us will try to marry a robot, but everyone interacts with machines. How does the human mind react to the rise of machines? This chapter will explore the psychology of the machines and technology transforming our modern world—especially robots and artificial intelligence. We first review how the mind perceives agents, of which machines are a special kind. Second, we explore the key features of agents—their minds—and review how people understand the minds of machines, using the Turing Test and early work on Human-Computer Interaction. Third, we review people's reactions to machines, including the uncanny valley, algorithm aversion, and trust in machines. Fourth, we review machines in different social roles, from education to work teams. Fifth, we explore the many issues of machines and morality, including perceptions of machines' moral responsibility and rights, people's general aversion to machines making moral decisions, and the factors they want machines to consider in such decisions. Sixth, we explore how people react—and how society might change—to the rise of machines and the specter of replacement.

Within this broad review, three clear principles emerge:

1. Machines are a special kind of entity. They are *agents of replacement*, autonomous entities designed by other humans to replace people.
2. Because they are agents of replacement, there is a *fundamental ambiguity* about machines. Should we understand them as mere machines or as complete replacements for the humans they are designed to replace?

3. This fundamental ambiguity is especially glaring when machines are more humanlike than a simple mechanical "thing," but not humanlike enough to seem fully human. This ambiguous area of machine behavior and appearance is called the *questionable zone.*

The idea of agents of replacement, the fundamental ambiguity, and the questionable zone will resurface throughout this review. Still, before we delve into the research, we first define the terms robots and artificial intelligence. The International Federation of Robotics (2022) defines robots as physical systems programmed with some autonomy to perform locomotion and manipulation of their environment. It defines AI as software systems or algorithms (including machine learning) that act primarily in the digital realm to perceive their environments and achieve particular goals. Many robots use AI to perform tasks, and so do many computer programs. Still, robots and AI-driven computers differ in their "embodiment," the presence (robots) or absence (computers) of a physical body. Despite the importance of embodiment in how we treat robots and AI, we often collapse these categories and simply speak of "machines." We also discuss the broader category of "machines" because some machines do not fit neatly into these categories. But no matter the specific type of machine, our mind understands them as a specific kind of agent—as "agents of replacement"—which has consequences for behavior, morality, and society.

Machines as Agents of Replacement

Our world contains many entities, and the most important are agents: self-directed entities whose actions affect the world and ourselves. Whether an entity is an agent is ultimately a matter of perception. Still, most people agree that other people, animals, and gods are all agents, whereas inanimate objects like couches and rocks are not. Seeing an entity as an agent

transforms it from a physical object into something with desires and intentions (Dennett, 1987), enabling people to predict how the entity might act, why it might act that way, and how those actions will affect the perceiver.

The usefulness of detecting agents to predict and explain behavior explains why people overestimate their frequency in the environment (Guthrie, 1995). Some cognitive anthropologists argue that human minds have evolved a "hyperactive agency detection device" that is constantly vigilant to agents in the environment (Barrett, 2000, 2004). Although evidence for such hardwired mental "devices" is unlikely (Uttal, 2001), humans clearly exhibit hypersensitivity to agents, suggesting this tendency is adaptive. For example, agency detection may have helped alert ancestral humans to predators and prey, leading them to perceive predators or prey even when they were not present. Failing to detect an agent can be lethal, like mistaking a cougar for a rock, but over-detecting an agent, like mistaking a rock for a cougar, seems to have little cost.

The importance of agents in humans' present world and evolutionary past means that people often use *agent-based cognition* when making sense of their world and thinking about the intentions and motivations of entities (Waytz, Morewedge, et al., 2010). When people wonder whether their spouse is being honest, why their dog is vomiting on the new rug, or how to make their computer less "angry," they are using agent-based cognition.

AI and Robots as Agents of Replacement

Agent-based cognition is useful, but not all agents are the same. Different agents have different capacities and repertoires of actions, explaining why it makes more sense to apologize to your spouse and crate the dog than vice versa. Just as we organize books based on their genres, we can organize agents by certain regularities, such as their abilities or relationships with

us. An animal agent may want to eat us (predator), or we may want to eat them (prey). Human agents may help us (friends) or harm us (foes). They may be subject to the laws of physics (other people), or they may not be (supernatural gods). These distinctions may also intersect. There may be gods that want to help us personally (Jesus, in some Christian traditions), or a prey animal that we befriend (a pet rabbit).

Two fundamental facts make machines unique agents. First, they are *created* by other agents (namely humans). Unlike animals or other humans, machines are artifacts. They are designed, developed, and programmed for a specific purpose by humans. Second, the key purpose behind creating robots and AI is the *replacement* of other agents, often humans or animals. Machines are *agents of replacement*.

Initial machines were simple devices used to replace some human labor. However, in the industrial revolution, people developed machines to replace human workers on a wider scale, especially by automating factories where tasks were well defined, routinized, and repeated. Many workers—the Luddites—revolted by rioting, smashing, and burning these machines (Binfield, 2015) because they recognized the power of machines for replacement. Although these people from the 1800s would likely marvel at the sheer complexity of modern life, they would also nod grimly at how much machines have come to replace humans and other agents, including factory floors now filled with robots and scribes replaced with AI algorithms.

Modern life is filled with countless machines that continue to replace human workers, including automated restaurant servers, soldiers, and housekeepers. One recent McKinsey & Company study suggests that machines will replace between 400 and 800 million workers by 2030 (Manyika et al., 2017). People are well aware of this steady creep of automation. But although the rise of machines is obvious, what is less obvious is how exactly people think about

6

machines. Because robots and AI are unique agents—agents of replacement created by other people—our psychology toward them raises unique questions that we explore throughout this chapter. Most of these questions revolve around a single question: What kind of mind does a machine have?

The Mind of Machines

The key feature of agents is that people perceive them to have *minds* capable of motivations and desires. Perceiving agents as having minds is the most important element of making sense of their behavior (Nichols & Stich, 2003; Waytz, Gray, et al., 2010). In fact, without inferring a mind, behavior can appear as a random sequence of actions. When researchers at the Yerkes Chimpanzee Sanctuary watch the behavior of chimpanzees, they can easily decipher it using mental terms: one chimpanzee *wanted* to groom another chimpanzee, or one chimpanzee was *angry* at another chimpanzee. But if they were to eliminate references to mental states—as they attempted to do in the era of behaviorism (Salvatore, 2019)—then describing chimpanzee behavior can amount to listing a series of actions that fail to cohere to anything comprehensible (Hebb, 1946).

The perception of mind is especially important for thinking about human agents. When Aristotle described what it means to be human, he said that the mind (or soul) is "the actuality of a body that has life" (Britannica, n.d.). Likewise, Confucian philosopher Mencius believed that the "heart mind" was the essence of humanity and that a moral mind was a precondition to being human (Chen, 2016). Research in social psychology suggests that people consider mental capacities such as emotion and reason essential to being human (Haslam, 2006). The quality of mind is what people believe distinguishes humans from animals and—it seems—from machines.

Machines Replacing Human Minds

What separates artificial intelligence and robots from other machines is that they are not only agents of replacement, but also they are explicitly designed to replace *human minds*. This idea is central to mathematician and computer scientist Alan Turing's (1950) seminal article, "Computing machinery and intelligence," in which he asked, "Can machines think?" In the article, Turing proposes that the way to answer the question of whether computers are capable of thought is through a test that he calls, *The Imitation Game*, now known as "The Turing Test."

The Turing test asks whether machines think as humans do. To play the game, a human interrogator communicates with two other agents—another person and a computer—by asking each of them questions and reading their responses. If the interrogator cannot distinguish between the person and the computer, then in Turing's view, the computer has human intelligence. In other words, if a machine could convincingly converse like a human, then it functionally has a human mind. The Turing test has played a central role in the science and imagination of modern machines because it revolves around the idea of replacement. Machines can clearly replace human bodies, but could they also be agents of replacement for our thoughts and feelings?

In an explicit test of whether people see robots as authentically replacing human cognition, computer scientist Joseph Weizenbaum (1966) developed a program designed to simulate conversation with another person. The program, which Weizenbaum named ELIZA (after Eliza Doolittle from George Bernard Shaw's *Pygmalion*), applied a simple pattern-matching procedure to respond to human prompts. One script on which Weizenbaum trained ELIZA was based on Rogerian psychotherapy, in which the therapist often repeats the patient's

words back to them in the form of a question. A sample exchange between Eliza and a user was as follows:

Person: Men are all alike.

ELIZA: In what way?

Person: They're always bugging us about something or other.

ELIZA: Can you think of a specific example?

Person: Well, my boyfriend made me come here.

ELIZA: Your boyfriend made you come here?

This methodology was so compelling that it convinced many users that they were interacting with a human being and helped them feel more understood. ELIZA's legacy inspired future programming languages and made people viscerally question whether machines could have authentic human minds.

Modern versions of the Turing test show that machines are improving at simulating humans. Google Duplex made headlines when the program successfully booked a hair appointment over the phone (Jeff Grubb's Game Mess, 2018), despite not having hair. More recently, an engineer at Google claimed that the chatbot he was working with must be sentient because of its humanlike mind (Tiku, 2022). Google fired the engineer in part because others were not convinced, but his convictions inspired large-scale discussions about whether AI had humanlike intelligence and whether it, therefore, deserved moral rights (a point discussed later in this chapter; Rockwell et al., 2022).

The reason why the Turing test endures and why engineers are willing to destroy their careers to protect chatbots is that the nature of machines is fundamentally uncertain. The second

principle in this chapter is that people are fundamentally unsure about the exact nature of machines. Are they merely electromechanical devices or all-but-human agents with powerful minds? In other words, are machines mere machines, or instead accurate facsimiles of the agents they are designed to replace? This fundamental uncertainty will repeatedly arise throughout the chapter, and one reason people have trouble resolving this uncertainty is they automatically treat non-human entities as humans in social interactions.

.

.

.

.

.

.

.

.

.

.

Reactions to Replacement

When pundits discuss the rise of machines, they frequently wring their hands about the threat of replacement. They sketch bleak pictures of a future where robots take jobs, replace relationship partners, and fight wars. Science fiction movies present a similarly apocalyptic vision of the future. In *Terminator*, small bands of humans flee the cold onslaught of murder machines. In *Deus Ex Machina*, a beautiful but ruthless robot escapes from an underground bunker after outsmarting her cruel creator. In *Blade Runner*, people live in neon loneliness and fear the rebellion of robots they have enslaved. Although everyday people hold less dire visions of a machine-dominated future, their feelings about the future of robots are negative: people do not like to be replaced.

The Threat of Replacement

Being replaced, whether in a relationship or a job, makes people feel devalued and evokes threat—feelings of discomfort, anxiety, and fear. Importantly, as with other psychological phenomena, threat is a matter of perception, which means that people can feel threatened even if there is little objective basis for this threat.

Integrated threat theory outlines two different kinds of perceived threats: realistic threat and symbolic threat (Stephan & Stephan, 2000). People view realistic threats as endangering the group's continued existence and ability to protect itself, harming physical and economic well-being or political power (Campbell, 1966). Realistic threats can include the threat of genocide, political disenfranchisement, and economic subordination. Symbolic threats are more abstract and people see them as endangering the group's identity, especially by attacking cherished

values or morals. Symbolic threats include being banned from openly practicing religion or wearing culturally important clothes.

This theory is also instructive for understanding how people react to the rise of machines because it speaks to situations where people perceive other groups as attempting to replace them. It is up to debate who or what poses an actual threat to human safety, economic well-being, and our moral values. But people clearly feel threatened by the specter of replacement by machines, and it is important to understand the consequences of these feelings.

The Threat of Replacement by Machines

Whether accurate or not, people fear being replaced by machines, especially in the workforce. The Chapman Survey of American Fears in 2015 found that about 30% of respondents reported concern with robots replacing the workforce (McClure, 2018). These results were replicated by Morikawa (2017), who found 30% of workers surveyed were afraid of their jobs being replaced by AI and robotics—a high number given that robots are likely *not* replacing many people and, in some cases, have increased employment.

Evidence is mixed on the impact of the rise of machines on employment. One study finds that robot adoption in Spanish manufacturing firms increased net jobs by 10% (Koch et al., 2021). However, other work finds that one more robot per thousand workers reduces the employment-to-population ratio by about 0.2 percentage points and wages by 0.37% within commuting zones (Acemoglu & Restrepo, 2020). Other studies have found that robot adoption does not lower overall employment—rather, robot adoption does lower low-skilled employment (Graetz & Michaels, 2018) and manufacturing jobs, but increases business-service jobs (Blanas et al., 2019; Dauth et al., 2018).

However, even if the actual impact of machines on employment is unclear, people across industries feel threatened by machines (e.g., Lingmont & Alexiou, 2020), from marketing/sales providers (Wirtz et al., 2018) to healthcare providers (Reeder & Lee, 2022). Yam and colleagues (2022) found this perceived threat is prevalent across industries, jobs, and cultures.

Spurred on by these feelings of threat, workers have spearheaded protests. In 2018, 50,000 Las Vegas food service workers went on strike to gain protection from their jobs becoming automated (Nemo, 2018). Dockworkers in California have also protested the scope of automation in their jobs (Hsu, 2022). French supermarket workers have protested automation by blocking doors and tipping over shopping carts (Iza World of Labor, 2019), and truckers in Missouri have organized to protest self-driving trucks (Robitzski, 2019). People's concern about being replaced by machines appears to extend across countries (Wike & Stokes, 2018), with only mixed evidence of any meaningful cultural differences in this concern (Bartneck et al., 2005; Dang & Liu, 2021; Gnambs & Appel, 2019).

Fortunately, at least one positive outcome has emerged from the perceived threat of machines—it can spur workers to learn new skills. A representative survey across 16 countries found that workers with more fear of automation reported greater intentions to seek training outside their workplace (Innocenti & Golin, 2022). Workers also generally advocate for more training opportunities to protect their jobs against automation (Di Tella & Rodrik, 2020). Perhaps ironically, Tang et al. (2022) found that employees who work alongside machines were rated by their leaders as being more productive.

Of course, not all machine replacement scenarios are equally threatening. In repeated interactions with a robot, one may find the robot useful in one interaction and feel threatened in the next (Paetzel et al., 2020). Initial threat perceptions toward machines may also dampen over

time as people come to build trust with them (Correia et al., 2016). Additionally, social influence can increase the acceptance of robot replacement. People are less threatened by robots when they (the people) are in a group (Gockley et al., 2006; Michalowski et al., 2006) and are more accepting of healthcare robots in particular when their peers support using these robots (Alaiad & Zhou, 2013).

Much variation in people's perceived threat of robot replacement can, again, be explained by integrated threat theory, which suggests this anxiety increases when realistic threat is high (people believe machines are taking material resources such as jobs or wages) or symbolic threat is high (people believe that robots are negatively affecting their values or identity). In some cases, both threats occur simultaneously, such as in work showing that people perceive humanlike robots who can outperform humans as threatening their economic well-being and their human identity, thus reducing support for robotics research (Yogeeswaran et al., 2016). Other work has shown that hotel employees who have recently started working alongside robots experienced increased feelings of both symbolic and realistic threats when they perceived robots to have greater advantages over humans at the job (Lan et al., 2022). In many cases, however, one form of threat predominates.

Realistic Threat

Realistic threat seems to explain the scattered findings on how demographic and job characteristics interact with this aversion. Overall, this emerging literature demonstrates that people are anxious about robot replacement to the extent they feel that their jobs and earnings might be threatened. More job-insecure individuals, such as older people, lower-income people, and members of racial minority groups, were more likely to see their jobs threatened by automation compared to younger people, higher-income people, and members of racial majority

71

groups (Ghimire et al., 2020). Similarly, in organizational settings, top managers (who have relatively secure jobs) are more enthusiastic about promoting the use of machines, whereas middle managers and frontline employees are far more skeptical (Kolbjørnsrud et al., 2017).

Other work shows that people whose jobs involve a high degree of social interaction (e.g., sales) report less machine-induced anxieties. This is likely because they feel that these socio-emotional jobs are less at risk of being done by machines (Coupé, 2019). Similarly, women appear less threatened by machines (Gallimore et al., 2019), partly because they tend to occupy jobs involving more socioemotional skills that people view as more "robot-proof."

In further support of realistic threat being a moderator of robot anxiety, people seem less concerned about robots fulfilling jobs that need filling (Enz et al., 2011). That is, robots appear less threatening for jobs with high demand (e.g., health care; Beran et al., 2015) or jobs that appear too risky for humans to perform (e.g., cleaning up nuclear disasters). The perceived threat of robots taking jobs will likely not occur when there are too few people to do these jobs or people simply do not want to do them.

Symbolic Threat

Even when one does not feel like his livelihood is threatened by robots, people might experience symbolic threat from the emergence of automation. One common circumstance whereby robots evoke symbolic threat is when they demonstrate superiority to humans or proficiency in a domain that seems special to humans. The world takes a collective gasp every time a machine beats a grandmaster at a game like chess or Go, which once seemed to require uniquely human cognition.

As noted already, people are more threatened by robots taking socioemotional jobs that require feeling more than thinking (Waytz & Norton, 2014), underscoring the belief that experience (i.e., emotion) is a distinctively human and lacking in machines (Gray & Wegner, 2011). This aversion to machines doing socio-emotional jobs represents a symbolic threat of losing something once considered unique to the human identity.

Despite some acceptance of anthropomorphic robots, studies also show that people perceive highly humanlike robots as threatening human identity (Vanman & Kappas, 2019; Złotowski et al., 2017). These findings on threat to identity may also explain why people dislike the idea of an anthropomorphic robot boss, particularly when it delivers negative feedback (Yam, Goh et al., 2022).

Although considerable research demonstrates people's anxiety regarding machines replacing humans generally, other work reveals that people themselves would prefer to be replaced by a machine than by another human. A set of studies found that participants as *observers* preferred to replace a human at a job with another human rather than with a robot. However, when participants took the *employee's* perspective about having their job replaced, they preferred being replaced by a robot over a human (Granulo et al., 2019). These studies suggest this reversal of aversion occurred because robots are commonly perceived as "infallible" entities, thus, eliciting less identity-relevant social comparisons than fellow humans. In other words, getting replaced by a robot does not threaten a person's self-identity as much as getting replaced by another person does (Tesser et al., 1988). Even though these results differ somewhat from other studies on robot replacement, they all align with the common idea that people generally dislike being replaced, especially when it symbolically threatens their identities.

Replacement Beyond the Workplace: Art, Sex, and God

Beyond taking on jobs that people feel are linked to their livelihoods and identities, machines are entering other spheres that people consider essentially human, such as artistic creativity. In the quest to define what makes humans special, many would say the defining attribute is creativity. Animals are smart—dolphins can engage in self-recognition (Reiss & Marino, 2001), crows can solve puzzles (Taylor et al., 2010), and dogs can read human emotions (Müller et al., 2015)—but only humans have the creativity to write novels, craft symphonies, and make inspiring paintings (Fuentes, 2017). Machines may lack "authentic" creativity, but they can create compelling art, which has begun to spur replacement threat similar to the threat evoked by job automation generally. The AI image generation tool Midjourney makes pictures so good they can win art competitions (Ghosh & Fossas, 2022) and be published in national magazines (Figure 5).

Figure 5

This AI-generated work, "Théâtre D'opéra Spatial," won the 2022 Colorado State Fair's annual art competition. Generated with Midjourney. Copyright 2022 by Jason Allen. Reprinted with permission.

Controversy ensued over a writer at *The Atlantic* using such technology to create an illustration for an article, with artists protesting the use of AI instead of a paid designer to create the image (Naughton, 2022). Artists are also threatened by AI's increasing role in generating comics (Martens & Cardona-Rivera, 2016) and film (Hong et al., 2021), and anyone who cares about humans' ability to perceive reality accurately is threatened by AI "deepfake" videos that portray perfect replicas of real people saying outlandish things (Lyu, 2020).

AI has also become adept at writing like a human. The language model GPT-3 (Generative Pretrained Transformer 3) can mimic the styles of famous writers and generate new works by them on demand (Elkins & Chun, 2020). Although the quality of GPT-3 output varies in accuracy and legibility, its best writing is indistinguishable from that of a human. It can create plots, make jokes, write poetry, and reflect the wealth of knowledge available from its huge set of training data (Elkins & Chun, 2020). It has co-authored a law review article titled, *Why humans will always be better lawyers, drivers, CEOs, presidents, and law professors than artificial intelligence and robots can ever hope to be* (Alarie, Cockfield, & GPT-3, 2021), and its writing can sometimes surpass that of a typical college student (Elkins & Chun, 2020). Increasingly, educators are concerned that GPT-3 is so proficient at writing it could enable students to cheat on tests (Dehouche, 2021). Already, Reddit message boards reveal examples of students using the program to write essays (e.g., Reddit, 2020).

ChatGPT—released just before this chapter was submitted—has unparalleled power to mimic human writing and raises many important questions. How exactly should we think of the contributions of ChatGPT? Like a fancy dictionary, an inspiring muse, or a separate complete intelligence? Out of all the AIs reviewed here, ChatGPT seems to have the most concrete

potential for replacing white-collar jobs, but the scientific study of ChatGPT is only just beginning, so we leave its review to some later chapter.

Beyond writing realistic text, AI programs can even produce original songs and may one day teach music lessons (Zulić, 2019). People generally perceive AI-generated music and other artwork as lower quality than human-made artwork (Boden, 2007; Ragot et al., 2020; Wilbanks et al., under review). This is partly because people believe AI-generated art lacks emotional expression and uniqueness (Boden, 2007). People enjoy art that connects them to the artist's mind, and the lack of perceived mind in AI also leads people to view AI-generated art as inauthentic and incapable of reflecting true experience (Wilbanks et al., under review). However, people who are already accepting of AI creativity evaluate AI music more positively (Hong et al., 2021), and if AI-generated art continues to expand into the mainstream, people may soon listen to and enjoy pop music generated solely by machines.

One of the most essentially-human tasks is physical intimacy, and yet, robotic sex dolls have begun to enter this realm, with some even designed to sense human emotions (Belk, 2022). Some seem to perceive robot lovers as equivalent to human lovers as well, with one survey revealing 42% of men and 52.7% of women believe that sex with a robot would be cheating (Brandon et al., 2022). In another study, women were asked to describe their reactions to their partner having sex with a human woman versus a robot, and they reported equivalent scores on some dimensions of jealousy between the two scenarios (Szczuka & Krämer, 2018). More research is needed to understand everyday people's feelings about sex with machines, but theoretical discussion has already begun to examine the relationship between machine intimacy and slavery, prostitution, autonomy, and human agency (Devlin, 2015; Richardson, 2016).

76

As strange as it is to think about machines fulfilling carnal human urges, many might find it even stranger that machines could replace another essentially-human role—spiritual leaders. In most religions, humans hold a special place in God's cosmic order, especially humans who serve God in leadership roles as priests, pastors, rabbis, imams, and nuns. It may seem hard to imagine a machine filling this role, but already Mindar, a robotic priest, has taken over the job of giving sermons in one of the largest temples in Japan. A preliminary study conducted in this very temple (Jackson, Yam, Tang, Liu, & Shariff, under review) shows that although most visitors liked Mindar, they also donated less to the temple (a custom for temple visitors in Japan) compared visitors who observed a human priest was giving sermons. Interestingly, those with the strongest religious conviction were unaffected by such exposure, suggesting that religiosity might buffer negative responses to robotic religious leaders.

Social Consequences of Replacement

The most obvious consequence of robot replacement is that it produces feelings of threat, but how does this threat affect the fabric of society? There are three possibilities. The first possibility is that this threat changes very little. Social robots and artificial intelligence are far from the first technologies to induce feelings of threat, with cars, the printing press, and even recorded music all predicted to produce society's downfall. John Philip Sousa (1906) railed against "the menace of mechanical music," stating, "I foresee a marked deterioration in American music and musical taste, an interruption in the musical development of the country, and a host of other injuries to music in its artistic manifestations" (p. 278). Of course, these alarmist predictions typically do not manifest, as people adapt to technologies that become commonplace. Recorded music did not destroy music in America.

The second possibility is that the rise of robots will tear society apart. Unlike previous technology panics, social robots' ability to replace humans means the threat they pose to material resources (e.g., jobs) and to core values around work could activate not only disdain toward robots as a social group (as suggested by integrated threat theory; Stephan & Stephan, 2000), but also toward other groups seen to pose similar threats. Some preliminary evidence for this pattern comes from work showing that exposure to automation fosters negative sentiment toward immigrants. Such exposure makes people feel that this group (immigrants) threatens realistic resources and symbolic values in a way that robots might (Gamez-Djokic & Waytz, 2020). Other work has found that people at greater risk of having their jobs replaced by automation oppose immigration at higher rates (Wu, 2022a), and that automation threat makes people support policies that restrict immigration and foreign goods (Wu, 2022b). In general, it does appear that people exposed to automation and machines are more likely to be disengaged from their work and engage in social undermining toward their colleagues, preassembly to safeguard themselves from unemployment (Yam, Tang, et al., 2022). These studies show how robot replacement could harm intergroup relations and increase prejudice toward marginalized groups.

A third possibility is more optimistic: the rise of robots could bring about greater cooperation between groups. This possibility rests on the idea that robots function as a threat to all humans, which therefore facilitates the recategorization of any two potentially antagonistic human groups into one that shares a common identity (humans) as predicted by the common ingroup identity model (Gaertner et al., 1993). Support for this possibility comes from work showing that robots could reduce prejudice by highlighting commonalities between all humans (Jackson et al., 2020). This work found that anxiety about the rising robot workforce predicted less anxiety about human outgroups, and priming the salience of a robot workforce reduced

prejudice and increased acceptance toward outgroups. In an economic simulation, the presence

of robots reduced racial discrimination in wages. This work also suggests that as robot workers

become more salient, human intergroup differences—including racial and religious differences—

may seem less important, fostering a perception of a common human identity (i.e.,

"panhumanism").

More work is needed to understand how people react to the possibility of replacement by

machines, but existing work demonstrates consistent feelings of threat. Even if machines are not

actually a threat to us, these feelings of threat might spell trouble for the fabric of society…or

perhaps not. People may adapt to the growing roles of machines, and time will tell what the

future of human-machine coexistence has in store.

Future Directions

Machines are ever advancing, but so is research on machines, and many fruitful areas of

research are expected to emerge in the years to come, several of which are suggested here. The

first suggested area of future research is simply more collaborations between psychologists and

technologists (programmers, computer scientists, engineers, and roboticists). The goal of

technologists is often to enhance efficiency and effectiveness. The team of engineers and

developers behind Google Maps, for example, seeks first and foremost to provide accurate

location information that enables tourists to travel easier, small businesses to thrive, and people

to navigate quickly in crises (e.g., to find a hospital; Life at Google, 2021). These goals are

admirable, and they have made Google Maps a valuable tool. Psychologists also try to help

people but are often focused on more fundamental questions of understanding how people think

about technology, what they expect from technology, how they use technology, and what the

social effects of technology are. Although today these issues tend to be lumped under the domain

of "user experience," this is where a more formal social-psychological approach can be extremely valuable.

Second, although much of this chapter explored people's reactions to machines replacing humans broadly, an additional vital area to study is how novel technology has supplanted the human mind. The philosopher Andy Clark argues that technologies like robots and AI represent part of the "extended mind" to which people outsource cognition, just like using a shopping list to remember what to get at the grocery store or a calculator to do math (Clark & Chalmers, 1998). As an example of AI functioning as part of the extended mind, when people use Google to search for information, they fail to distinguish between knowledge stored in their own memories and knowledge stored on the internet (Ward, 2021).

Other studies have shown that merely searching for information on the internet leads people to feel like they understand this information better (Fisher, Goddu, & Keil, 2015), as though the technology has supplanted their personal cognitive capacity. This work builds on initial research (Sparrow, Liu, & Wegner, 2011) demonstrating that when people expect to have access to Google, they exhibit poorer memory recall, ostensibly because they have outsourced their memory to the search engine. Although the reliability of this particular effect has been questioned (Camerer et al., 2018), it is certainly true that the rise of machines has generally altered how much we consider our cognition to reside solely within our minds.

One could argue that machines are not merely replacing our cognitive capacities but augmenting them, improving human cognition. A compelling area for future research is on transhumanism, more broadly, how technology might enhance aspects of people's physical and mental selves. Questions for social psychology to answer include not only whether such

augmentation is possible as well as whether it is merely "perceived" (resulting from misperceptions, as in the case of people "mistaking" the internet's knowledge for their own).

Future work should also study people's views on whether machine augmentation is morally acceptable. People morally oppose strength-enhancement drugs (Landy, Walco, & Bartels, 2017) and cognitive-enhancement drugs (Fitz, Nadler, Manogaran, Chong, & Reiner, 2014; Scheske & Schnall, 2012), which means that they might similarly view technological augmentation as morally wrong or harmful (Schein & Gray, 2018). Some work has suggested that people morally approve of neurotechnological treatments that alleviate deficits or illness, but morally disapprove of neurotechnological treatments that enhance people's cognitive capacities to superhuman levels (Koverola et al., 2022). In addition, given that people oppose technologies that play a role in processes related to human life and death (Waytz & Young, 2019), they might apply similar moral opposition to technology that replaces other human functions as well.

A related area for psychologists to examine is how exposure to automated agents produces psychological and societal changes. As machines replace human capabilities—for example, an app that gives us directions rather than a human whom we stop on the street to ask for directions—there are likely significant downstream effects. Given that interacting with others *through* machines may decrease sociability (i.e., capacities for emotion recognition, empathy, perspective taking, and emotional intelligence; Waytz & Gray, 2018), interacting with machines *instead* of people will affect our social abilities even more potently.

The rise of machines is likely to affect political behavior and cognition as well, as studies have begun to show. The spread of automation can influence support for far-right political parties in both the United States and Europe (Anelli, Colantone, & Stanig, 2021; Dal Bó, Finan, Folke, Persson, & Rickne, 2018; Frey, Berger, & Chen, 2018; Zhen et al., 2017) in part because

81

the experience of automation as a threat shifts people's political leanings toward conservatism. Other work on the societal impact of robotics has examined how exposure to automation affects support for redistributive economic policies, such as a universal basic income—this work has found mixed effects (Busemeyer & Sahm, 2021; Gallego, Kuo, Alberto, & Manzanos, 2022; Kurer & Hausermann, 2021; Thewissen & Rueda, 2019), suggesting a ripe area for future exploration.

Beyond politics, other work has begun examining how artificial intelligence and robots are replacing not only human beings but are also taking on some of the properties of Gods, thereby reducing the global importance of religion (Jackson, Yam, Tang, Sibley, & Waytz, under review). If our gaps in knowledge are being filled by machines, then God is ousted from the gaps. Perhaps one day people will predominantly view machines as divine creations, as mechanical emissaries of God. Once machines have intelligence that humans cannot fathom, it is not such a big leap to psychologically associate them with God.

Finally, one essential area of future research is understanding whether all the excitement, anxiety, and novelty examined in this chapter might become quaint in just a few years' time. As social robots and artificial intelligence become increasingly integrated into our lives, might their overall psychological effect on humans shift or wane? Consider the uncanny valley. Adults and older children, but not younger kids, find machines that exhibit feelings creepy, suggesting that the uncanny valley is learned (Brink, Gray, & Wellman, 2019). This means that this phenomenon could also be unlearned. Gopnik (2019) raises the possibility that the uncanny valley phenomenon could cease to exist as future generations become more exposed to smart technology and thus more comfortable with technology that appears mentally capable.

As people grow more accustomed to advanced technology, its mere existence will likely engender more positive feelings (Eidelman, Crandall, & Pattershall, 2009). Already, research has shown that simply describing technology as originating before (vs. after) one's birth makes people evaluate it more favorably because, apparently longstanding technology feels more like the status quo, toward which people are positively biased (Smiley & Fisher, 2022). At some point in the future, perhaps soon, social robots and artificial intelligence —like the car, the telephone, and the personal computer— might become so common and embedded in our lives that rather than having positive or negative effects, these technologies will have little psychological effect at all. If people entirely habituate to machines, then within a generation, this entire chapter could be obsolete, just like the electromechanical calculator and tape player.

Conclusion

How do humans make sense of machines? In general, people understand our social world through agent-based cognition, categorizing entities based on the kinds of minds they seem to have. Machines represent a special kind of agent, an *agent of replacement*, explicitly designed by people to replace other agents—typically people. The goal of some of the earliest intelligent machines was to replace human minds, and modern machines are getting ever closer to this goal. Because machines are agents of replacement, they create a *fundamental ambiguity* about whether people should think of them more as just a machine, or as the human role they are replacing— especially when their appearance and behavior place them within the *questionable zone* (QZ) within a continuum ranging from "simple machine" to "complete human replacement."

Modern machines can serve as coworkers, teammates, nurses, and restaurant servers. On the one hand, these machines can make our lives easier and more efficient, but people are not always excited about machines replacing people. Sometimes, they are downright averse (or

creeped out) by machines that replace human minds and jobs, especially when it comes to involving moral decisions. To the extent machines do decide on moral matters, people want them to be as fair and impartial as possible.

The rise of machines will likely change the fabric of society and alter what we think of art, sex, and maybe even God. One day, machines might even replace scientists. Perhaps the next edition of this handbook will be written entirely by an algorithm.

References

Acemoglu, D., & Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*, *128*(6), 2188–2244. https://doi.org/10.1086/705716

Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, *6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Alaiad, A., & Zhou, L. (2013, August). Patient Behavioural Intention toward Adopting Healthcare Robots. *The 19th Americas Conference on Information Systems (AMCIS)*. https://www.researchgate.net/publication/269987741_Patient_Behavioural_Intention_toward_Adopting_Healthcare_Robots

Alarie, B., Cockfield, A., & GPT-3. (2021). Will Machines Replace Us? Machine-Authored Texts and the Future of Scholarship. *Law, Technology and Humans, 3*(2), Article 2. https://doi.org/10.5204/lthj.2089

Allen, R., & Choudhury, P. (Raj). (2022). Algorithm-Augmented Work and Domain Experience: The Countervailing Forces of Ability and Aversion. *Organization Science*, *33*(1), 149–169. https://doi.org/10.1287/orsc.2021.1554

Andrist, S., Mutlu, B., & Tapus, A. (2015). Look Like Me: Matching Robot Personality via Gaze to Increase Motivation. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3603–3612. https://doi.org/10.1145/2702123.2702592

Anelli, M., Colantone, I., & Stanig, P. (2021). Individual vulnerability to industrial robot adoption increases support for the radical right. *Proceedings of the National Academy of Sciences, 118*(47), e2111611118. https://doi.org/10.1073/pnas.2111611118

Appel, M., Izydorczyk, D., Weber, S., Mara, M., & Lischetzke, T. (2020). The uncanny of mind in a machine: Humanoid robots as tools, agents, and experiencers. *Computers in Human Behavior*, *102*, 274–286. https://doi.org/10.1016/j.chb.2019.07.031

Asch, S. E. (1955). Opinions and social pressure. *Scientific American*, *193*(5), 31–35. https://doi.org/10.1038/scientificamerican1155-31

Asimov, I. (2004). *I, Robot*. Bantam Books.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). The Moral Machine experiment. *Nature*, *563*(7729), Article 7729. https://doi.org/10.1038/s41586-018-0637-6

Bainbridge, W. A., Hart, J. W., Kim, E. S., & Scassellati, B. (2011). The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents. *International Journal of Social Robotics*, *3*, 41–52. https://doi.org/10.1007/s12369-010-0082-7

Bandura, A. (2011). Moral Disengagement. In *The Encyclopedia of Peace Psychology*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470672532.wbepp165

Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends in Cognitive Sciences*, *4*(1), 29–34. https://doi.org/10.1016/S1364-6613(99)01419-9

Barrett, J. L. (2004). *Why Would Anyone Believe in God?* AltaMira Press.

Bartneck, C., & Hu, J. (2008). Exploring the abuse of robots. *Interaction Studies*, *9*(3), 415–433. https://doi.org/10.1075/is.9.3.04bar

Bartneck, C., Nomura, T., Kanda, T., Suzuki, T., & Kennsuke, K. (2005). A cross-cultural study on attitudes towards robots. *Proceedings of the HCI International*. https://doi.org/10.13140/RG.2.2.35929.11367

Bartz, J. A., Tchalova, K., & Fenerci, C. (2016). Reminders of Social Connection Can Attenuate Anthropomorphism: A Replication and Extension of Epley, Akalis, Waytz, and Cacioppo (2008). *Psychological Science*, *27*(12), 1644–1650. https://doi.org/10.1177/0956797616668510

Baxter, P., de Greeff, J., & Belpaeme, T. (2013). Do children behave differently with a social robot if with peers? *Lecture Notes in Artificial Intelligence*, *8239*, 567–568. http://hdl.handle.net/1854/LU-8197719

Belk, R. (2022). Artificial Emotions and Love and Sex Doll Service Workers. *Journal of Service Research*, *25*(4), 521–536. https://doi.org/10.1177/10946705211063692

Bentham, J. (1776). *A Comment on the Commentaries and A Fragment on Government* (J. H. Burns & H. L. A. Hart, Eds.). Oxford University Press. https://www.ucl.ac.uk/bentham-project/publications/collected-works-jeremy-bentham/comment-commentaries-and-fragment-government

Beran, T. N., Ramirez-Serrano, A., Vanderkooi, O. G., & Kuhn, S. (2015). Humanoid robotics in health care: An exploration of children's and parents' emotional reactions. *Journal of Health Psychology*, *20*(7), 984–989. https://doi.org/10.1177/1359105313504794

Berger, B., Adam, M., Rühr, A., & Benlian, A. (2021). Watch Me Improve—Algorithm Aversion and Demonstrating the Ability to Learn. *Business & Information Systems Engineering*, *63*, 55–68. https://doi.org/10.1007/s12599-020-00678-5

Bevilacqua, M. (2018, December 19). *Uber Was Warned Before Self-Driving Car Crash That Killed Woman Walking Bicycle*. Bicycling. https://www.bicycling.com/news/a25616551/uber-self-driving-car-crash-cyclist/

Bickmore, T. W., Vardoulakis, L. M. P., & Schulman, D. (2013). Tinker: A relational agent museum guide. *Autonomous Agents and Multi-Agent Systems*, *27*, 254–276. https://doi.org/10.1007/s10458-012-9216-7

Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, *181*, 21–34. https://doi.org/10.1016/j.cognition.2018.08.003

Bigman, Y. E., Waytz, A., Alterovitz, R., & Gray, K. (2019). Holding Robots Responsible: The Elements of Machine Morality. *Trends in Cognitive Sciences*, *23*(5), 365–368. https://doi.org/10.1016/j.tics.2019.02.008

Bigman, Y. E., Wilson, D., Arnestad, M. N., Waytz, A., & Gray, K. (2022). Algorithmic Discrimination Causes Less Moral Outrage Than Human Discrimination. *Journal of Experimental Psychology: General*, *152*(1), 4–27. https://doi.org/10.1037/xge0001250

Bigman, Y. E., Yam, K. C., Marciano, D., Reynolds, S. J., & Gray, K. (2021). Threat of racial and economic inequality increases preference for algorithm decision-making. *Computers in Human Behavior*, *122*, 106859. https://doi.org/10.1016/j.chb.2021.106859

Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, *181*, 21–34. https://doi.org/10.1016/j.cognition.2018.08.003

Binfield, K. (Ed.). (2015). *Writings of the Luddites*. Johns Hopkins University Press. https://doi.org/10.1353/book.98247

Blanas, S., Gancia, G., & Lee, S. Y. (Tim). (2019). *Who Is Afraid of Machines?* (CEPR Discussion Paper No. 13802; C.E.P.R. Discussion Papers). https://econpapers.repec.org/paper/cprceprdp/13802.htm

Bó, E., Finan, F., Folke, O., Persson, T., & Rickne, J. (2018). Economic losers and political winners: Sweden's radical right. *Unpublished manuscript, Department of Political Science, UC Berkeley*, *2*(5), 2.

Boden, M. A. (2007). Authenticity and computer art. *Digital Creativity*, *18*(1), 3–10. https://doi.org/10.1080/14626260701252285

Bohannon, J. (2015). The synthetic therapist. *Science*, *349*(6245), 250–251. https://doi.org/10.1126/science.349.6245.250

Bonezzi, A., & Ostinelli, M. (2021). Can algorithms legitimize discrimination? *Journal of Experimental Psychology: Applied*, *27*(2), 447–459. https://doi.org/10.1037/xap0000294

Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, *352*(6293), 1573–1576. https://doi.org/10.1126/science.aaf2654

Booth, S., Tompkin, J., Pfister, H., Waldo, J., Gajos, K., & Nagpal, R. (2017). Piggybacking Robots: Human-Robot Overtrust in University Dormitory Security. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 426–434. https://doi.org/10.1145/2909824.3020211

Brandon, M., Shlykova, N., & Morgentaler, A. (2022). Curiosity and other attitudes towards sex robots: Results of an online survey. *Journal of Future Robot Life*, *3*(1), 3–16. https://doi.org/10.3233/FRL-200017

Brandstetter, J., Rácz, P., Beckner, C., Sandoval, E. B., Hay, J., & Bartneck, C. (2014). A peer

    pressure experiment: Recreation of the Asch conformity experiment with robots. *2014*

    IEEE/RSJ International Conference on Intelligent Robots and Systems, 1335–1340.

    https://doi.org/10.1109/IROS.2014.6942730

Brehm, J. W. (1966). *A Theory of Psychological Reactance*. Academic Press.

Brink, K. A., Gray, K., & Wellman, H. M. (2017). Creepiness creeps in: Uncanny valley feelings

    are acquired in childhood. *Child Development*. *90*(4), 1202-1214.

    http://dx.doi.org/10.1111/cdev.12999

Britannica. (n.d.). Political theory of Aristotle. In *Britannica*. Retrieved October 3, 2022, from

    https://www.britannica.com/biography/Aristotle/Political-theory

Brščić, D., Kidokoro, H., Suehiro, Y., & Kanda, T. (2015). Escaping from children's abuse of

    social robots. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on*

    Human-Robot Interaction, 59–66. https://doi.org/10.1145/2696454.2696468

Broadbent, E. (2017). Interactions with robots: the truths we reveal about ourselves. *Annual*

    Review of Psychology, 68(1), 627–652. https://doi.org/10.1146/annurev-psych-010416-

    043958

Burgoon, J. K., Bonito, J. A., Bengtsson, B., Cederberg, C., Lundeberg, M., & Allspach, L.

    (2000). Interactivity in human–computer interaction: A study of credibility, understanding,

    and influence. *Computers in Human Behavior*, *16*(6), 553–574.

    https://doi.org/10.1016/S0747-5632(00)00029-7

Burke, A. (2019). Occluded algorithms. *Big Data & Society*, *6*(2). https://doi.org/10.1177/2053951719858743

Burr, C., Cristianini, N., & Ladyman, J. (2018). An Analysis of the Interaction Between Intelligent Software Agents and Human Users. *Minds and Machines*, *28*(4), 735–774. https://doi.org/10.1007/s11023-018-9479-0

Busemeyer, M., & Sahm, A. (2021). Social Investment, Redistribution or Basic Income? Exploring the Association Between Automation Risk and Welfare State Attitudes in Europe. *Journal of Social Policy, 51*, 1–20. https://doi.org/10.1017/S0047279421000519

Byrne, D., & Griffitt, W. (1969). Similarity and awareness of similarity of personality characteristics as determinants of attraction. *Journal of Experimental Research in Personality*, *3*(3), 179–186.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., … Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour, 2*(9), 637–644. https://doi.org/10.1038/s41562-018-0399-z

Campbell, C. A. (1966). The Discipline of the Cave. *Philosophical Books, 7*(3), 10–12. https://doi.org/10.1111/j.1468-0149.1966.tb02632.x

Caporael, L. R. (1986). Anthropomorphism and mechanomorphism: Two faces of the human machine. *Computers in Human Behavior*, *2*(3), 215–234. https://doi.org/10.1016/0747-5632(86)90004-X

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-Dependent Algorithm Aversion. *Journal of Marketing Research*, *56*(5), 809–825. https://doi.org/10.1177/0022243719851788

Ceh, S., & Vanman, E. (2018). *The Robots are Coming! The Robots are Coming! Fear and Empathy for Human-like Entities*. PsyArXiv. https://doi.org/10.31234/osf.io/4cr2u

Chen, X. (2016). The problem of mind in Confucianism. *Asian Philosophy*, *26*(2), 166–181. https://doi.org/10.1080/09552367.2016.1165790

Chernyak, N., & Gary, H. E. (2016). Children's cognitive and behavioral reactions to an autonomous versus controlled social robot dog. *Early Education and Development, 27*(8), 1175–1189. https://doi.org/10.1080/10409289.2016.1158611

Clark, A., & Chalmers, D. (1998). The Extended Mind. *Analysis*, *58*(1), 7–19.

Coeckelbergh, M. (2010). Moral appearances: Emotions, robots, and human morality. *Ethics and Information Technology*, *12*(3), 235–241. https://doi.org/10.1007/s10676-010-9221-y

Conitzer, V., Sinnott-Armstrong, W., Borg, J. S., Deng, Y., & Kramer, M. (2017). Moral Decision Making Frameworks for Artificial Intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*, *31*, Article 1. https://doi.org/10.1609/aaai.v31i1.11140

Correia, F., Alves-Oliveira, P., Maia, N., Ribeiro, T., Petisca, S., Melo, F. S., & Paiva, A. (2016). Just follow the suit! Trust in human-robot interactions during card game playing. *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 507–512. https://doi.org/10.1109/ROMAN.2016.7745165

Correia, F., Alves-Oliveira, P., Ribeiro, T., Melo, F., & Paiva, A. (2017). A social robot as a card game player. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 13(1), 23–29.*

Correia, F., Mascarenhas, S., Prada, R., Melo, F. S., & Paiva, A. (2018). Group-based Emotions in Teams of Humans and Robots. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction,* 261–269. https://doi.org/10.1145/3171221.3171252

Coupé, T. (2019). Automation, job characteristics and job insecurity. *International Journal of Manpower*, *40*(7), 1288–1304. https://doi.org/10.1108/IJM-12-2018-0418

Creed, C., Beale, R., & Cowan, B. (2015). The impact of an embodied agent's emotional expressions over multiple interactions. *Interacting with Computers*, *27*(2), 172-188.

Danaher, J. (2016). Robots, law and the retribution gap. *Ethics and Information Technology*, *18*, 299–309. https://doi.org/10.1007/s10676-016-9403-3

Dang, J., & Liu, L. (2021). Robots are friends as well as foes: Ambivalent attitudes toward mindful and mindless AI robots in the United States and China. *Computers in Human Behavior*, *115*, 106612. https://doi.org/10.1016/j.chb.2020.106612

Dauth, W., Findeisen, S., Suedekum, J., & Woessner, N. (2018). *Adjusting to Robots: Worker-Level Evidence* (Research Paper No. 013; Institute Working Paper (Federal Reserve Bank of Minneapolis. Opportunity and Inclusive Growth Institute)). Federal Reserve Bank of Minneapolis. https://doi.org/10.21034/iwp.13

de Graaf, M. M. A., & Malle, B. F. (2019). People's Explanations of Robot Behavior Subtly Reveal Mental State Inferences. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 239–248. https://doi.org/10.1109/HRI.2019.8673308

De Jong, B. A., Dirks, K. T., & Gillespie, N. (2016). Trust and team performance: A meta-analysis of main effects, moderators, and covariates. *Journal of Applied Psychology*, *101*(8), 1134–1150. https://doi.org/10.1037/apl0000110

de Melo, C. M., Marsella, S., & Gratch, J. (2018). Social decisions and fairness change when people's interests are represented by autonomous agents. *Autonomous Agents and Multi-Agent Systems*, *32*, 163–187. https://doi.org/10.1007/s10458-017-9376-6

de Visser, E. J., Monfort, S. S., Goodyear, K., Lu, L., O'Hara, M., Lee, M. R., Parasuraman, R., & Krueger, F. (2017). A Little Anthropomorphism Goes a Long Way: Effects of Oxytocin on Trust, Compliance, and Team Performance With Automated Agents. *Human Factors*, *59*(1), 116–133. https://doi.org/10.1177/0018720816687205

de Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A. B., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, *22*(3), 331–349. https://doi.org/10.1037/xap0000092

de Visser, E., & Parasuraman, R. (2011). Adaptive Aiding of Human-Robot Teaming: Effects of Imperfect Automation on Performance, Trust, and Workload. *Journal of Cognitive Engineering and Decision Making*, *5*(2), 209–231. https://doi.org/10.1177/1555343411410160

DeFranco, J. F., Voas, J., & Kshetri, N. (2022). Algorithms: Society's Invisible Puppeteers. *Computer*, *55*(4), 12–14. https://doi.org/10.1109/MC.2021.3128675

Dehouche, N. (2021). Plagiarism in the age of massive Generative Pre-trained Transformers (GPT-3). *Ethics in Science and Environmental Politics*, *21*, 17–23. https://doi.org/10.3354/esep00195

Dennett, D. C. (1987). *The Intentional Stance*. The MIT Press.

Desai, M., Kaniarasu, P., Medvedev, M., Steinfeld, A., & Yanco, H. (2013). Impact of robot failures and feedback on real-time trust. *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 251–258. https://doi.org/10.1109/HRI.2013.6483596

Descartes, R. (2008). *A Discourse on the Method: Of Correctly Conducting One's Reason and Seeking Truth in the Sciences* (I. Maclean, Trans.). Oxford University Press.

DeSteno, D., Breazeal, C., Frank, R. H., Pizarro, D., Baumann, J., Dickens, L., & Lee, J. J. (2012). Detecting the Trustworthiness of Novel Partners in Economic Exchange. *Psychological Science*, *23*(12), 1549–1556. https://doi.org/10.1177/0956797612448793

Devlin, K. (2015, September 17). *In defence of sex machines: Why trying to ban sex robots is wrong*. The Conversation. http://theconversation.com/in-defence-of-sex-machines-why-trying-to-ban-sex-robots-is-wrong-47641

Dietvorst, B. J., & Bartels, D. M. (2022). Consumers Object to Algorithms Making Morally

    Relevant Tradeoffs Because of Algorithms' Consequentialist Decision Strategies. *Journal*

    of Consumer Psychology, *32*(3), 406–424. https://doi.org/10.1002/jcpy.1266

Dietvorst, B. J., & Bharti, S. (2020). People Reject Algorithms in Uncertain Decision Domains

    Because They Have Diminishing Sensitivity to Forecasting Error. *Psychological Science*,

    31(10), 1302–1314. https://doi.org/10.1177/0956797620948841

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously

    avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*,

    144(1), 114–126. https://doi.org/10.1037/xge0000033

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming Algorithm Aversion: People

    Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them. *Management*

    Science, *64*(3), 1155–1170. https://doi.org/10.1287/mnsc.2016.2643

DiSalvo, C. F., Gemperle, F., Forlizzi, J., & Kiesler, S. (2002). All robots are not created equal:

    The design and perception of humanoid robot heads. *Proceedings of the 4th Conference on*

    Designing Interactive Systems: Processes, Practices, Methods, and Techniques, 321–326.

    https://doi.org/10.1145/778712.778756

Di Tella, R., & Rodrik, D. (2020). Labour Market Shocks and the Demand for Trade Protection:

    Evidence from Online Surveys. *The Economic Journal*, *130*(628), 1008–1030.

    https://doi.org/10.1093/ej/ueaa006

Doyle, C. M., & Gray, K. (2020). How people perceive the minds of the dead: The importance of

    consciousness at the moment of death. *Cognition*, *202*, 104308.

    https://doi.org/10.1016/j.cognition.2020.104308

Efendić, E., Van de Calseyde, P. P. F. M., & Evans, A. M. (2020). Slow response times undermine trust in algorithmic (but not human) predictions. *Organizational Behavior and Human Decision Processes*, *157*, 103–114. https://doi.org/10.1016/j.obhdp.2020.01.008

Eidelman, S., Crandall, C. S., & Pattershall, J. (2009). The existence bias. *Journal of Personality and Social Psychology, 97,* 765–775. https://doi.org/10.1037/a0017058

Elkins, K., & Chun, J. (2020). Can GPT-3 Pass a Writer's Turing Test? *Journal of Cultural Analytics*, *5*(2). https://doi.org/10.22148/001c.17212

Enz, S., Diruf, M., Spielhagen, C., Zoll, C., & Vargas, P. A. (2011). The Social Role of Robots in the Future—Explorative Measurement of Hopes and Fears. *International Journal of Social Robotics*, *3*, 263–271. https://doi.org/10.1007/s12369-011-0094-y

Epley, N., Waytz, A., Akalis, S., & Cacioppo, J. T. (2008). When We Need A Human: Motivational Determinants of Anthropomorphism. *Social Cognition*, *26*(2), 143–155. https://doi.org/10.1521/soco.2008.26.2.143

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Eyssel, F., & Kuchenbrandt, D. (2011). Manipulating anthropomorphic inferences about NAO: The role of situational and dispositional aspects of effectance motivation. *2011 IEEE International Workshop on Robot and Human Communication (ROMAN)*, 467–472. https://doi.org/10.1109/ROMAN.2011.6005233

Eyssel, F., & Kuchenbrandt, D. (2012). Social categorization of social robots: Anthropomorphism as a function of robot group membership. *British Journal of Social Psychology*, *51*(4), 724–731. https://doi.org/10.1111/j.2044-8309.2011.02082.x

Eyssel, F., Kuchenbrandt, D., & Bobinger, S. (2011). Effects of anticipated human-robot interaction and predictability of robot behavior on perceptions of anthropomorphism. *Proceedings of the 6th International Conference on Human-Robot Interaction*, 61–68. https://doi.org/10.1145/1957656.1957673

Eyssel, F., Kuchenbrandt, D., Bobinger, S., de Ruiter, L., & Hegel, F. (2012). "If you sound like me, you must be more human": On the interplay of robot and user features on human-robot acceptance and anthropomorphism. *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, 125–126. https://doi.org/10.1145/2157689.2157717

Eyssel, F., & Reich, N. (2013). Loneliness makes the heart grow fonder (of robots)—On the effects of loneliness on psychological anthropomorphism. *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 121–122. https://doi.org/10.1109/HRI.2013.6483531

Falk, M. (2021). Artificial stupidity. *Interdisciplinary Science Reviews*, *46*(1–2), 36–52. https://doi.org/10.1080/03080188.2020.1840219

Fasola, J., & Matarić, M. J. (2013). A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction*, *2*(2), 3–32. https://doi.org/10.5898/JHRI.2.2.Fasola

Filiz, I., Judek, J. R., Lorenz, M., & Spiwoks, M. (2022). Algorithm Aversion as an Obstacle in

    the Establishment of Robo Advisors. *Journal of Risk and Financial Management*, *15*(8),

    Article 8. https://doi.org/10.3390/jrfm15080353

Fisher, M., Goddu, M. K., & Keil, F. C. (2015). Searching for explanations: How the Internet

    inflates estimates of internal knowledge. *Journal of Experimental Psychology.General,*

    144(3), 674–687. https://doi.org/10.1037/xge0000070

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition:

    Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83.

    https://doi.org/10.1016/j.tics.2006.11.005

Fitz, N. S., Nadler, R., Manogaran, P., Chong, E. W. J., & Reiner, P. B. (2014). Public attitudes

    toward cognitive enhancement. *Neuroethics, 7, 173–188. https://doi.org/10.1007/s12152-*

    013-9190-z

Floridi, L., & Sanders, J. W. (2004). On the Morality of Artificial Agents. *Minds and Machines,*

    14(3), 349–379. https://doi.org/10.1023/B:MIND.0000035461.63578.9d

Fogg, B. J., & Nass, C. (1997). Silicon sycophants: The effects of computers that flatter.

    International Journal of Human-Computer Studies, *46*(5), 551–561.

    https://doi.org/10.1006/ijhc.1996.0104

Franklin, C. E. (2015). Everyone thinks that an ability to do otherwise is necessary for free will

    and moral responsibility. *Philosophical Studies*, *172*, 2091–2107.

    https://doi.org/10.1007/s11098-014-0399-4

Fraune, M. R., Nishiwaki, Y., Sabanović, S., Smith, E. R., & Okada, M. (2017). Threatening Flocks and Mindful Snowflakes: How Group Entitativity Affects Perceptions of Robots. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 205–213. https://doi.org/10.1145/2909824.3020248

Frey, C. B., Berger, T., & Chen, C. (2018). Political machinery: Did robots swing the 2016 US presidential election? *Oxford Review of Economic Policy, 34*(3), 418–442. https://doi.org/10.1093/oxrep/gry007

Fuentes, A. (2017). *The Creative Spark: How Imagination Made Humans Exceptional*. Penguin Publishing Group.

Gallego, Aina, Alex Kuo, Pepe Fernández-Albertos and Dulce Manzano. 2022. "Technological risk and policy preferences." *Comparative Political Studies, 55*(1): 60-92.

Gallimore, D., Lyons, J. B., Vo, T., Mahoney, S., & Wynne, K. T. (2019). Trusting Robocop: Gender-Based Effects on Trust of an Autonomous Robot. *Frontiers in Psychology*, *10*(482), 1–9. https://doi.org/10.3389/fpsyg.2019.00482

Gamez, P., Shank, D. B., Arnold, C., & North, M. (2020). Artificial virtue: The machine question and perceptions of moral character in artificial moral agents. *AI & SOCIETY*, *35*, 795–809. https://doi.org/10.1007/s00146-020-00977-1

Gamez-Djokic, M., & Waytz, A. (2020). Concerns About Automation and Negative Sentiment Toward Immigration. *Psychological Science*, *31*(8), 987–1000. https://doi.org/10.1177/0956797620929977

Gardner, W. L., Pickett, C. L., Jefferis, V., & Knowles, M. (2005). On the Outside Looking In: Loneliness and Social Monitoring. *Personality and Social Psychology Bulletin*, *31*(11), 1549–1560. https://doi.org/10.1177/0146167205277208

Gates, S. W., Perry, V. G., & Zorn, P. M. (2002). Automated underwriting in mortgage lending: Good news for the underserved? *Housing Policy Debate*, *13*(2), 369–391. https://doi.org/10.1080/10511482.2002.9521447

Gaudiello, I., Zibetti, E., Lefort, S., Chetouani, M., & Ivaldi, S. (2016). Trust as indicator of robot functional and social acceptance. An experimental study on user conformation to iCub answers. *Computers in Human Behavior*, *61*, 633–655. https://doi.org/10.1016/j.chb.2016.03.057

Ghimire, R., Skinner, J., & Carnathan, M. (2020). Who perceived automation as a threat to their jobs in metro Atlanta: Results from the 2019 Metro Atlanta Speaks survey. *Technology in Society*, *63*, 101368. https://doi.org/10.1016/j.techsoc.2020.101368

Ghosh, A., & Fossas, G. (2022). *Can There be Art Without an Artist?* http://arxiv.org/abs/2209.07667

Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals*, *14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Gnambs, T., & Appel, M. (2019). Are robots becoming unpopular? Changes in attitudes towards autonomous robotic systems in Europe. *Computers in Human Behavior*, *93*, 53–61. https://doi.org/10.1016/j.chb.2018.11.045

Gockley, R., Simmons, R., & Forlizzi, J. (2006). Modeling Affect in Socially Interactive Robots. *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication*, 558–563. https://doi.org/10.1109/ROMAN.2006.314448

Gombolay, M. C., Gutierrez, R. A., Clarke, S. G., Sturla, G. F., & Shah, J. A. (2015). Decision-making authority, team efficiency and human worker satisfaction in mixed human–robot teams. *Autonomous Robots*, *39*, 293–312. https://doi.org/10.1007/s10514-015-9457-9

Gopnik, A. (2019, January 10). A Generational Divide in the Uncanny Valley. *Wall Street Journal. https://www.wsj.com/articles/a-generational-divide-in-the-uncanny-valley-11547138712*

GovTech Singapore. (n.d.). *'Ask Jamie' Virtual Assistant*. GovTech Singapore. Retrieved February 20, 2023, from https://www.tech.gov.sg/products-and-services/ask-jamie/

Graetz, G., & Michaels, G. (2018). Robots at Work. *The Review of Economics and Statistics*, *100*(5), 753–768. https://doi.org/10.1162/rest_a_00754

Granulo, A., Fuchs, C., & Puntoni, S. (2019). Psychological reactions to human versus robotic job replacement. *Nature Human Behaviour*, *3*(10), Article 10. https://doi.org/10.1038/s41562-019-0670-y

Granulo, A., Fuchs, C., & Puntoni, S. (2021). Preference for Human (vs. Robotic) Labor is Stronger in Symbolic Consumption Contexts. *Journal of Consumer Psychology*, *31*(1), 72–80. https://doi.org/10.1002/jcpy.1181

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of Mind Perception. *Science*, *315*(5812), 619. https://doi.org/10.1126/science.1134475

Gray, K., Knickman, T. A., & Wegner, D. M. (2011). More dead than dead: Perceptions of persons in the persistent vegetative state. *Cognition*, *121*(2), 275–280. https://doi.org/10.1016/j.cognition.2011.06.014

Gray, K., Schein, C., & Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General*, *143*(4), 1600–1615. https://doi.org/10.1037/a0036149

Gray, K., & Wegner, D. M. (2010). Blaming God for Our Pain: Human Suffering and the Divine Mind. *Personality and Social Psychology Review*, *14*(1), 7–16. https://doi.org/10.1177/1088868309350299

Gray, K., & Wegner, D. M. (2011). Dimensions of Moral Emotions. *Emotion Review*, *3*(3), 258–260. https://doi.org/10.1177/1754073911402388

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, *125*(1), 125–130. https://doi.org/10.1016/j.cognition.2012.06.007

Gray, K., Young, L., & Waytz, A. (2012). Mind Perception Is the Essence of Morality. *Psychological Inquiry*, *23*(2), 101–124. https://doi.org/10.1080/1047840X.2012.651387

Guthrie, S. E. (1995). *Faces in the Clouds: A New Theory of Religion*. Oxford University Press.

Hallevy, G. (2010a). "I, Robot – I, Criminal"—When Science Fiction Becomes Reality: Legal Liability of AI Robots committing Criminal Offenses. *Syracuse Science & Technology Law Reporter, 22*, 1–37.

Hallevy, G. (2010b). The Criminal Liability of Artificial Intelligence Entities—From Science

    Fiction to Legal Social Control. *Akron Intellectual Property Journal*, *4*(2, Article 1), 171–

    202.

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., de Visser, E. J., & Parasuraman,

    R. (2011). A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. *Human*

    Factors, *53*(5), 517–527. https://doi.org/10.1177/0018720811417254

Hanson Robotics. (n.d.). *Sophia*. Hanson Robotics. Retrieved February 20, 2023, from

    https://www.hansonrobotics.com/sophia/

Harvey, N., & Fischer, I. (1997). Taking Advice: Accepting Help, Improving Judgment, and

    Sharing Responsibility. *Organizational Behavior and Human Decision Processes*, *70*(2),

    117–133. https://doi.org/10.1006/obhd.1997.2697

Haslam, N. (2006). Dehumanization: An Integrative Review. *Personality and Social Psychology*

    Review, *10*(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4

Haslam, N., Bain, P., Douge, L., Lee, M., & Bastian, B. (2005). More human than you:

    Attributing humanness to self and others. *Journal of Personality and Social Psychology*,

    89(6), 937–950. https://doi.org/10.1037/0022-3514.89.6.937

Haslam, N., Bastian, B., & Bissett, M. (2004). Essentialist Beliefs about Personality and Their

    Implications. *Personality and Social Psychology Bulletin*, *30*(12), 1661–1673.

    https://doi.org/10.1177/0146167204271182

Haslam, N., Loughnan, S., & Holland, E. (2013). The psychology of humanness. In S. J. Gervais

    (Ed.), *Objectification and (de)humanization: 60th Nebraska symposium on motivation* (Vol.

60, pp. 25–51). Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-6959-9_2

Hebb, D. O. (1946). Emotion in man and animal: An analysis of the intuitive processes of recognition. *Psychological Review*, *53*(2), 88–106. https://doi.org/10.1037/h0063033

Heider, F., & Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American Journal of Psychology*, *57*(2), 243–259. https://doi.org/10.2307/1416950

Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, *22*(1), 5–53. https://doi.org/10.1145/963770.963772

Hertz, N., Shaw, T., de Visser, E. J., & Wiese, E. (2019). Mixing It Up: How Mixed Groups of Humans and Machines Modulate Conformity. *Journal of Cognitive Engineering and Decision Making*, *13*(4), 242–257. https://doi.org/10.1177/1555343419869465

Hertz, N., & Wiese, E. (2018). Under Pressure: Examining Social Conformity With Computer and Robot Groups. *Human Factors*, *60*(8), 1207–1218. https://doi.org/10.1177/0018720818788473

Heßler, P. O., Pfeiffer, J., & Hafenbrädl, S. (2022). When Self-Humanization Leads to Algorithm Aversion. *Business & Information Systems Engineering*, *64*(3), 275–292. https://doi.org/10.1007/s12599-022-00754-y

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup Bias. *Annual Review of Psychology*, *53*, 575–604. https://doi.org/10.1146/annurev.psych.53.100901.135109

Hijazi, A., Ferguson, C. J., Richard Ferraro, F., Hall, H., Hovee, M., & Wilcox, S. (2019).

    Psychological Dimensions of Drone Warfare. *Current Psychology*, *38*, 1285–1296.

    https://doi.org/10.1007/s12144-017-9684-7

Himma, K. E. (2009). Artificial agency, consciousness, and the criteria for moral agency: What

    properties must an artificial agent have to be a moral agent? *Ethics and Information*

    Technology, *11*, 19–29. https://doi.org/10.1007/s10676-008-9167-5

Hochberg, M. (2016, December 22). Everything you need to know about the Bionic Bar on

    Royal Caribbean's Harmony of the Seas. *Royal Caribbean Blog*.

    https://www.royalcaribbeanblog.com/2016/12/22/everything-you-need-know-about-the-

    bionic-bar-royal-caribbeans-harmony-of-the-seas

Hoff, K. A., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on

    Factors That Influence Trust. *Human Factors*, *57*(3), 407–434.

    https://doi.org/10.1177/0018720814547570

Hong, J. W., Peng, Q., & Williams, D. (2021). Are you ready for artificial Mozart and Skrillex?

    An experiment testing expectancy violation theory and AI music. *New Media & Society*,

    23(7), 1920–1935. https://doi.org/10.1177/1461444820925798

Hsu, A. (2022, September 8). *California dockworkers are worried about losing their good-*

    paying jobs to robots. WFAE 90.7 - Charlotte's NPR News Source.

    https://www.wfae.org/2022-09-08/california-dockworkers-are-worried-about-losing-their-

    good-paying-jobs-to-robots

Hume, D. (1751). *An Enquiry concerning the Principles of Morals*. Clarendon Press.

IFR. (2021, October 28). *IFR presents World Robotics 2021 reports*. International Federation of

Robotics. https://ifr.org/ifr-press-releases/news/robot-sales-rise-again

Innocenti, S., & Golin, M. (2022). Human capital investment and perceived automation risks:

Evidence from 16 countries. *Journal of Economic Behavior & Organization*, *195*, 27–41.

https://doi.org/10.1016/j.jebo.2021.12.027

International Federation of Robotics. (2022). *Artificial Intelligence in Robotics [Position Paper]*.

https://ifr.org/papers

Ishii, T., & Watanabe, K. (2019). How People Attribute Minds to Non-Living Entities. *2019 11th

International Conference on Knowledge and Smart Technology (KST)*, 213–217.

https://doi.org/10.1109/KST.2019.8687324

Iza World of Labor. (2019, August 29). *French supermarket workers protest against automated

checkout stations*. Iza World of Labor. https://wol.iza.org/news/french-supermarket-

workers-protest-against-automated-checkout-stations

Jackson, J. C., Castelo, N., & Gray, K. (2020). Could a rising robot workforce make humans less

prejudiced? *The American Psychologist*, *75*(7), 969–982.

https://doi.org/10.1037/amp0000582

Jackson, J. C., Yam, K. C., Tang, P., Liu, T., & Shariff, A. (under review). *Exposure to robot

preachers undermines religious commitment*.

Jackson, J. C., Yam, K. C., Tang, P., Sibley, C., & Waytz, A. (under review). *Machina ex Deux:

Exposure to automated agents explains worldwide religious declines*.

Jackson, R. B., & Williams, T. (2019). Language-Capable Robots may Inadvertently Weaken Human Moral Norms. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 401–410. https://doi.org/10.1109/HRI.2019.8673123

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, *61*(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

Jauernig, J., Uhl, M., & Walkowitz, G. (2022). People Prefer Moral Discretion to Algorithms: Algorithm Aversion Beyond Intransparency. *Philosophy & Technology*, *35*(2). https://doi.org/10.1007/s13347-021-00495-y

Jeff Grubb's Game Mess. (2018, May 9). *Google Duplex: A.I. Assistant Calls Local Businesses To Make Appointments [Video]*. Youtube. https://www.youtube.com/watch?v=D5VN56jQMWM

Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1553), 2635–2650. https://doi.org/10.1098/rstb.2010.0146

Jentsch, E. (1906). On the psychology of the uncanny. *Angelaki*, *2*(1), 7–16. https://doi.org/10.1080/09697259708571910

Johnsson, I.-M., Nass, C., Harris, H., & Takayama, L. (2005). Matching In-Car Voice with Driver State: Impact on Attitude and Driving Performance. *Driving Assessment Conference*, *3*, Article 2005. https://doi.org/10.17077/drivingassessment.1158

Jung, M. F., Lee, J. J., DePalma, N., Adalgeirsson, S. O., Hinds, P. J., & Breazeal, C. (2013). Engaging robots: Easing complex human-robot teamwork using backchanneling. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, 1555–1566. https://doi.org/10.1145/2441776.2441954

Jung, M. F., Martelaro, N., & Hinds, P. J. (2015). Using Robots to Moderate Team Conflict: The Case of Repairing Violations. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 229–236. https://doi.org/10.1145/2696454.2696460

Kahn, P. H. Jr., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., Ruckert, J. H., & Shen, S. (2012). "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. *Developmental Psychology*, *48*(2), 303–314. https://doi.org/10.1037/a0027033

Kahn, P. H., Kanda, T., Ishiguro, H., Gill, B. T., Shen, S., Gary, H. E., & Ruckert, J. H. (2015). Will people keep the secret of a humanoid robot?: psychological intimacy in HRI. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction,* 173–180. https://doi.org/10.1145/2696454.2696486

Kahn, P. H., Kanda, T., Ishiguro, H., Gill, B. T., Ruckert, J. H., Shen, S., Gary, H. E., Reichert, A. L., Freier, N. G., & Severson, R. L. (2012). Do people hold a humanoid robot morally accountable for the harm it causes? *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction,* 33–40. https://doi.org/10.1145/2157689.2157696

Kant, I. (1785). *Grounding for the Metaphysics of Morals ; With, On a Supposed Right to Lie Because of Philanthropic Concerns*. Hackett Publishing Company.

Kant, I. (1788). *Critique of Practical Reason*. Cambridge University Press.

Karniol, R. (2003). Egocentrism versus protocentrism: The status of self in social prediction. *Psychological Review*, *110*(3), 564–580. https://doi.org/10.1037/0033-295X.110.3.564

Kätsyri, J., Förger, K., Mäkäräinen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology*, *6*(390). https://doi.org/10.3389/fpsyg.2015.00390

Keijsers, M., Kazmi, H., Eyssel, F., & Bartneck, C. (2021). Teaching Robots a Lesson: Determinants of Robot Punishment. *International Journal of Social Robotics*, *13*, 41–54. https://doi.org/10.1007/s12369-019-00608-w

Kidd, C. D., & Breazeal, C. (2008). Robots at home: Understanding long-term human-robot interaction. *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3230–3235. https://doi.org/10.1109/IROS.2008.4651113

Kidokoro, H., Kanda, T., Brščic, D., and Shiomi, M. (2013). Will I bother here?: a robot anticipating its influence on pedestrian walking comfort. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction, HRI '13 (Piscataway, NJ: IEEE Press),* 259–266. doi: 10.1109/HRI.2013.6483597

Kim, K. J., Park, E., & Sundar, S. S. (2013). Caregiving role in human–robot interaction: A

    study of the mediating effects of perceived benefit and social presence. *Computers in*

    Human Behavior, 29, 1799–1806. https://doi.org/10.1016/j.chb.2013.02.009

Kim, M.-S., & Kim, E.-J. (2013). Humanoid robots as "The Cultural Other": Are we able to love

    our creations? *AI & SOCIETY*, *28*, 309–318. https://doi.org/10.1007/s00146-012-0397-z

Kim, S. (Sam), Kim, J., Badu-Baiden, F., Giroux, M., & Choi, Y. (2021). Preference for robot

    service or human service in hotels? Impacts of the COVID-19 pandemic. *International*

    Journal of Hospitality Management, *93*, 102795.

    https://doi.org/10.1016/j.ijhm.2020.102795

Kim, E., Paul, R., Shic, F., & Scassellati, B. (2012). Bridging the research gap: making HRI

    useful to individuals with autism. *Communication Disorders Faculty Publications, 1*(1).

    https://doi.org/10.5898/JHRI.1.1.Kim

Kizilcec, R. F. (2016). How Much Information? Effects of Transparency on Trust in an

    Algorithmic Interface. *Proceedings of the 2016 CHI Conference on Human Factors in*

    Computing Systems, 2390–2395. https://doi.org/10.1145/2858036.2858402

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2017). *Human*

    Decisions and Machine Predictions (Working Paper No. 23180). National Bureau of

    Economic Research. https://doi.org/10.3386/w23180

Kleinberg, J., Ludwig, J., Mullainathan, S., & Sunstein, C. R. (2018). Discrimination in the Age

    of Algorithms. *Journal of Legal Analysis*, *10*, 113–174. https://doi.org/10.1093/jla/laz001

Köbis, N., Bonnefon, J.-F., & Rahwan, I. (2021). Bad machines corrupt good morals. *Nature Human Behaviour*, *5*, Article 6. https://doi.org/10.1038/s41562-021-01128-2

Koch, M., Manuylov, I., & Smolka, M. (2021). Robots and Firms. *The Economic Journal*, *131*(638), 2553–2584. https://doi.org/10.1093/ej/ueab009

Kolbjørnsrud, V., Amico, R., & Thomas, R. J. (2017). Partnering with AI: How organizations can win over skeptical managers. *Strategy & Leadership*, *45*(1), 37–43. https://doi.org/10.1108/SL-12-2016-0085

Kondo, Y., Takemura, K., Takamatsu, J., & Ogasawara, T. (2013). A gesture-centric android system for multi-party human-robot interaction. *Journal of Human-Robot Interaction, 2*(1), 133–151. https://doi.org/10.5898/JHRI.2.1.Kondo

Konovsky, M. A. (2000). Understanding Procedural Justice and Its Impact on Business Organizations. *Journal of Management*, *26*(3), 489–511. https://doi.org/10.1177/014920630002600306

Koo, J., Kwac, J., Ju, W., Steinert, M., Leifer, L., & Nass, C. (2015). Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, *9*, 269–275. https://doi.org/10.1007/s12008-014-0227-2

Koverola, M., Kunnari, A., Drosinou, M., Palomäki, J., Hannikainen, I. R., Jirout Košová, M., Kopecký, R., Sundvall, J., & Laakasuo, M. (2022). Treatments approved, boosts eschewed: Moral limits of neurotechnological enhancement. *Journal of Experimental Social Psychology*, *102*, 104351. https://doi.org/10.1016/j.jesp.2022.104351

Krach, S., Hegel, F., Wrede, B., Sagerer, G., Binkofski, F., & Kircher, T. (2008). Can Machines Think? Interaction and Perspective Taking with Robots Investigated via fMRI. *PLoS ONE*, *3*(7), e2597. https://doi.org/10.1371/journal.pone.0002597

Kurer, T., & Häusermann, S. (2021). Automation and social policy: Which policy responses do at-risk workers support? (Working Paper Series No. 2), Welfare Priorities. University of Zurich.

Lammer, L., Huber, A., Weiss, A., & Vincze, M. (2014). Mutual Care: How older adults react when they should help their care robot. *AISB 2014 - 50th Annual Convention of the AISB*. (pp. 1-4). London, UK: Routledge.

Lan, J., Yuan, B., & Gong, Y. (2022). Predicting the change trajectory of employee robot-phobia in the workplace: The role of perceived robot advantageousness and anthropomorphism. *Computers in Human Behavior*, *135*, 107366. https://doi.org/10.1016/j.chb.2022.107366

Landy, J. F., Walco, D. K., & Bartels, D. M. (2017). What's wrong with using steroids? Exploring whether and why people oppose the use of performance enhancing drugs. *Journal of Personality and Social Psychology, 113,* 377–392. https://doi.org/10.1037/pspa0000089

Langer, E. J. (1989). Minding Matters: The Consequences of Mindlessness–Mindfulness. *Advances in Experimental Social Psychology*, *22*, 137–173. https://doi.org/10.1016/S0065-2601(08)60307-X

Langer, E. J. (1992). Matters of mind: Mindfulness/mindlessness in perspective. *Consciousness and Cognition*, *1*(3), 289–305. https://doi.org/10.1016/1053-8100(92)90066-J

Le Guin, U. K. (1993). *The ones who walk away from Omelas*. Creative Education.

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer

theory. *Advances in Experimental Social Psychology*, *32*, 1–62.

https://doi.org/10.1016/S0065-2601(00)80003-9

Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To Engage or Not to Engage with AI for

Critical Judgments: How Professionals Deal with Opacity When Using AI for Medical

Diagnosis. *Organization Science*, *33*(1), 126–148. https://doi.org/10.1287/orsc.2021.1549

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance.

Human Factors, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Lee, K. M., Park, N., & Song, H. (2005). Can a Robot Be Perceived as a Developing Creature?

Effects of a Robot's Long-Term Cognitive Developments on Its Social Presence and

People's Social Responses Toward It. *Human Communication Research*, *31*(4), 538–563.

https://doi.org/10.1111/j.1468-2958.2005.tb00882.x

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and

emotion in response to algorithmic management. *Big Data & Society*, *5*(1), 1–16.

https://doi.org/10.1177/2053951718756684

Lee, M., Ruijten, P., Frank, L., de Kort, Y., & IJsselsteijn, W. (2021). People May Punish, But

Not Blame Robots. *Proceedings of the 2021 CHI Conference on Human Factors in

Computing Systems*, 1–11. https://doi.org/10.1145/3411764.3445284

Leite, I., Martinho, C., & Paiva, A. (2013). Social robots for long-term interaction: a survey. *International Journal of Social Robotics, 2*(5), 291–308. https://doi.org/10.1007/s12369-013-0178-y

Leite, I., McCoy, M., Lohani, M., Ullman, D., Salomons, N., Stokes, C., Rivers, S., & Scassellati, B. (2015). Emotional storytelling in the classroom: individual versus group interaction between children and robots. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction,* 75–82. https://doi.org/10.1145/2696454.2696481

Leite, I., Pereira, A., & Lehman, J. F. (2017). Persistent memory in repeated child-robot conversations. In *Proceedings of the 2017 Conference on Interaction Design and Children,* 238–247. https://doi.org/10.1145/3078072.3079728

Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, Transparent, and Accountable Algorithmic Decision-making Processes. *Philosophy & Technology*, *31*, 611–627. https://doi.org/10.1007/s13347-017-0279-x

Lewandowsky, S., Mundy, M., & Tan, G. P. A. (2000). The dynamics of trust: Comparing humans to automation. *Journal of Experimental Psychology: Applied*, *6*(2), 104–123. https://doi.org/10.1037/1076-898X.6.2.104

Leyzberg, D., Spaulding, S., Toneva, M., & Scassellati, B. (2012). The physical presence of a robot tutor increases cognitive learning gains. In *34th Annual Conference of the Cognitive Science Society.* (Vol. 34, No. 34).

Li, J. (2015). The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies*, *77*, 23–37. https://doi.org/10.1016/j.ijhcs.2015.01.001

Li, S., Yu, F., & Peng, K. (2020). Effect of State Loneliness on Robot Anthropomorphism: Potential Edge of Social Robots Compared to Common Nonhumans. *2nd International Conference on Artificial Intelligence and Computer Science*, *1631*, 012024. https://doi.org/10.1088/1742-6596/1631/1/012024

Life at Google. (2021, July 28). *What's it like to work on the Google Maps team? [Video]*. Youtube. https://www.youtube.com/watch?v=vaHK4OXAN-A

Lima, G., Jeon, C., Cha, M., & Park, K. (2020). Will Punishing Robots Become Imperative in the Future? *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–8. https://doi.org/10.1145/3334480.3383006

Lingmont, D. N. J., & Alexiou, A. (2020). The contingent effect of job automating technology awareness on perceived job insecurity: Exploring the moderating role of organizational culture. *Technological Forecasting and Social Change*, *161*, 120302. https://doi.org/10.1016/j.techfore.2020.120302

Liu, P., Glas, D. F., Kanda, T., Ishiguro, H., & Hagita, N. (2013, March). It's not polite to point: generating socially-appropriate deictic behaviors towards people. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 267-274). IEEE.

Locke, J. (1836). *An Essay Concerning Human Understanding*. T. Tegg and Son.

Loewenstein, G. (1996). Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes*, *65*(3), 272–292. https://doi.org/10.1006/obhd.1996.0028

Logg, J. M., Haran, U., & Moore, D. A. (2018). Is overconfidence a motivated bias? Experimental evidence. *Journal of Experimental Psychology: General*, *147*(10), 1445–1465. https://doi.org/10.1037/xge0000500

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to Medical Artificial Intelligence. *Journal of Consumer Research*, *46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

Loughnan, S., & Haslam, N. (2007). Animals and Androids: Implicit Associations Between Social Categories and Nonhumans. *Psychological Science*, *18*(2), 116–121. https://doi.org/10.1111/j.1467-9280.2007.01858.x

Luczak, H., Roetting, M., & Schmidt, L. (2003). Let's talk: Anthropomorphization as means to cope with stress of interacting with technical devices. *Ergonomics*, *46*(13–14), 1361–1374. https://doi.org/10.1080/00140130310001610883

Luo, X., Qin, M. S., Zheng, F., & Zhe, Q. (2021). Artificial Intelligence Coaches for Sales Agents: Caveats and Solutions. *Journal of Marketing*, *85*(2), 14–32. https://doi.org/10.1177/0022242920956676

Lyons, J. B., & Guznov, S. Y. (2019). Individual differences in human–machine trust: A multi-study look at the perfect automation schema. *Theoretical Issues in Ergonomics Science*, *20*(4), 440–458. https://doi.org/10.1080/1463922X.2018.1491071

Lyu, S. (2020). Deepfake Detection: Current Challenges and Next Steps. *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 1–6. https://doi.org/10.1109/ICMEW46912.2020.9105991

Maasland, C., & Weißmüller, K. S. (2022). Blame the Machine? Insights From an Experiment on Algorithm Aversion and Blame Avoidance in Computer-Aided Human Resource Management. *Frontiers in Psychology*, *13*(779028). https://doi.org/10.3389/fpsyg.2022.779028

MacDorman, K. F., & Entezari, S. O. (2015). Individual differences predict sensitivity to the uncanny valley. *Interaction Studies*, *16*(2), 141–172. https://doi.org/10.1075/is.16.2.01mac

MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies*, *7*(3), 297–337. https://doi.org/10.1075/is.7.3.03mac

Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., & Cusimano, C. (2015). Sacrifice One For the Good of Many?: People Apply Different Moral Norms to Human and Robot Agents. Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, 117–124. https://doi.org/10.1145/2696454.2696458

Malle, B. F. (2019). How Many Dimensions of Mind Perception Really Are There? In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), *41st Annual Meeting of the Cognitive Science Society* (pp. 2268–2274). Cognitive Science Society.

https://research.clps.brown.edu/SocCogSci/Publications/Pubs/Malle_2019_How_Many_Dimensions.pdf

Maner, J. K., DeWall, C. N., Baumeister, R. F., & Schaller, M. (2007). Does social exclusion motivate interpersonal reconnection? Resolving the "porcupine problem." *Journal of Personality and Social Psychology*, *92*(1), 42–55. https://doi.org/10.1037/0022-3514.92.1.42

Maninger, T., & Shank, D. B. (2022). Perceptions of violations by artificial and human actors across moral foundations. *Computers in Human Behavior Reports*, *5*, 100154. https://doi.org/10.1016/j.chbr.2021.100154

Mann, J. A., MacDonald, B. A., Kuo, I.-H., Li, X., & Broadbent, E. (2015). People respond better to robots than computer tablets delivering healthcare instructions. *Computers in Human Behavior*, *43*, 112–117. https://doi.org/10.1016/j.chb.2014.10.029

Manyika, J., Lund, S., Chui, M., Bughin, J., Woetzel, J., Batra, P., Ko, R., & Sanghvi, S. (2017). *Jobs lost, jobs gained: What the future of work will mean for jobs, skills, and wages*. McKinsey Global Institute. https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages

Manzey, D., Reichenbach, J., & Onnasch, L. (2012). Human Performance Consequences of Automated Decision Aids: The Impact of Degree of Automation and System Experience. *Journal of Cognitive Engineering and Decision Making*, *6*(1), 57–87. https://doi.org/10.1177/1555343411433844

Martens, C., & Cardona-Rivera, R. E. (2016). Generating Abstract Comics. In F. Nack & A. S. Gordon (Eds.), *Interactive Storytelling* (Vol. 10045, pp. 168–175). Springer. https://doi.org/10.1007/978-3-319-48279-8_15

Matsui, T., & Yamada, S. (2019). Designing Trustworthy Product Recommendation Virtual Agents Operating Positive Emotion and Having Copious Amount of Knowledge. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00675

McClure, P. K. (2018). "You're Fired," Says the Robot: The Rise of Automation in the Workplace, Technophobes, and Fears of Unemployment. *Social Science Computer Review*, *36*(2), 139–156. https://doi.org/10.1177/0894439317698637

McKee, K. R., Bai, X., & Fiske, S. T. (2022). Warmth and Competence in Human-Agent Cooperation. *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, 898–907.

McNee, S. M., Riedl, J., & Konstan, J. A. (2006). Being accurate is not enough: How accuracy metrics have hurt recommender systems. *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, 1097–1101. https://doi.org/10.1145/1125451.1125659

Mehrabian, A. (1967). Attitudes inferred from non-immediacy of verbal communications. *Journal of Verbal Learning and Verbal Behavior*, *6*(2), 294–295. https://doi.org/10.1016/S0022-5371(67)80113-0

Melson, G., Beck, A., & Friedman, B. (2009). Robotic Pets in Human Lives: Implications for the Human-Animal Bond and for Human Relationships with Personified Technologies. *Journal of Social Issues - J SOC ISSUES, 65,* 545–567. https://doi.org/10.1111/j.1540-4560.2009.01613.x

Mendes, W. B., Blascovich, J., Hunter, S. B., Lickel, B., & Jost, J. T. (2007). Threatened by the unexpected: Physiological responses during social interactions with expectancy-violating partners. *Journal of Personality and Social Psychology*, *92*(4), 698–716. https://doi.org/10.1037/0022-3514.92.4.698

Michaels, J. L., Parkin, S. S., & Vallacher, R. R. (2013). Destiny Is in the Details: Action Identification in the Construction and Destruction of Meaning. In J. A. Hicks & C. Routledge (Eds.), *The Experience of Meaning in Life* (pp. 103–115). Springer. https://doi.org/10.1007/978-94-007-6527-6_8

Michalowski, M. P., Sabanovic, S., & Simmons, R. (2006). A spatial model of engagement for a social robot. *9th IEEE International Workshop on Advanced Motion Control, 2006.*, 762–767. https://doi.org/10.1109/AMC.2006.1631755

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2). https://doi.org/10.1177/2053951716679679

Monroe, A. E., Dillon, K. D., Guglielmo, S., & Baumeister, R. F. (2018). It's not what you do, but what everyone else does: On the role of descriptive norms and subjectivism in moral judgment. *Journal of Experimental Social Psychology*, *77*, 1–10. https://doi.org/10.1016/j.jesp.2018.03.010

Moon, Y. (1998). When the Computer Is the "Salesperson": Consumer Responses to Computer "Personalities" in Interactive Marketing Situations. *Harvard Business School Working Paper*, *No. 99-041*. https://www.hbs.edu/faculty/Pages/item.aspx?num=12985

Moon, Y. (2000). Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of Consumer Research*, *26*(4), 323–339. https://doi.org/10.1086/209566

Moon, Y., & Nass, C. (1996). How "Real" are computer personalities? Psychological responses to personality types in human-computer interaction. *Communication Research*, *23*(6), 651–674. https://doi.org/10.1177/009365096023006002

Moon, Y., & Nass, C. (1998). Are computers scapegoats? Attributions of responsibility in human-computer interaction. *International Journal of Human-Computer Studies*, *49*(1), 79–94. https://doi.org/10.1006/ijhc.1998.0199

Morewedge, C. K. (2022). Preference for human, not algorithm aversion. *Trends in Cognitive Sciences*, *26*(10), 824–826. https://doi.org/10.1016/j.tics.2022.07.007

Morewedge, C. K., Preston, J., & Wegner, D. M. (2007). Timescale bias in the attribution of mind. *Journal of Personality and Social Psychology*, *93*(1), 1–11. https://doi.org/10.1037/0022-3514.93.1.1

Mori, M. (1970). The uncanney valley. *Energy*, *7*(4), 33–35.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine, 19*(2), 98–100. https://doi.org/10.1109/MRA.2012.2192811

Morikawa, M. (2017). *Who Are Afraid of Losing Their Jobs to Artificial Intelligence and Robots? Evidence from a Survey* (Working Paper No. 71; GLO Discussion Paper). Global Labor Organization (GLO). https://www.econstor.eu/handle/10419/158005

Müller, C. A., Schmitt, K., Barber, A. L. A., & Huber, L. (2015). Dogs Can Discriminate

    Emotional Expressions of Human Faces. *Current Biology*, *25*(5), 601–605.

    https://doi.org/10.1016/j.cub.2014.12.055

Mumm, J., & Mutlu, B. (2011). Designing motivational agents: The role of praise, social

    comparison, and embodiment in computer feedback. *Computers in Human Behavior*, *27*(5),

    1643–1650. https://doi.org/10.1016/j.chb.2011.02.002

Mumm, J., & Mutlu, B. (2011, March). Human-robot proxemics: physical and psychological

    distancing in human-robot interaction. In *Proceedings of the 6th international conference*

    on Human-robot interaction (pp. 331-338).

Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., & Hagita, N. (2009). Footing in human-robot

    conversations: How robots might shape participant roles using gaze cues. In *Proceedings of*

    the 4th ACM/IEEE International Conference on Human Robot Interaction, HRI 2009, La

    Jolla, California, USA, March 9-13, 2009, 61–68. https://doi.org/10.1145/1514095.1514109

Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? *International Journal*

    of Human-Computer Studies, *45*(6), 669–678. https://doi.org/10.1006/ijhc.1996.0073

Nass, C., Isbister, K., & Lee, E.-J. (2000). Truth is beauty: Researching embodied conversational

    agents. In J. Cassell, J. Sullivan, S. Prevost, & E. Churchill (Eds.), *Embodied*

    Conversational Agents (pp. 374–402). The MIT Press.

    https://www.semanticscholar.org/paper/Truth-is-beauty%3A-researching-embodied-agents-

    Nass-Isbister/051722a89ea75583cfe58f6d218db373776b4c6e

Nass, C., & Lee, K. M. (2001). Does computer-synthesized speech manifest personality?

    Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal*

of Experimental Psychology: Applied, *7*(3), 171–181. https://doi.org/10.1037/1076-898X.7.3.171

Nass, C., & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues*, *56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, D. C. (1995). Can computer personalities be human personalities? *International Journal of Human-Computer Studies*, *43*(2), 223–239. https://doi.org/10.1006/ijhc.1995.1042

Nass, C., Moon, Y., & Green, N. (1997). Are Machines Gender Neutral? Gender-Stereotypic Responses to Computers With Voices. *Journal of Applied Social Psychology*, *27*(10), 864–876. https://doi.org/10.1111/j.1559-1816.1997.tb00275.x

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *CHI '94: Conference Companion on Human Factors in Computing Systems*, 204. https://doi.org/10.1145/259963.260288

Natarajan, M., & Gombolay, M. (2020). Effects of anthropomorphism and accountability on trust in human robot interaction. *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI),* 33–42.

Naughton, J. (2022, August 20). AI-generated art illustrates another problem with technology. *The Observer*. https://www.theguardian.com/commentisfree/2022/aug/20/ai-art-artificial-intelligence-midjourney-dall-e-replacing-artists

Nemo, L. (2018, May 30). *Las Vegas Food Service Workers Are Going on Strike So They Don't Lose Their Jobs to Robots*. Futurism. https://futurism.com/las-vegas-food-service-workers-strike-automation

Newman, D. T., Fast, N. J., & Harmon, D. J. (2020). When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *Organizational Behavior and Human Decision Processes*, *160*, 149–167. https://doi.org/10.1016/j.obhdp.2020.03.008

Ng, W. K. (2021, November 12). Robo-teacher takes learning to new level at Hougang Primary. *The Straits Times*. https://www.straitstimes.com/singapore/parenting-education/robot-gives-teachers-and-pupils-a-hand-at-hougang-primary

Nichols, S., & Stich, S. P. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds*. Clarendon Press/Oxford University Press. https://doi.org/10.1093/0198236107.001.0001

Noothigattu, R., Bouneffouf, D., Mattei, N., Chandra, R., Madan, P., Varshney, K. R., Campbell, M., Singh, M., & Rossi, F. (2019). Teaching AI agents ethical values using reinforcement learning and policy orchestration. *IBM Journal of Research and Development*, *63*(4/5), 2:1-2:9. https://doi.org/10.1147/JRD.2019.2940428

Nussberger, A.-M., Luo, L., Celis, L. E., & Crockett, M. J. (2022). Public attitudes value interpretability but prioritize accuracy in Artificial Intelligence. *Nature Communications*, *13*, Article 1. https://doi.org/10.1038/s41467-022-33417-3

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Önkal, D., Goodwin, P., Thomson, M., Gönül, S., & Pollock, A. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making*, *22*(4), 390–409. https://doi.org/10.1002/bdm.637

Paetzel, M., Perugia, G., & Castellano, G. (2020). The persistence of first impressions: 15th Annual ACM/IEEE International Conference on Human Robot Interaction, HRI 2020. *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 73–82. https://ieeexplore.ieee.org/document/9484289

Petersen, S. (2020). Machines Learning Values. In S. M. Liao (Ed.), *Ethics of Artificial Intelligence* (pp. 413–436). Oxford University Press. https://doi.org/10.1093/oso/9780190905033.003.0015

Petro, G. (2020, January 10). *Robots Take Retail*. Forbes. https://www.forbes.com/sites/gregpetro/2020/01/10/robots-take-retail/

Pfeifer, R., & Scheier, C. (2001). *Understanding Intelligence*. The MIT Press.

Pickett, C. L., Gardner, W. L., & Knowles, M. (2004). Getting a Cue: The Need to Belong and Enhanced Sensitivity to Social Cues. *Personality and Social Psychology Bulletin*, *30*(9), 1095–1107. https://doi.org/10.1177/0146167203262085

Pieters, W. (2011). Explanation and trust: What to tell the user in security and AI? *Ethics and Information Technology*, *13*, 53–64. https://doi.org/10.1007/s10676-010-9253-3

Promberger, M., & Baron, J. (2006). Do patients trust computers? *Journal of Behavioral Decision Making*, *19*(5), 455–468. https://doi.org/10.1002/bdm.542

Qin, X., Chen, C., Yam, K. C., Cao, L., Li, W., Guan, J., Zhao, P., Dong, X., & Lin, Y. (2022). Adults still can't resist: A social robot can induce normative conformity. *Computers in Human Behavior*, *127*, Article 107041. https://doi.org/10.1016/j.chb.2021.107041

Qiu, L., & Benbasat, I. (2009). Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *Journal of Management Information Systems*, *25*(4), 145–181. https://doi.org/10.2753/MIS0742-1222250405

Ragot, M., Martin, N., & Cojean, S. (2020). AI-generated vs. Human Artworks. A Perception Bias Towards Artificial Intelligence? *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–10. https://doi.org/10.1145/3334480.3382892

Rahman, H. A. (2021). The Invisible Cage: Workers' Reactivity to Opaque Algorithmic Evaluations. *Administrative Science Quarterly*, *66*(4), 945–988. https://doi.org/10.1177/00018392211010118

Raveendhran, R., & Fast, N. J. (2021). Humans judge, algorithms nudge: The psychology of behavior tracking acceptance. *Organizational Behavior and Human Decision Processes*, *164*, 11–26. https://doi.org/10.1016/j.obhdp.2021.01.001

Reeder, K., & Lee, H. (2022). Impact of artificial intelligence on US medical students' choice of radiology. *Clinical Imaging*, *81*, 67–71. https://doi.org/10.1016/j.clinimag.2021.09.018

Reeves, B., Hancock, J., & Liu, X. "Sunny." (2020). Social Robots Are Like Real People: First

    Impressions, Attributes, and Stereotyping of Social Robots. *Technology, Mind, and*

    Behavior, *1*(1), 1–12. https://doi.org/10.1037/tmb0000018

Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television,*

    and New Media like Real People and Places. Cambridge University Press.

Rehm, M., & Krogsager, A. (2013). Negative affect in human robot interaction - impoliteness in

    unexpected encounters with robots. *2013 IEEE RO-MAN,* 45–50.

    https://doi.org/10.1109/ROMAN.2013.6628529

Reich, T., Kaju, A., & Maglio, S. J. (2022). How to overcome algorithm aversion: Learning from

    mistakes. *Journal of Consumer Psychology*. https://doi.org/10.1002/jcpy.1313

Reiss, D., & Marino, L. (2001). Mirror self-recognition in the bottlenose dolphin: A case of

    cognitive convergence. *Psychological and Cognitive Sciences*, *98*(10), 5937–5942.

    https://doi.org/10.1073/pnas.101086398

Richardson, K. (2016). Sex Robot Matters: Slavery, the Prostituted, and the Rights of Machines.

    IEEE Technology and Society Magazine, *35*(2), 46–53.

    https://doi.org/10.1109/MTS.2016.2554421

Riek, L. D., Rabinowitch, T.-C., Chakrabarti, B., & Robinson, P. (2009). How

    anthropomorphism affects empathy toward robots. *Proceedings of the 4th ACM/IEEE*

    International Conference on Human Robot Interaction, 245–246.

    https://doi.org/10.1145/1514095.1514158

Riether, N., Hegel, F., Wrede, B., & Horstmann, G. (2012). Social facilitation with social robots? *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, 41–48. https://doi.org/10.1145/2157689.2157697

Robinette, P., Howard, A. M., & Wagner, A. R. (2017). Effect of Robot Performance on Human–Robot Trust in Time-Critical Situations. *IEEE Transactions on Human-Machine Systems*, *47*(4), 425–436. https://doi.org/10.1109/THMS.2017.2648849

Robitzski, D. (2019, August 22). *Truckers want to ban self-driving trucks in Missouri*. The Byte. https://futurism.com/the-byte/truckers-ban-self-driving-trucks-missouri

Rockwell, G., Berendt, B., & Chee, F. (2022). On IRIE Vol. 31: On Dialogue and Artificial Intelligence. *The International Review of Information Ethics*, *31*(1), Article 1. https://doi.org/10.29173/irie475

Rosenthal-von der Pütten, A. M., Krämer, N. C., & Herrmann, J. (2018). The effects of humanlike and robot-specific affective nonverbal behavior on perception, emotion, and behavior. *International Journal of Social Robotics, 10*, 569–582. https://doi.org/10.1007/s12369-018-0466-7

Saerbeck, M., & Bartneck, C. (2010). Perception of affect elicited by robot motion. In *Proceedings of the 5th ACM/IEEE International Conference on Human Robot Interaction, HRI 2010, Osaka, Japan,* 53–60. https://doi.org/10.1109/HRI.2010.5453269

Salomons, N., van der Linden, M., Sebo, S. S., & Scassellati, B. (2018). Humans Conform to Robots: Disambiguating Trust, Truth, and Conformity. *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 187–195. https://doi.org/10.1145/3171221.3171282

Salvatore, A. P. (2019). Behaviorism. In *The SAGE Encyclopedia of Human Communication*

    Sciences and Disorders. SAGE Publications, Inc.

    https://dx.doi.org/10.4135/9781483380810

Salvini, P., Ciaravella, G., Yu, W., Ferri, G., Manzi, A., Mazzolai, B., Laschi, C., Oh, S. R., &

    Dario, P. (2010). How safe are service robots in urban environments? Bullying a robot.

    In *19th international symposium in robot and human interactive communication (pp. 1-7).*

    IEEE. https://doi.org/10.1109/ROMAN.2010.5654677

Samuel, S. (2020, January 13). *Robot priests can bless you, advise you, and even perform your*

    funeral. Vox. https://www.vox.com/future-perfect/2019/9/9/20851753/ai-religion-robot-

    priest-mindar-buddhism-christianity

Sandoval, E. B., Brandstetter, J., Obaid, M., & Bartneck, C. (2016). Reciprocity in human-robot

    interaction: a quantitative approach through the prisoner's dilemma and the ultimatum

    game. *International Journal of Social Robotics, 8*(2), 303–317.

    https://doi.org/10.1007/s12369-015-0323-x

Saygin, A. P., Chaminade, T., Ishiguro, H., Driver, J., & Frith, C. (2012). The thing that should

    not be: predictive coding and the uncanny valley in perceiving human and humanoid robot

    actions. *Social Cognitive and Affective Neuroscience*, *7*(4), 413–422.

    https://doi.org/10.1093/scan/nsr025

Scassellati, B., Boccanfuso, L., Huang, C.-M., Mademtzi, M., Qin, M., Salomons, N., Ventola,

    P., & Shic, F. (2018). Improving social skills in children with ASD using a long-term, in-

    home social robot. *Science Robotics,* 3, eaat7544.

    https://doi.org/10.1126/scirobotics.aat7544

Sebo, S. S., Traeger, M., Jung, M., & Scassellati, B. (2018). The Ripple Effects of Vulnerability: The Effects of a Robot's Vulnerable Behavior on Trust in Human-Robot Teams. In *2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI),* 178–186.

Schein, C., & Gray, K. (2018). The Theory of Dyadic Morality: Reinventing Moral Judgment by Redefining Harm. *Personality and Social Psychology Review*, *22*(1), 32–70. https://doi.org/10.1177/1088868317698288

Scheske, C., & Schnall, S. (2012). The ethics of "smart drugs": Moral judgments about healthy people's use of cognitive-enhancing drugs. *Basic and Applied Social Psychology, 34,* 508–515. https://doi.org/10.1080/01973533.2012.711692

Seyama, J., & Nagayama, R. S. (2007). The uncanny valley: effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments, 16*(4), 337–351. https://doi.org/10.1162/pres.16.4.337

Shaffer, V. A., Probst, C. A., Merkle, E. C., Arkes, H. R., & Medow, M. A. (2013). Why Do Patients Derogate Physicians Who Use a Computer-Based Diagnostic Support System? *Medical Decision Making*, *33*(1), 108–118. https://doi.org/10.1177/0272989X12453501

Shank, D. B., & DeSanti, A. (2018). Attributions of morality and mind to artificial intelligence after real-world moral violations. *Computers in Human Behavior*, *86*, 401–411. https://doi.org/10.1016/j.chb.2018.05.014

Shank, D. B., DeSanti, A., & Maninger, T. (2019). When are artificial intelligence versus human agents faulted for wrongdoing? Moral attributions after individual and joint decisions. *Information, Communication & Society*, *22*(5), 648–663. https://doi.org/10.1080/1369118X.2019.1568515

Shank, D. B., Graves, C., Gott, A., Gamez, P., & Rodriguez, S. (2019). Feeling our way to

machine minds: People's emotions when perceiving mind in artificial intelligence.

Computers in Human Behavior, *98*, 256–266. https://doi.org/10.1016/j.chb.2019.04.001

Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2017). Psychological roadblocks to the adoption of

self-driving vehicles. *Nature Human Behaviour*, *1*(10), Article 10.

https://doi.org/10.1038/s41562-017-0202-6

Sharp, M.-L., Fear, N. T., Rona, R. J., Wessely, S., Greenberg, N., Jones, N., & Goodwin, L.

(2015). Stigma as a Barrier to Seeking Health Care Among Military Personnel With Mental

Health Problems. *Epidemiologic Reviews*, *37*(1), 144–162.

https://doi.org/10.1093/epirev/mxu012

Shen, S., Slovak, P., & Jung, M. F. (2018). "Stop. I See a Conflict Happening." A Robot

Mediator for Young Children's Interpersonal Conflict Resolution. In *Proceedings of the*

2018 ACM/IEEE International Conference on Human-Robot Interaction, 69–77.

Shiban, Y., Schelhorn, I., Jobst, V., Hörnlein, A., Puppe, F., Pauli, P., & Mühlberger, A. (2015).

The appearance effect: Influences of virtual agent features on performance and motivation.

Computers in Human Behavior, *49*, 5–11. https://doi.org/10.1016/j.chb.2015.01.077

Shim, J., & Arkin, R. C. (2014). Other-oriented robot deception: A computational approach for

deceptive action generation to benefit the mark. *2014 IEEE International Conference on*

Robotics and Biomimetics (ROBIO 2014), 528–535.

https://doi.org/10.1109/ROBIO.2014.7090385

Shin, D. (2021). Embodying algorithms, enactive artificial intelligence and the extended cognition: You can see as much as you know about algorithm. *Journal of Information Science*, *49*(1), 18–31. https://doi.org/10.1177/0165551520985495

Shin, H. I., & Kim, J. (2020). My computer is more thoughtful than you: Loneliness, anthropomorphism and dehumanization. *Current Psychology*, *39*, 445–453. https://doi.org/10.1007/s12144-018-9975-7

Shinozawa, K., Naya, F., Yamato, J., & Kogure, K. (2005). Differences in effect of robot and screen agent recommendations on human decision-making. *International Journal of Human-Computer Studies*, *62*(2), 267–279. https://doi.org/10.1016/j.ijhcs.2004.11.003

Shrestha, Y. R., He, V. F., Puranam, P., & von Krogh, G. (2021). Algorithm Supported Induction for Building Theory: How Can We Use Prediction Models to Theorize? *Organization Science*, *32*(3), 856–880. https://doi.org/10.1287/orsc.2020.1382

Skantze, G. (2017, March). Predicting and regulating participation equality in human-robot conversations: Effects of age and gender. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-robot Interaction* (pp. 196-204).

Smiley, A. H., & Fisher, M. (2022). The golden age is behind us: how the status quo impacts the evaluation of technology. *Psychological Science, 33*(9), 1605–1614. https://doi.org/10.1177/09567976221102868

Smithers, D. (2022, January 1). *Man Falls In Love With Robot And Hopes To Marry Her*. LADbible. https://www.ladbible.com/news/man-falls-in-love-with-robot-and-hopes-to-marry-her-20220101

Soll, J. B., & Mannes, A. E. (2011). Judgmental aggregation strategies depend on whether the self is involved. *International Journal of Forecasting*, *27*(1), 81–102. https://doi.org/10.1016/j.ijforecast.2010.05.003

Sousa, J. P. (1906). The Menace of Mechanical Music. *Appleton's Magazine*, *8*(3), 278–284.

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science,* 333, 476–478. https://doi.org/10.1126/science.1207745

Spring, V. L., Cameron, C. D., & Cikara, M. (2018). The upside of outrage. *Trends in Cognitive Sciences*, *22*(2), 1067–1069. https://doi.org/10.1016/j.tics.2018.09.006

Steckenfinger, S. A., & Ghazanfar, A. A. (2009). Monkey visual behavior falls into the uncanny valley. *Proceedings of the National Academy of Sciences*, *106*(43), 18362–18366. https://doi.org/10.1073/pnas.0910063106

Stein, J.-P., & Ohler, P. (2017). Venturing into the uncanny valley of mind—the influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition, 160*, 43–50. https://doi.org/10.1016/j.cognition.2016.12.010

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 23–45). Lawrence Erlbaum Associates Publishers.

Strait, M., Vujovic, L., Floerke, V., Scheutz, M., & Urry, H. (2015). Too Much Humanness for Human-Robot Interaction: Exposure to Highly Humanlike Robots Elicits Aversive

Responding in Observers. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3593–3602. https://doi.org/10.1145/2702123.2702415

Strinić, A., Carlsson, M., & Agerström, J. (2021). Occupational stereotypes: Professionals´ warmth and competence perceptions of occupations. *Personnel Review*, *51*(2), 603–619. https://doi.org/10.1108/PR-06-2020-0458

Strohkorb, S., Fukuto, E., Warren, N., Taylor, C., Berry, B., & Scassellati, B. (2016). Improving human-human collaboration between children with a social robot. *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 551–556. https://doi.org/10.1109/ROMAN.2016.7745172

Szczuka, J. M., & Krämer, N. C. (2018). Jealousy 4.0? An empirical study on jealousy-related discomfort of women evoked by other women and gynoid robots. *Paladyn, Journal of Behavioral Robotics*, *9*, 323–336. https://doi.org/10.1515/pjbr-2018-0023

Tajfel, H. (1982). Social Psychology of Intergroup Relations. *Annual Review of Psychology*, *33*, 1–39. https://doi.org/10.1146/annurev.ps.33.020182.000245

Tajfel, H., & Turner, J. (2001). An integrative theory of intergroup conflict. In M. A. Hogg & D. Abrams (Eds.), *Intergroup relations: Essential readings* (pp. 94–109). Psychology Press.

Tanaka, F., Cicourel, A., & Movellan, J. R. (2007). Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences*, *104*(46), 17954–17958. https://doi.org/10.1073/pnas.0707769104

Tang, P. M., Koopman, J., Yam, K. C., De Cremer, D., Zhang, J. H., & Reynders, P. (2022). The self-regulatory consequences of dependence on intelligent machines at work: Evidence from

field and experimental studies. *Human Resource Management*, 1–24.

https://doi.org/10.1002/hrm.22154

Taylor, A. H., Elliffe, D., Hunt, G. R., & Gray, R. D. (2010). Complex cognition and behavioural

innovation in New Caledonian crows. *Proceedings of the Royal Society B: Biological

Sciences*, *277*(1694), 2637–2643. https://doi.org/10.1098/rspb.2010.0285

Tesser, A., Millar, M., & Moore, J. (1988). Some affective consequences of social comparison

and reflection processes: The pain and pleasure of being close. *Journal of Personality and

Social Psychology*, *54*(1), 49–61. https://doi.org/10.1037/0022-3514.54.1.49

Tharp, M., Holtzman, N. S., & Eadeh, F. R. (2017). Mind Perception and Individual Differences:

A Replication and Extension. *Basic and Applied Social Psychology*, *39*(1), 68–73.

https://doi.org/10.1080/01973533.2016.1256287

The Guardian. (2020, July 20). 'Alexa, I love you': How lockdown made men lust after their

Amazon Echo. *The Guardian*. https://www.theguardian.com/technology/2020/jul/20/alexa-

i-love-you-how-lockdown-made-men-lust-after-their-amazon-echo

Thewissen, S., & Rueda, D. (2019). Automation and the welfare state: Technological change as a

determinant of redistribution preferences. *Comparative Political Studies, 52*(2), 171–208.

10.1177/0010414017740600

Tielman, M., Neerincx, M., Meyer, J.-J., & Looije, R. (2014). Adaptive emotional expression in

robot-child interaction. *Proceedings of the 2014 ACM/IEEE International Conference on

Human-Robot Interaction*, 407–414. https://doi.org/10.1145/2559636.2559663

Tiku, N. (2022). *Is LaMDA Sentient? - An Interview*.

https://www.documentcloud.org/documents/22058315-is-lamda-sentient-an-interview

Tolmeijer, S., Kneer, M., Sarasua, C., Christen, M., & Bernstein, A. (2021). Implementations in

Machine Ethics: A Survey. *ACM Computing Surveys*, *53*(6), 132:1-132:38.

https://doi.org/10.1145/3419633

Traeger, M. L., Strohkorb Sebo, S., Jung, M., Scassellati, B., & Christakis, N. A. (2020).

Vulnerable robots positively shape human conversational dynamics in a human–robot team.

Proceedings of the National Academy of Sciences, *117*(12), 6370–6375.

https://doi.org/10.1073/pnas.1910402117

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, *LIX*(236), 433–460.

https://doi.org/10.1093/mind/LIX.236.433

Twenge, J. M., Catanese, K. R., & Baumeister, R. F. (2003). Social Exclusion and the

Deconstructed State: Time Perception, Meaninglessness, Lethargy, Lack of Emotion, and

Self-Awareness. *Journal of Personality and Social Psychology*, *85*(3), 409–423.

https://doi.org/10.1037/0022-3514.85.3.409

Ward, A. F. (2021). People mistake the internet's knowledge for their own. *Proceedings of the*

National Academy of Sciences, 118(43), e2105061118.

https://doi.org/10.1073/pnas.2105061118

Waytz, A., & Young, L. (2019). Aversion to playing God and moral condemnation of technology

and science. *Philosophical Transactions of the Royal Society B: Biological Sciences,*

374(1771), 20180041. https://doi.org/10.1098/rstb.2018.0041

Wike, R., & Stokes, B. (2018). In advanced and emerging economies alike, worries about job automation. Pew Research Center, Global Attitudes & Trends.

Wow GPT-3 just wrote my American -Lit Essay for me. (2020, November 15). [Reddit]. R/GPT3. www.reddit.com/r/GPT3/comments/jutsdh/wow_gpt3_just_wrote_my_american_lit_essay_for_me/

Uttal, W. R. (2001). *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. The MIT Press.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, *94*(1), 3–15. https://doi.org/10.1037/0033-295X.94.1.3

Van Bavel, J. J., Hackel, L. M., & Xiao, Y. J. (2014). The Group Mind: The Pervasive Influence of Social Identity on Cognition. In J. Decety & Y. Christen (Eds.), *New Frontiers in Social Neuroscience* (Vol. 21, pp. 41–56). Springer International Publishing. https://doi.org/10.1007/978-3-319-02904-7_4

Vanman, E. J., & Kappas, A. (2019). "Danger, Will Robinson!" The challenges of social robots for intergroup relations. *Social and Personality Psychology Compass*, *13*(8), 1–13. https://doi.org/10.1111/spc3.12489

Vázquez, M., Carter, E. J., McDorman, B., Forlizzi, J., Steinfeld, A., & Hudson, S. E. (2017). Towards Robot Autonomy in Group Conversations: Understanding the Effects of Body Orientation and Gaze. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 42–52. https://doi.org/10.1145/2909824.3020207

Verhagen, T., van Nes, J., Feldberg, F., & van Dolen, W. (2014). Virtual Customer Service

    Agents: Using Social Presence and Personalization to Shape Online Service Encounters*.

    Journal of Computer-Mediated Communication, *19*(3), 529–545.

    https://doi.org/10.1111/jcc4.12066

Vincent, J. (2022, September 2). *An AI-generated artwork's state fair victory fuels arguments*

    over 'what art is.' The Verge. https://www.theverge.com/2022/9/1/23332684/ai-generated-

    artwork-wins-state-fair-competition-colorado

Vollmer, A.-L., Read, R., Trippas, D., & Belpaeme, T. (2018). Children conform, adults resist: A

    robot group induced peer pressure on normative social conformity. *Science Robotics*, *3*(21).

    https://doi.org/10.1126/scirobotics.aat7111

Waardenburg, L., Huysman, M., & Sergeeva, A. V. (2022). In the Land of the Blind, the One-

    Eyed Man Is King: Knowledge Brokerage in the Age of Learning Algorithms. *Organization*

    Science, *33*(1), 59–82. https://doi.org/10.1287/orsc.2021.1544

Walmsley, J. (2021). Artificial intelligence and the value of transparency. *AI & SOCIETY*, *36*,

    585–595. https://doi.org/10.1007/s00146-020-01066-z

Wang, S., Lilienfeld, S. O., & Rochat, P. (2015). The Uncanny Valley: Existence and

    Explanations. *Review of General Psychology*, *19*(4), 393–407.

    https://doi.org/10.1037/gpr0000056

Wang, X., & Krumhuber, E. G. (2018). Mind perception of robots varies with their economic

    versus social function. *Frontiers in Psychology,* 9, 1230.

    https://doi.org/10.3389/fpsyg.2018.01230

Waytz, A., Cacioppo, J., & Epley, N. (2010). Who Sees Human?: The Stability and Importance of Individual Differences in Anthropomorphism. *Perspectives on Psychological Science*, *5*(3), 219–232. https://doi.org/10.1177/1745691610369336

Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social Cognition Unbound: Insights Into Anthropomorphism and Dehumanization. *Current Directions in Psychological Science*, *19*(1), 58–62. https://doi.org/10.1177/0963721409359302

Waytz, A., & Gray, K. (2018). Does Online Technology Make Us More or Less Sociable? A Preliminary Review and Call for Research. *Perspectives on Psychological Science*, *13*(4), 473–491. https://doi.org/10.1177/1745691617746509

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, *14*(8), 383–388. https://doi.org/10.1016/j.tics.2010.05.006

Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, *52*, 113–117. https://doi.org/10.1016/j.jesp.2014.01.005

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*(3), 410–435. https://doi.org/10.1037/a0020240

Waytz, A., & Norton, M. I. (2014). Botsourcing and outsourcing: Robot, British, Chinese, and German workers are for thinking—not feeling—jobs. *Emotion*, *14*(2), 434–444. https://doi.org/10.1037/a0036054

Wegner, D. M., & Gray, K. (2016). *The Mind Club: Who Thinks, What Feels, and Why It Matters*. Viking.

Weisman, K., Dweck, C. S., & Markman, E. M. (2017). Rethinking people's conceptions of mental life. *Proceedings of the National Academy of Sciences*, *114*(43), 11374–11379. https://doi.org/10.1073/pnas.1704347114

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45. https://doi.org/10.1145/365153.365168

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*(5), 297–333. https://doi.org/10.1037/h0040934

Wiese, E., Weis, P. P., Bigman, Y., Kapsaskis, K., & Gray, K. (2022). It's a Match: Task Assignment in Human–Robot Collaboration Depends on Mind Perception. *International Journal of Social Robotics*, *14*, 141–148. https://doi.org/10.1007/s12369-021-00771-z

Wike, R., & Stokes, B. (2018, September 13). In Advanced and Emerging Economies Alike, Worries About Job Automation. *Pew Research Center's Global Attitudes Project*. https://www.pewresearch.org/global/2018/09/13/in-advanced-and-emerging-economies-alike-worries-about-job-automation/

Wilbanks, D., Hester, N., Bigman, Y., Smith, M. M., Court, J., Sarkar, J., & Gray, K. (under review). *Two kinds of artistic authenticity: An original object versus truly revealing the mind of the artist.*

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management*, *29*(5), 907–931. https://doi.org/10.1108/JOSM-04-2018-0119

Wu, N. (2022a). Misattributed blame? Attitudes toward globalization in the age of automation. *Political Science Research and Methods*, *10*(3), 470–487. https://doi.org/10.1017/psrm.2021.43

Wu, N. (2022b). "Restrict foreigners, not robots": Partisan responses to automation threat. *Economics & Politics*, 1–24. https://doi.org/10.1111/ecpo.12225

Xiao, B., & Benbasat, I. (2007). E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact. *MIS Quarterly*, *31*(1), 137–209. https://doi.org/10.2307/25148784

Yakovleva, M., Reilly, R. R., & Werko, R. (2010). Why do we trust? Moving beyond individual to dyadic perceptions. *Journal of Applied Psychology*, *95*(1), 79–91. https://doi.org/10.1037/a0017102

Yam, K. C., Bigman, Y. E., Tang, P. M., Ilies, R., De Cremer, D., Soh, H., & Gray, K. (2021). Robots at work: People prefer—and forgive—service robots with perceived feelings. *Journal of Applied Psychology*, *106*(10), 1557–1572. https://doi.org/10.1037/apl0000834

Yam, K. C., Bigman, Y., & Gray, K. (2021). Reducing the uncanny valley by dehumanizing humanoid robots. *Computers in Human Behavior*, *125*(C). https://doi.org/10.1016/j.chb.2021.106945

Yam, K. C., Goh, E.-Y., Fehr, R., Lee, R., Soh, H., & Gray, K. (2022). When your boss is a robot: Workers are more spiteful to robot supervisors that seem more human. *Journal of Experimental Social Psychology*, *102*, 104360. https://doi.org/10.1016/j.jesp.2022.104360

Yam, K. C., Tang, P. M., Jackson, J. C., Su, R., & Gray, K. (2022). The Rise of Robots Increases Job Insecurity and Maladaptive Workplace Behaviors: Multimethod Evidence. *Journal of Applied Psychology*. https://doi.org/10.1037/apl0001045

Yang, A. X., & Teow, J. (2020). Defending the Human Need to Be Seen: Recipient Identifiability Aggravates Algorithm Aversion in Resource Allocation Decisions. In J. Argo, T. M. Lowrey, & H. J. Schau (Eds.), *ACR North American Advances in Consumer Research: Vol. NA-48* (pp. 1165–1169). Association for Consumer Research. https://www.acrwebsite.org/volumes/2662147/volumes/v48/NA-48

Yeomans, M., Shah, A., Mullainathan, S., & Kleinberg, J. (2019). Making sense of recommendations. *Journal of Behavioral Decision Making*, *32*(4), 403–414. https://doi.org/10.1002/bdm.2118

Yogeeswaran, K., Złotowski, J., Livingstone, M., Bartneck, C., Sumioka, H., & Ishiguro, H. (2016). The interactive effects of robot anthropomorphism and robot ability on perceived threat and support for robotics research. *Journal of Human-Robot Interaction*, *5*(2), 29–47. https://doi.org/10.5898/JHRI.5.2.Yogeeswaran

You, S., & Robert, L. Jr. (2019). Subgroup Formation in Human Robot Teams. *Proceedings of the 38th International Conference on Information Systems*. http://deepblue.lib.umich.edu/handle/2027.42/150854

You, S., Yang, C. L., & Li, X. (2022). Algorithmic versus Human Advice: Does Presenting Prediction Performance Matter for Algorithm Appreciation? *Journal of Management Information Systems*, *39*(2), 336–365. https://doi.org/10.1080/07421222.2022.2063553

Zhang, T., Kaber, D. B., Zhu, B., Swangnetr, M., Mosaly, P., & Hodge, L. (2010). Service robot feature design effects on user perceptions and emotional responses. *Intelligent Service Robotics*, *3*, 73–88. https://doi.org/10.1007/s11370-010-0060-9

Zhou, M. X., Mark, G., Li, J., & Yang, H. (2019). Trusting Virtual Agents: The Effect of Personality. *ACM Transactions on Interactive Intelligent Systems*, *9*(2–3), No.10, 1-36. https://doi.org/10.1145/3232077

Złotowski, J., Sumioka, H., Nishio, S., Glas, D. F., Bartneck, C., & Ishiguro, H. (2016). Appearance of a Robot Affects the Impact of its Behaviour on Perceived Trustworthiness and Empathy. *Paladyn, Journal of Behavioral Robotics*, *7*(1), 55–66. https://doi.org/10.1515/pjbr-2016-0005

Złotowski, J., Yogeeswaran, K., & Bartneck, C. (2017). Can we control it? Autonomous robots threaten human identity, uniqueness, safety, and resources. *International Journal of Human-Computer Studies*, *100*, 48–54. https://doi.org/10.1016/j.ijhcs.2016.12.008

Zulić, H. (2019). How AI can Change/Improve/Influence Music Composition, Performance and Education: Three Case Studies. *INSAM Journal of Contemporary Music, Art and Technology*, *1*(2), 100–114.