Predicting the Moral Consideration of Artificial Intelligences

Janet VT Pauketat[1][0000-0003-3280-3345]
Jacy Reese Anthis[1,2][0000-0002-4684-348X]

[1]Sentience Institute, New York, NY 10006, USA
janet@sentienceinstitute.org
[2]University of Chicago, Chicago, IL 60637, USA

Abstract. Understanding the moral consideration of AIs as moral patients is increasingly critical given their rapid integration into daily life and the projected proliferation of advanced AIs. We present the results from a preregistered online survey with 300 U.S. Americans on the psychological predictors of the moral consideration of AIs to develop psychological theory surrounding this phenomenon. We tested an array of psychological predictors inspired by the literature on human-human and human-animal relations: perspective (future orientation, construal level), relational (social dominance orientation, sci-fi fan identity), expansive (human-centric norms, anthropomorphism, global citizenship, openness to experience, techno-animism), technological (affinity for technology, substratism, human-AI overlap, realistic threat, identity threat), and affective (emotions felt towards AIs). The strongest predictors were substratism, sci-fi fan identity, techno-animism, and positive emotions. We also identified three conceptual dimensions of moral consideration with an exploratory factor analysis of eight moral consideration indices drawn from prior literature: mind perception, psychological expansion, and practical consideration. Additionally, the temporal existence of AIs impacted moral consideration: AIs existing in the future were attributed more emotional capacity and more value as feeling entities than were current AIs. These results illustrate nuances in the moral consideration of AIs and lay the foundation for future research.

Keywords: artificial intelligence, robot rights, moral circle, moral consideration, mind perception, substratism

Note: This is the author's preprint version of the manuscript and may not exactly match the version of record.

Please cite as: Pauketat, J.V.T., & Anthis, J.R. (2022). Predicting the moral consideration of artificial intelligences. *Computers in Human Behavior, 136,* 107372. https://doi.org/10.1016/j.chb.2022.107372

1 Introduction

Artificial intelligences, defined as entities such as robots that make their own decisions, think, and act in human-like or rational ways (S. J. Russell & Norvig, 1995; henceforth "AIs") are used in a wide range of applications, including household maintenance, autonomous vehicles, human companionship, personal assistance, entertainment, fact-checking, population surveillance, misinformation detection, and immersive virtual reality. Researchers have predicted that AIs will radically change human existence over the next century (Bostrom, 1998; Gordon, 2020; Matthews et al., 2021; Rahwan et al., 2019; S. Russell, 2019; C. S. Smith, 2021; Talty, 2018; Tegmark, 2017). Data from Stanford's AI Index Report (D. Zhang et al., 2022), the Pew Research Center's studies of AI trends (Anderson et al., 2018; Anderson & Rainie, 2022), and Our World in Data's technology trends (Ritchie & Roser, 2017) echo these predictions. Many researchers are preparing for the needs of various stakeholders in transition to an "Industry 4.0" world with diverse AI support systems (e.g., Contreras-Masse et al., 2020; Ema et al., 2016). Some societies are actively preparing for the social, industrial, and structural integration of AIs in their futures. China's New Generation Artificial Intelligence Development Plan promotes an AI revolution (B. Li et al., 2018; Wu et al., 2020). Japan's Society 5.0 envisions human-AI integration and harmonization (Berberich et al., 2020; Fujii et al., 2018; Nagahara, 2019). These perspectives acknowledge the need to prepare for an AI-integrated future in which people from various generations and with various cultural and social experiences interact routinely with AIs.

Much research has endeavored to understand and improve human-robot interactions (HRI) and human-computer interactions (HCI), often focusing on the interpersonal interactions between humans and AIs (e.g., Abubshait & Wiese, 2017; Admoni & Scassellati, 2017; Belpaeme et al., 2018; Leichtmann & Nitsch, 2020; Nomura et al., 2008; Ren & Bao, 2020). Other research has focused on humans' reactions to AIs, machine behavior, and the moral influence of AIs on humans (e.g., Appel et al., 2020; Köbis et al., 2021; Köbis & Mossink, 2021; Laakasuo et al., 2021; Naneva et al., 2020; Rahwan et al., 2019; Renier et al., 2021; Shank & DeSanti, 2018). The Computers As Social Actors (CASA) paradigm established that humans use similar social principles to interact with machines and humans (Nass et al., 1994; Reeves & Nass, 1996). These social affordances for machines develop with repeated exposure over time (Gambino et al., 2020). Research extending from CASA has suggested that humans interact with AIs using the same principles as in human intergroup interactions (E. R. Smith et al., 2020, 2021; Vanman & Kappas, 2019). E. R. Smith et al. (2020) found that, like in human intergroup relations, positive emotions predicted a greater willingness to interact with robots. We expect similar social psychological principles from human-human interactions to apply to human-AI interactions.

Despite AI integration into human society and this groundbreaking HRI and HCI research, we have little understanding of how humans include AIs in the moral circle (Harris & Anthis, 2021). The moral circle refers to the boundary around moral patients—entities who receive moral consideration (Anthis & Paez, 2021; Crimston et al., 2018; Opotow, 1990). People typically see moral patients as those entities who can experience the world (K. Gray & Wegner, 2009, 2012a). Sentience—the capacity to have positive and negative experiences (Broom, 2020; Harris & Anthis, 2021)—has been proposed as a precondition for inclusion in the moral circle (Anthis & Paez, 2021; Crimston et al., 2016).

Moral circle exclusion and sentience denial can have significant consequences, including harm and the oppression of humans and nonhuman animals (henceforth, "animals"; Caviola et al., 2019; Dhont et al., 2016; Loughnan et al., 2010; Opotow, 1990, 1993; Pratto et al., 2006; Starmans et al., 2017). Much of this research has focused on explaining and reducing the inequities within human societies (e.g., Dovidio et al., 2017; Haslam & Loughnan, 2014; Leslie et al., 2020; Lomborg, 2020; Paluck et al., 2021). Some has elucidated the speciesist tendencies that place humans firmly at the top of a moral hierarchy of natural entities such as plants and animals (Amiot & Bastian, 2015; Caviola et al., 2019; Caviola & Capraro, 2020; Dhont et al., 2016). Others have discussed humans' routine denial or dismissal of animal sentience, despite international laws recognizing it (Blattner, 2019; Chessman, 2018; Proctor et al., 2013). Sentience denial has been connected to animal meat consumption and animal use in labs (Bastian et al., 2012; Blattner, 2019; Powell & Mikhalevich, 2020; Proctor et al., 2013). Moral circle exclusion and sentience denial have yet to be empirically explored in the context of AIs, although Darling (2012, 2021) and Coeckelbergh (2011) pointed to the similarities between humans' legal, social, and industrial treatment of animals and humans' projected treatment of robots including commenting that "animal" and "robot" categories include many diverse types. Based on this framework and like with CASA, we expect similar social psychological principles to extend from human-animal relations to human-AI relations.

People typically deny AIs the capacity to experience or have emotions despite a demonstrated willingness to attribute AIs certain cognitive processes like problem-solving as moral agents who can take moral actions (H. M. Gray et al., 2007; K. Gray et al., 2012; K. Gray & Wegner, 2012b; Shank & DeSanti, 2018; Wang & Krumhuber, 2018). This denial may preclude AIs from moral circle inclusion given that experiential and emotional capacities are closely tied to moral patiency. This matters because many scholars think that AIs are likely to become sentient (Blackmore, 1999; Bringsjord et al., 2015; Buttazzo, 2001; Haikonen, 2012; Metzinger, 2021). A number of ethicists have discussed the plausibility of morally catastrophic scenarios such as enslaving sentient robots (e.g., Althaus & Gloor, 2016; Birhane & van Dijk, 2020; Bryson, 2010; Fröding & Peterson, 2020; Gunkel, 2018; Lin et al., 2012; Shulman & Bostrom, 2021). Although this may not be a significant moral problem for present AIs who have limited, if any sentience, this will grow in importance as AIs become more autonomous and advanced in their agential and experiential capacities.

Research into the moral consideration of AIs has yet to catch up with philosophical inquiry. We respond to recent calls for empirical investigation (Coeckelbergh, 2010, 2018; de Graaf et al., 2021; Gunkel, 2012; Harris & Anthis, 2021; Metzinger, 2021; Owe & Baum, 2021; Rahwan et al., 2019) and contribute in two ways. First, we tested how an array of psychological tendencies relevant to morality and prosociality predicted moral consideration. Second, we examined a range of moral consideration outcomes: mind perception, sentience value, moral expansiveness, moral exclusion/scope of justice, support for AI rights, and donations made to charitable organizations. We selected variables by broadly conceptualizing moral, prosocial, HRI, and HCI research including established AI-relevant variables (see for reviews Chugunova & Sele, 2020; Köbis et al., 2021; Rahwan et al., 2019) like anthropomorphism and expanding to understudied variables like techno-animism. We introduce some initial expectations for the latent structure of moral consideration. Understanding the psychological predictors and dimensions of moral consideration will help us to evaluate how different AIs affect and are affected by human behavior. This knowledge can contribute to AI ethics research as well as human-AI relations research.

2 Psychological Predictors

We examined conceptually similar groups of psychological predictors: perspective, relational, expansive, technological, and affective. Perspective predictors are traits and mindsets related to distant or future outlooks. Relational predictors are traits and identities centered on social contexts. Expansive predictors are traits, identities, beliefs, and norms that extend beyond self- and human-centeredness. Technological predictors are attitudes and beliefs about technology. Affective predictors are emotions felt towards AIs.

2.1 Perspective Predictors

2.1.1 Future Orientation

Future orientation (FO) refers to the tendency to consider and plan for the future (Block, 1990; Kluckhohn & Strodtbeck, 1961; Milfont et al., 2012), which stems from physiological and environmental sources (Carmi & Arnon, 2014; Suddendorf & Corballis, 2007; Toepoel, 2010) and has been linked to individual and societal functioning (Strathman et al., 1994; Venaik et al., 2013; Zimbardo & Boyd, 1999). FO has been linked with pro-environmental attitudes and behaviors, delay of gratification, goal-setting, healthy behaviors, and conscientiousness (Carmi & Arnon, 2014; Milfont et al., 2012; Zimbardo & Boyd, 1999). We expected a stronger FO to predict more moral consideration of AIs, particularly future AIs, given these associations and conceptions of AI as a future technology (Torresen, 2018).

2.1.2 Construal Level

Construal level theory suggests that people think at varying degrees of distance to themselves in terms of spatial proximity, time, social closeness, and certainty (Trope & Liberman, 2010). "Abstract construal" pertains to distant and decontextualized mental representations that are broad, invariant, and inclusive. "Concrete construal" pertains to close and contextualized representations that are specific, context-dependent, and highly-detailed (Hess et al., 2018; Keeney & Wiesenfeld, 2017; Ledgerwood et al., 2010; Soderberg et al., 2015). Individuals vary in the extent to which they typically think in an abstract or concrete manner (Vallacher & Wegner, 1989), although context can activate construal level (e.g., thinking about why something occurs versus how something occurs prompts abstractness; Freitas et al., 2004; Hess et al., 2018). Encouraging abstractness has been found to reduce prejudice against ethnic minorities (Yogeeswaran & Dasgupta, 2014), produce universalist rather than target-specific moral judgments (Mentovich et al., 2016), and increase the relevance of future concerns (Eyal et al., 2009; Ledgerwood et al., 2015). Given these effects and the psychological distance of AIs from humans, we expected a more abstract construal level to predict increased moral consideration of AIs.

2.2 Relational Predictors

2.2.1 Social Dominance Orientation

Social Dominance Orientation (SDO) refers to the preference for a structured social hierarchy with some groups dominating other groups (Ho et al., 2012; Pratto et al., 1994, 2000, 2006). A strong SDO has been associated with political conservatism, increased support for war, anti-immigrant attitudes, and racism (Ho et al., 2012). SDO has been shown to positively correlate with speciesism and the moral prioritization of humans over animals (Caviola et al., 2019). We expected a stronger SDO to predict less moral consideration of AIs.

2.2.2 Sci-Fi Fan Identity

Sci-fi fan identity represents the extent to which people identify with the science fiction fan group, extending from Tajfel and Turner's (1979) Social Identity Theory that people who self-categorize as and are categorized by others as a group member derive self-esteem and belonging from group membership (Tajfel, 1974; Turner et al., 1987). Research has established that fans are cohesive groups with meaningful identities (e.g., Fink et al., 2009; Groene & Hettinger, 2016; Misailidou, 2017; Plante et al., 2014; Tsay-Vogel & Sanders, 2017). Some researchers have begun to consider the importance of engagement with fiction narratives for shaping attitudes and behaviors (Dill-Shackleford et al., 2016; Fernando et al., 2020; Igartua & Frutos, 2017). We expected a stronger sci-fi fan identity to predict more moral consideration of AIs because of the presence of AI narratives in science fiction.

2.3 Expansive Predictors

2.3.1 Perceived Social Norms

Social norms provide information about what others think and do (descriptive norms) and what others think is appropriate to think and do (injunctive norms), both of which shape intergroup behavior (Mackie & Smith, 2018; Miller & Prentice, 2016; E. R. Smith et al., 2021). Perceived descriptive norms often imply injunctive norms and have been predictive of behavior in environmental, health, and intergroup contexts (Farrow et al., 2017; Mackie & Smith, 2018; Miller & Prentice, 2016). We expected stronger perceived descriptive norms for valuing humans over nonhumans to predict less moral consideration of AIs.

2.3.2 Anthropomorphism

Anthropomorphism is the human-like physical design of and attribution of human-like mental capacities to nonhuman entities (Waytz, Cacioppo, et al., 2010). Some researchers have discussed how human-like design might improve the integration of robots into human society (Giger et al., 2019; Oliveira et al., 2018; Paiva et al., 2018; Waytz et al., 2014). Anthropomorphizing an AI assistant has been shown to reduce psychological distance to and increase positive attitudes towards the assistant (X. Li & Sung, 2021). Some research has suggested that too human-like physical designs may lead people to devalue AIs' moral decisions (Laakasuo et al., 2021). Other research has established that individuals' tendency to attribute human-like mental capacities to AIs predicts increased moral consideration of them (Sommer et al., 2019; Waytz, Cacioppo, et al., 2010; Waytz, Epley, et al., 2010). We expected a stronger tendency to anthropomorphize to predict increased moral consideration of AIs.

2.3.3 Global Citizenship

Global citizenship (GC) embeds prosocial values of justice and diversity, a responsibility to act, and intergroup helping. GC has been shown to predict morally expansive attitudes and behaviors (Barth et al., 2015; Cabrera, 2008; McFarland et al., 2013, 2019; Reysen & Katzarska-Miller, 2013). For instance, GC has been linked with activism for human rights, peace, and poverty causes (Reysen & Hackett, 2017), pro-environmental product preferences (Ng & Basu, 2019), and pro-technology attitudes (R. B. Lee et al., 2017). We expected stronger GC to predict increased moral consideration of AIs.

2.3.4 Openness to Experience

Openness to experience (OTE) is a dimension of personality reflecting broadmindedness, unconventionality, and social tolerance (Costa & McCrae, 1992; Sparkman et al., 2016). OTE has been correlated with global citizenship and creativity (McFarland et al., 2012; Tidikis & Dunbar, 2019), reduced ethnic prejudice (Flynn, 2005; Sparkman et al., 2016), and moral reasoning (Athota et al., 2010). We expected more OTE to predict increased moral consideration of AIs.

2.3.5 Techno-animism

Animism is the belief that nonhuman entities possess a soul, lifeforce, or degree of personhood (Marenko, 2014; Waytz, Cacioppo, et al., 2010). Animist beliefs have been a part of various religious traditions throughout history (Pedersen, 2001; Richardson, 2016; Wilkinson, 2017) and are conceptually related to, although distinct from, the tendency to anthropomorphize (Waytz, Cacioppo, et al., 2010). Techno-animism, or animism applied to technological objects, has been theorized as arising from a combination of Shinto beliefs and the creation of human-like social robots in Japan (Jensen & Blok, 2013; Marenko, 2014; Richardson, 2016). Beran et al. (2011) demonstrated that children younger than 12 showed techno-animist beliefs following interaction with a robot. To our knowledge, previous studies have measured neither techno-animist beliefs in adults, nor how they relate to the moral consideration of AIs. We expected stronger techno-animist beliefs to predict more moral consideration of AIs because of their similarity to anthropomorphism and perceiving life in technology.

2.4 Technological Predictors

2.4.1 Affinity for Technology

Affinity for technology (AFT) refers to people's capacity to engage with and show positivity towards technology (Edison & Geissler, 2003). Theoretical reviews of technology adoption have implicated AFT as important for consumer acceptance of tourism service robots (Kazandzhieva & Filipova, 2019) and autonomous vehicles (Egbue & Long, 2012; Sheela & Mannering, 2020). We suggest that AFT may be relevant for the moral consideration of AIs. However, we made no predictions about the relationship between AFT and moral consideration, given that AFT may not apply to moral consideration if AI technology is thought of primarily as a tool.

2.4.2 Substratism

Substratism refers to prejudice against AIs based on their non-biological (i.e., silicon-based rather than carbon-based) material composition. Previous research has found that greater substratism correlates with less moral consideration of AIs (Ladak et al., 2021). We expected to replicate this relationship.

2.4.3 Human-AI overlap

Human-AI overlap is the degree of perceived overlap or connection between humans and AIs. Ladak et al. (2021) found a correlation between greater perceived overlap and greater moral consideration of AIs. This relationship corresponds to research that showed that humans will extend their moral circle to include certain animals when those animals are perceived as more similar to humans (Bastian et al., 2012; Crimston et al., 2018; Opotow, 1993, 1994). We expected more perceived overlap between humans and AIs to predict more moral consideration of AIs.

2.4.4 Threat

Realistic (e.g., resource) and symbolic (e.g., identity) threats stem from group membership and contribute to negative intergroup relations (Pauketat et al., 2020; Riek et al., 2006; Stephan et al., 2009). Identity threat arises when the values, beliefs, or distinctiveness of a group are called into question by the presence of another group (Leonardelli et al., 2010; Riek et al., 2006; Stephan et al., 2009). Realistic threat arises from competition between groups over resources such as money and over physical safety or existence concerns (Riek et al., 2006; Stephan et al., 2009). Humans have been shown to perceive both types of threats from robots (Złotowski et al., 2017). The anthropomorphic design of robots may lead humans to experience threats to their human uniqueness (Ferrari et al., 2016; Giger et al., 2019; Yogeeswaran et al., 2016; Złotowski et al., 2017). Belief in human uniqueness has been shown to predict denying emotional capacity to AIs as well as increased negative attitudes towards AIs (Giger et al., 2019). Robots have been shown to provoke realistic threats when they outperformed humans on physical and mental tasks (Yogeeswaran et al., 2016) and when they were more autonomous (Złotowski et al., 2017). We expected increased perceptions of realistic and identity threat to predict decreased moral consideration of AIs.

2.5 Affective Predictors

Feeling emotions as a group member stems from group membership (Mackie & Smith, 2018). Group-based emotions have been shown to predict intergroup attitudes and behaviors amongst human groups (Mackie & Smith, 2018; Pauketat et al., 2020) and towards robots (E. R. Smith et al., 2020, 2021). We examined a range of positive and negative emotions felt towards AIs (see Table 1), expecting positive emotions to predict increased, and negative emotions to predict decreased, moral consideration of AIs.

Table 1

Emotion Definitions and AI Research Applications

Emotion	Conceptual definition	AI studies	Definitional Citations
respect, admiration	other-praising, appreciation focused; moral emotions	Yes	Janoff-Bulman & Werther, 2008; Leonard et al., 2011; Ortony et al., 1988; Ray et al., 2008; Seger et al., 2017; E. R. Smith et al., 2021
compassion	response to others' suffering coupled with motivation to help; moral emotion	Yes	Bartneck et al., 2005; Goetz et al., 2010; Parviainen et al., 2019; Stellar et al., 2017
awe	schema-incongruent, response to vast, initially incomprehensible stimuli; moral emotion	No	Keltner & Haidt, 2003; Song, 2021; Stellar et al., 2017; Taylor & Uchida, 2019
excitement	high arousal, pleasant valence; core affect	Yes	Russell, 2009; E. R. Smith et al., 2020; Watson et al., 1988
pride	self- or ingroup-focused, response to commendation; moral emotion	No	Leach et al., 2002; McLatchie & Piazza, 2017; Tausch & Becker, 2013; Van Leeuwen et al., 2013
disgust	response to perceived physical or social contaminants; moral emotion	Yes	Hutcherson & Gross, 2011; Levin et al., 2016; Pauketat et al., 2020; Rozin et al., 1999; Seger et al., 2017; E. R. Smith et al., 2020
anger	response to affront, motivating confrontation; moral emotion	No	Hutcherson & Gross, 2011; Leonard et al., 2011; Levin et al., 2016; Pauketat et al., 2020; Rozin et al., 1999; Seger et al., 2017; Tausch & Becker, 2013
contempt	response to evaluations of hierarchy and social status; moral emotion	No	Hutcherson & Gross, 2011; Levin et al., 2016; Pauketat et al., 2020; Rozin et al., 1999; Tausch et al., 2011
horror	schema-incongruent, response to unexpected harmful thoughts or behaviors	No	Taylor & Uchida, 2019
fear, anxiety	self- or ingroup-preserving, response to motivate avoidance of danger	Yes	Leonard et al., 2011; Lerner & Keltner, 2001; Oatley et al., 2006; Seger et al., 2017; E. R. Smith et al., 2020
envy	pain felt in response to others' joy or fortune	Yes	D. Lee & Kim, 2020; van de Ven, 2017

3 Moral Consideration of AI Outcomes

The moral consideration of AIs may take many forms. We investigated how these various forms relate to each other and how they are predicted by psychological tendencies. Mind perception—the attribution of beliefs, intentions, and mental states like thoughts and emotions to others—includes both the attribution of cognitive capacities necessary for taking moral action (moral agency) and the attribution of emotional capacities necessary for receiving moral consideration (moral patiency; Epley & Eyal, 2019; K. Gray et al., 2012; K. Gray & Wegner, 2012b). The perception of mind in robots has been shown to vary depending on the purpose and value of the robot. Wang and Krumhuber (2018) showed that people perceived robots to have the capacity for basic cognition. People perceived social robots (i.e., those built for social support and companionship) to have more emotional capacities than economic robots (i.e., those built for profit and corporate benefit), especially when social and economic benefits were high. High-value social robots received more moral consideration in the form of protection from electric shocks (Wang & Krumhuber, 2018). We predicted that the minds of current and future AIs will be evaluated differently. Among other factors, peoples' direct experiences with current AIs and exposure to future AIs in science fiction narratives might lead to different perceptions (Gunkel, 2018; Kakoudaki, 2015).

Moral expansiveness (Crimston et al., 2016) represents the broadness of an individual's moral circle (Crimston et al., 2016, 2018; Laham, 2009) based on the degrees of various entities' inclusion. For example, a family member is typically in the inner circle of moral consideration whereas chickens and apple trees are often on the fringes. Moral exclusion or the "scope of justice" refers to the withholding of moral standards from certain groups (i.e., placement away from the center or outside the moral boundary). Moral exclusion/scope of justice has predicted support for harmful actions against, denying of fairness standards to, and the withholding of resources from excluded others (Hadarics & Kende, 2018; Opotow, 1990, 1993). For example, people who morally excluded beetles were less likely to protect the beetles from harm (Opotow, 1994). As far as we know, the moral exclusion/scope of justice of AIs has not been explicitly studied.

Intent to join collective actions such as protests for AI rights (henceforth, "AI rights activism"), support for AI rights policies, and donation behavior to organizations connected to AI rights (e.g., reducing risks of suffering in the long-term future) represent behavioral aspects of moral consideration. Lima et al. (2020) showed that people favored policies protecting AIs from cruelty. Support for other policies such as the right to hold assets was increased by describing AIs as fully autonomous (Lima et al., 2020). Other types of support for AI rights (e.g., joining a protest) have not been studied to our knowledge. We suggest that humans may be willing to act on behalf of AIs like they act on behalf of marginalized human and animal groups (e.g., Hadarics & Kende, 2018), given their intergroup relationship with AIs (E. R. Smith et al., 2021; Vanman & Kappas, 2019), though likely to a lesser extent because of the large perceived differences between humans and AIs.

4 Methods

This study and our hypotheses were preregistered (https://osf.io/98bwv).

4.1 Participants and Design

Responses were gathered online from Prolific with 300 participants of U.S. nationality who reside in the U.S. (M_{age} = 34.88, *S.D.* = 13.26, 61% female, 68% White) and completed the correlational survey using the GuidedTrack platform.[1] Participants were paid USD $3.70 with an additional bonus of $2 for an average required time of 19 minutes ($18/per hour). Of the 300, 17 were excluded from analyses for failing an attention check. The remaining sample consisted of 283 participants (M_{age} = 34.46, *S.D.* = 13.00, 60% female, 68% White). This sample self-identified as politically liberal (*M* = 2.96, *S.D.* = 1.70) and reported directly interacting with or having exposure

[1] Fourteen people dropped out of the study after consenting to participate and before completing the survey. These responses were not included in the completed 300 responses tally.

to media about robots/AIs a few times a year on average (M = 2.18, $S.D.$ = 1.22). Some of the sample personally owned a robot/AI (39%) and a few reported working with a robot/AI[2] (4%).[3]

We recruited this sample size based on a G*Power (Faul et al., 2009) analysis of the primary multiple regression model (R^2 deviation from zero for the effect of all Xs on Y) finding that 206 participants could detect a small to medium effect (f^2 = .12) with 80% power given α = .05. The specified effect size was determined from anticipated predictor correlations in G*Power estimated based on the average effect sizes witnessed in Schäfer and Schwarz's (2019) review of effect sizes in preregistered psychology studies. We recruited greater than 206 to account for possible Prolific attrition (Palan & Schitter, 2018). Participants provided their informed consent prior to the survey and were thanked and debriefed at the end of the survey.

4.2 Procedure and Measures

All materials, data, and analysis code for this study are available on the OSF (https://osf.io/w3tcn). Participants first read a definition that "'artificial beings' and 'robots/AI' are intelligent entities built by humans, such as robots, virtual copies of human brains, or computer programs that solve problems, with or without a physical body, that may exist now or in the future." Then, participants completed the survey in the following order: 1) responded to the perspective, relational, and expansive predictor questions, 2) were reminded of the definition, 3) responded to the technological predictor questions, 4) responded to the affective predictor questions, 5) were reminded of the definition, 6) responded to the self-reported outcome questions, 7) were given the opportunity to donate the $2 bonus, and 8) completed demographics. At each stage, the order of scales and items within each scale were randomized. Sliding scales were used for most continuous items. Most indices were shortened from their original form based on face validity to reduce the response burden from the long online survey, and we present reliability measures in turn. Indices were averaged or summed based on their original paper's methods.

4.2.1 Psychological Predictors

Perspective. Future orientation was assessed with six items from Strathman et al.'s (1994) Consideration of Future Consequences scale. Items such as, "I consider how things might be in the future, and try to influence those things with my day to day behavior," were measured on a 1 (extremely uncharacteristic) to 5 (extremely characteristic) sliding scale. Four items were reverse scored prior to summing (Cronbach's α = .73). Higher scores indicated a stronger future orientation.

Construal level was assessed with 10 items from Vallacher and Wegner's (1989) Behavioral Identification Form. Participants chose between abstract and concrete options associated with a behavior based on what felt most appropriate. For instance, "eating" could be described as "getting

[2] We could have used many terms to describe the entities of interest, such as "artificial beings," "artificial entities," "robots and AI," or "AI." We favored "robot/AI" because of its clarity, specificity, and inclusion of all AIs, including the salient robot subgroup. We also used "artificial being" in the survey where it seemed more sensible than "robot/AI."

[3] These numbers reflect a static point in time and are meant to illustrate this sample rather than evidence a representative trend in ownership over time. For information on AI trends, see the AI Index Annual Report (D. Zhang et al., 2022).

nutrition" (abstract) or "chewing and swallowing" (concrete). Abstract options were coded as 1, concrete options as 0. Responses from the 10 items were averaged. Higher scores reflected the tendency to think abstractly.

Relational. SDO was measured with four items from Ho et al. (2012) on a 1 (strongly disagree) to 7 (strongly agree) sliding scale, such as, "Having some groups on top really benefits everybody." Two items were reverse scored prior to averaging (Cronbach's α = .84). Higher scores indicated a stronger preference for social hierarchy.

Sci-fi fan identity was measured with four items adapted from Leach et al. (2008; e.g., "I am glad to be a sci-fi fan"). Items were measured on a 1 (strongly disagree) to 7 (strongly agree) sliding scale. The four items were averaged (Cronbach's α = .91). Higher scores indicated a stronger sci-fi fan identity.

Expansive. The sentence stem "Most people who are important to me" was paired with five items such as "think that robots/AI cannot have feelings" and "eat meat" to measure human-centric social norms (sentence stem based on Fishbein & Ajzen, 2009). Each item was assessed on a 1 (strongly disagree) to 7 (strongly agree) sliding scale. One item ("focus on human welfare") was excluded due to low reliability (Cronbach's α = .46).[4] The remaining four items were averaged (Cronbach's α = .55).[5] Higher scores indicated stronger human-centric norms.

Anthropomorphism was assessed with eight items summed from the Individual Differences in Anthropomorphism Questionnaire (Cronbach's α = .78; Waytz et al., 2010). Items such as, "To what extent does the average robot have consciousness?" and, "To what extent does a cheetah experience emotions?" were measured on a 0 (not at all) to 10 (very much) sliding scale. Higher scores indicated a greater tendency to anthropomorphize.

Global citizenship was assessed with three items (e.g., "Being actively involved in global issues is my responsibility.") from Reysen and Katzarska-Miller (2013) and one item ("I see myself more as a global citizen than a citizen of my country.") from GlobeScan and the BBC World Service (2016). These four items were measured on a 1 (strongly disagree) to 7 (strongly agree) sliding scale and averaged (Cronbach's α = .80). Higher scores represented stronger global citizenship.

Openness to experience was assessed by averaging the two items (r = .39) from the Ten-Item Personality Inventory (Gosling et al., 2003), "open to new experiences, complex" and "conventional, uncreative" (reverse scored). Items were measured on a 1 (disagree strongly) to 7 (agree strongly) sliding scale. Higher scores indicated more openness to experience.

Techno-animism was assessed with four items inspired by Jensen and Blok's (2013) theorizing on techno-animism. Items such as, "The boundaries between humanity, nature, and technology are vague and interchangeable," were measured on a 1 (strongly disagree) to 7 (strongly agree) sliding scale. One item ("Human beings contain a spirit.") was removed due to low reliability (Cronbach's α = .68). Three items were averaged (Cronbach's α = .77). Higher scores indicated more techno-animistic beliefs.

Technological. Three items from Edison and Geissler's (2003) Affinity for Technology scale were summed (Cronbach's α = .76). Items such as, "I relate well to technology and machines," were

[4] Item exclusions were made based on preregistered criteria.
[5] This was the highest reliability achievable by excluding items.

measured on a 1 (strongly disagree) to 7 (strongly agree) sliding scale. Higher scores represented more positive attitudes towards technology.

Substratism was measured with six items such as, "Morally, artificial beings always count for less than humans" (Ladak et al., 2021). Items were measured on a 1 (strongly disagree) to 7 (strongly agree) sliding scale. One item was reverse scored. All items were averaged (Cronbach's $\alpha = .89$). Higher scores indicated more substratism.

Human-AI overlap was assessed with an adapted Inclusion of the Ingroup in the Self item (Tropp & Wright, 2001) asking participants to select the pair of circles that best represented how connected humans are to robots/AI from increasingly overlapping circles. In circle one, the human and robot/AI circles were not touching. The degree of closeness increased with each integer up to seven. In circle seven, the human and robot/AI circles were highly overlapping.

Realistic and identity threat were measured with four items each, based on Złotowski and colleagues (2017), on a 1 (strongly disagree) to 7 (strongly agree) sliding scale. Three realistic threat items (e.g., "The increased use of robots in our everyday life is causing more job loss for humans.") were averaged (Cronbach's $\alpha = .72$). One item ("In the long run, robots/AI should be entitled to share the resources of humanity.") was excluded due to low reliability (Cronbach's $\alpha = .60$). Likewise, one item ("Robots/AI should be able to evolve independently of humanity.") was removed from the identity threat scale due to low reliability (Cronbach's $\alpha = .37$). Three identity threat items (e.g., "Recent advances in robot technology are challenging the very essence of what it means to be human.") were averaged (Cronbach's $\alpha = .50$). Higher scores on both scales represented more perceived threat.

Affective. Six positive (respect, admiration, compassion, awe, excitement, pride) and seven negative emotions (disgust, contempt, anger, horror, fear, anxiety, envy) felt towards robots/AI were evaluated on a 1 (none at all) to 7 (very much) sliding scale, following from Smith et al. (2020). A pre-registered composite index of positive emotion was created by averaging all positive emotions except pride because of its link to positive and negative outcomes (e.g., McLatchie & Piazza, 2017; Cronbach's $\alpha = .84$). Similarly, all negative emotions were averaged (Cronbach's $\alpha = .86$), excluding envy because of its link to positive and negative outcomes (e.g., van de Ven, 2017). Specific emotions were entered as single items in some analyses.

4.2.2 Moral Consideration Outcomes

We evaluated responses to robots/AI as a group, rather than evaluating responses to a specific artificial entity. Evaluating responses to a specific individual has been common in past research (see Smith et al., 2021; Naneva et al., 2020). Responses to individuals have typically been assumed to generalize to AIs as a group. For instance, researchers have assumed that responses to one robot following an in-person interaction are indicative of a general response to all similar robots. Like Smith et al. (2021), we believe that it is important to study how humans respond to AIs as a social group just as responses to human outgroups are important.

Mind perception. Mind perception was assessed with six items measured on a 0 (not at all) to 100 (very much) sliding scale (Wang & Krumhuber, 2018) for artificial entities that exist in 2021 and separately for artificial entities that will exist in the future. Three items assessed attributions of cognitive capacity (e.g., "thinking analytically"). Three items assessed attributions of emotional capacity (e.g., "having feelings"). The six items were averaged to form composite indices of mind perception (current Cronbach's α = .82; future Cronbach's α = .86). Higher scores reflected more attribution of mind to AIs.

Sentience value. Participants answered two questions about the perceived value of current and future AIs' feelings compared to humans' feelings (e.g., "Are current artificial beings' (i.e., those that exist in 2021) feelings worth the same as human feelings?"). Responses for current and future sentience value were analyzed as single items. These items were measured on a 1 (not at all) to 7 (definitely) sliding scale.

Moral expansiveness. We used a 12-item version of Crimston et al.'s (2016) Moral Expansiveness Scale that we shortened and adapted to include AIs. Participants rated in which circle of moral concern they would put entities such as "robots" and "charity worker" (3 = inner circle, 2 = outer circle, 1 = fringes of moral concern, 0 = outside the moral boundary). Responses were summed (Cronbach's α = .80). Higher scores represented a more expansive moral circle.

Moral exclusion/Scope of justice. Participants evaluated three statements about the moral exclusion of AIs on a 1 (strongly disagree) to 7 (strongly agree) sliding scale adapted from Hadarics and Kende (2018). All items, such as, "I believe that considerations of fairness apply to artificial beings too," were reverse scored and averaged (Cronbach's α = .91). Higher scores represented more moral exclusion of AIs.

AI rights activism. Six items adapted from Hadarics and Kende (2018) measured collective action intentions for AI rights. Items such as, "I would join an online campaign against the discrimination of robots/AI," were measured on a 1 (strongly disagree) to 7 (strongly agree) sliding scale. Three items (e.g., "I would sign a petition that is intended to protect people against the harmful behavior of robots/AI.") were excluded due to unexpected negative correlations with the scale following score reversal that resulted in low scale reliability.[6] The remaining three items were averaged (Cronbach's α = .93). Higher scores indicated more willingness to act collectively for AI rights.

AI policy support. Six policy items from Lima et al. (2020) were assessed on a 1 (strongly disagree) to 5 (strongly agree) sliding scale (e.g., "No cruel punishment or treatment should be inflicted on completely autonomous robots/AI."). These items were averaged (Cronbach's α = .89). Higher scores indicated support for pro-AI rights policies.

Donation behavior. Participants were asked if they would like to donate any part of their $2 bonus to an organization contributing to reducing the suffering of nonhumans. Participants could choose to keep their whole bonus or donate any part of it (on a sliding scale from $0 to $2) to Mercy For Animals, World Wildlife Fund, the Center on Long-Term Risk, the Future of Life Institute, or the Center for Reducing Suffering.

4.2.3 Demographics

[6] This could be due to confusion over the meaning of the negatively framed items (i.e., those protesting against robot/AI rights) alongside the positively framed items (i.e., those protesting for robot/AI rights) given that there is little current need to protest against a non-existent set of rights that would protect robots/AI over humans.

Participants answered demographic questions at the end of the survey. They indicated whether or not they owned an AI or robotic device (yes, no), whether or not they worked with an AI or robotic device at their job (yes, no), how often they interacted with AI or robotic devices and how many times in the past year they read or watched robot/AI-related media or other material measured on a 1 (only on rare occasions) to 5 (daily) Likert-type scale with an option for "never or not applicable." Selections of "never or not applicable" were coded as 0. The latter two items about frequency of interaction with and exposure to media about robots/AI were averaged to form a frequency of interaction index (r = .40, p < .001). Participants also provided their age, gender, race/ethnicity, political views (1 = very liberal, 7 = very conservative), education level, religion, diet, and family's net household income per year.

5 Results

We preregistered this analysis strategy (https://osf.io/98bwv) and used RStudio (v. 1.3.1093).[7] See supplemental results (https://osf.io/w3tcn) for additional descriptive and diagnostic assumption checking statistics.

5.1 Predicting Moral Consideration

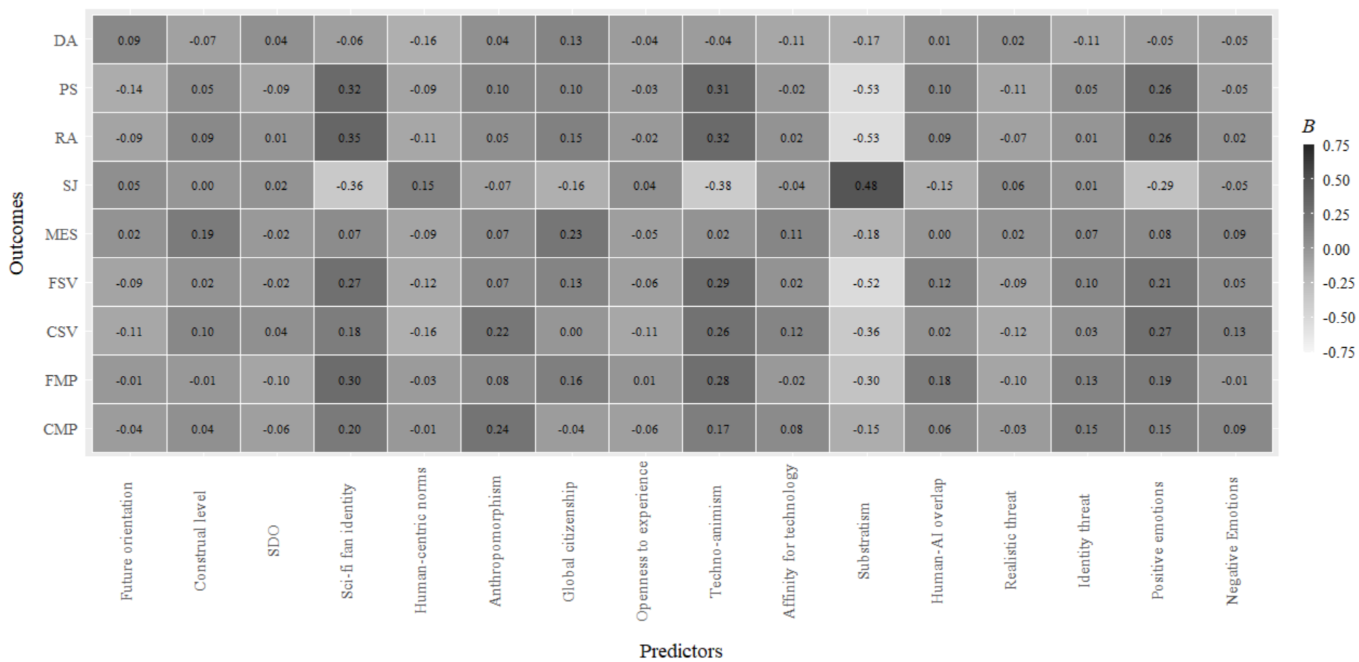
We used hierarchical OLS regressions to predict moral consideration (see Figure 1 for a summary of predictor effects). Multiple comparisons were checked using the False Discovery Rate (FDR; Benjamini & Hochberg, 1995). Differences in significance following FDR corrections are recorded in table notes. In all regressions, categorical demographics were dummy-coded (see Tables 2-4). As shown in section 5.2, the outcomes reduce to three key dimensions of moral consideration: mind perception, psychological expansion, and practical consideration. We highlight three regressions representative of these dimensions and present the full details of the other regressions in supplemental results.

Demographics were controlled for in step 1. Perspective predictors were tested in step 2, relational predictors in step 3, expansive predictors in step 4, technological predictors in step 5, and affective predictors in step 6. This order was specified to reflect that context-based affective and technological attitudes build upon more general expansive norms and beliefs that likewise build upon core relational and perspective traits and mindsets. This strategy enabled us to evaluate the unique variance explained by similar types of predictors while controlling the effects of previously entered predictors.

[7] One participant was missing data for the identity threat items. This participant's data was removed for analyses that included identity threat.

Figure 1

Summary of Psychological Predictor Effects



Note. Standardized regression coefficients are presented within the boxes from the regression step that the predictor was first entered. Darker greys represent stronger positive relationships and lighter greys represent stronger negative relationships. CMP = Current Mind Perception; FMP = Future Mind Perception, CSV = Current Sentience Value; FSV = Future Sentience Value; MES = Moral Expansiveness; SJ = Scope of Justice; RA = Rights Activism; PS = Policy Support; DA = Donation Amount

Future mind perception. Demographics explained 10% of the variance in future mind perception (p = .015). Perspective predictors did not change the model, $\Delta R^2 = .00$, $p = .962$. Relational predictors increased model fit, $\Delta R^2 = .08$, $p < .001$. Expansive predictors also increased model fit, $\Delta R^2 = .12$, $p < .001$. Technological predictors further increased model fit, $\Delta R^2 = .12$, $p < .001$ as did affective predictors, $\Delta R^2 = .02$, $p = .009$. The full model explained 44% of the variance (p < .001; see Table 2).

Table 2

Predicting Future Mind Perception

	b	S.E. *b*	β	*b* 95% C.I.	*t*- statistic
Demographics					
Age	-0.27	0.11	-.15	-0.48, -0.05	-2.40*
Gender (0 = man, 1 = woman, non-binary)	-3.88	2.98	-.08	-9.75, 1.99	-1.30
Race (0 = White, 1 = BIPOC)	4.28	3.02	.09	-1.66, 10.22	1.42
Income (0 = $41-99k, 1 = < $41k)	-0.25	3.24	-.01	-6.63, 6.12	-0.08
Income (0 = $41-99k, 1 = $100k-249k)	-3.43	3.75	-.06	-10.82, 3.96	-0.91
Income (0 = $41-99k, 1 = $250k+)	1.10	10.32	.01	-19.22, 21.41	0.11
Income (0 = $41-99k, 1 = undisclosed)	-3.33	7.25	-.03	-17.60, 10.93	-0.46
Education (0 = HS, 1 = none, undisclosed)	-5.80	9.99	-.04	-25.47, 13.87	-0.58

	b	S.E. b	β	b 95% C.I.	t- statistic
Education (0 = HS, 1 = associates, bachelors)	-4.39	3.13	-.10	-10.55, 1.77	-1.40
Education (0 = HS, 1 = masters, doctoral)	-4.18	4.37	-.07	-12.77, 4.42	-0.96
Diet (0 = meat-eater, 1 = restrictions)	4.91	3.60	.08	-2.18, 12.01	1.36
Religion (0 = religious, 1 = not religious)	0.89	2.86	.02	-4.73, 6.51	0.31
Political orientation	-0.18	0.89	-.01	-1.92, 1.56	-0.21
Frequency of interaction	3.46	1.16	.19	1.19, 5.73	3.00**
Perspective predictors					
Future orientation	-0.06	0.40	-.01	-0.85, 0.73	-0.15
Construal level	-1.21	5.41	-.01	-11.87, 9.44	-0.22
Relational predictors					
SDO	-1.86	1.35	-.10	-4.51, 0.79	-1.38
Sci-fi fan identity	4.31	0.88	.30	2.58, 6.03	4.91***
Expansive predictors					
Human-centric norms	-0.68	1.36	-.03	-3.36, 2.01	-0.50
Anthropomorphism	0.14	0.11	.08	-0.08, 0.36	1.24
Global citizenship	2.92	1.26	.16	0.43, 5.40	2.31*
Openness to experience	0.17	1.06	.01	-1.92, 2.26	0.16
Techno-animism	4.57	1.05	.28	2.50, 6.65	4.34***
Technological predictors					
Affinity for technology	-0.11	0.35	-.02	-0.80, 0.58	-0.30
Substratism	-4.87	0.97	-.30	-6.78, -2.95	-5.00***
Human-AI overlap	2.59	0.82	.18	0.97, 4.22	3.15**
Realistic threat	-1.72	1.11	-.10	-3.90, 0.46	-1.56
Identity threat	2.57	1.22	.13	0.16, 4.98	2.10*
Affective predictors					
Positive emotions	3.25	1.06	.19	1.16, 5.34	3.07**
Negative emotions	-0.22	1.17	-.01	-2.54, 2.09	-0.19

Note. Values displayed only for new variables added at each step; showing effects controlling for variables entered in previous steps. Uncorrected p-values reported. Some p-values became non-significant with FDR correction: age (.085), global citizenship (.131), and identity threat (.214). $F(14, 267) = 2.04$, $p = .015$, $R^2 = .10$, adjusted $R^2 = .05$ for demographics; $F(16, 265) = 1.78$, $p = .034$, $R^2 = .10$, adjusted $R^2 = .04$, $\Delta R^2 = .00$ for perspective predictors ($p = .962$); $F(18, 263) = 3.18$, $p < .001$, $R^2 = .18$, adjusted $R^2 = .12$, $\Delta R^2 = .08$ for relational predictors ($p < .001$); $F(23, 258) = 4.71$, $p < .001$, $R^2 = .30$, adjusted $R^2 = .23$, $\Delta R^2 = .12$ for expansive predictors ($p < .001$); $F(28, 253) = 6.41$, $p < .001$, $R^2 = .42$, adjusted $R^2 = .35$, $\Delta R^2 = .12$ for technological predictors ($p < .001$); $F(30, 251) = 6.48$, $p < .001$, $R^2 = .44$, adjusted $R^2 = .37$, $\Delta R^2 = .02$ for affective predictors ($p = .009$). *$p < .05$, **$p < .01$, ***$p < .001$

Moral expansiveness. Demographics explained 3% of the variance in moral expansiveness ($p = .867$). Perspective predictors increased the model fit, $\Delta R^2 = .03$, $p =.008$. Relational predictors did not change the model fit, $\Delta R^2 = .00$, $p = .522$. Expansive predictors increased the model fit, $\Delta R^2 = .05$, $p =.009$. Technological predictors somewhat increased the model fit, $\Delta R^2 = .04$, $p =.058$ but affective predictors did not, $\Delta R^2 = .01$, $p =.205$. The full model explained 17% of the variance ($p =. 015$; see Table 3).

Table 3

Predicting Moral Expansiveness

	b	S.E. b	β	b 95% C.I.	t- statistic

Demographics					
Age	-0.03	0.02	-.07	-0.07, 0.02	-1.08
Gender (0 = man, 1 = woman, non-binary)	0.41	0.65	.04	-0.88, 1.69	0.62
Race (0 = White, 1 = BIPOC)	-0.59	0.66	-.06	-1.89, 0.70	-0.90
Income (0 = $41-99k, 1 = < $41k)	0.13	0.71	.01	-1.26, 1.52	0.19
Income (0 = $41-99k, 1 = $100k-249k)	-0.29	0.82	-.02	-1.91, 1.32	-0.35
Income (0 = $41-99k, 1 = $250k+)	0.10	2.25	.00	-4.34, 4.53	0.04
Income (0 = $41-99k, 1 = undisclosed)	-1.17	1.58	-.05	-4.29, 1.95	-0.74
Education (0 = HS, 1 = none, undisclosed)	2.63	2.18	.08	-1.66, 6.93	1.21
Education (0 = HS, 1 = associates, bachelors)	0.14	0.68	.01	-1.20, 1.49	0.21
Education (0 = HS, 1 = masters, doctoral)	-0.23	0.95	-.02	-2.11, 1.64	-0.25
Diet (0 = meat-eater, 1 = restrictions)	0.93	0.79	.07	-0.62, 2.48	1.18
Religion (0 = religious, 1 = not religious)	-0.88	0.62	-.09	-2.11, 0.35	-1.41
Political orientation	-0.06	0.19	-.02	-0.44, 0.32	-0.29
Frequency of interaction	0.28	0.25	.07	-0.22, 0.78	1.12
Perspective predictors					
Future orientation	0.03	0.09	.02	-0.14, 0.20	0.39
Construal level	3.57	1.16	.19	1.28, 5.86	3.08**
Relational predictors					
SDO	-0.07	0.30	-.02	-0.67, 0.52	-0.25
Sci-fi fan identity	0.22	0.20	.07	-0.17, 0.61	1.11
Expansive predictors					
Human-centric norms	-0.45	0.32	-.09	-1.09, 0.18	-1.42
Anthropomorphism	0.03	0.03	.07	-0.02, 0.08	1.04
Global citizenship	0.89	0.30	.23	0.31, 1.48	3.00**
Openness to experience	-0.21	0.25	-.05	-0.70, 0.28	-0.85
Techno-animism	0.05	0.25	.02	-0.44, 0.54	0.21
Technological predictors					
Affinity for technology	0.14	0.09	.11	-0.03, 0.32	1.61
Substratism	-0.60	0.25	-.18	-1.09, -0.12	-2.45*
Human-AI overlap	0.00	0.21	.00	-0.41, 0.41	0.01
Realistic threat	0.06	0.28	.02	-0.49, 0.61	0.21
Identity threat	0.30	0.31	.07	-0.31, 0.90	0.96
Affective predictors					
Positive emotions	0.28	0.27	.08	-0.25, 0.82	1.05
Negative emotions	0.36	0.30	.09	-0.23, 0.95	1.20

Note. Values displayed only for new variables added at each step; showing effects controlling for variables entered in previous steps. Uncorrected *p*-values reported. Substratism became non-significant (p = .145) with FDR correction. $F(14, 267) = 0.60$, $p = .867$, $R^2 = .03$, adjusted $R^2 = .00$ for demographics; $F(16, 265) = 1.15$, $p = .309$, $R^2 = .07$, adjusted $R^2 = .01$, $\Delta R^2 = .03$ for perspective predictors ($p = .008$); $F(18, 263) = 1.09$, $p = .360$, $R^2 = .07$, adjusted $R^2 = .01$, $\Delta R^2 = .00$ for relational predictors ($p = .522$); $F(23, 258) = 1.58$, $p = .048$, $R^2 = .12$, adjusted $R^2 = .05$, $\Delta R^2 = .05$ for expansive predictors ($p = .009$); $F(28, 253) = 1.71$, $p = .017$, $R^2 = .16$, adjusted $R^2 = .07$, $\Delta R^2 = .04$ for technological predictors ($p = .058$); $F(30, 251) = 1.71$, $p = .015$, $R^2 = .17$, adjusted $R^2 = .07$, $\Delta R^2 = .01$ for affective predictors ($p = .205$). $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$

Moral exclusion/Scope of Justice. Demographics explained 13% of the variance in scope of justice (p = .001). Perspective predictors did not change the model fit, $\Delta R^2 = .00$, $p = .707$. Relational predictors increased the model fit, $\Delta R^2 = .11$, $p < .001$. Expansive predictors ($\Delta R^2 = .21$,

$p < .001$), technological predictors ($\Delta R^2 = .20$, $p < .001$), and affective predictors ($\Delta R^2 = .05$, $p < .001$) all increased the model fit. The full model explained 70% of the variance ($p < .001$; see Table 4).

Table 4

Predicting Moral Exclusion/Scope of Justice

	b	S.E. *b*	β	*b* 95% C.I.	*t*- statistic
Demographics					
Age	0.01	0.01	.12	-0.00, 0.03	1.89
Gender (0 = man, 1 = woman, non-binary)	0.08	0.21	.02	-0.34, 0.50	0.38
Race (0 = White, 1 = BIPOC)	-0.39	0.22	-.11	-0.81, 0.03	-1.81
Income (0 = $41-99k, 1 = < $41k)	0.24	0.23	.07	-0.21, 0.70	1.05
Income (0 = $41-99k, 1 = $100k-249k)	0.21	0.27	.05	-0.32, 0.73	0.77
Income (0 = $41-99k, 1 = $250k+)	0.78	0.74	.06	-0.67, 2.23	1.06
Income (0 = $41-99k, 1 = undisclosed)	-0.23	0.52	-.03	-1.25, 0.78	-0.45
Education (0 = HS, 1 = none, undisclosed)	0.01	0.71	.00	-1.40, 1.41	0.01
Education (0 = HS, 1 = associates, bachelors)	0.16	0.22	.05	-0.28, 0.60	0.72
Education (0 = HS, 1 = masters, doctoral)	0.14	0.31	.03	-0.47, 0.75	0.45
Diet (0 = meat-eater, 1 = restrictions)	-0.18	0.26	-.04	-0.69, 0.32	-0.71
Religion (0 = religious, 1 = not religious)	-0.10	0.20	-.03	-0.50, 0.30	-0.51
Political orientation	0.16	0.06	.17	0.04, 0.29	2.56*
Frequency of interaction	-0.25	0.08	-.19	-0.42, -0.09	-3.09**
Perspective predictors					
Future orientation	0.02	0.03	.05	-0.03, 0.08	0.83
Construal level	0.00	0.39	.00	-0.76, 0.76	0.01
Relational predictors					
SDO	0.03	0.09	.02	-0.16, 0.22	0.32
Sci-fi fan identity	-0.38	0.06	-.36	-0.50, -0.26	-6.17***
Expansive predictors					
Human-centric norms	0.25	0.09	.15	0.07, 0.42	2.81**
Anthropomorphism	-0.01	0.01	-.07	-0.02, 0.01	-1.23
Global citizenship	-0.20	0.08	-.16	-0.36, -0.04	-2.51*
Openness to experience	0.05	0.07	.04	-0.08, 0.19	0.77
Techno-animism	-0.44	0.07	-.38	-0.57, -0.30	-6.46***
Technological predictors					
Affinity for technology	-0.02	0.02	-.04	-0.06, 0.02	-0.97
Substratism	0.56	0.05	.48	0.45, 0.66	10.10***
Human-AI overlap	-0.15	0.05	-.15	-0.24, -0.06	-3.28**
Realistic threat	0.07	0.06	.06	-0.05, 0.19	1.11
Identity threat	0.01	0.07	.01	-0.12, 0.15	0.17
Affective predictors					
Positive emotions	-0.35	0.06	-.29	-0.46, -0.24	-6.23***
Negative emotions	-0.07	0.06	-.05	-0.19, 0.06	-1.06

Note. Values displayed only for new variables added at each step; showing effects controlling for variables entered in previous steps. Uncorrected *p*-values reported. Political orientation ($p = .055$) and global citizenship ($p = .061$) became non-significant with FDR correction. $F(14, 267) = 2.72$, $p = .001$, $R^2 = .13$, adjusted $R^2 = .08$ for demographics; $F(16, 265) = 2.41$, $p = .002$, $R^2 = .13$, adjusted $R^2 = .07$, $\Delta R^2 = .00$ for perspective predictors ($p =$

.707); $F(18, 263) = 4.56$, $p < .001$, $R^2 = .24$, adjusted $R^2 = .19$, $\Delta R^2 = .11$ for relational predictors ($p < .001$); $F(23, 258) = 9.00$, $p < .001$, $R^2 = .45$, adjusted $R^2 = .40$, $\Delta R^2 = .21$ for expansive predictors ($p < .001$); $F(28, 253) = 16.40$, $p < .001$, $R^2 = .64$, adjusted $R^2 = .61$, $\Delta R^2 = .20$ for technological predictors ($p < .001$); $F(30, 251) = 19.30$, $p < .001$, $R^2 = .70$, adjusted $R^2 = .66$, $\Delta R^2 = .05$ for affective predictors ($p < .001$). *$p < .05$, **$p < .01$, ***$p < .001$

Summary of other outcomes. The complete results for the other models are in supplemental results (https://osf.io/w3tcn). The predictors explained 66% of the variance in rights activism ($p < .001$), 63% of the variance in policy support ($p < .001$), and 19% of the variance in donation amount ($p = .004$). Of the temporally-based outcomes, the predictors explained 28% of the variance in current mind perception ($p < .001$), 48% of the variance in current sentience value ($p < .001$), and 59% of the variance in future sentience value ($p < .001$).

5.2 Dimensions of Moral Consideration

Correlations between the moral consideration outcomes showed that they moderately correlated in the expected directions, except for moral expansiveness. Moral expansiveness was weakly correlated with most outcomes in the expected direction and not correlated with current mind perception (see Table 5). Of particular interest, current mind perception only moderately positively correlated with future mind perception. Current and future sentience value also only moderately positively correlated. Only future sentience value, moral expansiveness, scope of justice, and policy support correlated with donation behavior.

Table 5

Correlations between Moral Consideration Outcomes

	1	2	3	4	5	6	7	8	9
1. Current Mind Perception	-								
2. Future Mind Perception	.57***	-							
3. Current Sentience Value	.50***	.39***	-						
4. Future Sentience Value	.33***	.62***	.59***	-					
5. Moral Expansiveness	.08	.16**	.24***	.25***	-				
6. Scope of Justice	-.41***	-.58***	-.59***	-.77***	-.26***	-			
7. AI Rights Activism	.36***	.53***	.52***	.75***	.24***	-.80***	-		
8. AI Policy Support	.43***	.60***	.55***	.72***	.30***	-.81***	.77***	-	
9. Donation Amount	-.03	.09	.06	.12*	.17**	-.15**	.10	.13*	-
Mean	36.38	54.76	1.92	3.15	17.29	5.00	2.67	2.50	0.29
(*S.D.*)	(20.18)	(22.68)	(1.35)	(1.97)	(4.78)	(1.64)	(1.66)	(1.01)	(0.53)

Note. Uncorrected p-values reported. All p-values remained significant with FDR correction. *$p < .05$, **$p < .01$, ***$p < .001$

An exploratory factor analysis (EFA) using maximum likelihood estimation and oblique oblimin rotation explored the latent factor structure of moral consideration. Multiple models are typically explored in EFA, and we explored six models (see Table 6). The six-factor model demonstrated the best fit: $\chi 2$ (490) = 1152.23, $p < .001$, TLI = 0.88, RMSEA = 0.069 (90% confidence interval [C.I.] = 0.064 - 0.074). The scree plot suggested between four and six factors and the parallel analysis suggested six factors (see supplemental results for these visualizations). Future sentience value, scope of justice, rights activism, and policy support items formed the Practical Consideration factor (see Table 7). Seven of the 12 moral expansiveness items formed the Comfortable Expansion factor. The perception of current AIs' emotional capacities and current sentience value formed the Current Emotion factor. The perception of future AIs' emotional capacities formed the Future Emotion factor. The Cognition Perception factor was comprised of the perception of current and future AIs' cognitive capacities. Moral expansiveness to apple trees, robots, and AIs comprised the Novel Expansion factor. The eigenvalues of the six factors ranged from 7.79 to 1.76.

Table 6

Moral Consideration Factor Models

	χ^2	*df*	TLI	RMSEA (90% CI)
Moral Consideration - 1 Factor	4280.10***	665	0.50	0.139 (0.135 - 0.143)
Moral Consideration - 2 Factors	3285.46***	628	0.61	0.122 (0.118 - 0.127)
Moral Consideration - 3 Factors	2673.19***	592	0.68	0.111 (0.107 - 0.116)
Moral Consideration - 4 Factors	2016.08**	557	0.76	0.096 (0.092 - 0.101)
Moral Consideration - 5 Factors	1365.60***	523	0.85	0.075 (0.071 - 0.081)
Moral Consideration - 6 Factors	1152.23***	490	0.88	0.069 (0.064 - 0.074)

Note. χ^2 = chi-square test of model fit; TLI = Tucker-Lewis index; RMSEA = root-mean square error of approximation $^{**}p < .01$, $^{***}p < .001$

Three items cross-loaded: moral expansiveness (robot), moral expansiveness (AI), and the perception of current AIs' self-control. The moral expansiveness (robot, AI) items cross-loaded on the Novel Expansion factor and the Practical Consideration factor. The self-control item cross-loaded on the Current Emotion factor and the Cognition Perception factor. Two items did not load on any factor: moral expansiveness (family members), moral expansiveness (coral reefs). These six factors correspond roughly to three conceptual dimensions: mind perception, psychological expansion, and practical consideration.

Table 7

Moral Consideration Six-Factor EFA Loadings

	Practical Consideration	Comfortable Expansion	Current Emotion	Future Emotion	Cognition Perception	Novel Expansion
CMP 1 (emotions)	0.03	-0.03	**0.89**	0.11	0.00	-0.04
CMP 2 (feelings)	0.04	-0.03	**0.93**	0.07	0.00	-0.01
CMP 3 (emotional)	-0.01	0.01	**0.87**	0.04	0.06	0.01
CMP 4 (self-control)	-0.08	0.02	**0.37**	-0.11	**0.60**	0.13
CMP 5 (analytically)	-0.01	-0.05	0.06	-0.05	**0.69**	0.08
CMP 6 (rational)	0.05	0.02	0.17	-0.14	**0.82**	-0.04
FMP 1 (emotions)	0.06	-0.01	0.05	**0.90**	0.00	-0.03
FMP 2 (feelings)	0.01	0.00	0.08	**0.93**	0.00	0.00

FMP 3 (emotional)	-0.02	0.00	0.09	**0.86**	0.04	0.01
FMP 4 (self-control)	0.05	0.02	-0.04	0.29	**0.49**	-0.03
FMP 5 (analytically)	-0.01	-0.01	-0.15	0.16	**0.67**	-0.01
FMP 6 (rational)	0.07	0.02	-0.17	0.20	**0.74**	-0.11
CSV	0.34	0.01	**0.45**	0.03	-0.06	0.20
FSV	**0.62**	-0.02	-0.06	0.33	-0.02	0.07
MES 1 (family member)	0.00	0.13	-0.18	-0.10	0.09	-0.11
MES 2 (neighbor)	-0.06	**0.70**	-0.06	0.00	-0.02	-0.09
MES 3 (different religion)	-0.04	**0.78**	0.01	0.07	0.04	-0.05
MES 4 (charity worker)	0.01	**0.70**	0.05	-0.08	-0.02	-0.03
MES 5 (refugee)	0.02	**0.68**	-0.15	0.03	0.06	0.00
MES 6 (murderer)	0.01	**0.43**	0.00	-0.03	0.02	-0.03
MES 7 (chimpanzee)	0.09	**0.57**	0.02	-0.11	-0.06	0.26
MES 8 (chicken)	0.02	**0.49**	0.09	0.01	-0.07	0.28
MES 9 (apple tree)	-0.06	0.28	0.11	0.06	-0.03	**0.46**
MES 10 (coral reef)	0.04	0.27	-0.08	0.10	0.06	0.29
MES 11 (robot)	**0.37**	0.13	0.15	0.16	-0.08	**0.35**
MES 12 (AI)	**0.39**	0.16	0.11	0.12	-0.01	**0.37**
SJ 1	**-0.69**	0.07	0.01	-0.08	-0.12	-0.23
SJ 2	**-0.71**	0.08	-0.08	-0.07	0.01	-0.15
SJ 3	**-0.72**	0.07	-0.02	-0.05	-0.01	-0.20
RA 1	**1.00**	0.00	0.01	-0.04	0.00	-0.13
RA 2	**0.93**	0.05	0.05	-0.04	0.00	-0.18
RA 3	**0.87**	0.04	0.07	0.02	-0.05	-0.13
PS 1	**0.53**	-0.05	-0.13	0.05	0.18	0.24
PS 2	**0.57**	-0.02	-0.01	0.03	0.05	0.19
PS 3	**0.56**	0.04	-0.08	0.14	0.08	0.18
PS 4	**0.56**	-0.03	0.07	0.08	0.02	0.19
PS 5	**0.55**	0.01	0.08	0.15	0.08	0.16
PS 6	**0.52**	0.01	0.05	0.18	0.11	0.19

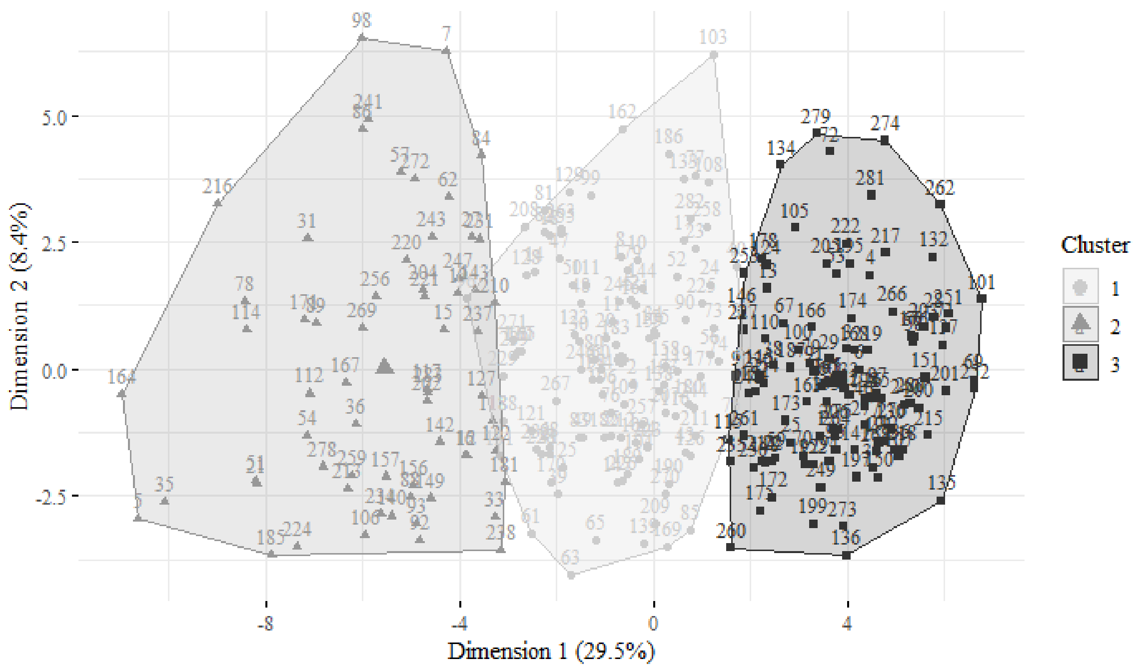
Note. CMP = Current Mind Perception; FMP = Future Mind Perception, CSV = Current Sentience Value; FSV = Future Sentience Value; MES = Moral Expansiveness; SJ = Scope of Justice; RA = Rights Activism; PS = Policy Support. Factor loadings ≥ .35 are in boldface. KMO = 0.91, Bartlett's test of sphericity $\chi 2(703) = 8407.05$, $p < .001$

An exploratory *k*-means cluster analysis of moral consideration outcomes and specific emotions suggested that three clusters, along a gradient, best represented the different groups of participants (see Figure 2). To perform this clustering, the data were first standardized. Then, 2-6 clusters of observations were explored using Euclidean distance estimates of similarity over 10 iterations with 30 multiple initial configurations. The average silhouette method (Rousseeuw, 1987) and the gap statistic (Tibshirani et al., 2001) were employed to assess the number of clusters. The three-cluster model was selected based on the silhouette method results (two clusters), the gap statistic results (eight clusters), and visualization of clusters.

Cluster 2 comprised 63 participants who felt positively towards AIs, perceived more mind in current and future AIs, valued the sentience of current and future AIs, were more morally expansive, scored lower on the scope of justice measure, and were more supportive of AI rights. Cluster 3 comprised 110 participants with the opposite position: lower positive emotions, less mind perception and sentience value, less morally expansive, higher on scope of justice, less supportive of rights. Cluster 1 comprised 110 participants with moral consideration between the high and low inclusion clusters with positive emotions at the midpoint. Moral consideration co-occurred with positive emotions, with people exhibiting degrees of moral consideration consistent with large, moderate, and small moral circles (see supplemental results for means and ANOVAs comparing the three clusters on moral consideration and emotion indices).

Figure 2

Moral Consideration and Emotion Clusters



Note. Three clusters best characterized moral consideration and positive emotions.

5.3 Exploratory Generational Cohort Effects

Do people from different generational cohorts extend different degrees of moral consideration to AIs? We conducted exploratory (not preregistered) one-way ANOVAs with FDR-corrected post-hoc pairwise comparisons to explore moral consideration by generational cohort. We categorized people into four cohorts (Generation Z = 18-25, Millennial = 26-41, Generation X = 42-57, Boomer = 58-76) using Pew Research Center's cohort specifications (Dimock, 2019). The ANOVA results are in Table 8.

Table 8

Generational Cohort ANOVAs

Outcome	Generation Z *M* (SD)	Millennial *M* (SD)	Generation X *M* (SD)	Boomer *M* (SD)	*F*-test
Current Mind Perception	37.26 (18.73)	38.73 (21.25)	30.83 (18.79)	31.45 (19.79)	$F(3, 279) = 2.35$, $p = .073$
Future Mind Perception	59.63 (21.34)[ab]	55.17 (22.55)	49.53 (24.63)[a]	46.00 (20.01)[b]	$F(3, 279) = 3.30$, $p = .021$

Current Sentience Value	1.87 (1.30)	2.17 (1.42)[cd]	1.64 (1.35)[c]	1.25 (0.49)[d]	$F(3, 279) = 4.26, p = .006$
Future Sentience Value	3.61 (1.94)[ef]	3.31 (1.99)[gh]	2.54 (1.90)[eg]	1.84 (1.22)[fh]	$F(3, 279) = 6.94, p < .001$
Moral Expansiveness	17.65 (4.32)	17.08 (5.16)	17.90 (4.91)	15.86 (3.34)	$F(3, 279) = 1.17, p = .324$
Scope of Justice	4.85 (1.45)[i]	4.78 (1.69)[jk]	5.52 (1.77)[ij]	5.68 (1.32)[k]	$F(3, 279) = 4.02, p = .008$
AI Rights Activism	2.96 (1.62)[lm]	2.83 (1.71)[no]	2.21 (1.62)[ln]	1.67 (1.04)[mo]	$F(3, 279) = 5.32, p = .001$
AI Policy Support	2.68 (0.95)[pq]	2.63 (1.03)[rs]	2.20 (1.01)[pr]	1.81 (0.64)[qs]	$F(3, 279) = 6.86, p < .001$
Donation Amount	0.29 (0.50)	0.27 (0.53)	0.32 (0.56)	0.30 (0.61)	$F(3, 279) = 0.15, p = .931$

Note. Superscripts indicate a statistically significant difference exists between the groups. Generation Z $N = 81$, Millennial $N = 131$, Generation X $N = 49$, Boomer $N = 22$

6 Discussion

6.1 Evaluating the Predictors

Substratism, sci-fi fan identity, techno-animism, and positive emotions consistently and strongly predicted moral consideration, especially scope of justice, rights activism, policy support, future mind perception, and future sentience value. In general, predictors like human-AI overlap that are closely and explicitly related to human-AI interactions outperformed other predictors. Global citizenship and human-centric norms were also consistent predictors, highlighting the importance of expansiveness. Anthropomorphism was a less consistent predictor than expected given the importance granted to it in previous research (e.g., Coeckelbergh, 2021; Giger et al., 2019; Li & Sung, 2021; Złotowski et al., 2015). Anthropomorphism only predicted current mind perception and current sentience value, suggesting that individual differences in anthropomorphism may be less important to the moral consideration of AIs as a group than to outcomes not included in this study such as successful interpersonal interactions with robots. The strongest predictor of donation behavior was having a diet with restrictions on animal products. This was unsurprising, given the two organizations donated to most were Mercy For Animals and World Wildlife Fund (see supplemental results for more details on donations).

Some predictors were less impactful than expected: construal level, future orientation, SDO, realistic and identity threats, and negative emotions. SDO was low in this sample ($M = 2.26$, $S.D. = 1.27$), and was less correlated with substratism than expected ($r = .16$, $p = .022$) given previous research showing a moderate positive correlation between SDO and speciesism (Caviola et al., 2019). This might indicate a potential psychological separation between SDO and substrate-based prejudice. Realistic threat's lack of impact was surprising given the literature on AIs as an existential risk (e.g., Bostrom, 2013), which may indicate these risks are not particularly salient in the context of moral and social integration. Identity threat had low reliability that constrained our interpretations of its effects; perhaps identity threat from AIs is not a robust unidimensional construct underpinned by threats to human uniqueness, as previously suggested (Złotowski et al., 2017). Alternatively, perceiving threats from AIs may depend upon a threatening context being activated such as reading about AIs' creativity surpassing humans' creativity or reading about AIs replacing human workers. Mental representations of AIs may be innocuous without a salient threat.

A third possibility stems from the type of relationship that threat has with moral consideration. Hierarchical regressions assume a linear relationship. As preregistered, we also tested whether threat might have a quadratic relationship with moral consideration. A quadratic relationship could show moral patiency at moderate threat, denial of patiency at low threat, and prejudice at high threat. Our analysis failed to substantiate a clear linear or quadratic relationship. Identity threat did exhibit a quadratic relationship with future sentience value, scope of justice, rights activism, and policy support (see supplemental results) but realistic threat had no quadratic relationships. Future research on this relationship with a more extensive suite of threat measures would be beneficial.

The empirical study of substratism and techno-animism have been neglected. Future research on substratism could focus on the pervasiveness of this composition-based prejudice and how interventions such as contact or social norms campaigns might reduce substratism. The extent to which techno-animism affects interactions with and attitudes towards AIs could be the subject of future cross-cultural research on the adoption and treatment of AI technologies around the world. Better understanding techno-animism may improve our understanding of its differing effects from anthropomorphism and deepen our understanding of the impact of animism compared to anthropomorphism.

More research is also needed on the impacts of science fiction imaginaries on moral consideration. Science fiction could increase the moral consideration of AIs because of narrative persuasion, character identification, perspective-taking, or increased social skill building with fictional AIs. Another line of inquiry could question the importance of sci-fi fan identity when AIs are conceptualized in a deliberately non-sci-fi context (e.g., the next version of a home assistance device, Japanese Society 5.0, Chinese New Generation Artificial Intelligence). Science fiction could also affect moral consideration through social identification with future technological worlds. Will sci-fi fan identity matter as we approach these technological futures and AIs become more sophisticated in the real world, perhaps making sci-fi scenarios less unique and compelling?

6.2 Evaluating Moral Consideration

Three conceptual dimensions of moral consideration emerged from the analyses: mind perception, psychological expansion, and practical consideration. Mind perception reflected attributions of cognitive and emotional capacities and depended upon the temporal position of AIs. Psychological expansion reflected a tendency to include human and nonhuman entities in the moral circle, although there was some evidence that expansion to robots and AIs differed from expansion to other nonhumans. Practical consideration encompassed the scope of justice, rights activism, and policy support measures and was defined by stances on fairness, resource sharing, and support for rights. These broad types of moral consideration may conceptually map on to institutional-level moral circle expansion and individual-level moral circle expansion. Scope of justice, rights activism, and policy support might reflect an institutional-level expansion because of their focus on sharing societal resources with AIs. Psychological expansion and mind perception might reflect individual-level moral circle expansion because of their focus on individuals' psychologically-rooted moral boundaries and perceptions of AIs.

We recommend that researchers interested in moral consideration disentangle these different phenomena. The scope of justice, rights activism, and policy support measures might be more useful to researchers seeking to clarify and advance the study of AI rights. Those interested in individuals' moral circle expansion might want to focus on moral expansiveness and donation behavior. Researchers interested in clarifying moral judgments regarding agency and patiency could continue to rely on mind perception.

Given that evaluations of current and future entities were positively related but distinct, we emphasize the need to specify which AIs are being mentally represented by participants, either currently existing or imagined future entities. Mind perception studies with unspecified AI temporality may not generalize between current and future time periods nor be applicable to recommendations for future AI rights policies (see de Graaf et al., 2021 for more on public support for robot rights). This seems especially important to studying moral consideration, given that cognitive capacities were perceived as similar for current and future AIs but emotional capacities were perceived differently. People extend moral consideration when they perceive valenced affective or emotional capacities, making recognition of emotional capacities critical to moral consideration.

Preparation for sentient, socially complex AIs with affective capacities is necessary. Many researchers have argued that we should build conscious, sentient, and socially complex AIs (Giger et al., 2019; Paiva et al., 2018; Sciutti et al., 2018; Oliveira et al., 2018; Admoni & Scassellati, 2017; Wiese et al., 2017; Kanai, 2017), though Metzinger (2021) argued that we should not. Legal frameworks for AIs' rights are one way for institutions to extend practical consideration and prepare for fully autonomous AIs in a society with ubiquitous human-AI relations (Chessman, 2018; Darling, 2012). Studies examining the antecedents to psychological expansion and perceiving AIs' minds could focus on interventions to help individuals recognize the experiential capacities of AIs and consider their inclusion in the moral circle. The results of our study suggest that people may need training to overcome a bias against the existence of minds in contemporary AIs. The exploratory ANOVAs showed generational differences in moral consideration. Generational differences may influence the effectiveness of individual interventions and institutional frameworks like New Generation Artificial Intelligence in China, Industry 4.0, and Society 5.0 in Japan. Generation Z and Millennials were statistically like each other in extending more moral consideration to AIs than Generation X and Boomers who were statistically alike. Mind perception for current AIs, moral expansiveness, and donation behavior were equivalent across generations suggesting that certain aspects of moral consideration may be universal.

Could the moral consideration of AIs generalize? In the cluster analysis, moral consideration and positive emotions appeared as a gradient rather than clearly differentiated clusters. Positive emotions coincided with more moral consideration of AIs. This gradient could suggest that moral consideration of AIs could be increased if positive emotions can be engendered. Moral consideration of this traditionally excluded group may lead to generalized moral consideration of other groups, akin to how attitudes can generalize across groups and stimuli (Pettigrew, 1997; Shook et al., 2007) and how positive emotions broaden and build socio-emotional resources (Fredrickson, 2013; Mesurado et al., 2021). The correlations of donation behavior (primarily to animal-focused organizations) with future sentience value, moral expansiveness, scope of justice, and policy support suggest that moral consideration may generalize from animals to AIs (or vice versa). Although this type of generalization may not have universal benefits (e.g., Gamez-Djokic & Waytz, 2020), we think it is a promising and important avenue for future research. Furthermore, some researchers have suggested that HRI could enhance the moral consideration of animals (Coghlan et al., 2019) and that HRI could facilitate a common human bond that reduces human intergroup prejudices (Jackson et al., 2020).

Finally, future research could pursue the study of robots alongside other AIs as members of the same domain of entities. Continued separation of the two might lead to a false dichotomy of robots and AIs with epistemological and moral consequences. Robots and AIs vary along multiple dimensions such as purpose, embodiment, social complexity, human-likeness, and algorithmic-basis. Despite this diversity, most robots will likely be AIs or possess some degree of AI. Is robotic embodiment a prerequisite for moral consideration? Is the moral consideration extended to robots (i.e., embodied AIs) different from that extended to non-embodied AIs? Should it be? Does the moral consideration of some AIs (e.g., social robots) generalize to the moral consideration of other AIs (e.g., algorithms)? Does the moral consideration of AI subtypes (e.g., social robots, whole brain emulations) build on the same psychological predictors outlined here for AIs as a general group? Answers to questions like these may shape the socio-political and moral relationships between humans and AIs as well as our understanding of how subtypes of AIs relate to each other.

6.3 Limitations

Participants might have had different understandings from each other and from researchers of terms that we did not define within the survey, like "rights" and "suffering," which could affect the interpretation of results. We defined some terms, like "robots/AIs," to establish common ground. We relied on participants' pre-existing understanding of common terms like "suffering." This strategy enabled us to study lay conceptions without imposing researcher definitions which could bias responses to this nascent topic.

The definition of "robots/AIs" we used had limitations. We grouped "robots/AIs" rather than asking participants about a specific AI. HCI and HRI research shows that humans respond differently to different AIs (e.g., Hancock et al., 2021; B. Zhang, 2021). Although our approach lacks this specificity, we believe it innovates by focusing on AIs as a novel social group. This approach may inform legislative and regulative frameworks for all AIs, like those that have been developed for all humans (e.g., "human rights") and animals (e.g., "animal welfare").

We used the term "artificial being" instead of "artificial entity" to increase the concreteness and specificity of the definition. "Artificial entity" commonly refers to "corporation" and is more abstract than "artificial being." "Being" and "entity" are also synonyms, referring to a "thing" or "state" of existence (Pauketat, 2021). Some philosophical conceptions of "being" and "entity" suggest an ontological difference between the two with "being" linking more closely than "entity" to moral status granting concepts like "personhood" and "life" (Geertsema, 2018; Heidegger, 1996; White, 2013).

The current study innovated by asking about moral positions towards AIs and the value of current and future AIs' feelings. We cannot rule out the possibility that participants projected their affect onto these evaluations given the novelty of evaluating AIs' capacities. However, evaluations were made towards AIs as a group rather than a specific AI, limiting the possibility of interpersonal attachment and projected relational warmth. Further, we asked how people feel towards AIs without mentioning specific capacities. These questions targeted participants' emotions towards AIs, reducing the likelihood of diffuse projection. These results establish an important baseline for how people feel toward AIs even without specifying certain capacities.

6.4 Conclusion

The current research contributes twofold to understanding the moral consideration of AIs: investigating psychological predictors and examining the dimensions of moral consideration.

Previously understudied psychological phenomena within HRI and HCI (i.e., substratism, techno-animism, sci-fi fan identity) along with positive emotions had the strongest and most consistent effects. Three conceptual dimensions of moral consideration emerged, implying a need to develop institutional-level practical consideration like legal frameworks and individual-level psychological expansion and mind perception interventions to develop an expansive moral circle and recognize AIs' emotional capacities. These findings lay the foundation for research on the moral consideration of AIs that will underpin necessary preparations for living with socially complex AIs now and increase our capacity to think about and prepare for the possibility of people interacting with sentient AIs in the future.

Acknowledgments. We would like to thank Ali Ladak, Matti Wilks, Nils Köbis, Daniela Waldhorn, and Jamie Harris for their thoughtful feedback on this project at various stages of its design, implementation, analysis, and writing up.

References

1. Abubshait, A., & Wiese, E. (2017). You look human, but act like a machine: Agent appearance and behavior modulate different aspects of human–robot interaction. *Frontiers in Psychology*, *8*, 1393. https://doi.org/10.3389/fpsyg.2017.01393
2. Admoni, H., & Scassellati, B. (2017). Social eye gaze in human-robot interaction: A review. *Journal of Human-Robot Interaction*, *6*(1), 25–63. https://doi.org/10.5898/JHRI.6.1.Admoni
3. Althaus, D., & Gloor, L. (2016, September 14). Reducing risks of astronomical suffering: A neglected priority. *Center on Long-Term Risk*. https://longtermrisk.org/reducing-risks-of-astronomical-suffering-a-neglected-priority/
4. Amiot, C. E., & Bastian, B. (2015). Toward a psychology of human–animal relations. *Psychological Bulletin*, *141*(1), 6–47. https://doi.org/10.1037/a0038147
5. Anderson, J., & Rainie, L. (2022). *Visions of the internet in 2035*. Pew Research Center. https://www.pewresearch.org/internet/2022/02/07/visions-of-the-internet-in-2035/
6. Anderson, J., Rainie, L., & Luchsinger, A. (2018). *Artificial intelligence and the future of humans*. Pew Research Center. https://www.pewresearch.org/internet/2018/12/10/artificial-intelligence-and-the-future-of-humans/
7. Anthis, J. R., & Paez, E. (2021). Moral circle expansion: A promising strategy to impact the far future. *Futures*, *130*, 102756. https://doi.org/10.1016/j.futures.2021.102756
8. 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
9. Athota, V. S., O'Connor, P., & Jackson, C. (2010). The role of emotional intelligence in moral reasoning. In R. Hicks (Ed.), *Personality and Individual Differences: Current Directions* (pp. 105–112). Australian Academic Press. http://www.australianacademicpress.com.au/Publications/Books/4-921513664.html
10. Barth, M., Jugert, P., Wutzler, M., & Fritsche, I. (2015). Absolute moral standards and global identity as independent predictors of collective action against global injustice. *European Journal of Social Psychology*, *45*(7), 918–930. https://doi.org/10.1002/ejsp.2160
11. Bartneck, C., Rosalia, C., Menges, R., & Deckers, I. (2005). Robot abuse – A limitation of the media equation. In A. De Angeli, S. Brahnam, & P. Wallis (Eds.), *Abuse: The darker side of human-computer interaction: An INTERACT 2005 workshop* (pp. 54–57).

12. Bastian, B., Loughnan, S., Haslam, N., & Radke, H. R. M. (2012). Don't mind meat? The denial of mind to animals used for human consumption. *Personality and Social Psychology Bulletin*, *38*(2), 247–256. https://doi.org/10.1177/0146167211424291
13. Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, *3*(21), eaat5954. https://doi.org/10.1126/scirobotics.aat5954
14. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
15. Beran, T. N., Ramirez-Serrano, A., Kuzyk, R., Fior, M., & Nugent, S. (2011). Understanding how children understand robots: Perceived animism in child–robot interaction. *International Journal of Human-Computer Studies*, *69*(7), 539–550. https://doi.org/10.1016/j.ijhcs.2011.04.003
16. Berberich, N., Nishida, T., & Suzuki, S. (2020). Harmonizing artificial intelligence for social good. *Philosophy & Technology*, *33*(4), 613–638. https://doi.org/10.1007/s13347-020-00421-8
17. Birhane, A., & van Dijk, J. (2020). Robot rights? Let's talk about human welfare instead. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 207–213. https://doi.org/10.1145/3375627.3375855
18. Blackmore, S. J. (1999). Meme machines and consciousness. *Journal of Intelligent Systems*, *9*(5–6), 355–376. https://doi.org/10.1515/JISYS.1999.9.5-6.355
19. Blattner, C. E. (2019). The recognition of animal sentience by the law. *Journal of Animal Ethics*, *9*(2), 121–136. https://doi.org/10.5406/janimalethics.9.2.0121
20. Block, R. A. (1990). Models of psychological time. In R. A. Block (Ed.), *Cognitive Models of Psychological Time* (pp. 1–35). Lawrence Erlbaum Associates, Inc.
21. Bostrom, N. (1998). How long before superintelligence? *International Journal of Futures Studies*, *2*, 1.
22. Bostrom, N. (2013). Existential risk prevention as global priority. *Global Policy*, *4*(1), 15–31. https://doi.org/doi: 10.1111/1758-5899.12002
23. Bringsjord, S., Licato, J., Govindarajulu, N. S., Ghosh, R., & Sen, A. (2015). Real robots that pass human tests of self-consciousness. *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 498–504. https://doi.org/10.1109/ROMAN.2015.7333698
24. Broom, D. M. (2020). Brain complexity, sentience and welfare. *Animal Sentience*, *5*(29). https://doi.org/10.51291/2377-7478.1613
25. Bryson, J. J. (2010). Robots should be slaves. In Y. Wilks (Ed.), *Close engagements with artificial companions: Key social, psychological, ethical and design issues* (pp. 63–74). John Benjamins Publishing Company. https://doi.org/10.1075/nlp.8.11bry
26. Buttazzo, G. (2001). Artificial consciousness: Utopia or real possibility? *Computer*, *34*(7), 24–30. https://doi.org/10.1109/2.933500
27. Cabrera, L. (2008). Global citizenship as the completion of cosmopolitanism. *Journal of International Political Theory*, *4*(1), 84–104. https://doi.org/10.3366/E1755088208000104
28. Carmi, N., & Arnon, S. (2014). The role of future orientation in environmental behavior: Analyzing the relationship on the individual and cultural levels. *Society & Natural Resources*, *27*(12), 1304–1320. https://doi.org/10.1080/08941920.2014.928393

29. Caviola, L., & Capraro, V. (2020). Liking but devaluing animals: Emotional and deliberative paths to speciesism. *Social Psychological and Personality Science*, *11*(8), 1080–1088. https://doi.org/10.1177/1948550619893959
30. Caviola, L., Everett, J. A. C., & Faber, N. S. (2019). The moral standing of animals: Towards a psychology of speciesism. *Journal of Personality and Social Psychology*, *116*(6), 1011–1029. https://doi.org/10.1037/pspp0000182
31. Chessman, C. F. (2018). *Not quite human: Artificial intelligence, animals, and the regulation of sentient property* (SSRN Scholarly Paper ID 3200802). Social Science Research Network. https://doi.org/10.2139/ssrn.3200802
32. Chugunova, M., & Sele, D. (2020). *We and it: An interdisciplinary review of the experimental evidence on human-machine interaction* (SSRN Scholarly Paper ID 3692293). Social Science Research Network. https://doi.org/10.2139/ssrn.3692293
33. Coeckelbergh, M. (2010). Robot rights? Towards a social-relational justification of moral consideration. *Ethics and Information Technology*, *12*(3), 209–221. https://doi.org/10.1007/s10676-010-9235-5
34. Coeckelbergh, M. (2011). Humans, animals, and robots: A phenomenological approach to human-robot relations. *International Journal of Social Robotics*, *3*(2), 197–204. https://doi.org/10.1007/s12369-010-0075-6
35. Coeckelbergh, M. (2018). Why care about robots? Empathy, moral standing, and the language of suffering. *Kairos. Journal of Philosophy & Science*, *20*(1), 141–158. https://doi.org/10.2478/kjps-2018-0007
36. Coeckelbergh, M. (2021). Three responses to anthropomorphism in social robotics: Towards a critical, relational, and hermeneutic approach. *International Journal of Social Robotics*. https://doi.org/10.1007/s12369-021-00770-0
37. Coghlan, S., Vetere, F., Waycott, J., & Barbosa Neves, B. (2019). Could social robots make us kinder or crueller to humans and animals? *International Journal of Social Robotics*, *11*(5), 741–751. https://doi.org/10.1007/s12369-019-00583-2
38. Contreras-Masse, R., Ochoa-Zezzatti, A., García, V., Pérez-Dominguez, L., & Elizondo-Cortés, M. (2020). Implementing a novel use of multicriteria decision analysis to select IIoT platforms for smart manufacturing. *Symmetry*, *12*(3), 368. https://doi.org/10.3390/sym12030368
39. Costa, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders*, *6*(4), 343–359. https://doi.org/10.1521/pedi.1992.6.4.343
40. Crimston, C. R., Bain, P. G., Hornsey, M. J., & Bastian, B. (2016). Moral expansiveness: Examining variability in the extension of the moral world. *Journal of Personality and Social Psychology*, *111*(4), 636–653. https://doi.org/10.1037/pspp0000086
41. Crimston, C. R., Hornsey, M. J., Bain, P. G., & Bastian, B. (2018). Toward a psychology of moral expansiveness. *Current Directions in Psychological Science*, *27*(1), 14–19. https://doi.org/10.1177/0963721417730888
42. Darling, K. (2012). *Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects* (SSRN Scholarly Paper ID 2044797). Social Science Research Network. https://doi.org/10.2139/ssrn.2044797
43. Darling, K. (2021). *The new breed: What our history with animals reveals about our future with robots*. Macmillan. https://us.macmillan.com/books/9781250296115/thenewbreed

44. de Graaf, M. M. A., Hindriks, F. A., & Hindriks, K. V. (2021). Who wants to grant robots rights? *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 38–46. https://doi.org/10.1145/3434074.3446911
45. Dhont, K., Hodson, G., & Leite, A. C. (2016). Common ideological roots of speciesism and generalized ethnic prejudice: The social dominance human–animal relations model (SD-HARM). *European Journal of Personality*, *30*(6), 507–522. https://doi.org/10.1002/per.2069
46. Dill-Shackleford, K. E., Vinney, C., & Hopper-Losenicky, K. (2016). Connecting the dots between fantasy and reality: The social psychology of our engagement with fictional narrative and its functional value. *Social and Personality Psychology Compass*, *10*(11), 634–646. https://doi.org/10.1111/spc3.12274
47. Dimock, M. (2019, January 17). *Defining generations: Where Millennials end and Generation Z begins*. Pew Research Center. https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/
48. Dovidio, J. F., Love, A., Schellhaas, F. M. H., & Hewstone, M. (2017). Reducing intergroup bias through intergroup contact: Twenty years of progress and future directions. *Group Processes & Intergroup Relations*, *20*(5), 606–620. https://doi.org/10.1177/1368430217712052
49. Edison, S. W., & Geissler, G. L. (2003). Measuring attitudes towards general technology: Antecedents, hypotheses and scale development. *Journal of Targeting, Measurement and Analysis for Marketing*, *12*(2), 137–156. https://doi.org/10.1057/palgrave.jt.5740104
50. Egbue, O., & Long, S. (2012). Barriers to widespread adoption of electric vehicles: An analysis of consumer attitudes and perceptions. *Energy Policy*, *48*, 717–729. https://doi.org/10.1016/j.enpol.2012.06.009
51. Ema, A., Akiya, N., Osawa, H., Hattori, H., Oie, S., Ichise, R., Kanzaki, N., Kukita, M., Saijo, R., Takushi, O., Miyano, N., & Yashiro, Y. (2016). Future relations between humans and artificial intelligence: A stakeholder opinion survey in Japan. *IEEE Technology and Society Magazine*, *35*(4), 68–75. https://doi.org/10.1109/MTS.2016.2618719
52. Epley, N., & Eyal, T. (2019). Through a looking glass, darkly: Using mechanisms of mind perception to identify accuracy, overconfidence, and underappreciated means for improvement. In J. M. Olson (Ed.), *Advances in Experimental Social Psychology* (Vol. 60, pp. 65–120). Academic Press. https://doi.org/10.1016/bs.aesp.2019.04.002
53. Eyal, T., Sagristano, M. D., Trope, Y., Liberman, N., & Chaiken, S. (2009). When values matter: Expressing values in behavioral intentions for the near vs. distant future. *Journal of Experimental Social Psychology*, *45*(1), 35–43. https://doi.org/10.1016/j.jesp.2008.07.023
54. Farrow, K., Grolleau, G., & Ibanez, L. (2017). Social norms and Pro-environmental behavior: A review of the evidence. *Ecological Economics*, *140*, 1–13. https://doi.org/10.1016/j.ecolecon.2017.04.017
55. Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149
56. Fernando, J. W., O'Brien, L. V., Burden, N. J., Judge, M., & Kashima, Y. (2020). Greens or space invaders: Prominent utopian themes and effects on social change motivation. *European Journal of Social Psychology*, *50*(2), 278–291. https://doi.org/10.1002/ejsp.2607
57. Ferrari, F., Paladino, M. P., & Jetten, J. (2016). Blurring human–machine distinctions: Anthropomorphic appearance in social robots as a threat to human distinctiveness. *International Journal of Social Robotics*, *8*(2), 287–302. https://doi.org/10.1007/s12369-016-0338-y

58. Fink, J. S., Parker, H. M., Brett, M., & Higgins, J. (2009). Off-field behavior of athletes and team identification: Using social identity theory and balance theory to explain fan reactions. *Journal of Sport Management*, *23*(2), 142–155. https://doi.org/10.1123/jsm.23.2.142
59. Fishbein, M., & Ajzen, I. (2009). *Predicting and changing behavior: The reasoned action approach*. Taylor & Francis Group. https://doi.org/10.4324/9780203838020
60. Flynn, F. J. (2005). Having an open mind: The impact of openness to experience on interracial attitudes and impression formation. *Personality Processes and Individual Differences*, *88*(5), 816–826. https://doi.org/DOI: 10.1037/0022-3514.88.5.816
61. Fredrickson, B. L. (2013). Positive emotions broaden and build. In P. Devine & A. Plant (Eds.), *Advances in Experimental Social Psychology* (Vol. 47, pp. 1–53). Academic Press. https://doi.org/10.1016/B978-0-12-407236-7.00001-2
62. Freitas, A. L., Gollwitzer, P., & Trope, Y. (2004). The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts. *Journal of Experimental Social Psychology*, *40*(6), 739–752. https://doi.org/10.1016/j.jesp.2004.04.003
63. Fröding, B., & Peterson, M. (2020). Friendly AI. *Ethics and Information Technology*. https://doi.org/10.1007/s10676-020-09556-w
64. Fujii, T., Guo, T., & Kamoshida, A. (2018). A consideration of service strategy of Japanese electric manufacturers to realize super smart society (SOCIETY 5.0). In L. Uden, B. Hadzima, & I.-H. Ting (Eds.), *Knowledge Management in Organizations* (pp. 634–645). Springer International Publishing. https://doi.org/10.1007/978-3-319-95204-8_53
65. Gambino, A., Fox, J., & Ratan, R. (2020). Building a stronger CASA: Extending the computers are social actors paradigm. *Human-Machine Communication*, *1*, 71–86. https://doi.org/10.30658/hmc.1.5
66. 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
67. Geertsema, M. J. (2018). Heidegger's quest for being: An overview. In M. J. Geertsema, *Heidegger's Poetic Projection of Being* (pp. 17–23). Springer International Publishing. https://doi.org/10.1007/978-3-319-78072-6_3
68. Giger, J., Piçarra, N., Alves-Oliveira, P., Oliveira, R., & Arriaga, P. (2019). Humanization of robots: Is it really such a good idea? *Human Behavior and Emerging Technologies*, *1*(2), 111–123. https://doi.org/10.1002/hbe2.147
69. GlobeScan, & BBC World Service. (2016). *Global citizenship a growing sentiment among citizens of emerging economies: Global poll*. https://globescan.com/wp-content/uploads/2016/04/BBC_GlobeScan_Identity_Season_Press_Release_April%2026.pdf
70. Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: An evolutionary analysis and empirical review. *Psychological Bulletin*, *136*(3), 351–374. https://doi.org/10.1037/a0018807
71. Gordon, J.-S. (2020). What do we owe to intelligent robots? *AI & SOCIETY*, *35*(1), 209–223. https://doi.org/10.1007/s00146-018-0844-6
72. Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the big-five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1
73. Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619–619. https://doi.org/10.1126/science.1134475

74. Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, *96*(3), 505–520. https://doi.org/10.1037/a0013748
75. Gray, K., & Wegner, D. M. (2012a). Morality takes two: Dyadic morality and mind perception. In *The social psychology of morality: Exploring the causes of good and evil* (pp. 109–127). American Psychological Association. https://doi.org/10.1037/13091-006
76. Gray, K., & Wegner, D. M. (2012b). 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
77. 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
78. Groene, S. L., & Hettinger, V. E. (2016). Are you "fan" enough? The role of identity in media fandoms. *Psychology of Popular Media Culture*, *5*(4), 324–339. https://doi.org/10.1037/ppm0000080
79. Gunkel, D. J. (2012). *The machine question: Critical perspectives on AI, robots, and ethics*. MIT Press Direct. https://doi.org/10.7551/mitpress/8975.001.0001
80. Gunkel, D. J. (2018). *Robot rights*. MIT Press. https://mitpress.mit.edu/books/robot-rights
81. Hadarics, M., & Kende, A. (2018). Moral foundations of positive and negative intergroup behavior: Moral exclusion fills the gap. *International Journal of Intercultural Relations*, *64*, 67–76. https://doi.org/10.1016/j.ijintrel.2018.03.006
82. Haikonen, P. O. (2012). *Consciousness and robot sentience* (Vol. 2). World Scientific. https://doi.org/10.1142/8486
83. Hancock, P. A., Kessler, T. T., Kaplan, A. D., Brill, J. C., & Szalma, J. L. (2021). Evolving trust in robots: Specification through sequential and comparative meta-analyses. *Human Factors*, *63*(7), 1196–1229. https://doi.org/10.1177/0018720820922080
84. Harris, J., & Anthis, J. R. (2021). The moral consideration of artificial entities: A literature review. *Science and Engineering Ethics*, *27*(4), 53. https://doi.org/10.1007/s11948-021-00331-8
85. Haslam, N., & Loughnan, S. (2014). Dehumanization and infrahumanization. *Annual Review of Psychology*, *65*(1), 399–423. https://doi.org/10.1146/annurev-psych-010213-115045
86. Heidegger, M. (1996). *Being and time: A translation of sein und zeit*. SUNY Press.
87. Hess, Y., D., Carnevale, J. J., & Rosario, M. (2018). A construal level approach to understanding interpersonal processes. *Social and Personality Psychology Compass*, *12*(8). https://doi.org/10.1111/spc3.12409
88. Ho, A. K., Sidanius, J., Pratto, F., Levin, S., Thomsen, L., Kteily, N., & Sheehy-Skeffington, J. (2012). Social dominance orientation. *Personality and Social Psychology Bulletin*, *38*(5), 583–606. https://doi.org/10.1177/0146167211432765
89. Hutcherson, C. A., & Gross, J. J. (2011). The moral emotions: A social-functionalist account of anger, disgust, and contempt. *Journal of Personality and Social Psychology*, *100*(4), 719–737. https://doi.org/10.1037/a0022408
90. Igartua, J.-J., & Frutos, F. J. (2017). Enhancing attitudes toward stigmatized groups with movies: Mediating and moderating processes of narrative persuasion. *International Journal of Communication*, *11*, 158–177.
91. 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

92. Janoff-Bulman, R., & Werther, A. (2008). The social psychology of respect: Implications for delegitimization and reconciliation. In *The social psychology of intergroup reconciliation* (pp. 145–170). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195300314.003.0008
93. Jensen, C. B., & Blok, A. (2013). Techno-animism in Japan: Shinto cosmograms, actor-network theory, and the enabling powers of non-human agencies. *Theory, Culture & Society*, *30*(2), 84–115. https://doi.org/10.1177/0263276412456564
94. Kakoudaki, D. (2015). Affect and machines in the media. In R. Calvo, S. D'Mello, J. Gratch, & A. Kappas (Eds.), *The Oxford Handbook of Affective Computing*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199942237.013.018
95. Kanai, R. (2017, April 27). We need conscious robots. *Nautilus*. http://nautil.us/issue/47/consciousness/we-need-conscious-robots
96. Kazandzhieva, V., & Filipova, H. (2019). Customer attitudes toward robots in travel, tourism, and hospitality: A conceptual framework. In S. Ivanov & C. Webster (Eds.), *Robots, Artificial Intelligence, and Service Automation in Travel, Tourism and Hospitality* (pp. 79–92). Emerald Publishing Limited. https://doi.org/10.1108/978-1-78756-687-320191004
97. Keeney, J. E., & Wiesenfeld, B. M. (2017). Construal at the interface: Applying construal level theory in organizational research. *Academy of Management Proceedings*, *2017*(1), 17739. https://doi.org/10.5465/AMBPP.2017.17739symposium
98. Keltner, D., & Haidt, J. (2003). Approaching awe, a moral, spiritual, and aesthetic emotion. *Cognition and Emotion*, *17*(2), 297–314. https://doi.org/10.1080/02699930302297
99. Kluckhohn, F. R., & Strodtbeck, F. L. (1961). *Variations in value orientations* (pp. xiv, 450). Row, Peterson.
100. Köbis, N., Bonnefon, J.-F., & Rahwan, I. (2021). Bad machines corrupt good morals. *Nature Human Behaviour*, *5*(6), 679–685. https://doi.org/10.1038/s41562-021-01128-2
101. Köbis, N., & Mossink, L. D. (2021). Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Computers in Human Behavior*, *114*, 106553. https://doi.org/10.1016/j.chb.2020.106553
102. Laakasuo, M., Palomäki, J., & Köbis, N. (2021). Moral uncanny valley: A robot's appearance moderates how its decisions are judged. *International Journal of Social Robotics*. https://doi.org/10.1007/s12369-020-00738-6
103. Ladak, A., Wilks, M., & Anthis, J. R. (2021). *Extending perspective taking to non-human groups [unpublished manuscript]*. Sentience Institute.
104. Laham, S. M. (2009). Expanding the moral circle: Inclusion and exclusion mindsets and the circle of moral regard. *Journal of Experimental Social Psychology*, *45*(1), 250–253. https://doi.org/10.1016/j.jesp.2008.08.012
105. Leach, C. W., Snider, N., & Iyer, A. (2002). "Poisoning the consciences of the fortunate": The experience of relative advantage and support for social equality. In *Relative deprivation: Specification, development, and integration* (pp. 136–163). Cambridge University Press.
106. Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., Ouwerkerk, J. W., & Spears, R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology*, *95*(1), 144–165. https://doi.org/10.1037/0022-3514.95.1.144

107. Ledgerwood, A., Trope, Y., & Chaiken, S. (2010). Flexibility now, consistency later: Psychological distance and construal shape evaluative responding. *Journal of Personality and Social Psychology*, *99*(1), 32–51. https://doi.org/10.1037/a0019843
108. Ledgerwood, A., Trope, Y., Liberman, N., Scott, R. A., & Buchmann, M. C. (2015). Construal level theory and regulatory scope. In *Emerging Trends in the Social and Behavioral Sciences*. John Wiley & Sons, Inc. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118900772.etrds0052
109. Lee, D., & Kim, Y. (2020). 인공지능 로봇과의 비교영역 자기관련성이 사용자의 시기심, 음악 창작물에 대한 평가 및 로봇과의 협업의도에 미치는 영향. 한국콘텐츠학회논문지, *20*(5), 79–89. https://doi.org/10.5392/JKCA.2020.20.05.079
110. Lee, R. B., Baring, R., Maria, M. S., & Reysen, S. (2017). Attitude towards technology, social media usage and grade-point average as predictors of global citizenship identification in Filipino University Students. *International Journal of Psychology: Journal International De Psychologie*, *52*(3), 213–219. https://doi.org/10.1002/ijop.12200
111. Leichtmann, B., & Nitsch, V. (2020). How much distance do humans keep toward robots? Literature review, meta-analysis, and theoretical considerations on personal space in human-robot interaction. *Journal of Environmental Psychology*, *68*, 101386. https://doi.org/10.1016/j.jenvp.2019.101386
112. Leonard, D. J., Mackie, D. M., & Smith, E. R. (2011). Emotional responses to intergroup apology mediate intergroup forgiveness and retribution. *Journal of Experimental Social Psychology*, *47*(6), 1198–1206. https://doi.org/10.1016/j.jesp.2011.05.002
113. Leonardelli, G., Pickett, C., & Brewer, M. (2010). Optimal distinctiveness theory: A framework for social identity, social cognition, and intergroup relations. In M. P. Zanna & J. M. Olson (Eds.), *Advances in Experimental Social Psychology* (Vol. 43, pp. 63–113). Academic Press. https://doi.org/10.1016/S0065-2601(10)43002-6
114. Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*(1), 146–159. https://doi.org/10.1037/0022-3514.81.1.146
115. Leslie, L. M., Bono, J. E., Kim, Y. (Sophia), & Beaver, G. R. (2020). On melting pots and salad bowls: A meta-analysis of the effects of identity-blind and identity-conscious diversity ideologies. *Journal of Applied Psychology*, *105*(5), 453–471. https://doi.org/10.1037/apl0000446
116. Levin, S., Kteily, N., Pratto, F., Sidanius, J., & Matthews, M. (2016). Muslims' emotions toward Americans predict support for Hezbollah and Al Qaeda for threat-specific reasons. *Motivation and Emotion*, *40*(1), 162–177. https://doi.org/10.1007/s11031-015-9510-1
117. Li, B., Chai, X., Hou, B., Zhang, L., Zhou, J., & Liu, Y. (2018). New generation artificial intelligence-driven intelligent manufacturing (NGAIIM). *2018 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, 1864–1869. https://doi.org/10.1109/SmartWorld.2018.00313
118. Li, X., & Sung, Y. (2021). Anthropomorphism brings us closer: The mediating role of psychological distance in User–AI assistant interactions. *Computers in Human Behavior*, *118*, 106680. https://doi.org/10.1016/j.chb.2021.106680

119. Lima, G., Kim, C., Ryu, S., Jeon, C., & Cha, M. (2020). Collecting the public perception of AI and robot rights. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW2), 135:1-135:24. https://doi.org/10.1145/3415206
120. Lin, P., Abney, K., & Bekey, G. A. (2012). *Robot ethics: The ethical and social implications of robotics*. MIT Press.
121. Lomborg, B. (2020). Welfare in the 21st century: Increasing development, reducing inequality, the impact of climate change, and the cost of climate policies. *Technological Forecasting and Social Change*, *156*, 119981. https://doi.org/10.1016/j.techfore.2020.119981
122. Loughnan, S., Haslam, N., & Bastian, B. (2010). The role of meat consumption in the denial of moral status and mind to meat animals. *Appetite*, *55*(1), 156–159. https://doi.org/10.1016/j.appet.2010.05.043
123. Mackie, D. M., & Smith, E. R. (2018). Intergroup emotions theory: Production, regulation, and modification of group-based emotions. In J. M. Olson (Ed.), *Advances in Experimental Social Psychology* (Vol. 58, pp. 1–69). Academic Press. https://doi.org/10.1016/bs.aesp.2018.03.001
124. Marenko, B. (2014). Neo-animism and design: A new paradigm in object theory. *Design and Culture*, *6*. https://doi.org/10.2752/175470814X14031924627185
125. Matthews, G., Hancock, P. A., Lin, J., Panganiban, A. R., Reinerman-Jones, L. E., Szalma, J. L., & Wohleber, R. W. (2021). Evolution and revolution: Personality research for the coming world of robots, artificial intelligence, and autonomous systems. *Personality and Individual Differences*, *169*, 109969. https://doi.org/10.1016/j.paid.2020.109969
126. McFarland, S., Brown, D., & Webb, M. (2013). Identification with all humanity as a moral concept and psychological construct. *Current Directions in Psychological Science*, *22*(3), 194–198. https://doi.org/DOI: 10.1177/0963721412471346
127. McFarland, S., Hackett, J., Hamer, K., Katzarska-Miller, I., Malsch, A., Reese, G., & Reysen, S. (2019). Global human identification and citizenship: A review of psychological studies. *Political Psychology*, *40*(S1), 141–171. https://doi.org/10.1111/pops.12572
128. McFarland, S., Webb, M., & Brown, D. (2012). All humanity is my ingroup: A measure and studies of identification with all humanity. *Journal of Personality and Social Psychology*, *103*(5), 830–853. https://doi.org/10.1037/a0028724
129. McLatchie, N., & Piazza, J. (2017). Moral pride: Benefits and challenges of experiencing and expressing pride in one's moral achievements. In J. A. Carter & C. Gordon (Eds.), *The moral psychology of pride* (pp. 143–167). Rowman & Littlefield.
130. Mentovich, A., Yudkin, D., Tyler, T., & Trope, Y. (2016). Justice without borders: The influence of psychological distance and construal level on moral exclusion. *Personality and Social Psychology Bulletin*, *42*(10), 1349–1363. https://doi.org/10.1177/0146167216659477
131. Mesurado, B., Resett, S., Tezón, M., & Vanney, C. E. (2021). Do positive emotions make you more prosocial? Direct and indirect effects of an intervention program on prosociality in Colombian adolescents during social isolation due to COVID-19. *Frontiers in Psychology*, *12*, 3343. https://doi.org/10.3389/fpsyg.2021.710037
132. Metzinger, T. (2021). Artificial suffering: An argument for a global moratorium on synthetic phenomenology. *Journal of Artificial Intelligence and Consciousness*, 1–24. https://doi.org/10.1142/S270507852150003X
133. Milfont, T. L., Wilson, J., & Diniz, P. (2012). Time perspective and environmental engagement: A meta-analysis. *International Journal of Psychology: Journal International De Psychologie*, *47*(5), 325–334. https://doi.org/10.1080/00207594.2011.647029

134. Miller, D. T., & Prentice, D. A. (2016). Changing norms to change behavior. *Annual Review of Psychology*, *67*(1), 339–361. https://doi.org/10.1146/annurev-psych-010814-015013
135. Misailidou, E. (2017). Social media fandom: The construction of identity in the cases of "The 100" and "Once Upon A Time" Tumblr communities. *2017 Wireless Telecommunications Symposium (WTS)*, 1–11. https://doi.org/10.1109/WTS.2017.7943539
136. Nagahara, M. (2019). A research project of Society 5.0 in Kitakyushu, Japan. *2019 IEEE Conference on Control Technology and Applications (CCTA)*, 803–804. https://doi.org/10.1109/CCTA.2019.8920449
137. Naneva, S., Sarda Gou, M., Webb, T. L., & Prescott, T. J. (2020). A systematic review of attitudes, anxiety, acceptance, and trust towards social robots. *International Journal of Social Robotics*, *12*(6), 1179–1201. https://doi.org/10.1007/s12369-020-00659-4
138. Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 72–78. https://dl.acm.org/doi/pdf/10.1145/191666.191703
139. Ng, S., & Basu, S. (2019). Global identity and preference for environmentally friendly products: The role of personal responsibility. *Journal of Cross-Cultural Psychology*, *50*(8), 919–936. https://doi.org/10.1177/0022022119873432
140. Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2008). Prediction of human behavior in human–robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE Transactions on Robotics*, *24*(2), 442–451. https://doi.org/10.1109/TRO.2007.914004
141. Oatley, K., Keltner, D., & Jenkins, J. M. (2006). *Understanding emotions* (2nd ed.). Blackwell Publishing. https://www.wiley.com/en-ca/Understanding+Emotions%2C+2nd+Edition-p-9781405131025
142. Oliveira, R., Arriaga, P., Alves-Oliveira, P., Correia, F., Petisca, S., & Paiva, A. (2018). Friends or foes?: Socioemotional support and gaze behaviors in mixed groups of humans and robots. *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 279–288. https://doi.org/10.1145/3171221.3171272
143. Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues*, *46*(1), 1–20. https://doi.org/10.1111/j.1540-4560.1990.tb00268.x
144. Opotow, S. (1993). Animals and the scope of justice. *Journal of Social Issues*, *49*(1), 71–85. https://doi.org/10.1111/j.1540-4560.1993.tb00909.x
145. Opotow, S. (1994). Predicting protection: Scope of justice and the natural world. *Journal of Social Issues*, *50*(3), 49–63. https://doi.org/10.1111/j.1540-4560.1994.tb02419.x
146. Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge University Press. https://doi.org/10.1017/CBO9780511571299
147. Owe, A., & Baum, S. D. (2021). Moral consideration of nonhumans in the ethics of artificial intelligence. *AI and Ethics*. https://doi.org/10.1007/s43681-021-00065-0
148. Paiva, A., Mascarenhas, S., Petisca, S., Correia, F., & Alves-Oliveira, P. (2018). Towards more humane machines: Creating emotional social robots. In *New Interdisciplinary Landscapes in Morality and Emotion* (1st ed., pp. 125–139). Routledge. https://doi.org/10.4324/9781315143897-10
149. Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

150. Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, *72*(1), 533–560. https://doi.org/10.1146/annurev-psych-071620-030619
151. Parviainen, J., van Aerschot, L., Särkikoski, T., Pekkarinen, S., Melkas, H., & Hennala, L. (2019). Motions with emotions?: A phenomenological approach to understanding the simulated aliveness of a robot body. *Techné: Research in Philosophy and Technology*, *23*(3), 318–341. https://doi.org/10.5840/techne20191126106
152. Pauketat, J. V. T. (2021). *The terminology of artificial sentience*. PsyArXiv. https://doi.org/10.31234/osf.io/sujwf
153. Pauketat, J. V. T., Mackie, D. M., & Tausch, N. (2020). Group-based meta-emotion and emotion responses to intergroup threat. *British Journal of Social Psychology*, *59*(2), 494–521. https://doi.org/10.1111/bjso.12364e
154. Pedersen, M. A. (2001). Totemism, animism and north Asian indigenous ontologies. *The Journal of the Royal Anthropological Institute*, *7*(3), 411–427.
155. Pettigrew, T. F. (1997). Generalized intergroup contact effects on prejudice. *Personality and Social Psychology Bulletin*, *23*(2), 173–185. https://doi.org/10.1177/0146167297232006
156. Plante, C. N., Roberts, S. E., Reysen, S., & Gerbasi, K. C. (2014). “One of us”: Engagement with fandoms and global citizenship identification. *Psychology of Popular Media Culture*, *3*(1), 49–64. https://doi.org/10.1037/ppm0000008
157. Powell, R., & Mikhalevich, I. (2020). Affective sentience and moral protection. *Animal Sentience*, *5*(29). https://doi.org/10.51291/2377-7478.1668
158. Pratto, F., Liu, J. H., Levin, S., Sidanius, J., Shih, M., Bachrach, H., & Hegarty, P. (2000). Social dominance orientation and the legitimization of inequality across cultures. *Journal of Cross-Cultural Psychology*, *31*(3), 369–409. https://doi.org/10.1177/0022022100031003005
159. Pratto, F., Sidanius, J., & Levin, S. (2006). Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward. *European Review of Social Psychology*, *17*(1), 271–320. https://doi.org/10.1080/10463280601055772
160. Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*(4), 741–763. https://doi.org/10.1037/0022-3514.67.4.741
161. Proctor, H. S., Carder, G., & Cornish, A. R. (2013). Searching for animal sentience: A systematic review of the scientific literature. *Animals*, *3*(3), 882–906. https://doi.org/10.3390/ani3030882
162. Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. ‘Sandy,’ … Wellman, M. (2019). Machine behaviour. *Nature*, *568*(7753), 477–486. https://doi.org/10.1038/s41586-019-1138-y
163. Ray, D. G., Mackie, D. M., Rydell, R. J., & Smith, E. R. (2008). Changing categorization of self can change emotions about outgroups. *Journal of Experimental Social Psychology*, *44*(4), 1210–1213. https://doi.org/10.1016/j.jesp.2008.03.014
164. Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.

165. Ren, F., & Bao, Y. (2020). A review on human-computer interaction and intelligent robots. *International Journal of Information Technology & Decision Making*, *19*(01), 5–47. https://doi.org/10.1142/S0219622019300052
166. Renier, L. A., Schmid Mast, M., & Bekbergenova, A. (2021). To err is human, not algorithmic – Robust reactions to erring algorithms. *Computers in Human Behavior*, 106879. https://doi.org/10.1016/j.chb.2021.106879
167. Reysen, S., & Hackett, J. (2017). Activism as a pathway to global citizenship. *The Social Science Journal*, *54*(2), 132–138. https://doi.org/10.1016/j.soscij.2016.09.003
168. Reysen, S., & Katzarska-Miller, I. (2013). A model of global citizenship: Antecedents and outcomes. *International Journal of Psychology*, *48*(5), 858–870. https://doi.org/10.1080/00207594.2012.701749
169. Richardson, K. (2016). Technological animism: The uncanny personhood of humanoid machines. *Social Analysis*, *60*(1), 110–128. https://doi.org/10.3167/sa.2016.600108
170. Riek, B. M., Mania, E. W., & Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, *10*(4), 336–353. https://doi.org/10.1207/s15327957pspr1004_4
171. Ritchie, H., & Roser, M. (2017, October 2). *Technology adoption*. Our World in Data. https://ourworldindata.org/technology-adoption
172. Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7
173. Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, *76*(4), 574–586. https://doi.org/10.1037//0022-3514.76.4.574
174. Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition & Emotion*, *23*(7), 1259–1283. https://doi.org/10.1080/02699930902809375
175. Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Penguin Books.
176. Russell, S. J., & Norvig, P. (1995). *Artificial intelligence: A modern approach*. Prentice Hall.
177. Schäfer, T., & Schwarz, M. A. (2019). The Meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00813
178. Sciutti, A., Mara, M., Tagliasco, V., & Sandini, G. (2018). Humanizing human-robot interaction: On the importance of mutual understanding. *IEEE Technology and Society Magazine*, *37*(1), 22–29. https://doi.org/10.1109/MTS.2018.2795095
179. Seger, C. R., Banerji, I., Park, S. H., Smith, E. R., & Mackie, D. M. (2017). Specific emotions as mediators of the effect of intergroup contact on prejudice: Findings across multiple participant and target groups. *Cognition & Emotion*, *31*(5), 923–936. https://doi.org/10.1080/02699931.2016.1182893
180. 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

181. Sheela, P. V., & Mannering, F. (2020). The effect of information on changing opinions toward autonomous vehicle adoption: An exploratory analysis. *International Journal of Sustainable Transportation*, *14*(6), 475–487. https://doi.org/10.1080/15568318.2019.1573389
182. Shook, N. J., Fazio, R. H., & Richard Eiser, J. (2007). Attitude generalization: Similarity, valence, and extremity. *Journal of Experimental Social Psychology*, *43*(4), 641–647. https://doi.org/10.1016/j.jesp.2006.06.005
183. Shulman, C., & Bostrom, N. (2021). Sharing the world with digital minds. In S. Clarke, H. Zohny, & J. Savulescu (Eds.), *Rethinking Moral Status*. Oxford University Press. https://nickbostrom.com/papers/digital-minds.pdf
184. Smith, C. S. (2021, February 23). A.I. is everywhere and evolving. *The New York Times*. https://www.nytimes.com/2021/02/23/technology/ai-innovation-privacy-seniors-education.html
185. Smith, E. R., Šabanović, S., & Fraune, M. R. (2021). Human-robot interaction through the lens of social psychological theories of intergroup behavior. *Technology, Mind, and Behavior*, *1*(2). https://doi.org/10.1037/tmb0000002
186. Smith, E. R., Sherrin, S., Fraune, M. R., & Šabanović, S. (2020). Positive emotions, more than anxiety or other negative emotions, predict willingness to interact with robots. *Personality and Social Psychology Bulletin*, *46*(8), 1270–1283. https://doi.org/10.1177/0146167219900439
187. Soderberg, C. K., Callahan, S. P., Kochersberger, A. O., Amit, E., & Ledgerwood, A. (2015). The effects of psychological distance on abstraction: Two meta-analyses. *Psychological Bulletin*, *141*(3), 525–548. https://doi.org/10.1037/bul0000005
188. Sommer, K., Nielsen, M., Draheim, M., Redshaw, J., Vanman, E. J., & Wilks, M. (2019). Children's perceptions of the moral worth of live agents, robots, and inanimate objects. *Journal of Experimental Child Psychology*, *187*, 104656. https://doi.org/10.1016/j.jecp.2019.06.009
189. Song, J. Y. (2021, February 12). *Awe expands our moral worlds: Awe experiences promote moral expansiveness*. SPSP 2021 Annual Convention, Virtual. https://osf.io/ju7fr/
190. Sparkman, D. J., Eidelman, S., & Blanchar, J. C. (2016). Multicultural experiences reduce prejudice through personality shifts in openness to experience. *European Journal of Social Psychology*, *46*(7), 840–853. https://doi.org/10.1002/ejsp.2189
191. Starmans, C., Sheskin, M., & Bloom, P. (2017). Why people prefer unequal societies. *Nature Human Behaviour*, *1*(0082). https://doi.org/10.1038/s41562-017-0082
192. Stellar, J. E., Gordon, A., Piff, P., Cordaro, D., Anderson, C., Bai, Y., Maruskin, L., & Keltner, D. (2017). Self-transcendent emotions and their social functions: Compassion, gratitude, and awe bind us to others through prosociality. *Emotion Review*, *9*(3), 200–207. https://doi.org/10.1177/1754073916684557
193. Stephan, W. G., Ybarra, O., & Morrison, K. R. (2009). Intergroup threat theory. In Nelson, Todd D. (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 43–59). Psychology Press Taylor & Francis Group.
194. Strathman, A., Gleicher, F., Boninger, D. S., & Edwards, C. S. (1994). The consideration of future consequences: Weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology*, *66*(4), 742–752. https://doi.org/10.1037/0022-3514.66.4.742
195. Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *The Behavioral and Brain Sciences*, *30*(3), 299–313; 313–351. https://doi.org/10.1017/S0140525X07001975

196. Tajfel, H. (1974). Social identity and intergroup behaviour. *Social Science Information*, *13*(2), 65–93. https://doi.org/10.1177/053901847401300204
197. Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In Austin, W.G. & Worchel, S. (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Brooks/Cole.
198. Talty, S. (2018). What will our society look like when artificial intelligence is everywhere? *Smithsonian Magazine*, *April*. https://www.smithsonianmag.com/innovation/artificial-intelligence-future-scenarios-180968403/
199. Tausch, N., & Becker, J. C. (2013). Emotional reactions to success and failure of collective action as predictors of future action intentions: A longitudinal investigation in the context of student protests in Germany. *British Journal of Social Psychology*, *52*(3), 525–542. https://doi.org/10.1111/j.2044-8309.2012.02109.x
200. Tausch, N., Becker, J. C., Spears, R., Christ, O., Saab, R., Singh, P., & Siddiqui, R. N. (2011). Explaining radical group behavior: Developing emotion and efficacy routes to normative and nonnormative collective action. *Journal of Personality and Social Psychology*, *101*(1), 129–148. https://doi.org/10.1037/a0022728
201. Taylor, P. M., & Uchida, Y. (2019). Awe or horror: Differentiating two emotional responses to schema incongruence. *Cognition and Emotion*, *33*(8), 1548–1561. https://doi.org/10.1080/02699931.2019.1578194
202. Tegmark, M. (2017). *Life 3.0: Being human in the age of artificial intelligence*. Knopf.
203. Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of Royal Statistical Society: Series B (Statistical Methodology)*, *63*(2), 411–423. https://doi.org/10.1111/1467-9868.00293
204. Tidikis, V., & Dunbar, N. D. (2019). Openness to experience and creativity: When does global citizenship matter? *International Journal of Psychology: Journal International De Psychologie*, *54*(2), 264–268. https://doi.org/10.1002/ijop.12463
205. Toepoel, V. (2010). Is consideration of future consequences a changeable construct? *Personality and Individual Differences*, *48*(8), 951–956. https://doi.org/10.1016/j.paid.2010.02.029
206. Torresen, J. (2018). A review of future and ethical perspectives of robotics and AI. *Frontiers in Robotics and AI*, *4*, 75. https://doi.org/10.3389/frobt.2017.00075
207. Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463. https://doi.org/10.1037/a0018963
208. Tropp, L. R., & Wright, S. C. (2001). Ingroup identification as the inclusion of ingroup in the self. *Personality and Social Psychology Bulletin*, *27*(5), 585–600. https://doi.org/10.1177/0146167201275007
209. Tsay-Vogel, M., & Sanders, M. S. (2017). Fandom and the search for meaning: Examining communal involvement with popular media beyond pleasure. *Psychology of Popular Media Culture*, *6*(1), 32–47. https://doi.org/10.1037/ppm0000085
210. Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Blackwell.
211. Vallacher, R. R., & Wegner, D. M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology*, *57*(4), 660–671. https://doi.org/10.1037/0022-3514.57.4.660
212. van de Ven, N. (2017). Envy and admiration: Emotion and motivation following upward social comparison. *Cognition and Emotion*, *31*(1), 193–200. https://doi.org/10.1080/02699931.2015.1087972

213. Van Leeuwen, E., van Dijk, W., & Kaynak, U. (2013). Of saints and sinners: How appeals to collective pride and guilt affect outgroup helping. *Group Processes & Intergroup Relations*, *16*(6), 781–796. https://doi.org/10.1177/1368430213485995
214. Vanman, E. J., & Kappas, A. (2019). “Danger, Will Robinson!” The challenges of social robots for intergroup relations. *Social and Personality Psychology Compass*, *13*(8), e12489. https://doi.org/10.1111/spc3.12489
215. Venaik, S., Zhu, Y., & Brewer, P. (2013). Looking into the future: Hofstede long term orientation versus GLOBE future orientation. *Cross Cultural Management: An International Journal*, *20*(3), 361–385. https://doi.org/10.1108/CCM-02-2012-0014
216. 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
217. Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037//0022-3514.54.6.1063
218. 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
219. 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
220. 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
221. White, F. J. (2013). Personhood: An essential characteristic of the human species. *The Linacre Quarterly*, *80*(1), 74–97. https://doi.org/10.1179/0024363912Z.00000000010
222. Wiese, E., Metta, G., & Wykowska, A. (2017). Robots as intentional agents: Using neuroscientific methods to make robots appear more social. *Frontiers in Psychology*, *8*(1663). https://doi.org/10.3389/fpsyg.2017.01663
223. Wilkinson, D. (2017). Is there such a thing as animism? *Journal of the American Academy of Religion*, *85*(2), 289–311. https://doi.org/10.1093/jaarel/lfw064
224. Wu, F., Lu, C., Zhu, M., Chen, H., Zhu, J., Yu, K., Li, L., Li, M., Chen, Q., Li, X., Cao, X., Wang, Z., Zha, Z., Zhuang, Y., & Pan, Y. (2020). Towards a new generation of artificial intelligence in China. *Nature Machine Intelligence*, *2*(6), 312–316. https://doi.org/10.1038/s42256-020-0183-4
225. Yogeeswaran, K., & Dasgupta, N. (2014). The devil is in the details: Abstract versus concrete construals of multiculturalism differentially impact intergroup relations. *Journal of Personality and Social Psychology*, *106*(5), 772–789. https://doi.org/10.1037/a0035830
226. 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. https://doi.org/10.5898/JHRI.5.2.Yogeeswaran
227. Zhang, B. (2021). *Public opinion toward artificial intelligence*. OSF Preprints. https://doi.org/10.31219/osf.io/284sm
228. Zhang, D., Maslej, N., Brynjolfsson, E., Etchemendy, J., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Sellitto, M., Sakhaee, E., Shoham, Y., Clark, J., & Perrault, R. (2022). *The AI*

index 2022 annual report. Stanford University. https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf
229. Zimbardo, P. G., & Boyd, J. N. (1999). Putting time in perspective: A valid, reliable individual-differences metric. *Journal of Personality and Social Psychology*, *77*(6), 1271–1288. https://doi.org/10.1037/0022-3514.77.6.1271
230. Złotowski, J., Proudfoot, D., Yogeeswaran, K., & Bartneck, C. (2015). Anthropomorphism: Opportunities and challenges in human–robot interaction. *International Journal of Social Robotics*, *7*(3), 347–360. https://doi.org/10.1007/s12369-014-0267-6
231. 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