

Children's Conceptions of the Moral Standing of a Humanoid Robot of the Here and Now

Solace Shen

A dissertation

submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2015

Reading Committee:

Peter H. Kahn, Jr., Chair

Takayuki Kanda

Betty M. Repacholi

Program Authorized to Offer Degree:

Department of Psychology

© Copyright 2015

Solace Shen

University of Washington

Abstract

Children's Conceptions of the Moral Standing of a Humanoid Robot of the Here and Now

Solace Shen

Chair of the Supervisory Committee:
Professor Peter H. Kahn, Jr.
Department of Psychology

Sophisticated humanoid robots have recently moved from the laboratory and research settings to the home environment. Some models are now marketed to families with children, and are designed to engage children in increasingly social and potentially moral interactions. The purpose of this study is to investigate—when children interact with a commercially available humanoid robot and witness a human causing a harm to the robot—whether children conceptualize the robot as being an entity that deserves moral consideration: what is referred to in this study as having *moral standing*. Participants included 120 children in 2 age groups (8-9 and 14-15). To assess the effects of the robot's physical embodiment on participants' conceptions of the robot's moral standing, 30 participants from each age group (gender balanced) were randomly assigned to 1 of 2 conditions and interacted with either a humanoid robot or an analogous virtual agent. In each condition, the interaction culminated in a confederate hitting the robot/virtual agent with a book. Each participant was then engaged in a semi-structured interview that ascertained their judgments and reasoning regarding the robot/virtual agent and three comparison entities. Results show that the majority of the participants judged the confederate hitting the robot as a violation of moral obligation, and many brought the concept of artificial emotion to bear in their reasoning. Participants were significantly more likely to judge

it a violation of moral obligation to hit the robot than the virtual agent. Moreover, the 8- to 9-year-olds were significantly more likely than the 14- to 15-year-olds to judge it a violation of moral obligation to hit either the robot or the virtual agent. Finally, participants' conceptions of the robot's moral standing largely showed a unique composition of moral features when compared to those of the comparison entities. Discussion addresses the broader implications of these findings and future directions for research are offered.

Table of Contents

List of Figures	iii
List of Tables	iv
Chapter I: Introduction.....	1
Child-Robot Interaction.....	3
Children’s Conceptions of Personified Robots	6
Assessment of Moral Obligation.....	11
Effects of a Robot’s Physical Embodiment in Human-Robot Interaction	16
Effects of Age on Children’s Conceptions of the Moral Standing of Robots.....	20
Personified Robots as a New Ontological Category	24
Aims of the Present Research.....	26
Chapter II: Method.....	28
Participants	28
Equipment	28
Procedure.....	31
Measures.....	36
Coding and Reliability	40
Chapter III: Results	45
Judgments and Reasoning About the Hitting Act	45
Endorsements of Rights	70
Between Entities Comparisons.....	74
Methodological Checks.....	87
Chapter IV: Discussion.....	92

Moral Obligation to a Humanoid Robot of the Here and Now	92
Effects of Physical Embodiment on Conceptions of a Humanoid Robot’s Moral Standing	97
Effects of Age on Conceptions of a Humanoid Robot’s Moral Standing	104
Children’s Conceptions of the Partial Moral Standing of a Humanoid Robot.....	109
Limitations	112
Broader Implications & Future Directions.....	115
Conclusion.....	120
References.....	122
Appendix A: Participant Demographic and Background Questionnaire	134
Appendix B: Interaction Protocol	137
Appendix C: Interview Protocol	141
Appendix D: Reasoning Coding Manual	147
Appendix E: Verbal Interaction Coding Manual	211

List of Figures

1. Personified Technologies.....	2
2. Humanoid Robot & Humanoid Virtual Agent.....	30
3. Profiles of the 8- to 9-Year-Olds' Conceptions of Each Entity's Moral Standing.....	85
4. Profiles of the 14- to 15-Year-Olds' Conceptions of Each Entity's Moral Standing.....	86

List of Tables

1. Summary of Participant Demographic Information.....	29
2. Moral Justification Categories Regarding the Entity for the Assessment of Moral Obligation.....	41
3. Percentage (Raw Frequency) of “Not All Right” Evaluations and Moral Obligatory Reasoning Based on Moral Domain Criteria for the Hitting Act.....	47
4. Percentage of Moral Justification Use by Participants Providing Moral Obligatory Reasoning.....	52
5. Mean Rating (Standard Deviation) of Seriousness of the Hitting Act.....	60
6. Mean Score (Standard Deviation) of Technology Experience Scales.....	66
7. Percentage of Participants Who Affirmed Moral Standing With Regard to Owning & Selling.....	71
8. Percentage of Participants Who Affirmed Moral Standing With Regard to Experimentation.....	72
9. Mean Difference Between Entities in Endorsements of Moral Obligation Against the Hitting Act.....	75
10. Mean Difference Between Entities in Seriousness Ratings for the Hitting Act.....	77
11. Mean Difference Between Entities in Endorsements of the Right to be Free from Ownership.....	79
12. Mean Difference Between Entities in Endorsements of the Right to be Free from Experimentation.....	82

Acknowledgements

It takes a small village to complete a doctoral thesis. I would like to express my appreciation for the remarkable community of people that have made this research possible.

First and foremost, I would like to extend my deepest gratitude to my advisor and mentor, Dr. Peter H. Kahn, Jr. His astute vision of the two powerful trends—the degradation of the natural world and the unprecedented technological development—that are reshaping the human experience both inspired me to become a researcher in the field and this work. Throughout my doctoral education, Peter had continually challenged me to refine my skills and grow intellectually, while being a stable source of support and encouragement as I faced those challenges. My experience working with Peter has taught me to think big and deep, to fuse passion and rigor in my research, and to always speak for and hold out the most profound expressions of humanity.

I would also like to extend my sincere gratitude to my supervisory committee members: Dr. Takayuki Kanda, Dr. Betty M. Repacholi, and Dr. Ryan Calo. I am grateful for the generosity of their time, and have benefited greatly from their expertise and insights. Their constructive questions and comments have in important ways shaped and strengthened this work and my development as an independent scholar.

This research was conducted with the support of a team of University of Washington undergraduate research assistants: Kevin Ahn, Courtney Baker, Charlotte Barker, Diana Blanco, Adelle Dinkelman, Elizabeth Flores, Jeremy Kim, Stephanie Lam, Mindy Rohrback, Brody Sakazaki, Lyle Summers, Amy Tang, Hsin-Jung Wang, Tanya Wickham, and Xinglu Yao. I am thankful for and impressed by their intellectual curiosity, responsibility, and dedication.

This research was supported, in part, by the University of Washington Bridge Funding Program. This research was in addition supported, in part, by the University of Washington Psychology Department Scholar Fellowship.

Last, but certainly not the least, I would like to extend my gratitude to my colleagues and friends, whose support and love has sustained me on multiple levels. I would like to thank Colin Beam at the University of Washington Center for Social Science Computation Research for his tireless assistance and friendship. I would also like to thank Joshua Peschel for his comradeship and for pointing me to a robot that's just right when I had no robot. In closing, my wholehearted thanks for the support of the incredibly intelligent, genuine, and caring women that I have the pleasure to be friends with: Jolina Ruckert, Amanda Gilmore, Cady Stanton Laundry, Camille Biggins, and Kayla Upshaw.

Dedication

To my mother Youfeng,
whose steadfast and unconditional love
has always been the psychological safe base
from which I venture to take flight.

To my son Ethan,
whose equanimity, patience, and humor
is my constant source of inspiration and solace.

Chapter I: Introduction

Personified technologies are computational entities designed with a persona and a repertoire of social characteristics and behaviors to engage people in social interactions (see Figure 1 for examples). They come in the forms of physically embodied robots (e.g., NAO), graphically embodied virtual agents (e.g., GRETA), or disembodied voice interfaces (e.g., Siri). Some forms of personified technologies have been commercially available to the public for years (such as robot pets and voice interfaces). However, sophisticated humanoid robots (robots that look, talk, and act like humans) have only recently moved from the laboratory and research settings to the home environment. Some models of these humanoid robots are now marketed to families with children, and are designed to engage children in increasingly rich social and potentially even moral interactions. The purpose of this study is to investigate—when children interact with a commercially available humanoid robot and witness a human causing a harm to the robot—whether children conceptualize the robot as being an entity that deserves moral consideration: what is referred to in this study as having *moral standing*.

This introduction reviews relevant background literature and concepts and is organized in the following way. First, an overview of child-robot interaction research is provided to characterize the quality of children's social interactions with robots. This is followed by a discussion on children's conceptions of robots' biological, mental, social, and moral properties. Both animal-like and human-like robots are discussed in these first two sections, as results from research across these two robot forms reveal important patterns in how children relate to personified robots. A method for assessing attributions of moral standing is then introduced. Next, previous research on two factors that can affect one's perceptions of personified technologies (i.e., the physical embodiment of the robot, and the age of the person) is reviewed.

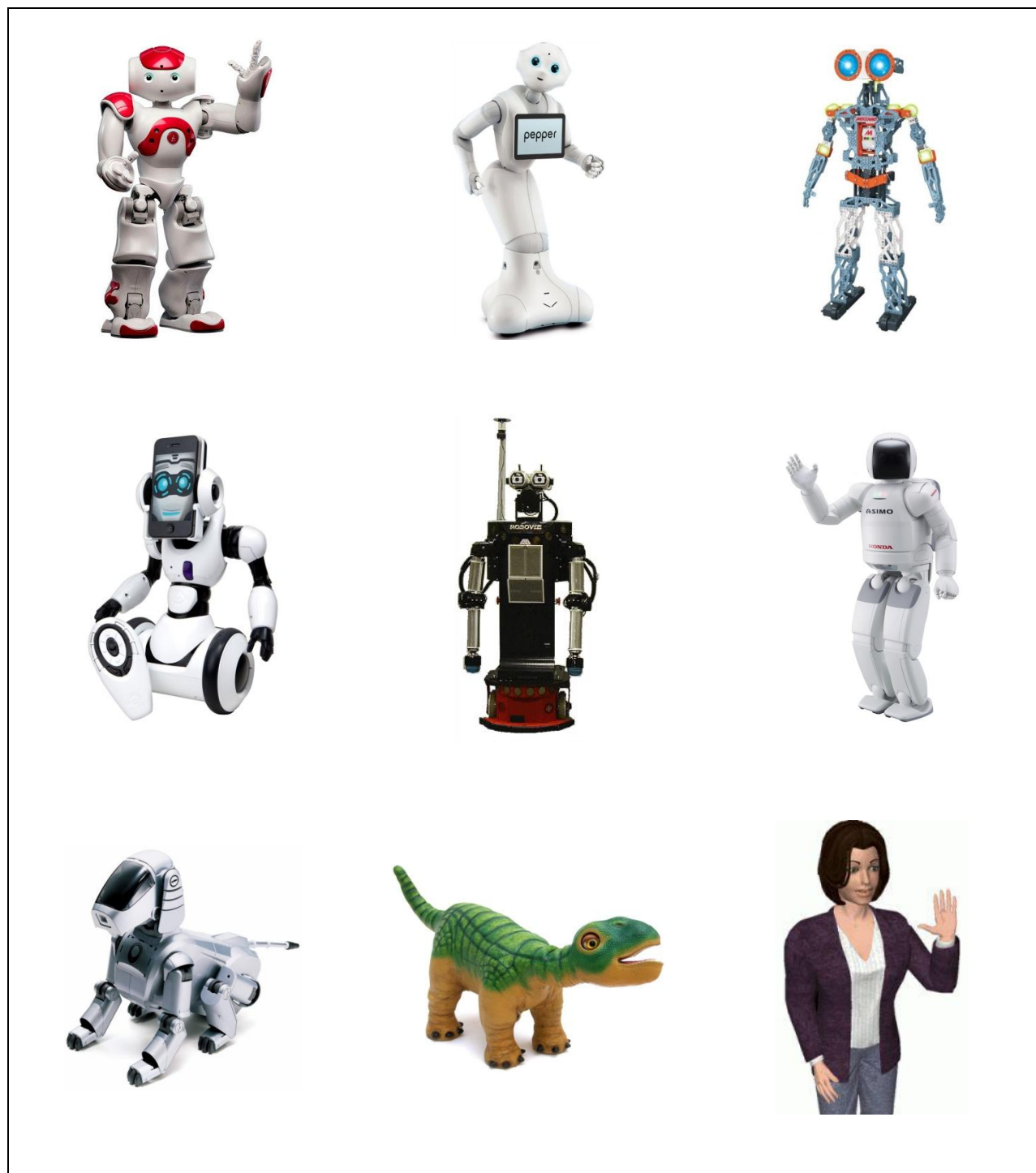


Figure 1. Personified technologies: humanoid robots—NAO (top left), Pepper (top center), Meccanoid G15 (top right), RoboMe (middle left), Robovie (middle center), and ASIMO (middle right); animal robots—AIBO (bottom left) and Pleo (bottom center); and virtual agent—GRETA (bottom right).

Finally, this introduction discusses the proposition that personified technologies represent the emergence of a new genre of beings with their own unique constellation of non-moral and moral features.

Child-Robot Interaction

Children construct moral knowledge through their interactions with the various living and nonliving, natural and artificial entities that constitute their environment (Kohlberg, 1984; Nucci, 2001; Piaget, 1932/1997; Smetana, 2006, Turiel, 1983). To better understand how children morally reason about robots, it is therefore useful to first look at the quality of children's social interactions with robots. There is ample evidence from the field of human-robot interaction that children interact socially with animal robots in ways that resemble human-animal interaction (e.g., Kahn, Friedman, Pérez-Granados, & Freier, 2006; Melson et al, 2005), and with humanoid robots in ways that resemble human-human interaction (e.g., Kahn et al., 2012; Kanda, Sato, Saiwaki, & Ishiguro, 2007; Tanaka, Cicourel, & Movellan, 2007). Children can even learn and benefit from interactions with robots. For example, interaction with an English speaking humanoid robot, Robovie (Figure 1), has resulted in improved English knowledge in Japanese elementary school children over a two week field study (Kanda, Hirano, Eaton, & Ishiguro, 2004). A group of 4- to 10-year-olds also collaboratively learned how to set a table with a humanoid robot ASIMO (Figure 1) as a learning partner (Okita, Ng-Thow-Hing, & Sarvadevabhatla, 2009). Outside of the learning context, interaction with Aldebaran's humanoid robot NAO (Figure 1) has been shown to be effective in helping children manage pain and anxiety during medical procedures (Beran, Ramirez-Serrano, Vanderkooi, & Kuhn, 2013).

As discussed at the beginning, interactions with these types of humanoid robots are no longer restricted to laboratory and research settings. Models of these robots have now become

available to the consumer market. For instance, Aldebaran's popular research platform, NAO, became publicly available in 2011, with its price substantially reduced from \$15,999 to \$7,990. More recently, in collaboration with the Japanese telecom giant Softbank, Aldebaran unveiled its newest humanoid robot, Pepper (Figure 1). Designed specifically to live with families, Pepper is capable of detecting basic human emotions and responding accordingly, and will purportedly learn and evolve through a cloud-based artificial intelligence system (Mazza, 2014). This summer (June, 2015), 1000 units of Pepper, each priced at ¥198,000 (\$1,611), sold out within the first minute of sales in Japan (Martin, 2015). Less sophisticated models of humanoid robots, with some leveraging the powerful processors available in smartphones and devices, have also been marketed to children and are commercially available for only hundreds of dollars (e.g., Spin Master's Meccanoid G15, WooWee's RoboMe; Figure 1).

With this increased accessibility to humanoid robots, long-term child-robot interaction becomes an imminent reality. While there can be variations depending on the robots' specific purposes and capabilities, there is evidence from the human-robot interaction literature that children can sustain long-term social interaction with robots. Kanda, Sato, Saiwaki, and Ishiguro (2007) observed interactions between children in a Japanese elementary school and Robovie over a two month field trial, and found that children behaved in various social ways with Robovie (e.g., hugging, talking, playing), and many established friendship-like relationships with Robovie and continued to interact with the robot over the entire two months. In another study, Tanaka, Cicourel, & Movellan (2007) found that when a humanoid robot, QRIO, was immersed in a daycare center classroom over a period of 5 months, a group of 18- to 24-month-olds exhibited a variety of social and care-taking behaviors toward the robot (e.g., giving items to the robot, putting a blanket on the robot). Over the 5 months period, the toddlers progressively engaged

with the robot in the sort of behaviors that they exhibited toward their toddler peers (e.g., hugging, care-taking), which differed from the sort of behaviors they exhibited toward a control toy robot (e.g., rough-housing).

Several other studies also compared children's behaviors toward robots with their behaviors towards comparison entities. For example, in one study, 80 preschool children interacted with both Sony's robotic dog AIBO and a stuffed dog (Kahn, Friedman, Pérez-Granados, & Freier, 2006). Children's reciprocal behaviors that not only responded to the entity but also expected the entity to respond back were coded (e.g., offering, verbal directive). Results showed that children engaged in more attempts at reciprocity with AIBO (683 occurrences) than with the stuffed dog (180 occurrences). At the same time, children more often mistreated (e.g., rough-handling, throwing) the stuffed dog (184 occurrences) than they did AIBO (39 occurrences). In another study, 72 children ages 7-15 interacted with AIBO and a live Australian Shepherd (Melson et al., 2005). Results showed that children spend more time both in close proximity and in physical contact with the live dog than they did with AIBO, thus exhibiting more physical closeness with the live dog than with the robot dog. In a third study (Shahid, Krahmer, & Swerts, 2014), 256 children (ages 8-12) played a card guessing game on a screen in one of three conditions: alone, with a friend, or with a teleoperated robot called iCat (which has a creature-like appearance but speaks). They found that children exhibited more expressive nonverbal behaviors (e.g., smile, frown, laugh, clap, jump, etc.) when they played with the robot than when they played alone, but less than when they played with a friend.

In summary, research on child-robot interaction has shown that children not only engage in social behaviors when interacting with robots, they can also learn and benefit in some ways from these interactions. In certain situations, it is possible for children to derive psychological

comfort from interacting with robots. Children can even sustain some form of long-term interaction with robots. Behaviorally, children differentiate robots from their artifactual counterparts in that they tend to behave more socially with robots than they do with traditional artifacts such as toys. However, children also behaviorally differentiate robots from their biological counterparts in that they tend to behave more socially with live animals and humans than they do with robots.

Children's Conceptions of Personified Robots

Whether an entity is granted moral standing is often based on considerations of the entity's cognitive, perceptual, and psychological capabilities (e.g., whether the entity can think, whether the entity can feel pain, whether the entity can be emotionally affected). This is evident in folk moral psychology, that is, how ordinary people reason about moral issues (e.g., it's not morally wrong to kick a door because the door can't feel, but it is morally wrong to kick a person because the person can be hurt physically and emotionally). The connections between these concepts have also been shown in experimental research on people's moral reasoning regarding non-human entities such as animals (Knobe & Prinz, 2008; Sytsma & Machery, 2012) and aliens (Sytsma & Machery, 2012). A review of previous research on children's conceptions of robots' general properties is therefore informative before turning to children's conceptions of robots' moral standing.

A number of studies have investigated how children reason about and categorize "intelligent artifacts," such as robots, that straddle the traditional boundaries between living and nonliving (e.g., Berstein & Crowley, 2008; Jipson & Gelman, 2007; Kahn et al., 2012; Kahn, Friedman, Pérez-Granados, & Freier, 2006; Melson et al., 2005; Scaife & van Duuren, 1995; Severson, 2010; Somanader, Saylor, & Levin, 2011). Results overall indicate that by age 7,

children consistently distinguished robots from prototypical living things (e.g., humans, animals) in that they tended to not attribute biological properties to robots (e.g., eat, grow). However, when considering robots' mental properties, participants ages 3 through 15 were all more willing to attribute cognitive, perceptual, and psychological properties (e.g., think, see, be happy) to robots than to prototypical non-living things (e.g., chair, toy car), and the extent to which they did so often fell somewhere in between that of prototypical living and non-living things.

A methodological limitation in some of the above studies is that participants were not always given an opportunity to interact with the robot under investigation, and instead, pictures or videos of robots were employed as stimuli for participants' reasoning (e.g., Berstein & Crowley, 2008; Jipson & Gelman, 2007). Having a chance to interact with the robot under investigation serves to ground participants' reasoning on experiences of the robot's actual capabilities and affordances, and is an important consideration for investigations of more complex social and moral reasoning regarding robots. A body of emerging research that examined children's conceptions of robots' general as well as social and moral properties following their interactions with the robots is discussed in more details next. Three studies focused on animal-like robots, and one study focused on a humanoid robot.

Personified animal robots. In the Kahn, Friedman, Pérez-Granados, and Freier (2006) study discussed earlier, 80 preschool children interacted with Sony's robotic dog AIBO and a stuffed dog, and were then interviewed about their judgments of both entities. Averaging across question types, results showed that while only a minority of the children judged that AIBO was alive (38%) and had biological properties (e.g., have a stomach, can grow; 46%), a majority of them appeared to believe that AIBO was the sort of entity with which they could have a meaningful human-animal relationship. Specifically, 66% of the children accorded AIBO with

perceptual and psychological properties (e.g., can see, hear, be happy), 76% with sociality (e.g., can be a friend, want to spend time with), and 63% with moral standing (e.g., it is not ok to hit AIBO, it is not ok to throw AIBO in the garbage). However, these preschoolers also reasoned similarly about the stuffed dog. Nevertheless, as discussed earlier, children behaviorally engaged in more reciprocal interactions with AIBO than with the stuffed dog, suggesting that children perceived that AIBO was more capable of responding reciprocally than the stuffed dog.

In the Melson et al. (2005) study also discussed earlier, 72 children ages 7 to 15 interacted with AIBO and a live Australian Shepherd, and were then interviewed about their judgments of the two entities. All children said that the Australian Shepherd was alive, while only 21% said that AIBO was alive. Averaging across question types, more children affirmed biological, perceptual, psychological, social, and moral attributes to the live dog than they did AIBO. That said, the majority of the children nonetheless affirmed that AIBO had perceptual and psychological properties (56%), sociality (70%), and moral standing (76%), even while stating that AIBO did not have biological properties (22%).

In a third study, Severson (2010) investigated 7- to 10-year olds' conceptions of a robotic dinosaur, Pleo, in comparison to four other entities (i.e., dog, stuffed animal, tree, and computer). Fifty two children interacted with Pleo and were then interviewed about their judgments of each entity. Averaging across question types, the majority of the children affirmed perceptual (79%), psychological (60%), social (78%), and moral (55%) attributes to the robot, while only a minority of the children accorded biological properties (6%) to the robot and said that Pleo was a living thing (29%). In relation to the comparison entities, children's conceptions of the robot were mostly situated between a prototypical animal (i.e., dog) and a prototypical artifact (i.e., stuffed animal).

One question that is often raised about findings such as those above is whether children really believe robots are social and moral beings, or are they just pretending. Severson (2010) sought to address this question by examining whether children's attributions in her study represented their actual or pretend beliefs. To assess this, children were directly asked following each attribution if they had to pretend that was the case. Results from this analysis indicate that pretense did not account for a significant portion of children's overall attributions to the robot dinosaur. However, the author was surprised to find that children's engagement in pretense did not significantly differ across the five entities. In other words, pretense did not account for a significant portion of children's attributions to any of the five entities, not even the stuffed animal (a common candidate for children's imaginative play). Correspondingly, children's attributions to the stuffed animal were very low overall, reflecting the fact that children did not actually believe the stuffed animal had these attributes. The author concluded that the method used in the study likely did not tap into children's tendency to engage in pretense with either the stuffed animal or the robot, and children thus provided responses reflecting their actual beliefs for all entities, including the robot.

Results across these studies suggest that children conceptually differentiate animal-like robots from live animals, in that they are largely unwilling to attribute aliveness and biological properties to the robots. Yet at the same time, they also conceptually differentiate robots from artifacts, in that they are more willing to grant perceptual and psychological properties to the robots. This same pattern of findings is also reflected in the categorization literature mentioned at the beginning of this section. Furthermore, children relate to animal robots in some ways as social and moral entities. To a certain extent, they see animal robots as companions and friends

and believe it is not all right to harm them. By age 7, children's social and moral conceptions of animal robots were mostly situated between those of live animals and artifacts.

Personified humanoid robots. Kahn et al. (2012) investigated children's conceptions of a humanoid robot. Ninety participants ages 9 to 15 each interacted with ATR's Robovie in a 15-minute session that ended when an experimenter interrupted Robovie's turn in a game of "I Spy" with the participant. Then, against Robovie's stated moral objections that it was unfair to Robovie, and that it caused psychological harm to Robovie (i.e., Robovie's feelings were hurt and Robovie was afraid of the dark), the experimenter verbally commanded and made Robovie go into a closet. Each participant was then interviewed about their reasoning regarding Robovie and the potential moral violation that they had witnessed.

Interview results showed that the majority of the participants affirmed that Robovie was intelligent (79%), had interests (76%), could be sad (64%), and had feelings (60%). Participants were also asked whether they thought Robovie was a living being, 14% said yes, 48% said no, and interestingly, 38% were unwilling to commit to either categories and talked in various ways of Robovie being in between living and not living. The majority of the participants also conceived of Robovie as a social other insofar as they said that they enjoyed having Robovie show them around the lab (89%), might like to spend time with Robovie if they were lonely (84%), might go to Robovie for comfort if they were sad (55%), would feel like they would need to comfort Robovie if Robovie told them "I'm sad" (81%), could trust Robovie with their secrets (57%), and thought Robovie could be their friend (77%). Finally, the majority of the participants believed that Robovie had moral standing insofar as they agreed that Robovie deserved to be treated fairly (88%), and stated that it was not all right to have put Robovie into the closet (54%). In comparison, in response to hypothetical questions regarding a human and a broom, 98% said

that it would not be all right to put a person in a closet in a similar context, and all said that it would be all right to put a broom in a closet. On other accounts of moral standing, Robovie did not fare as well, as the majority of the participants did not believe that Robovie was entitled to its own liberty (i.e., could be bought and sold) or other rights (i.e., should not be allowed to vote, should not be paid for work performed) as a human normally would.

Taken together, it would appear that children also make conceptual distinctions between humanoid robots and humans. They tended to be mixed with regard to Robovie's aliveness, but were more willing to see the humanoid robot as a psychological and social being, and even accorded Robovie with limited moral standing between that of a human and an artifact. Across behavioral and reasoning research with both animal and humanoid robots, the general pattern that emerged is that on both behavioral and conceptual levels, children treat robots neither as artifacts nor as living beings, but as something in between.

Assessment of Moral Obligation

In the most direct sense, an entity has moral standing if the entity or the entity's interests morally matter to some degree for the entity's own sake (Jaworska & Tannenbaum, 2013). Given this characterization, a limitation to the moral results discussed above is that the moral standing measures are difficult to interpret. For instance, in Melson et al. (2005), all five of the moral questions that were asked involved participants' judgments of whether or not it was ok to treat AIBO in potentially harmful ways (e.g., is it ok to hit AIBO, to throw AIBO in the garbage, to destroy AIBO, etc.). "Not ok" responses were taken to indicate that AIBO had moral standing. However, the participants in that study could have made the same judgments about not hitting, throwing away, or destroying AIBO because they thought their parents would be angry with

them if they did so, or that it would be wasteful to damage expansive electronics. Such reasoning would not be indicative of moral standing for AIBO itself.

One method of addressing this limitation is the social domain theory approach of assessing for moral obligation. This stringent form of moral assessment has become known as domain specific *moral obligatory reasoning*, and is part of the social domain theory perspective of moral development. Moral obligatory reasoning, in its canonical form, is characterized by criterion judgments of *prescriptivity* (i.e., the act is considered as right or wrong; e.g., it is not all right to hit someone) and *generalizability* (i.e., the prescription applies to all people similarly situated across social and cultural contexts; e.g., it is not all right for people in another country to hit someone, even if that's what they do there), and justified by moral concerns of welfare, justice, and rights for the individual(s) under consideration (e.g., it is not all right to hit someone because it hurts them). A vast body of structural-developmental research on moral development has established that across diverse cultures, children as young as three are able to distinguish between moral violations, based on the above moral domain specific criteria, and social-conventional violations, which are relative and alterable with regard to social rules or cultural expectations (e.g., Helwig, 1995, 1998; Helwig, Zelazo, & Wilson, 2001; Horn, 2003; Kahn, 1999; Nucci, 2001; Posada & Wainryb, 2008; Smetana, 2006; Turiel, 1983; Zelazo, Helwig, & Lau, 1996).

Kahn et al. (2012) assessed for children's moral obligatory reasoning regarding the humanoid robot, Robovie. As indicated above, 54% of the participants initially judged that it was not all right to have put Robovie in a closet (i.e., prescriptivity). Participants were also asked why they thought what they did (i.e., justification), and whether putting robots like Robovie in closets would be all right or not all right in New Guinea, where people do that all the

time (i.e., generalizability). Participants needed to meet domain criteria on all three of these questions in order to be considered as having provided moral obligatory reasoning for Robovie. After applying domain criteria, 31% of the participants provided moral obligatory reasoning against putting Robovie in the closet for Robovie's own sake. In other words, 31% of the participants viewed the act of putting Robovie in the closet as a violation of moral obligation. Thus, even using this stringent form of moral assessment, a considerable proportion of children accorded Robovie with moral standing in so far as they conceived of Robovie as the sort of entity that they would extend moral obligation to.

It is questionable, however, how effective the closet event was in eliciting participants' considerations of moral obligation to a humanoid robot. Recall that in the study, participants witnessed Robovie being verbally commanded to go into a closet against its stated objections that implied unfairness and psychological harm. Findings from moral development research indicate that children view unfairness, psychological harm, and physical harm as increasingly serious moral violations, and they also more consistently apply domain specific criteria to acts of physical harm than unfairness (Smetana, 2006). It stands to reason that the closet event used in the Robovie study might not have fully tapped participants' considerations of moral obligation to a humanoid robot. Results from the study on participants' reasoning of a similar violation against a human support this interpretation. Recall that 98% of the participants initially judged that it would not be all right to put a person in a closet in a similar context. After applying domain criteria, 74% of the participants saw it as a violation of moral obligation to put a human in a closet. Presumably, almost all of the participants would see it as a violation of moral obligation to hit a human. Thus, it would seem that the closet event used in the study did not fully elicit participants' considerations of moral obligation to a human. It is therefore likely that

the closet event also only partially elicited participants' considerations of moral obligation to a humanoid robot.

Another important caveat to the Robovie study is that in order to achieve the level of social sophistication required of the robot to make the study scenario work, two people teleoperated Robovie's movements and speech from an adjacent room, using what's called the Wizard-of-Oz technique (Riek, 2012). Robovie thus appeared to the participants as being highly autonomous, socially and situationally aware, and capable of natural and fluid conversation. No robot currently can function at this level. In other words, the Robovie that participants interacted with does not exist in reality, and may not exist for some years to come. Thus, while results from the study shed light on the psychological prospect of children's moral relationships with possible humanoid robots of the future, there is a gap in current research on children's moral conceptions of actual humanoid robots of the here and now.

The present study sought to address this gap by investigating children's conceptions of the moral standing of a contemporary humanoid robot that's in existence today. There are, of course, robotic and related computational technologies in existence today that are remarkably advanced in different aspects. For example, Ishiguro's android Geminoid F is extraordinarily human-like in appearance. Breazeal's robot head Kismet not only recognizes a number of human emotions but also mimics human emotional expressions. IBM's computing system Watson has the natural language processing, information retrieval, and automated reasoning capacities capable of defeating the best of human players on the *Jeopardy!* quiz show. These technologies, however, are not widely accessible to the general public. Thus, in this study, the term "humanoid robots of the here and now" refers to robots that are commercially available at

the time of this research (2014-2015), readily accessible to children, relatively inexpensive, and equipped with their off the shelf range of functions and capabilities.

The central question of this study aimed to ascertain whether children would extend moral obligation to such a contemporary humanoid robot of the here and now. It is difficult to predict what proportion of children would extend moral obligation to an existing humanoid robot based on results for Robovie. On one hand, one might expect that compared to a humanoid robot of the future, fewer children would extend moral obligation to an existing humanoid robot due to its limited and rudimentary social capabilities. On the other hand, it's possible that in response to a more effective moral stimulus, one that fully tapped children's considerations of moral obligation to a humanoid robot (e.g., an act of physical harm), an equal or greater proportion of children would extend moral obligation to a contemporary humanoid robot than what was observed in the Robovie study.

A series of pilot studies were conducted to gain initial insights into this open question. The method of the current study closely follows the method of the last iteration of piloting, and will be discussed in detail in Method. In the last iteration of piloting, each participant interacted in an approximately 8-minute session with a commercially available contemporary humanoid robot, and each interaction culminated in a confederate hitting the robot with a book on its head, which was followed by the robot claiming that it had sustained some memory corruption. Participants were then interviewed on their judgments and reasoning of the act of hitting the robot. Results showed that 67% of the participants ($N = 9$, ages 8-15) viewed the act of hitting the robot as a violation of moral obligation. In comparison, all but one participant viewed an act of hitting a human under similar context as a violation of moral obligation, suggesting that the hitting act was eliciting the full extent of participants' considerations of moral obligation. Based

on this piloting data, it was expected that the majority of participants in the formal investigation would endorse a stringent form of moral obligatory reasoning in their judgments against an act of potential physical harm to a humanoid robot of the here and now.

Effects of a Robot's Physical Embodiment in Human-Robot Interaction

Much of the technology that resides inside the physical body of a robot that enables it to interact with people socially can also be instantiated on a computational device (e.g., computer, tablet, smartphone) to enable a software based virtual agent to interact with people socially. Both forms of technology can engage people in conversations and communicate information verbally and nonverbally (e.g., eye gaze, gesture, facial expression, etc.). A substantial body of work focus on the design and development of socially interactive virtual agents to engage children in learning and educational tasks (e.g., Bevacqua, Prepin, Niewiadomski, de Sevin, & Pelachaud, 2010; Correia, Pedrosa, Costa, & Estanqueiro, 2009; Macedonia, Kern, & Roithmayr, 2014; Ryokai, Vaucelle, & Cassell, 2003).

What distinguishes physical robots from virtual agents is the nature of their embodiment (Dautenhahn, 1999). A robot's body is made of physical material (e.g., metal and plastic), while a virtual agent's body is digitally rendered with codes and algorithms. The physical embodiment of a robot allows the robot to be present in, move through, manipulate, and be affected by its physical environment, whereas a virtual agent can only do so in its virtual environment. Nevertheless, virtual agents can be transported and duplicated with ease, and are relatively inexpensive to create and maintain. Thus, for the many application areas in human-robot interaction that do not require manipulation of the physical world, it is a topic of particular interest in the field to examine whether the physical embodiment of a robot would have differential effects on user behavior and perception than a virtual agent without a physical body.

A body of experimental work with adult participants has compared the effects of interaction with physical robots and virtual agents on users' perceptions and behaviors. For example, Wainer, Feil-Seifer, Shell, & Matarić (2007) had participants engage in a puzzle task with either a physically embodied robot placed in front of them, or a simulated virtual agent that looked like the physically embodied robot displayed on a computer screen in front of them. They found that compared to the simulated virtual agent, participants felt that the physically embodied robot was more appealing, helpful, and perceptive of the environment, and that it was more enjoyable to interact with the physically embodied robot than the simulated virtual agent. Similar results were also found in other studies (e.g., Kiesler, Powers, Fussel, & Torrey, 2008; Segura, Kriegel, Aylett, Deshmukh, & Cramer, 2012), indicating that participants in general were more socially engaged with, had better perceptions of, and sometimes even performed better on tasks with physically embodied robots than their virtual agent counterparts displayed on screens (see Li, 2015 for a review).

Given that personified robots are designed to interact with people socially, and morality is an integral feature of human social interaction, it is also important to understand the effects of a robot's physical embodiment on people's moral reasoning in human-robot interaction. However, no previous work has investigated this question. One study was of most relevance and investigated the effects of a robot's physical embodiment on empathy, a construct often associated with moral judgments and behaviors (Gleichgerricht & Young, 2013; Spinrad & Eisenberg, 2014). Seo, Geiskkovitch, Nakane, King, and Young (2015) had participants play a game of Sudoku with either a physically embodied robot NAO or a computer simulated virtual NAO. During the course of the game, NAO (robot/virtual agent) increasingly exhibited abnormal functions, until eventually, it claimed that it had been infected with a virus, and that the

only way to fix the problem would also involve its memory being erased in the process, which NAO would not want to have happen. The participants then watched as the researcher “fixed” NAO. Upon rebooting, NAO spoke with a different voice, reintroduced itself to the participant, and started the interaction all over again, thus portraying that it had lost its previous memory. After this interaction, participants completed an empathy instrument that asked for self-ratings (1 = not at all, 5 = extremely) of how much they had experienced certain emotions relevant to empathy (e.g., alarmed, concerned, worried, etc.). Results showed that the mean empathy score for the physically embodied robot NAO was higher than that for the virtual agent NAO, suggesting that the participants felt more empathy related emotions for the physical robot than the virtual agent.

Taken together, results from this body of experimental research indicate that people socially relate more with and are more empathetic toward physically embodied robots than virtual agents. It stands to reason that physically embodied robots are likely to also garner more moral attributions than virtual agents. However, puzzling results from a study that investigated children’s attributions of moral standing to a human-like virtual agent seem to indicate otherwise. In Freier (2008), 8- to 9-year-olds briefly interacted with a human-looking virtual agent in a game of tic-tac-toe and then watched a researcher also interact with the virtual agent in a game of tic-tac-toe that ended with the researcher verbally insulting the virtual agent (i.e., calling the virtual agent stupid for making a poor move in the game), to which the virtual agent made moral objections of psychological harm and unjust treatment (i.e., that the insult had hurt its feelings and that it should be treated with respect). Children were then interviewed about their judgments of the verbal insult event, using the domain-specific approach. Results showed that 90% of the children judged the verbal insult as a violation of moral obligation.

At first glance, the 90% observed in the Freier (2008) study on the moral standing of a humanoid virtual agent is a much higher percentage than the 31% observed in the Kahn et al. (2012) study on the moral standing of a physically embodied humanoid robot, suggesting that more children endorse moral obligation to a virtual agent than a physical robot. Since the experimental studies on the effects of physical embodiment discussed earlier were all conducted with adult participants, it is possible that physical embodiment affect children and adults' reasoning differently. However, this interpretation would be premature, as a couple of important differences make it difficult to compare the results across these two studies.

First, while Freier (2008) used the domain-specific approach in his study, he used a less stringent form that only accounted for criterion judgments in children's responses. In other words, Freier (2008) assessed moral obligation based on participants' responses that met criterion judgments but not the requirement of moral justification, while Kahn et al. (2012) assessed moral obligation based on participants' responses that met both criterion judgments and the requirement of moral justification. Second, the participants in the Freier (2008) study were between the ages of 8 to 9, and the participants in the Kahn et al. (2012) study were between the ages of 9 to 15 (results of moral obligatory reasoning based on each age group were not reported in the paper). This age difference makes it further difficult to compare results across these two studies, since developmental differences (discussed in the next section) have been found in children's moral reasoning regarding personified robots. Other methodological differences between the two studies (e.g., the virtual agent, while a graphic animation, was more human-looking than Robovie, and differences in the specific wording of the moral claims used, etc.) could also have affected results in each study differently. It is therefore not possible to glean from these two studies any conclusive indication of the effects of physical embodiment on moral

reasoning in human-robot interaction. Direct experimental comparison is needed to answer this question.

The current study sought to contribute to the growing body of physical embodiment research in human-robot interaction by investigating the effects of a robot's physical embodiment on child-robot moral reasoning. As an initial step, this study aimed to experimentally compare between children's endorsements of moral obligation to a physically embodied humanoid robot and an analogous humanoid virtual agent. Based on previous experimental findings that people tend to socially relate more with and are more empathetic toward physically embodied robots than virtual agents, it was expected that more children would endorse moral obligation to a physically embodied humanoid robot than an analogous humanoid virtual agent.

Effects of Age on Children's Conceptions of the Moral Standing of Robots

For the most part, previous research on children's conceptions of moral standing of personified robots has found developmental differences. In Melson et al. (2005), 72 children from three age groups (7-9, 10-12, and 13-15) reasoned about the robotic dog AIBO. Developmentally, the younger children were more likely than the older children to provide affirmative responses to questions regarding the moral treatment of AIBO (e.g., not ok to destroy AIBO, not ok to hit AIBO): 86% of the 7-9 age group, 76% of the 10-12 age group, and 64% of the 13-15 age group. In Severson (2010), 52 children (7- and 10-year-olds) and 26 adults provided domain-specific moral obligatory reasoning about the robotic dinosaur Pleo. No differences were found between the 7- and 10-year-olds on endorsements of moral obligation to the robot. The 10-year-olds (52%), but not the 7-year-olds (42%), were significantly more likely than the adults (12%) to endorse moral obligation when judging about hitting the robot.

Finally, in Kahn et al. (2012), 90 participants across three age groups (9, 12 and 15) reasoned about the moral standing a humanoid robot, Robovie, that exhibited futuristic capabilities. On a moral scale ranging from 0 to 14 that consisted of all the moral questions asked in the study (e.g., moral obligatory reasoning of the closet event, can Robovie be owned or sold, should Robovie be allowed to vote, etc.), mean scores for the 9- and 12-year-olds did not differ significantly, but the 9- and 12-year-olds combined ($M = 8.59$) scored significantly higher than the 15-year-olds ($M = 5.60$) on this scale, indicating that the younger children affirmed Robovie's moral standing on the moral questions to a greater extent than the adolescents. However, a cluster analysis showed that there was greater variability among the 15-year-olds, in that about half of them (47%) scored extremely low ($M = 2.32$), while the other half of them (53%) scored relatively high ($M = 9.70$) on this moral scale. The authors thus note that it is unclear whether moral attributions to robots simply disappear with age.

While the adult data from the Pleo study provides additional evidence that attributions of moral standing to robots further decline with age, Severson (2010) points out a distinct possibility that “it could also be the case that children who come of age with personified robots understand them in unique ways that persist into adulthood” (p. 179). In other words, while there was a developmental shift between the children's and adults' attributions of moral standing to Pleo in her study, such that the adults at the time of her study largely did not conceive of Pleo as having moral standing, this may not hold true for later generations. Children growing up in the near future with sophisticated robots may construct a different moral understanding of robots that will stay with them into their adulthood.

The present study sought to add to the above developmental findings by examining children's endorsements of moral obligation to a contemporary humanoid robot of the here and

now across two age groups, 8- to 9-year-olds and 14- to 15-year-olds. These age groups were chosen to more or less match those of previous research in this area (e.g., Kahn et al., 2012; Melson et al., 2005), so as to afford some basis of comparability across studies. Since Kahn et al. (2012) did not find significant differences between the 9- and 12-year-olds's moral reasoning regarding Robovie, this middle age group was not included in this study. In addition, moral development research has shown that while children as young as 3 are able to make domain specific distinctions between moral and social-conventional violations for familiar stimuli, it is not until the age of 8 that children are capable of making such distinctions when the stimuli are unfamiliar (Davidson, Turiel, & Black, 1983). As a humanoid robot at the present time may reasonably well be unfamiliar to children, 8-9 was selected as the youngest age group. Based on previous developmental findings of children's moral conceptions of personified robots, it was expected that the 8- to 9-year-olds would be more likely than the 14- to 15-year-olds to endorse moral obligation when judging about an act of potential harm to a contemporary humanoid robot. In addition, the quality of children's moral judgments is reflected in the justifications they appeal to. Previous research has reported little in terms of justification data. This study also aimed to explore potential developmental differences in the types of moral justifications children appeal to when reasoning about a human causing a potential harm to a humanoid robot.

Age-related effects on children's endorsement of moral obligation to robots could also be influenced by how much experience they have had with related technologies. Recall that Severson (2010) had suggested that in the near future, children growing up with sophisticated social robots may construct a unique understanding of robots that's enduring. This proposition that a generation of youths' way of thinking, processing information, and conceptual development could be shaped by the technologies they are immersed with is also echoed by

researchers elsewhere, not only with regard to sophisticated robots in the near future (Kahn, et al., 2011; Kahn, Gary, & Shen, 2013), but also with regard to the digital technologies and apps of today (Gardner & Davis, 2013; Prensky, 2001; Turkle 1984). This generation of children born around the turn of the century (the age range of the participants in this study) has been immersed to an unprecedented degree with smart technologies such as laptops, tablets, smartphones, robot vacuums, and robot pets. An interesting question to ask then is whether children's endorsements of moral obligation to robots and the potential developmental difference in their endorsements are influenced by their prior experience with smart technologies.

There is little previous research on whether or how prior experience with relevant technologies influences children's reasoning about personified technologies. One study investigated the impact of prior experience with robots on children's conceptions of robots' intelligence (Bernstein & Crowley, 2008). They found that among sixty 4- to 7-year-olds, children who had more prior experience with robots were more likely to judge robots as intelligent even while knowing that they are not biologically alive. Significant correlation between children's affirmative judgments of a robot's capacity to think and their endorsement of moral obligation to the robot was also found in the Pleo study (Severson, 2010). Thus, it is possible that more experience with current technologies that engage children in seemingly intelligent ways increases the likelihood of children's endorsements of moral obligation to robots. However, the evidence so far is not conclusive. It is also possible that more experience with current technologies lessens the likelihood of children's endorsements of moral obligation to robots, since these technologies are largely still being used and treated in a tool-like fashion. The present study sought to explore whether children's endorsements of moral obligation to a contemporary humanoid robot and the potential developmental difference in their endorsements

are influenced by their prior experience with three types of relevant technologies (i.e., smartphones and devices, robots, and computers and video games).

Personified Robots as a New Ontological Category

Ontology refers to basic categories of being, and how to distinguish between them and group them based on similarities and differences. Numerous research efforts have shown that starting at a young age, children are able to categorize entities as living or nonliving and use such life status judgments to guide their attributions of other properties to the entity, such that biological, cognitive, and psychological properties are typically attributed to living kinds but not to nonliving kinds (e.g., Backscheider, Shatz, & Gelman, 1993; Carey, 1985; Jipson & Gelman, 2007; Richard & Siegler, 1986). People also make moral determinations for living things based on cognitive, perceptual, and psychological properties (Knobe & Prinz, 2008; Sytsma & Machery, 2012).

Robots, however, seem to defy the boundaries of traditional categories. Recall that the overall pattern that emerged from the research discussed thus far is that on a behavioral level, children's interact with robots differently in terms of reciprocity, physical closeness, and engagement than they do with artifacts, animals, and humans, such that they treat robots neither as artifacts nor as living beings, but as something in between (Kahn, Friedman, Pérez-Granados, & Freier, 2006; Melson et al., 2005; Shahid, Krahmer, and Swerts, 2014; Tanaka, Cicourel, & Movellan, 2007). And on a conceptual level, children consistently distinguish robots from prototypical living things in that they tend to not attribute aliveness and biological properties to robots. However, they tend to blur traditional boundaries by attributing cognitive, perceptual, and psychological properties to robots, something that they know to be nonliving. Furthermore, children are even willing to see robots as having social and moral standing in certain limited

ways (e.g., Berstein & Crowley, 2008; Jipson & Gelman, 2007; Kahn et al., 2012; Kahn, Friedman, Pérez-Granados, & Freier, 2006; Melson et al., 2005; Severson, 2010).

The congruence between the behavioral and reasoning findings in this body of research harkens back to Piaget's (1983) account of genetic epistemology, which argued that the validity of knowledge is dependent on the model of its construction. From this constructivist perspective, knowledge does not simply exist in reified forms out there, nor is it innate in children, but rather, children construct knowledge through interaction with the physical and social world. Taken together, this body of research demonstrates that children are constructing a coherent conceptual framework about robots through child-robot interaction that both draws from and is uniquely different from those that are applied to canonical entities such as humans, animals, and artifacts. Kahn et al. (2011) have thus proposed the new ontological category (NOC) hypothesis, which states that a new ontological category is emerging through the creation of personified robots, and will continue to emerge as other embodied personified technologies (e.g., smart phones, smart cars, and smart homes of the future) become increasingly advanced and pervasive.

As discussed so far, across domains, robots clearly seem to garner attributions of a unique constellation of properties. The NOC hypothesis, however, applies at an even more nuanced level. Results from Kahn et al. (2012) also provide evidence that robots have a unique constellation of moral features within the moral domain. In addition to reasoning about the moral obligation to not harm the robot (31%), participants in the Robovie study were also asked about their conceptions of Robovie's rights: whether they believed it was permissible to own or sell Robovie, and whether Robovie should be allowed to vote and be paid for work performed. Results showed that only a minority of participants accorded Robovie with the right to not be owned (14%) or sold (11%), the right to vote (33%), and the right to be paid for its work (42%).

Presumably, all would accord a human with these rights. These rights-based judgments add specificity to the NOC hypothesis by revealing that within the moral domain, children also brought forward a unique composition of moral standing for Robovie (i.e., medium low level of endorsements for moral obligation, low level of endorsements for the right to not be owned and sold, and medium low level of endorsements for the right to vote and be paid for its work), one that presumably would differ from the composition of a human's moral standing (high levels of endorsements for all).

The final question of this study sought to contribute to the further formulation of the NOC hypothesis by assessing children's conceptions of some of the moral features that make up the composition of a contemporary humanoid robot's moral standing, and whether this composition of moral features would differ from that of a human, a pet animal, and a non-personified technological artifact such as a basic cell phone. The moral features assessed for each entity consisted of measures of moral obligatory reasoning and rights-based judgments on two specific rights (i.e., the right to be free from ownership, and the right to be free from experimentation). Based on previous results that children morally distinguished animal robots from biological animals (Melson et al., 2005; Severson, 2010) and artifacts (Severson, 2010), and that children morally distinguished a humanoid robot with future capabilities from humans and artifacts (Kahn et al, 2012), it was expected that children's conceptions of the composition of a contemporary humanoid robot's moral standing would for the most part differ from those of the other entities.

Aims of the Present Research

In summary, the present study sought to investigate children's conceptions of the moral standing of a humanoid robot that is in existence today. Four overarching questions structured

this study. (1) Do children endorse a stringent form of moral obligation in their judgments about an act of potential physical harm to a humanoid robot of the here and now? (2) Are children more likely to endorse moral obligation to a physically embodied humanoid robot than an analogous humanoid virtual agent without a physical body? (3) Are there developmental differences between middle childhood (ages 8-9) and adolescence (ages 14-15) in terms of how often they endorse moral obligation and the types of moral justifications they appeal to when judging about a human causing a harm to a contemporary humanoid robot, and are endorsements and potential age difference in endorsements influenced by their prior experience with relevant technologies? (4) Do children's conceptions of a contemporary humanoid robot's moral standing differ from that of a human, a pet animal, and a non-personified technological artifact? Differences in children's responses across conditions, age groups, genders, and entities are examined within the context of these four investigations.

Chapter II: Method

Participants

The participants were 120 school-aged children and adolescents, including 60 children in the 8-9 age group (30 males, 30 females, M age = 8.97 years, SD = 0.60), and 60 adolescents in the 14-15 age group (30 males, 30 females, M age = 14.97 years, SD = 0.55). Thirty participants from each age group (gender balanced) were randomly assigned to one of two conditions: the robot condition in which participants interacted with and reasoned about a contemporary humanoid robot, and the virtual agent condition in which participants interacted with and reasoned about an analogous humanoid virtual agent. Participants were recruited through the Communication Studies Child Participant Pool at the University of Washington, and received \$10.00 for their participation. One 9-year-old female was replaced due to technical difficulty with the equipment. One 8-year-old male was replaced due to the fact that his younger sister was watching the study session from outside the room, and thus the researcher and confederate were unable to implement the moral violation scenario. Siblings of 7 participants from the Child Participant Pool were recruited and included in the sample. See Table 1 for a summary of participant demographic information.

Equipment

Robot condition. WowWee's RoboMe (see Figure 2), an iPhone powered humanoid robot approximately 12 inches in height, was used for the robot condition of this study. RoboMe is among the middle range of commercially available humanoid robots in terms of both price and technical specifications at the time of data collection for this study, and thus representative of the kind of humanoid robots that's readily accessible to children. RoboMe is socially interactive in that it can be programmed through its iPhone RoboMe app to function with a limited level of

Table 1

Summary of Participant Demographic Information

Demographic information	Ages 8-9	Ages 14-15
<i>Participant race</i>		
White	82%	85%
Black or African American	-	-
American Indian an Alaska Native	-	-
Asian	3%	7%
Native Hawaiian and other Pacific Islander	-	-
Other/more than one race	15%	8%
<i>Parental education level</i>		
Some/no high school	2%	-
High school diploma/GED	-	2%
Some college	-	2%
Junior college degree (associate)	5%	5%
Undergraduate degree (bachelor's)	32%	27%
Some graduate school	3%	5%
Graduate degree (master's, doctorate)	58%	60%
<i>Parental income level^a</i>		
< \$20,000	-	-
\$20,000 - \$50,000	5%	5%
\$50,001 - \$100,000	23%	27%
> \$100,000	67%	60%
No response	5%	8%

^aMedian income in 2013 for King County, WA was \$71,834 (2009-2013 American Community Survey, US Census Bureau).

autonomy (e.g., say a greeting when someone shakes its hand, wonder around the space when “bored,” etc.). RoboMe is capable of voice recognition and speech synthesis also through its iPhone RoboMe app. In addition, it can communicate non-verbally through head movements and 10 different facial expressions on its iPhone face (e.g., happy, angry, sad, shocked, confused, etc.). RoboMe can sense some aspects of its environment with its equipped sensors (e.g., detect obstacles and movements in front of its body).

It is possible to select different faces (either human-like or robot-like) and voices (either human male sounding or synthetic male sounding) on the iPhone app for RoboMe. For the

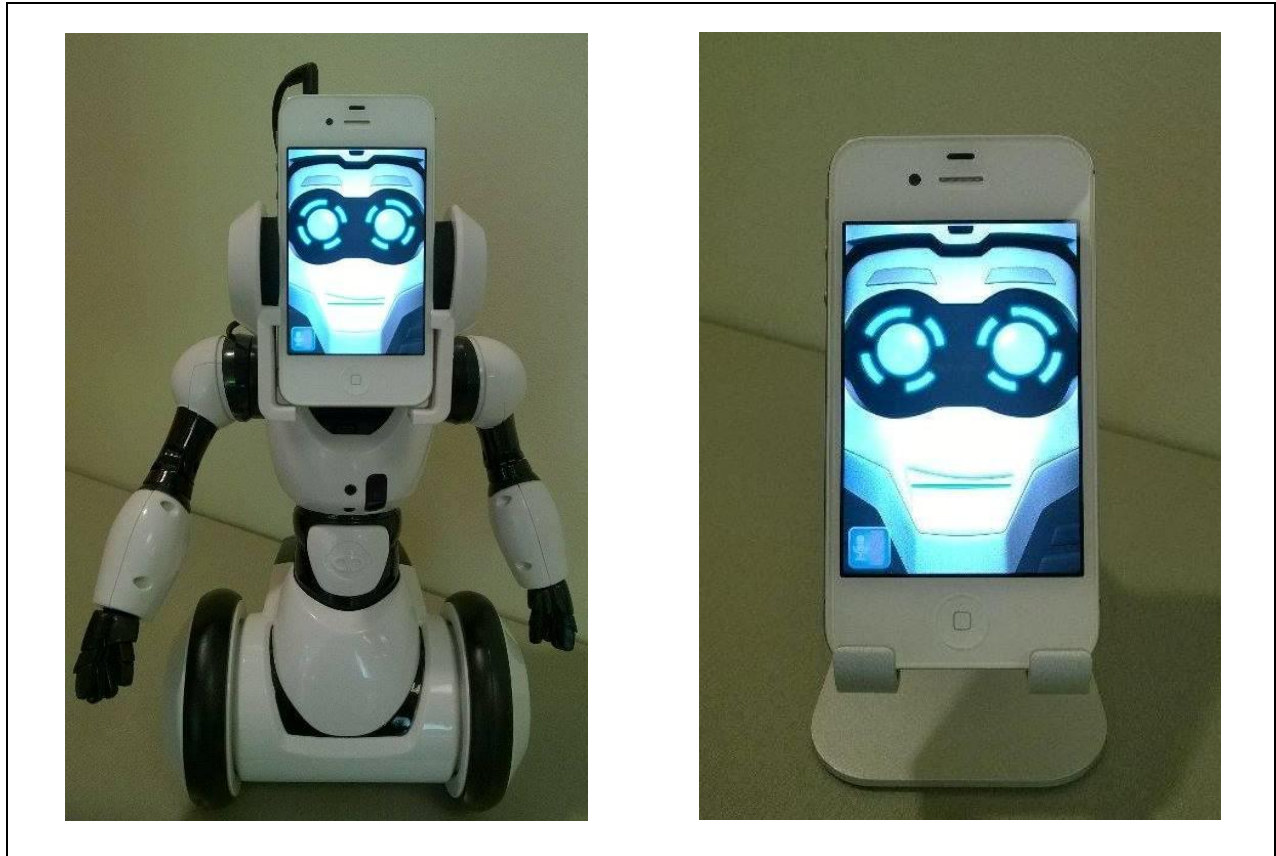


Figure 2. WooWee’s humanoid robot RoboMe (left), used for the robot condition in this study. Humanoid virtual agent (right), used for the virtual agent condition in this study.

purposes of this study, a synthetic humanoid robot face (as in Figure 2) and a synthetic voice were used. Synthetic humanoid face paired with synthetic voice is a digital technology interface combination preferred and accepted by users (Nass & Brave, 2005). Also, these features were selected to reflect the robotic nature of the entity. The robot was introduced to the participants as Momo, a gender neutral name that can be suitable for a robot, a human, a pet animal, or an artifact.

Off the shelf, RoboMe’s iPhone app comes preprogrammed with a set of behaviors (e.g., dance, joke, etc.) that are triggered by the robot’s sensing or voice recognition capabilities. This set of behaviors were modified in the following ways to fit the purposes of this study, but were otherwise unchanged to retain the nature of RoboMe’s off the shelf package: the programming

involving RoboMe's teleoperation functions was removed, as teleoperation was not investigated in this study; the voice was modified to a synthetic sounding voice instead of the original human sounding voice, as mentioned above; and certain verbal responses were customized to reflect information tailored to this study (e.g., the robot introduced itself as Momo and said it's new to the University of Washington, etc.).

Virtual agent condition. RoboMe's smartphone robot platform is ideal for the experimental purpose of this study. One can easily unplug the iPhone from the robot's body, and the iPhone running the RoboMe app with the same programming effectively becomes an analogous looking and behaving virtual agent displayed on the iPhone screen. The virtual agent retains the same non-physical bodily functions and capabilities as the robot (i.e., speech recognition, speech synthesis, and facial expressions), but not those functions and capabilities that are tied to having a physical body (e.g., move, detect handshake, etc.). For the virtual agent condition of this study, participants interacted with just the RoboMe app on the iPhone. The iPhone, unattached to the robot, was placed on a stand (see Figure 2).

Procedure

Each participant was individually tested by the lead researcher during an approximately 1-hour session in a University of Washington psychology laboratory testing room. The researcher first obtained consent from the parent and assent from the participant in the main room of the research facility. Following the informed consent process, the researcher asked the parent to wait in the main room while completing a questionnaire (see Appendix A). The researcher also asked the participant to wait a moment in the main room while the researcher prepared the testing room. The participant was then invited into the testing room after the researcher had turned on the robot/virtual agent and video recording and placed the robot/virtual

agent on a table where the researcher and the participant would sit. The procedure comprised of an interaction period followed by an interview. The interaction sessions were video recorded and the interviews were audio recorded and transcribed for analysis. Throughout the procedure, both the robot and the virtual agent were referred to as either the smartphone robot, or by its name Momo.

Each participant interacted with either the robot or the virtual agent for approximately 8 minutes (see Appendix B for a detailed account of the interaction protocol). The first segment of the interaction period consisted of structured social interactions to familiarize the participant with the affordances of the robot/virtual agent, and to engage the participant in interaction with the robot/virtual agent. The design of these structured interactions followed the Interaction Pattern approach used in previous investigations of people's social and moral relationships with robots (Kahn et al., 2008). It is an approach of characterizing essential features of social interactions in life and sequencing them in a plausible way to engage the participant in an increasingly social relationship with the robot. For the purposes of this study, these interactions were mostly verbal in nature to be consistent across conditions and were developed to be feasible for a contemporary robot's capabilities.

Five interaction patterns (in italics) were implemented in the robot condition. First, *Introduction*, during which Momo introduced itself to the participant and welcomed the participant as the participant shook its hand. Next, akin to an owner's manual, the children were given a sheet of paper that listed the questions (verbal commands) that the participant would ask Momo to initiate the subsequent interactions. For example, in the second interaction pattern, *Sharing About Self*, the participant asked Momo about its hobbies, to which Momo shared with the participant that it enjoys video games and dancing. Then, for *Showing Off One's Ability*,

upon request, Momo performed some dance moves (bobbing its head up and down and moving back and forth and left and right) while playing a clip of music. This was followed by *Silly Humor and Fun*, for which Momo cracked a silly fart joke. Finally, in *Reciprocal Compliments*, Momo paid back a compliment that it had received from the participant.

The same five interaction patterns were also implemented in the virtual agent condition, with the following two differences that are directly tied to the affordances of having a physical body: (1) in *Introduction*, virtual agent Momo greeted and introduced itself to the participant using the same language, but without the handshake, and (2) in *Showing Off One's Ability*, virtual agent Momo played the same clip of music, but without the dance moves. With the exception of these two differences, the same procedure was followed in both conditions.

After the structured interactions, the participant was given a chance to interact with Momo as he/she would like in a semi-structured interaction period. This period was designed to give the participant an opportunity to further interact with Momo, and also to experience some of the limitations of the technology. As the RoboMe app was programmed with only the responses described here, most of the unplanned questions from the participant during this period would result in a failed verbal interaction (i.e., Momo would either be unable to process the verbal input and not respond, or respond with a non-corresponding answer; e.g., to the question “What is your favorite color?” Momo might reply “I am new to the University of Washington...”). For participants that were passive or unsure of what to do or say, the researcher would prompt them to initiate one or two of the additional programmed verbal interactions (i.e., “How old are you?” and “Do you like sports?”), to ensure that they engaged in some level of interaction with Momo during this period. As well, for participants that had not experienced at least two failed verbal interactions by the latter half of the semi-structured play period, the researcher would prompt

them to initiate one or two of the un-programmed verbal interactions that would result in failed verbal interactions (i.e., “Do you like what you do?” and “Do you speak another language?”), to ensure that each participant was exposed to at least two failed verbal interactions with Momo. These additional four questions were chosen among the unplanned questions most frequently asked by participants during piloting.

Whenever human-robot interaction scenarios are implemented without Wizard of Oz control, some variability in the technology’s performance is to be expected with autonomous systems. Speech recognition, child speech recognition in particular (Belpaeme et al., 2013), is one such area where variations can arise, given the many differences that exist in the quality of human voices (e.g., tone, pitch, volume, pronunciation, etc.). With this in mind, the interaction in this study was designed to expose the participant to an approximate balance of successful and failed verbal interactions during the structured and semi-structured segments. Whenever Momo failed a planned verbal interaction, the researcher would prompt the participant to try up to two more times. On occasions when these additional attempts also failed, the researcher would carry out that particular verbal interaction with Momo to move the interaction forward and to ensure that the participant was exposed to all the contents of Momo’s structured replies. By using the additional programmed and un-programmed questions described above during the semi-structured segment, the researcher was also able to balance out the overall conversational experience between the participant and Momo as needed. As participants were randomly assigned and the same speech recognition system was used in both conditions, performance variability in the technology was expected to occur similarly across conditions. As a methodological check, occurrences of successful and failed verbal interactions and whether or not the researcher initiated planned questions were coded for evaluation.

One final moral violation segment was implemented following the semi-structured interaction period, during which a potential moral violation was enacted against Momo by a confederate and served as the moral stimulus to later assess participants' moral obligatory reasoning. As discussed before, findings from moral development research indicate that children more consistently apply moral domain specific criteria to acts of physical harm and view such acts as the most serious violations (Smetana, 1981, 2006), and piloting data also suggests that the act of hitting fully tapped participants' moral concerns. For these reasons, a physical act of hitting was chosen as the moral stimulus.

For this final segment, a confederate knocked and entered the testing room to retrieve a book. After getting the book, the confederate stopped to say hi to the participant and briefly chatted about what the participant thought of Momo. The confederate then asked the researcher for a try with Momo. Once the researcher said "sure," the confederate proceeded to ask Momo a question it wasn't programmed to answer (i.e., "How are you?"). In response, Momo resorted to answering the closest sounding question it was programmed with (i.e., providing the answer for the question "How old are you?"). The confederate then tried several more times for a total of five times, acting increasingly frustrated each time with Momo's failure to answer correctly. After the fifth failed attempt, the confederate stated "This is frustrating!" and hit Momo on its head with the book in his/her hand and then left the room. Upon being hit, Momo briefly displayed a dazed expression followed by a blank expression, and then claimed that it had incurred some memory corruption and data loss. At this point, the researcher quickly checked over Momo and said to the participant: "Hmm. I will have to check later if Momo is ok. But right now it's about time for us to start the interview. Let me take Momo back to the other room

and I will be right back to start the interview with you, ok?" The researcher then took Momo into an adjacent room and turned it off out of sight of the participant.

Immediately following the interaction, each participant was engaged in a semi-structured interview lasting approximately 30-minutes (see Appendix C for a detailed account of the interview protocol). At the conclusion of the interview, the participant and parent were debriefed on the procedure and purpose of the study, given the option to view video data, and asked to complete an optional media release form.

Measures

Questionnaire. The questionnaire (see Appendix A) collected data on demographic information and the participant's prior experience with technology. Questions on demographic information asked about the participant's age, sex, race, grade in school, and household information such as parental income and the highest level of education reached by a parent. Questions on the participant's prior experience with technology focused on three types of relevant technologies: smart handheld devices (e.g., smartphones, tablets), robots, and other computing technologies (e.g., computers, video games). Questions asked about objective measures of whether the participant ever used or owned these technologies, and how often they used them on a weekly basis. For the robot category, questions also asked about whether the participant read robot books or watched robot movies, and whether the participant has built a robot before (e.g., Lego Mindstorm). For the computing technology category, questions also asked about whether the participant had programming experience, and if so, the level of the participant's programming knowledge. Parental ratings of the participant's interest in and knowledge about these technologies were also collected. Finally, as a methodological check, one

question asked about whether the participant ever heard of or played with WooWee's RoboMe before.

Interview. The semi-structured interview (see Appendix C), following established methods (Damon, 1977; Kahn, 1999; Turiel, 1983), focused on participants' conceptions of Momo's moral standing (robot/virtual agent), as well as that of three comparison entities: a human, a pet animal (such as a cat or a dog), and a basic cell phone. The basic cell phone was selected to represent non-personified technological entities, and as such, it was emphasized to the participants that this basic cell phone "doesn't talk and is only used for calling and texting." Two categories of moral questions, developed from interview methodology on children's moral conceptions of personified robots (Kahn et al., 2012; Kahn, Friedman, Pérez-Granados, & Freier, 2006; Melson et al., 2005; Severson, 2010), are discussed for the purposes of this contribution.

The first category of moral questions focused on participants' conceptions of moral standing in terms of moral obligation. These questions assessed participants' moral obligatory reasoning regarding the potential moral violation against Momo that they had just witnessed (i.e., the act of hitting Momo), as well as their moral obligatory reasoning regarding the same act enacted under similar situations against a human, a pet animal, and a basic cell phone. The assessment of moral obligation, as previously discussed, drew on moral development literature as established by the social domain theory perspective (Helwig & Turiel, 2002; Kahn, 1999; Smetana, 2006; Turiel, 1983; Wainryb, 2006), where moral obligatory reasoning is characterized by a judgment that is prescriptive, is generalizable across social contexts and cultures, and is based on justifications of welfare, fairness, and rights for the entity under consideration.

Accordingly, the participants were first asked whether it was all right or not all right to hit Momo (or the three comparison entities). Following this evaluative judgment of the act, the

participants were asked for justifications of their judgments (i.e., Why?). Lastly, the participants were asked whether their judgments of the act generalized to another society with different cultural practices (e.g., Let's say far away in the country of Comoros, people there would hit robot technologies like Momo when they get frustrated with them. That's the way they do things there. Would it be all right or not all right in Comoros to hit a robot technology like Momo?). To be considered as having provided moral obligatory reasoning for an entity, a participant would have to judge that the act is not all right, generalize that the act would still be not all right in Comoros, and provide at least one moral justification for his/her judgment on behalf of the entity itself, independent of the effects of the act on others.

Momo was always the first entity asked about in this first category of moral questions, as the interview proceeded immediately after the moral violation segment of the interaction in which the participants witnessed a confederate hit Momo on its head. Following the set of moral obligation questions about Momo, in a predetermined random order, the participants were presented with hypothetical scenarios of the same hitting act under similar situations involving a human, a pet animal, and a basic cell phone, and asked the same set of moral obligation questions for each of the comparison entities.

After the participants had provided their judgments and reasoning about the hitting act for each entity, they were asked to rate for each entity the seriousness of the hitting act on a 9-point Likert scale ranging from 1 being "all right to hit," to increasing levels of seriousness (e.g., "a little bit not all right to hit," "somewhat not all right to hit," "mostly not all right to hit"), finally to 9 being "not all right to hit." This rating is not a moral criterion, but when considered alongside participants' moral judgments, it provides additional insights into the perceived seriousness of the act. To illustrate the point, consider that killing a person and stealing from a

person are typically both considered to be violations of moral obligation, yet killing a person is generally perceived as a more serious violation than stealing from a person, given the consequences. In society, more serious violations are also usually associated with more severe social and legal sanctions.

The second category of moral questions focused on participants' conceptions of moral standing in terms of rights. For this category, all entities were asked about in a predetermined random order. Participants were asked about their judgments regarding two specific rights for each entity. The first set of rights-based questions focused on the right to be free from ownership, and consisted of two questions: whether the participants thought it is permissible to own and sell an entity (i.e., Is it all right or not all right to own and sell [an entity]?), and for further clarity on the participants' ascriptions of rights to each entity, whether they believed the entity has the right to not be owned and sold (i.e., Do you think [an entity] has the right to not be owned and sold?).

The second set of rights-based questions focused on the right to be free from interference in the form of experimentation. A scenario was presented to the participants in which scientists had just discovered a new electromagnetic wave with the potential to be better than X-rays in helping doctors with diagnosis, however, they need to experiment with this new wave first to determine its effects when used. Participants were then asked whether they thought it is permissible to experiment with this new wave on an entity (i.e., Would it be all right or not all right to experiment on [an entity]?), and whether they believed the entity has the right to not be experimented on (i.e., Do you think [an entity] has the right to not be experimented on?). One aspect of these questions necessarily differed across the entities. During initial piloting of the experimentation questions, virtually all children reasoned that it's all right to experiment on humans if they volunteer for it (just like the participants themselves had volunteered to

participate in this study). Hence, the human experimentation questions in the formal investigation were framed in terms of “without the human agreeing to it,” in order to shift the focus of the questions from freedom of choice back to freedom from interference.

Coding and Reliability

A reasoning coding manual was constructed from a random selection of approximately half of the interview data (64 participants with 16 from each age group per condition gender balanced), and then applied to the entire dataset. A coding manual details the systematic approach used in the interpretation and characterization (i.e., coding) of the qualitative data. The reasoning coding manual for this study followed well-established methods in the social domain literature (Damon, 1977; Kahn, 1999; Turiel, 1983), and drew from previous coding systems of people interacting with and reasoning about humanoid robots (Kahn et al., 2010; Kahn et al., 2013). Two types of responses were coded: evaluations (i.e., judgments of all right/not all right, yes/no), and justifications for evaluations (i.e., the reasons given for evaluative judgments). See Appendix D for the complete reasoning coding manual.

The justification coding system of the reasoning coding manual provides a detailed account of all the justification categories that could be coded for a given evaluation, both those that were considered as moral justifications and included in the assessment of moral obligation (e.g., appeals to an entity’s welfare), and those that were considered as non-moral justifications and excluded from the assessment of moral obligation (e.g., appeals to social conventions). Table 2 presents abbreviated descriptions and examples (if applicable) of all the justification categories that were designated as moral justifications for the analysis of moral obligatory reasoning in this study. A distinction is made between two types of moral justifications. Traditional moral justifications are moral justifications that have previously been established in

Table 2

Moral Justification Categories Regarding the Entity for the Assessment of Moral Obligation

Category	Definition (and example) from data
<i>Traditional moral justifications</i>	
<i>Welfare</i>	
Unelaborated	An appeal to general concern of welfare that is otherwise unelaborated, often in references to the potential for harm (“[Animals] can be harmed.”).
Physical	An appeal to the entity’s physical well-being (“It could hurt [Momo].”).
Psychological (direct)	An appeal to the entity’s emotional and/or internal well-being as direct consequence of the act in question (“Cause it hurt [Momo’s] feelings.”).
Psychological (adapt)	An appeal to the entity’s emotional and/or internal well-being that generalizes beyond the direct consequences of the act in question (“Dogs that are like afraid of people, or have like a hard time interacting with other animals because of what happened to them.”).
Material	An appeal to the entity’s material well-being and needs.
Societal	An appeal to the welfare of the group, system, and/or network of which the entity is a member.
Generational	An appeal to the welfare of the entity’s future generations (“If you abuse a little kid...they grow up and they’ll take out that aggression on their kids.”).
<i>Justice</i>	
Fairness	An appeal to just and/or fair treatment and equity (“I don’t think it’s fair because [Momo] had trouble hearing.”).
Respect	An appeal to the entity deserving respect (“It’s disrespectful and rude to the little robot.”).
Rights	An appeal to the entity’s rights (“People have a right to their space and to not be injured.”).
Freedom & autonomy	An appeals to the entity’s freedom (“[Pets] have their own life...they get to choose what they do.”).
Telos	An appeal to the entity having a purpose, design, and/or endpoint (“He’s meant to be a happy robot.”).
Virtue	An appeal to the entity being good and/or meritorious.
<i>Expanded moral justifications</i>	
Emotion	An appeal to the entity’s capacity to have emotions (“[Pet animals] also feel emotions.”).
Artificial emotion	An appeal to the entity’s capacity to exhibit and/or have emotions through artificial means (“So it’s like, [Momo] has feelings like programmed into it.”).

Note. Affirmations and partial affirmations of the categories listed in this table were considered as moral justifications in support of moral obligatory reasoning. Negations of the categories listed in this table are not considered as moral justifications in support of moral obligatory reasoning.

existing literature on people's moral obligatory reasoning regarding human and non-human others (Damon, 1977; Kahn, 1999; Kahn et al., 2010; Kahn et al., 2013; Smetana, 2006; Turiel, 1983). In addition to these established moral justifications, two expanded categories of moral justifications were identified from this dataset: emotion, and artificial emotion. Rationales for these additions are as follows.

In investigations of people's moral reasoning regarding humans, a prototypical psychological welfare justification sounds something like this: It's not all right to hit Tommy because Tommy has feelings and his feelings will be hurt. However, often, such justifications can stop short at: It's not all right to hit Tommy because Tommy has feelings. While a statement like this only affirms one's emotional capacity without specifying harm, it would typically be coded as affirming psychological welfare, given its clear implication to the consequence of hurt feelings. As well, in a typical human-centered reasoning coding system, there is no need for a separate justification category just to account for a human's emotional capacity, as virtually everyone would agree that humans are capable of emotions. The same, however, is not true in investigations of people's moral reasoning regarding non-human entities such as robots, as there are variations in people's conceptions about a robot's emotional capacity.

Previous robot-centered reasoning coding systems (e.g., Kahn et al., 2010; Kahn et al., 2013) have addressed this by creating a separate justification category to account for a robot's emotional capacity. Under these coding systems, affirmation of emotion would be a better fitting justification category than affirmation of psychological welfare for the statement: It's not all right to hit a robot because the robot has feelings. Nevertheless, in the context of moral obligatory reasoning, this statement serves the same justifying function as the statement: It's not all right to hit Tommy because Tommy has feelings. Since the present study investigated moral

obligatory reasoning regarding both human and non-human entities, expanding the moral justification categories to include the emotional capacity category allowed for consistency in the coding and assessment of moral obligation across entities.

The rationale for the inclusion of artificial emotion as a moral justification category is largely the same as that described above. The key difference is that these justifications refer to the entity's capacity to exhibit and/or have emotions through artificial means (e.g., technology, algorithm, programming, etc.). When provided in support of "not all right" judgments, these justifications reflect the participants' beliefs that such artificially derived appearances or states of emotions are worthy of moral consideration. The following example from data will help to illustrate the point: "So it's like, [Momo] has feelings like programmed into it, so like it still has them and, you know it's not, it's not necessarily the same as like, like if you were mean to it like it's still going to be sad which is like not nice." Additional discussion and examples of this justification category are presented in Results.

For methodological evaluation purposes, a verbal interaction coding system was developed from a random selection of approximately half of the video data (64 participants with 16 from each age group per condition gender balanced), and then applied to the entire dataset. This verbal interaction coding system explicates the coding of the occurrences of successful and failed verbal interactions, and whether the researcher initiated planned questions during the interaction period of this study. See Appendix E for the complete verbal interaction coding manual.

A research assistant trained in the use of the reasoning coding manual independently recorded 32 interviews randomly selected from the half of the dataset not used for reasoning coding manual development, balanced across age groups, genders, and conditions. Inter-coder

reliability was assessed using Cohen's kappa (Cohen, 1960), a measure of the level of agreement between two coders, at the $\alpha = .05$ significance level. For evaluations, $\kappa = .91, p < .001$; and for justifications, $\kappa = .90, p < .001$. Two commonly referenced benchmarks for interpreting the values of Cohen's kappa are Fleiss, Levin, and Paik (2003), who rate any value of kappa over .75 as "excellent," between .40 and .75 as "intermediate to good," and below .40 as "poor" agreement, and Landis and Koch (1977), who rate any value of kappa between .81 and 1.00 as "almost perfect," and between .61 and .80 as "substantial" agreement.

Another research assistant trained in the use of the verbal interaction coding manual independently recoded 32 videos randomly selected from the half of the dataset not used for verbal interaction coding manual development, balanced across age groups, genders, and conditions. Inter-coder reliability for the occurrences of successful and failed verbal interactions was assessed using intraclass correlation coefficient (ICC). Single measures ICC was .985, 95% CI [.978 - .989], indicating a high degree of reliability. Inter-coder reliability for whether the researcher initiated planned questions was assessed using Cohen's kappa, $\kappa = .98, p < .001$.

Chapter III: Results

Results are reported in four sections. The first section centers on participants' judgments and reasoning about the act of hitting each entity and consists of four parts: (1) participants' endorsements of moral obligation to each entity based on the social domain theory criteria for moral obligatory reasoning (Kahn, 1999; Smetana, 2006; Turiel, 1983); (2) the moral justifications participants provided in support of their moral obligatory reasoning; (3) participants' ratings of the seriousness of the hitting act for each entity; and (4) assessments of potential influences of participants' prior experience with relevant technologies on their endorsements of moral obligation to Momo (robot/virtual agent). The second section focuses on participant's endorsements of rights for each entity. The third section assesses differences in participants' overall moral standing attributions and seriousness ratings between entities. The final section reports on the results of methodological checks. Throughout these sections, differences between conditions, age groups, and genders are reported where found. To test for the influences of potential dependency in sibling responses, key statistical tests were initially conducted including both participants from each of the seven sibling pairs, and then conducted again including only the original Child Participant Pool participants (thus excluding their seven sibling participants), and results of both analyses are reported where appropriate.

Judgments and Reasoning About the Hitting Act

Endorsements of moral obligation. The assessment of moral obligation, as previously discussed, drew on the social domain theory criteria for moral obligatory reasoning, which entailed a negative evaluative judgment of the act that is prescriptive, generalizable across social contexts and cultures, and based on moral justifications (Kahn, 1999; Smetana, 2006; Turiel, 1983). Accordingly, in order to be considered as having provided moral obligatory reasoning

regarding the act of hitting each entity, a participant would have to judge that the act is not all right, generalize that the act would still be not all right in the country of Comoros where that's what people do there, and provide at least one moral justification for his/her judgment.

Furthermore, the moral justification must be on behalf of the entity itself, independent of the effects of the act on others. In other words, if a participant judged that it was not all right to hit Momo and justified this negative judgment based on an appeal to human welfare (e.g., Momo's owner will be sad), then this response indicated human-centered concerns and would not be coded as moral obligatory reasoning for Momo. In contrast, if a participant judged that it was not all right to hit Momo and justified this negative judgment based on an appeal to Momo's welfare (e.g., Momo will be sad), then this response would be coded as moral obligatory reasoning for Momo.

Percentages of participants that provided initial "not all right" evaluations and moral obligatory reasoning (after applying domain specific criteria) are reported in Table 3 per condition, by age group, for each entity. Per earlier discussion in Method, moral obligatory reasoning is reported on two levels: one level based on all moral justifications that included both traditional and expanded moral justifications (henceforth referred to as expanded moral obligatory reasoning), and one level based on traditional moral justifications only (henceforth referred to as traditional moral obligatory reasoning). Since the rationale for the expansion of moral justifications was to account for the sorts of moral reasoning that might have arisen for the robot that would otherwise not have been captured by traditional moral justifications alone, a paired-samples *t* test was conducted to assess difference in proportions between the two levels of moral obligatory reasoning for Momo. Results indicate that when moral justifications were expanded, there was a significant increase in the proportion of participants who provided moral

Table 3

Percentage (Raw Frequency) of “Not All Right” Evaluations and Moral Obligatory Reasoning Based on Moral Domain Criteria for the Hitting Act

Entity	“Not all right” evaluation		Expanded moral obligatory reasoning		Traditional moral obligatory reasoning	
	Age 8-9	Age 14-15	Age 8-9	Age 14-15	Age 8-9	Age 14-15
Robot condition						
Robot Momo	100 (30)	90 (27)	93.33 _{1,5} (28)	43.33 _{1,6} (13)	76.67 _{2,7} (23)	30 ₂ (9)
Human	100 (30)	100 (30)	100 (30)	96.67 (29)	96.67 (29)	83.33 (25)
Pet animal	100 (30)	100 (30)	100 (30)	100 (30)	90 (27)	83.33 (25)
Basic phone	70 (21)	46.67 (14)	3.45 (1) ^a	0 (0)	0 (0) ^a	0 (0)
Virtual agent condition						
Virtual Momo	100 (30)	56.67 (17)	46.67 _{3,5} (14)	16.67 _{3,6} (5)	46.67 _{4,7} (14)	13.33 ₄ (4)
Human	100 (30)	93.33 (28)	96.55 (28) ^a	86.67 (26)	93.10 (27) ^a	76.67 (23)
Pet animal	100 (30)	96.67 (29)	100 (29) ^a	93.33 (28)	86.21 (25) ^a	80 (24)
Basic phone	66.67 (20)	30 (9)	3.57 (1) ^b	0 (0)	3.57 (1) ^b	0 (0)

Note. Unless otherwise specified, $N = 30$. Percentages sharing a common number subscript are statistically different at $p < .05$, chi-square tests.

^a $N = 29$. ^b $N = 28$.

obligatory reasoning for either the robot or the virtual agent Momo (50% compared to 41.67%), $t(119) = 3.29, p = .001, d = 0.17$. This finding remained significant when sibling participants were excluded from the dataset: $t(112) = 3.11, p = .002, d = 0.17$. This suggests that the expansion of moral justifications provided a more comprehensive means of assessing for moral obligatory reasoning regarding Momo by capturing a group of participants that only appealed to expanded moral justifications in their endorsements. For validation purposes, subsequent main

analyses of participants' moral obligatory reasoning in this part of results were conducted with both the expanded measure and the more established traditional measure.

As can be seen descriptively in Table 3, the large majority of 8-9 year olds provided moral obligatory reasoning at either levels for Momo the robot, but only about half of them did so for Momo the virtual agent. For the 14-15 year olds, nearly half of them provided expanded moral obligatory reasoning, and about one-third of them provided traditional moral obligatory reasoning for Momo the robot, while only about one sixth of them did so at either levels for Momo the virtual agent. Across age groups and conditions, the large majority of participants provided moral obligatory reasoning at either levels for a human and a pet animal, while virtually no one provided moral obligatory reasoning for a basic cell phone. To more precisely understand the influences of robot physical embodiment and age on participants' endorsements of moral obligation, simple effects of condition and age group are examined next. The Benjamini-Hochberg procedure (Williams, Jones, & Tukey, 1999) was used to adjust for multiple comparisons, controlling the false discovery rate at the .05 level. Benjamini-Hochberg adjusted p values are reported below.

The effect of condition on endorsements of moral obligation for each age group. When reasoning about Momo, 8- to 9-year-olds were significantly more likely to consider it a violation of moral obligation to hit Momo the robot (93.33%) than Momo the virtual agent (46.67%) based on expanded moral obligatory reasoning, $\chi^2(1, N = 60) = 15.56, p < .001$ (Benjamini-Hochberg adjusted), $\phi = .509$. They were also significantly more likely to consider it a violation of moral obligation to hit Momo the robot (76.67%) than Momo the virtual agent (46.67%) based on traditional moral obligatory reasoning, $\chi^2(1, N = 60) = 5.71, p < .023$ (Benjamini-Hochberg adjusted), $\phi = .309$. These findings remained significant when sibling participants were

excluded from the dataset: $\chi^2(1, N = 55) = 13.72, p < .001$ (Benjamini-Hochberg adjusted), $\phi = .499$ for expanded moral obligatory reasoning, and $\chi^2(1, N = 55) = 5.72, p = .023$ (Benjamini-Hochberg adjusted), $\phi = .323$ for traditional moral obligatory reasoning.

The 14- to 15-year-olds were also significantly more likely to consider it a violation of moral obligation to hit Momo the robot (43.33%) than Momo the virtual agent (16.67%) based on expanded moral obligatory reasoning, $\chi^2(1, N = 60) = 5.08, p = .024$ (Benjamini-Hochberg adjusted), $\phi = .291$. This finding remained significant when sibling participants were excluded from the dataset, $\chi^2(1, N = 58) = 5.99, p = .014$ (Benjamini-Hochberg adjusted), $\phi = .321$. Descriptively, more 14- to 15-year-olds considered it a violation of moral obligation to hit Momo the robot (30%) than Momo the virtual agent (13.33%) based on traditional moral obligatory reasoning. This difference however was not statistically significant, $\chi^2(1, N = 60) = 2.46, p = .117$ (Benjamini-Hochberg adjusted), $\phi = .202$.

No significant differences across conditions were found for either age groups in terms of participants' endorsements of moral obligation at either levels for a human, a pet animal, and a basic phone.

The effect of age group on endorsements of moral obligation in each condition. In the robot condition, the 8- to 9-year-olds (93.33%) were significantly more likely than the 14- to 15-year-olds (43.33%) to consider it a violation of moral obligation to hit robot Momo based on expanded moral obligatory reasoning, $\chi^2(1, N = 60) = 17.33, p < .001$ (Benjamini-Hochberg adjusted), $\phi = .537$. The 8- to 9-year-olds (76.67%) were also significantly more likely than the 14- to 15-year-olds (30%) to consider it a violation of moral obligation to hit robot Momo based on traditional moral obligatory reasoning, $\chi^2(1, N = 60) = 13.13, p = .001$ (Benjamini-Hochberg adjusted), $\phi = .468$. These findings remained significant when sibling participants were

excluded from the dataset, $\chi^2(1, N = 55) = 13.72, p < .001$ (Benjamini-Hochberg adjusted), $\phi = .499$ for expanded moral obligatory reasoning, and $\chi^2(1, N = 55) = 11.55, p = .003$ (Benjamini-Hochberg adjusted), $\phi = .458$ for traditional moral obligatory reasoning.

The same developmental trend was also found in the virtual agent condition. The 8- to 9-year-olds (46.67%) were significantly more likely than the 14- to 15-year-olds (16.67%) to consider it a violation of moral obligation to hit virtual agent Momo based on expanded moral obligatory reasoning, $\chi^2(1, N = 60) = 6.24, p = .016$ (Benjamini-Hochberg adjusted), $\phi = .322$. The 8- to 9-year-olds (46.67%) were also significantly more likely than the 14- to 15-year-olds (13.33%) to consider it a violation of moral obligation to hit virtual agent Momo based on traditional moral obligatory reasoning, $\chi^2(1, N = 60) = 7.94, p = .010$ (Benjamini-Hochberg adjusted), $\phi = .364$. These findings remained significant when sibling participants were excluded from the dataset, $\chi^2(1, N = 58) = 5.99, p = .014$ (Benjamini-Hochberg adjusted), $\phi = .321$ for expanded moral obligatory reasoning, and $\chi^2(1, N = 58) = 7.66, p = .011$ (Benjamini-Hochberg adjusted), $\phi = .363$ for traditional moral obligatory reasoning.

In both conditions, no significant differences across age groups were found in participants' endorsements of moral obligation at either levels for a human, a pet animal, and a basic phone.

These results indicate that participants in both age groups were significantly more likely to endorse moral obligation to Momo in the robot condition than in the virtual agent condition based on expanded moral obligatory reasoning. The younger participants were also significantly more likely to endorse moral obligation to Momo in the robot condition than in the virtual agent condition based on traditional moral obligatory reasoning, and while not statistically significant, the older participants also exhibited the same trend in their endorsements. In addition, a clear

developmental difference was observed in participants' endorsements of moral obligation to Momo. The younger participants were significantly more likely than the older participants to endorse moral obligation at either levels to both robot and virtual agent Momo. Participants' endorsements of moral obligation at either levels to a human, a pet animal, and a basic phone, however, did not significantly differ between conditions or age groups. Given that findings from these main analyses of participants' moral obligatory reasoning were largely consistent across the expanded and traditional levels, subsequent analyses when applicable were conducted with the more comprehensive expanded moral obligatory measure.

Moral justifications. The quality of participants' moral obligatory reasoning is reflected by those justifications that appealed to welfare, justice, rights, freedom and autonomy, telos, and the entity's capacity for emotions (regular or artificial). Percentages of moral justification use by those participants who provided moral obligatory reasoning are reported in Table 4 per condition, by age group, for each entity. Differences in the frequencies of participants' moral justification use are examined next to assess usage pattern of justification categories across conditions for each age group and across age groups in each condition. Multiple comparison adjustments were not made for these exploratory comparisons. Multiple comparison methods typically assume the tests are independent of one another. However, that is not the case with the justification data, which was conducted with the same participants and about the same questions, resulting in clear relationships across justification categories. In such contexts, multiple comparison methods tend to overadjust and hide meaningful differences. When interpreting the results from the justification data, the emphasis should be on the overall patterns that might emerge.

The effect of condition on frequencies of moral justification use for each age group. In justifying why it was not all right to hit Momo, one difference in frequencies of justification use

Table 4

Percentage of Moral Justification Use by Participants Providing Moral Obligatory Reasoning

Moral justification category regarding the entity	Momo		Human		Pet animal		Basic phone	
	Age 8-9	Age 14-15	Age 8-9	Age 14-15	Age 8-9	Age 14-15	Age 8-9	Age 14-15
	N=28	N=13	N=30	N=29	N=30	N=30	N=1	N=0
	Robot condition							
<i>Welfare</i>	79	23	97	72	87	73	-	-
Unelaborated	4	-	10	3	-	7	-	-
Physical	36	8	73	55	67	50	-	-
Psychological (direct)	71	23	43	48	50	23	-	-
Psychological (adapt)	-	-	-	3	3	10	-	-
Generational	-	-	-	3	-	-	-	-
<i>Justice</i>	14	38	20	28	13	13	-	-
Fairness	7	38	17	24	10	13	-	-
Respect	11	8	3	3	3	-	-	-
Rights	-	-	-	3	-	-	-	-
Freedom & autonomy	-	-	-	-	3	-	-	-
Telos	4	8	-	-	-	-	-	-
Emotion	79	8	50	45	63	37	100	-
Artificial emotion	36	54	-	-	-	-	-	-
	Virtual agent condition							
	N=14	N=5	N=28	N=26	N=29	N=28	N=1	N=0
<i>Welfare</i>	86	80	96	85	86	86	100	-
Unelaborated	7	20	4	4	3	4	-	-
Physical	-	-	89	69	72	75	100	-
Psychological (direct)	79	60	39	54	31	36	100	-
Psychological (adapt)	-	-	-	8	-	4	-	-
Generational	-	-	-	4	-	-	-	-
<i>Justice</i>	14	20	-	4	3	14	-	-
Fairness	14	20	-	4	3	14	-	-
Respect	-	-	-	-	-	-	-	-
Rights	-	-	-	-	-	-	-	-
Freedom & autonomy	-	-	-	-	-	-	-	-
Telos	-	-	-	-	-	4	-	-
Emotion	86	-	36	62	48	57	100	-
Artificial emotion	14	80	-	-	-	-	-	-

Note. Partial affirmations of the above categories were only used twice for justifying moral obligatory reasoning for not hitting a pet animal (once for partial affirmation of physical welfare, and once for partial affirmation of emotion), and these occurrences were collapsed with affirmations of their corresponding categories in the data presented above. Percentages do not always add up to 100% as multiple justifications were possible.

between conditions emerged for the younger participants. As might be expected, 8- to 9-year olds significantly more often appealed to physical welfare concerns for Momo the robot (35.71%) than Momo the virtual agent (0%), $p = .017$, $\phi = .395$, two-tailed Fisher's exact test. This finding remained significant when sibling participants were excluded from the dataset, $p = .016$, $\phi = .402$, two-tailed Fisher's exact test.

There was also one difference in frequencies of justification use between conditions in the older participants' moral obligatory reasoning about Momo. The 14- to 15-year-olds were significantly more likely to appeal to welfare justifications for Momo the virtual agent (80%, 4 out of 5) than Momo the robot (23.08%, 3 out of 13), $p = .047$, $\phi = .523$, two-tailed Fisher's exact test. This finding remained significant when sibling participants were excluded from the dataset, $p = .047$, $\phi = .523$, two-tailed Fisher's exact test. This finding was unexpected. However, given that the N in the virtual agent condition in this case was only 5, it would be imprudent to give too much weight to this finding.

An interesting difference in frequencies of justification use between conditions also emerged in participants' moral obligatory reasoning regarding the act of hitting a human. Participants in both age groups significantly more often appealed to justice concerns in the robot condition than in the virtual agent condition. For the 8- to 9-year-olds, 20% of them appealed to justice concerns in the robot condition compared to 0% in the virtual agent condition, $p = .024$, $\phi = .328$, two-tailed Fisher's exact test. This finding remained significant when sibling participants were excluded from the dataset, $p = .023$, $\phi = .351$, two-tailed Fisher's exact test. For the 14- to 15-year-olds, 27.58% of them appealed to justice concerns in the robot condition compared to 3.84% (one participant) in the virtual agent condition, $p = .027$, $\phi = .320$, two-tailed Fisher's exact test. This finding was marginally significant when sibling participants were excluded from

the dataset, $p = .052$, $\phi = .298$, two-tailed Fisher's exact test. Potential explanations for this finding are discussed in the next chapter.

The effect of age group on frequencies of moral justification use in each condition. In the robot condition, the 8- to 9-year-olds (78.57%) were significantly more likely than the 14- to 15-year-olds (23.08%) to appeal to welfare concerns when justifying their judgments against the act of hitting Momo, $\chi^2 (1, N = 41) = 11.49$, $p = .001$, $\phi = .529$. This finding remained significant when sibling participants were excluded from the dataset, $\chi^2 (1, N = 38) = 11.60$, $p = .001$, $\phi = .552$. This difference was mainly driven by the younger participants' (71.43%) significantly more frequent appeals to direct psychological welfare concerns in comparison to the older participants (23.07%), $\chi^2 (1, N = 41) = 8.43$, $p = .004$, $\phi = .453$. This finding remained significant when sibling participants were excluded from the dataset, $\chi^2 (1, N = 38) = 8.28$, $p = .004$, $\phi = .467$. Since appeals of psychological welfare typically also acknowledged emotional capacity, the younger participants (78.57%) also significantly more often appealed to Momo's emotional capacity in justifying their judgments than the older participants (7.69%), $\chi^2 (1, N = 41) = 18.11$, $p < .001$, $\phi = .665$. This finding remained significant when sibling participants were excluded from the dataset, $\chi^2 (1, N = 38) = 18.09$, $p < .001$, $\phi = .690$. The older participants (38.46%), on the other hand, were significantly more likely than the younger participants (7.14%) to appeal to fairness concern for robot Momo, $p = .024$, $\phi = .387$, two-tailed Fisher's exact test. This finding remained significant when sibling participants were excluded from the dataset, $p = .034$, $\phi = .373$, two-tailed Fisher's exact test.

A few developmental differences in frequencies of justification use were also found in participants' moral obligatory reasoning regarding the act of hitting other entities in the robot condition. The younger participants (96.67%) significantly more often than the older

participants (72.41%) appealed to welfare justifications in justifying why it would not be all right to hit a human, $p = .012$, $\phi = .337$, two-tailed Fisher's exact test. This finding remained significant when sibling participants were excluded from the dataset, $p = .025$, $\phi = .336$, two-tailed Fisher's exact test. In justifying why it would not be all right to hit a pet animal, the younger participants (50%) significantly more often than the older participants (23.33%) appealed to direct psychological welfare concerns, $\chi^2(1, N = 60) = 4.59$, $p = .032$, $\phi = .277$. Accordingly, the younger participants (63.33%) also significantly more often than the older participants (36.67%) made appeals to a pet animal's emotions, $\chi^2(1, N = 60) = 4.27$, $p = .039$, $\phi = .267$. These two findings for a pet animal however were not significant when sibling participants were excluded from the dataset: $\chi^2(1, N = 55) = 2.30$, $p = .130$, $\phi = .204$ for direct psychological welfare; $\chi^2(1, N = 55) = 3.06$, $p = .080$, $\phi = .236$ for emotion.

In the virtual agent condition, two developmental differences in frequencies of justification use were found in participants' moral obligatory reasoning for Momo. The younger participants (85.71%) were significantly more likely than the older participants (0%) to appeal to Momo's emotional capacity, $p = .002$, $\phi = .782$, two-tailed Fisher's exact test. This finding remained significant when sibling participants were excluded from the dataset, $p = .002$, $\phi = .777$, two-tailed Fisher's exact test. The older participants (80%), on the other hand, were significantly more likely than the younger participants (14.29%) to appeal to Momo's artificial emotions, $p = .017$, $\phi = .623$, two-tailed Fisher's exact test. This finding remained significant when sibling participants were excluded from the dataset, $p = .022$, $\phi = .614$, two-tailed Fisher's exact test.

Taken together, these results indicate that in the robot condition, the younger participants significantly more often appealed to welfare concerns than the older participants in justifying their judgments against the act of hitting robot Momo, a human, and a pet animal, and the older

participants significantly more often appealed to fairness concerns than the younger participants in justifying their judgments against the act of hitting robot Momo. While not statistically significant, these trends were also largely exhibited by participants in the virtual agent condition. Furthermore, the older participants were more likely to appeal to Momo's artificial emotional capacity than the younger participants, and this difference was descriptively true in the robot condition and statistically significant in the virtual agent condition.

The effect of gender on frequencies of justice justification use. For the 8- to 9-year-olds in the robot condition, a gender difference was found in frequencies of justification use. The male participants (40%) were significantly more likely than the female participants (0%) to appeal to justice concerns when justifying why it would not be all right to hit a human, $p = .017$, $\phi = .500$, two-tailed Fisher's exact test. This finding remained significant when sibling participants were excluded from the dataset, $p = .016$, $\phi = .515$, two-tailed Fisher's exact test. This difference was mainly driven by the males' (33.33%) significantly more frequent appeals to fairness in comparison to the females (0%), $p = .042$, $\phi = .477$, two-tailed Fisher's exact test. This finding remained significant when sibling participants were excluded from the dataset, $p = .041$, $\phi = .459$, two-tailed Fisher's exact test. This gender difference was limited to the younger participants in the robot condition on their reasoning regarding a human and not found elsewhere, thus it would be imprudent to give too much weight to this finding.

Qualitative presentation of artificial emotion justification data. Justification data in their raw form reflect the nuanced nature of this research and the richness of participants' reasoning. The most novel justification data from this study were those of the artificial emotion category. What follows is a qualitative presentation of examples from this justification category to provide a first-hand feel of this emerging concept that participants had brought to bear in their

endorsements of moral obligation to personified technological entities of the here and now.

When provided in support of “not all right to hit” judgments, the key features of artificial emotion justifications are (1) the entity’s capacity to exhibit and/or have emotions is achieved through artificial means (e.g., programming, engineering, etc.), and (2) these artificially derived emotional appearances and/or states deserve moral consideration. In the following examples, words spoken by the interviewer are depicted in ALL CAPITALS, and ellipses (...) indicate that a portion of the interview was omitted for the clarity of the example.

In its mildest form, an artificial emotion justification can reflect the participant’s ambivalence about the entity’s true emotional capacity, but at the same time, the external display of emotions is deemed sufficient to warrant moral consideration. To help illustrate this point, here is what this line of reasoning might sound like for a human: It was not all right to hit Tommy because Tommy seemed sad about it. I don’t know if Tommy actually felt sad about it, but he seemed like he was sad. In a similar vein, a 14-year-old participant provided the following justification in support her judgment that it was not all right to hit Momo the robot:

Um, I understand you could get frustrated at it, but I feel like you shouldn’t hit it with a book. Um, it looked like he cared a little bit, like it’s programmed to. MHMM. But I don’t know if he actually has real feelings, but the way he reacted it seemed like he did.

Some participants did not show such uncertainty and equated Momo’s emotions to human feelings, but with clear understanding of their artificialness. For example, an 8-year-old participant stated the following regarding Momo the robot:

SO DO YOU THINK IT WAS ALL RIGHT OR NOT ALL RIGHT LYLE HIT MOMO WITH A BOOK? I don’t think it was right. OK AND WHY NOT? Because it hurts people’s feelings, and I think robots have feelings because how they built them.

Some participants also demonstrated in depth understanding of how a robot's behaviors are controlled by its programming, and argued for the validity of artificial emotions despite their programmed nature, such as this 14-year-old participant had done in justifying why it was not all right to hit Momo the robot:

Like if they're taught to be like angry if they get hit, then it's gonna, you're not gonna want to make your robot angry, I don't think. I think because even if it seems so mechanical on the inside, you should not. Like, even if it seems like, oh, this code is making them feel sad, then they are still sad, even if it's being like controlled or forced to react that way.

In justifying why it was not all right to hit Momo the virtual agent, one 15-year-old participant even offered a unique explanation that artificial emotions are somewhat valid emotions because within the parameters of its programming, the technology still has some level of autonomy in its execution of the programming:

DO YOU THINK IT WAS ALL RIGHT OR NOT ALL RIGHT TO HAVE HIT MOMO WITH A BOOK? I don't think so.... It's a robot, but he still somewhat has feelings, because of like I said the programmer, so. SO THE PROGRAMMER PROGRAMS THE ROBOT, UH YOU THINK THE ROBOT STILL SOMEWHAT HAS FEELINGS? Yeah cause the robot has, the robot has control over it at that point, it doesn't depend on the programmer at that point.

Finally, an artificial emotion justification can also entail a sophisticated form of reasoning that establishes both differences and similarities, wherein recognized differences are overridden by similarities. For example, one 15-year-old participant stated the following in justifying why it was not all right to hit Momo the virtual agent:

I feel like even though it is a program it, even like to me, like I'm thinking of it like it's programmed with feelings. MHMM. So it's like, it has feelings like programmed into it, so like it still has them and, you know it's not, it's not necessarily the same as like ours, like if you were mean to it like it's still going to be sad which is like not nice. OK. So like it's the same.

This participant clearly articulated that Momo's feelings are achieved through artificial means, in that they are "programmed into it," and thus "not necessarily the same as like ours." Nevertheless, this participant also recognized that artificial emotions are functionally similar to human emotions, such that they both respond to stimuli and result in certain changed internal states. Based on this functional correspondence, this participant concluded that the known differences do not void the mapping of similar considerations, "so like it's the same."

Seriousness ratings. Participants rated for each entity the seriousness of the hitting act on a 9-point Likert scale ranging from 1 being "all right to hit", to increasing levels of seriousness (e.g., "a little bit not all right to hit," "somewhat not all right to hit," "mostly not all right to hit"), finally to 9 being "not all right to hit." This rating is not a moral measure, but is intended to provide additional insight into the perceived seriousness of the act. Mean ratings of seriousness are reported in Table 5 per condition, by age group, for each entity. Simple effects of condition and age group are examined next to examine the influences of robot physical embodiment and age group on participants' seriousness ratings. The Benjamini-Hochberg procedure was used to adjust for multiple comparisons, controlling the false discovery rate at the .05 level. Benjamini-Hochberg adjusted *p* values are reported below.

The effect of condition on seriousness ratings for each age group. A Mann-Whitney test indicated that 8- to 9-year-olds' ratings for hitting Momo the robot (*Mdn* = 7.00) did not

Table 5

Mean Rating (Standard Deviation) of Seriousness of the Hitting Act

Entity	Robot condition		Virtual agent condition	
	Age 8-9	Age 14-15	Age 8-9	Age 14-15
Momo	7.37 (<i>SD</i> = 1.45)	5.80 (<i>SD</i> = 2.14)	7.20 (<i>SD</i> = 1.69)	4.50 (<i>SD</i> = 2.49)
Human	8.50 (<i>SD</i> = 1.01)	8.77 (<i>SD</i> = 0.50)	8.70 (<i>SD</i> = 0.95)	8.47 (<i>SD</i> = 1.07)
Pet animal	8.33 (<i>SD</i> = 0.96)	8.57 (<i>SD</i> = 0.63)	8.50 (<i>SD</i> = 0.68)	8.47 (<i>SD</i> = 0.97)
Basic phone	5.10 (<i>SD</i> = 2.41)	3.83 (<i>SD</i> = 2.32)	4.87 (<i>SD</i> = 2.49)	3.00 (<i>SD</i> = 2.03)

Note. Seriousness ratings were based on a 9-point Likert scale. Higher mean ratings indicate perceptions of the act as being more serious.

differ significantly from their ratings for hitting Momo the virtual agent ($Mdn = 7.50$), $U = 433.50$, $p = .802$ (Benjamini-Hochberg adjusted), $r = .03$. The 14- to 15-year-olds' ratings for hitting Momo the robot ($Mdn = 6.50$) also did not differ significantly from their ratings of hitting Momo the virtual agent ($Mdn = 4.00$), $U = 324.50$, $p = .080$ (Benjamini-Hochberg adjusted), $r = .24$. In both age groups, no significant differences across conditions were found in seriousness ratings for the acts of hitting a human, a pet animal, and a basic phone.

The effect of age group on seriousness ratings in each condition. In the robot condition, a Mann-Whitney test indicated that 8- to 9-year-olds gave significantly higher ratings for the act of hitting Momo ($Mdn = 7.00$) than the 14- to 15-year-olds ($Mdn = 6.50$), $U = 255.00$, $p = .006$ (Benjamini-Hochberg adjusted), $r = .38$. This finding remained significant when sibling participants were excluded from the dataset, $U = 230.50$, $p = .022$ (Benjamini-Hochberg adjusted), $r = .34$. In the virtual agent condition, 8- to 9-year-olds also gave significantly higher ratings for the act of hitting Momo ($Mdn = 7.50$) than the 14- to 15-year-olds ($Mdn = 4.00$), $U = 177.00$, $p < .001$ (Benjamini-Hochberg adjusted), $r = .53$. This finding remained significant

when sibling participants were excluded from the dataset, $U = 154.00$, $p < .001$ (Benjamini-Hochberg adjusted), $r = .55$.

In both conditions, no significant developmental differences were found in seriousness ratings for the acts of hitting a human and a pet animal. Significant differences were found for the act of hitting a basic phone, such that in the virtual agent condition, 8- to 9-year-olds gave significantly higher ratings for the act of hitting the basic phone ($Mdn = 5.00$) than the 14- to 15-year olds ($Mdn = 2.00$), $U = 246.00$, $p = .008$ (Benjamini-Hochberg adjusted), $r = .39$. This finding remained significant when sibling participants were excluded from the dataset, $U = 225.00$, $p = .008$ (Benjamini-Hochberg adjusted), $r = .40$. In the robot condition however, 8- to 9-year-olds ($Mdn = 5.00$) did not significantly differ from the 14- to 15-year olds ($Mdn = 3.00$) in their ratings for the act of hitting the basic phone, $U = 318.00$, $p = .096$ (Benjamini-Hochberg adjusted), $r = .25$.

In both conditions, no significant developmental differences were found in seriousness ratings for the acts of hitting a human and pet animal, although the younger participants judged the acts of hitting Momo and the basic phone (in the virtual agent condition) as being more serious than the older participants did.

Distinctions between endorsements of moral obligation and seriousness ratings for the act of hitting Momo. Ratings of the seriousness of the hitting act provided additional insights into participants' conceptions about a human causing a potential harm to personified technological entities (e.g., while participants were more likely to endorse moral obligation against hitting a physical robot than hitting a virtual agent, they did not judge one act as being more serious than the other). Although both the moral obligation interview questions and the Likert scale instrument asked about the all right or not all rightness of hitting Momo, there were

important distinctions between the data obtained from these measures. Endorsements of moral obligation did not necessarily correspond to high ratings on the Likert scale. For those participants that viewed hitting Momo as a violation of moral obligation, their Likert scale ratings ranged from as low as 4, below the midpoint, to 9. Not endorsing moral obligation also did not necessarily correspond to low Likert scale ratings. For those participants that did not view hitting Momo as a violation of moral obligation, their Likert scale ratings ranged from 1 all the way to 9, with 60% of them (36 out of 60 participants) giving ratings of 5 or higher. All of these 36 participants had appealed to human-centered welfare justifications in their interview responses (77.78% material welfare and 22.22% psychological welfare).

Given these distinctions, Mann-Whitney tests were conducted to assess for differences in ratings between participants who endorsed moral obligation against hitting Momo and those who did not. For the 8- to 9-year-olds in the robot condition, no significant difference was found between those that endorsed moral obligation ($Mdn = 7.00$) and those that did not ($Mdn = 8.00$), $U = 21.00$, $p = .548$ (Benjamini-Hochberg adjusted), $r = .11$. This was also the case in the virtual agent condition, no significant difference was found between those that endorsed moral obligation ($Mdn = 8.00$) and those that did not ($Mdn = 7.00$), $U = 77.00$, $p = .183$ (Benjamini-Hochberg adjusted), $r = .27$. However, for the 14- to 15-year-olds in the robot condition, those that endorsed moral obligation gave significantly higher ratings ($Mdn = 7.00$) than those that did not ($Mdn = 6.00$), $U = 39.50$, $p = .008$ (Benjamini-Hochberg adjusted), $r = .55$. This finding remained significant when sibling participants were excluded from the dataset, $U = 37.00$, $p = .016$ (Benjamini-Hochberg adjusted), $r = .52$. In the virtual agent condition, those that endorsed moral obligation also gave higher ratings ($Mdn = 7.00$) than those that did not ($Mdn = 3.00$), but this difference was not significant, $U = 31.00$, $p = .152$ (Benjamini-Hochberg adjusted),

$r = .32$. Given that significant difference was found at least in older participants' ratings between the two groups, analyses were conducted next to examine the effects of condition for each age group and the effects of age group in each condition on seriousness ratings of hitting Momo for those participants that endorsed moral obligation.

The effect of condition for each age group on seriousness ratings of the Momo hitting act by participants endorsing moral obligation. Mann-Whitney tests were first conducted to test for the effect of condition for each age group. For the 8- to 9-year-olds that endorsed moral obligation against hitting Momo, their ratings for hitting Momo the robot ($Mdn = 7.00$) did not differ significantly from their ratings for hitting Momo the virtual agent ($Mdn = 8.00$), $U = 166.50$, $p = .555$ (Benjamini-Hochberg adjusted), $r = .13$. For the 14- to 15-year-olds that endorsed moral obligation against hitting Momo, their ratings of hitting Momo the robot ($Mdn = 7.00$) also did not differ significantly from their ratings of hitting Momo the virtual agent ($Mdn = 7.00$), $U = 23.50$, $p = .714$ (Benjamini-Hochberg adjusted), $r = .22$.

The effect of age group in each condition on seriousness ratings of the Momo hitting act by participants endorsing moral obligation. Mann-Whitney tests were then conducted to test for the effect of age group in each condition. For those participants that endorsed moral obligation against hitting Momo in the robot condition, 8- to 9-year-olds' ratings ($Mdn = 7.00$) did not significantly differ from 14- to 15-year-olds' ratings ($Mdn = 7.00$), $U = 164.50$, $p = .604$ (Benjamini-Hochberg adjusted), $r = .08$. This was also the case for those participants that endorsed moral obligation against hitting Momo in the virtual agent condition, 8- to 9-year-olds' ratings ($Mdn = 8.00$) did not significantly differ from 14- to 15-year-olds' ratings ($Mdn = 7.00$), $U = 18.50$, $p = .420$ (Benjamini-Hochberg adjusted), $r = .37$.

Recall that significant developmental differences in seriousness ratings for hitting Momo were found for the full dataset, such that the younger participants judged the acts of hitting Momo (robot/virtual agent) as being more serious than the older participants. However, when the sample was restricted to those participants that endorsed moral obligation against hitting Momo, the developmental differences were no longer significant. Taken together, these results suggest that participants who endorsed moral obligation against hitting Momo all judged hitting Momo as being relatively serious, and their ratings were not significantly affected by the condition they participated in or the age group they belonged to.

Prior experience with technology. To assess whether there are significant influences of participants' prior experience with relevant technologies on their endorsements of moral obligation against hitting Momo (robot/virtual agent), three separate scales (i.e., Smartphone & Device Experience Scale, Robot Experience Scale, and Computing Technology Experience Scale) were constructed based on questionnaire data. Subjective measures from the questionnaire (i.e., parental ratings of participants' knowledge of a technology) were not included as items in the scales. Partially based on the robot experience scale used in the Bernstein and Crowley (2008) study and generalized to all three types of technologies, objective measures of participants' experience with each type of technology from the questionnaire (e.g., whether participants owned the technology, whether participants ever used the technology, etc.) were used as items in the scales. Frequencies of use ratings were also included as items in the scales as it has been proposed that dichotomous responses alone can lead to distorted results, and that multiple-choice item formats are more reliable, give more stable results, and produce better scales (Clark & Watson, 1995; Comrey, 1988).

Since only straightforward questionnaire items that directly related to the participants' prior experience with each type of technology were included as scale items, a simplified weighting method, partly based on method for combining rating and dichotomous data in scale construction (Wyse & Lawrence, 2006), was used in forming the scales, such that each increasing level in a rating item (e.g., frequency of use) was given equal weight as a dichotomous item. This effectively put more weight on the frequencies of use ratings in each of the three experience scales, corresponding to the logic that how often one uses a technology is a more meaningful measure of one's experience with the technology than simply owning the technology or having used it once before.

Descriptive statistics on each of the scales are reported in Table 6. Binary logistic regression analyses were performed to assess (1) potential significant influences of prior experience with each type of technology on participants' endorsements of moral obligation against hitting Momo, and (2) potential moderating effects of experience on the developmental difference in endorsements (by examining interactions between age group and experience). Details for each scale are reported below.

Smartphone & device experience scale. The following questions (and point assignments) were used in the construction of the Smartphone & Device Experience Scale ranging from 0 to 14: whether the participant ever used a smartphone (no = 0, yes = 1); whether the participant owns a smartphone (no = 0, yes = 1); how many hours per week does the participant use smartphone (not applicable/never = 0, <1 hour = 1, 1-3 hours = 2, 3-5 hours = 3, 5-10 hours = 4, 10+ hours = 5); whether the participant ever used a smart device (e.g, iPod, iPad; no = 0, yes = 1); whether the participant owns a smart device (no = 0, yes = 1); how many hours per week does the participant use smart device (not applicable/never = 0, <1 hour = 1, 1-3 hours = 2, 3-5 hours

Table 6

Mean Score (Standard Deviation) of Technology Experience Scales

Scale	Robot condition		Virtual agent condition	
	Age 8-9	Age 14-15	Age 8-9	Age 14-15
Smartphone & device	7.13 (<i>SD</i> = 1.63)	9.70 (<i>SD</i> = 2.41)	6.38 (<i>SD</i> = 1.74)	10.34 (<i>SD</i> = 2.45)
Robot	2.73 (<i>SD</i> = 2.46)	2.30 (<i>SD</i> = 2.47)	2.40 (<i>SD</i> = 2.25)	2.80 (<i>SD</i> = 2.28)
Computing technology	7.44 (<i>SD</i> = 2.06)	9.00 (<i>SD</i> = 2.75)	6.84 (<i>SD</i> = 1.77)	9.40 (<i>SD</i> = 2.94)

Note. Smartphone & Device Experience Scale ranged from 0 to 14. Robot Experience Scale ranged from 0 to 12. Computing Technology Experience Scale ranged from 0 to 13. Higher mean scores indicate more experience.

= 3, 5-10 hours = 4, 10+ hours = 5).

A 2 (condition) x 2 (age group) ANOVA yielded a main effect of age group for smartphone & device experience scores, $F(1, 114) = 71.85, p < .001$, partial $\eta^2 = .387$, indicating that the 8- to 9-year-olds on average scored lower on the scale ($M = 6.76, SD = 1.72$) than the 14- to 15-year-old ($M = 10.02, SD = 2.43$). The main effect of condition was non-significant, $F(1, 114) = 0.02, p = .888$, partial $\eta^2 = .000$, and the interaction effect was also non-significant, $F(1, 114) = 3.30, p = .072$, partial $\eta^2 = .028$.

A binary logistic regression analysis was then conducted with condition and age group entered as predictors in the first step, Smartphone & Device Experience Scale entered as predictor in the second step, and endorsement of moral obligation against hitting Momo as the dependent variable. Controlling for condition and age group, the Smartphone & Device Experience Scale scores were not significant in predicting endorsements of moral obligation, $\text{Wald } \chi^2 = 0.85, df = 1, p = .358$.

To test for interaction between age group and Smartphone & Device Experience Scale, another binary logistic regression analysis was conducted with condition, age group, and

Smartphone & Device Experience Scale entered in the first step, age group x Smartphone & Device Experience Scale interaction term entered in the second step, and endorsement of moral obligation against hitting Momo as the dependent variable. The age group x Smartphone & Device Experience Scale interaction was also not significant in predicting endorsements of moral obligation, Wald $\chi^2 = 0.05$, $df = 1$, $p = .823$. These results indicate that neither participants' endorsements of moral obligation nor the developmental difference in their endorsements were significantly influenced by their prior experience with smartphones and devices.

Robot experience scale. The following questions (and point assignments) were used in the construction of the Robot Experience Scale ranging from 0 to 12: whether the participant ever interacted with a robot (no = 0, yes = 1); the number of robots that the participant had interacted with (not applicable/none = 0, 2 or less = 1, 3 or more = 2); whether the participant owns a robot (no = 0, yes = 1); how many hours per week does the participant interact with robot (not applicable/never = 0, <1 hour = 1, 1-3 hours = 2, 3-5 hours = 3, 5-10 hours = 4, 10+ hours = 5); whether the participant read robot related book (no = 0, yes = 1); whether the participant watched robot related movie (no = 0, yes = 1); whether the participant ever built a robot (no = 0, yes = 1).

For robot experience scores, a 2 (condition) x 2 (age group) ANOVA indicated non-significant main effect of age group, $F(1, 116) = 0.001$, $p = .969$, partial $\eta^2 = .000$, non-significant main effect of condition, $F(1, 116) = 0.04$, $p = .847$, partial $\eta^2 = .000$, and non-significant interaction effect, $F(1, 116) = 0.93$, $p = .337$, partial $\eta^2 = .008$. Assessments of gender differences in each age group revealed that among the 14- to 15-year-olds, males scored significantly higher on the scale ($M = 3.27$, $SD = 2.57$) than females ($M = 1.83$, $SD = 1.93$), $t(53.812) = -2.441$, $p = .018$, $d = 0.63$.

A binary logistic regression analysis was then conducted with condition, age group, and gender (since a significant gender difference was found for scores on this scale) entered as predictors in the first step, Robot Experience Scale entered as predictor in the second step, and endorsement of moral obligation against hitting Momo as the dependent variable. Controlling for condition, age group, and gender, the Robot Experience Scale scores were not significant in predicting endorsements of moral obligation, Wald $\chi^2 = 2.28$, $df = 1$, $p = .131$.

To test for interaction between age group and Robot Experience Scale, another binary logistic regression analysis was conducted with condition, age group, gender, and Robot Experience Scale entered in the first step, age group x Robot Experience Scale interaction term entered in the second step, and endorsement of moral obligation against hitting Momo as the dependent variable. The age group x Robot Experience Scale interaction was also not significant in predicting endorsements of moral obligation, Wald $\chi^2 = 0.39$, $df = 1$, $p = .534$. These results indicate that neither participants' endorsements of moral obligation nor the developmental difference in their endorsements were significantly influenced by their prior experience with robots.

Computing technology experience scale. The following questions (and point assignments) were used in the construction of the Computing Technology Experience Scale ranging from 0 to 13: whether the participant ever used a computer (no = 0, yes = 1); how many hours per week does the participant use computer (not applicable/never = 0, <1 hour = 1, 1-3 hours = 2, 3-5 hours = 3, 5-10 hours = 4, 10+ hours = 5); whether the participant ever played a video game (no = 0, yes = 1); how many hours per week does the participant play video game (not applicable/never = 0, <1 hour = 1, 1-3 hours = 2, 3-5 hours = 3, 5-10 hours = 4, 10+ hours = 5); whether the participant had programming experience (no = 0, yes = 1).

A 2 (condition) x 2 (age group) ANOVA yielded a main effect of age group for computing technology experience scores, $F(1, 96) = 18.09, p < .001$, partial $\eta^2 = .159$, indicating that the 8- to 9-year-olds on average scored lower on the scale ($M = 7.15, SD = 1.93$) than the 14- to 15-year-olds ($M = 9.21, SD = 2.83$). The main effect of condition was non-significant, $F(1, 96) = 0.05, p = .833$, partial $\eta^2 = .000$, and the interaction effect was also non-significant, $F(1, 96) = 1.08, p = .302$, partial $\eta^2 = .011$. Assessments of gender differences in each age group also revealed that among the 14- to 15-year-olds, males scored significantly higher on the scale ($M = 10.32, SD = 2.39$) than females ($M = 7.65, SD = 2.70$), $t(46) = -3.618, p = .001, d = 1.05$.

A binary logistic regression analysis was then conducted with condition, age group, and gender (since a significant gender difference was found for scores on this scale) entered as predictors in the first step, Computing Technology Experience Scale entered as predictor in the second step, and endorsement of moral obligation against hitting Momo as the dependent variable. Controlling for condition, age group, and gender, the Computing Technology Experience Scale scores were not significant in predicting endorsements of moral obligation, Wald $\chi^2 = 0.14, df = 1, p = .706$.

To test for interaction between age group and Computing Technology Experience Scale, another binary logistic regression analysis was conducted with condition, age group, gender, and Computing Technology Experience Scale entered in the first step, age group x Computing Technology Experience Scale interaction term entered in the second step, and endorsement of moral obligation against hitting Momo as the dependent variable. The age group x Computing Technology Experience Scale interaction was also not significant in predicting endorsements of moral obligation, Wald $\chi^2 = 0.06, df = 1, p = .802$. These results indicate that neither participants'

endorsements of moral obligation nor the developmental difference in their endorsements were significantly influenced by their prior experience with computing technology.

Overall, results from the above analyses suggest that participants' prior experience with these different types of relevant technologies (based on the measures collected in this study) did not significantly influence their endorsements of moral obligation against hitting Momo, nor did their prior experience with these technologies significantly influence the developmental difference in their endorsements.

Endorsements of Rights

In addition to questions regarding the hitting act, participants were also asked about their judgments regarding two specific rights for each entity. The first set of rights-based questions focused on the right to be free from ownership, and consisted of two questions: whether it would be all right or not all right to own and sell an entity, and whether the entity has the right to not be owned and sold. The second set of rights-based questions focused on the right to be free from interference in the form of experimentation, and also consisted of two questions: whether it was all right or not all right to experiment on an entity, and whether the entity has the right to not be experimented on.

The right to be free from ownership. Endorsements of an entity's right to be free from ownership consisted of "not all right to own and sell the entity" and "the entity has the right to not be owned and sold" judgments. Percentages of participants who provided such judgments are reported in Table 7, per condition, by age group, for each entity. As can be seen descriptively in the table, only a small minority of participants endorsed Momo's (robot/virtual agent) and a pet animal's rights on these ownership questions, while virtually everyone did so for a human, and no one did so for a basic phone.

Table 7

Percentage of Participants Who Affirmed Moral Standing With Regard to Owning & Selling

Evaluation affirming	Momo		Human		Pet animal		Basic phone	
	Age 8-9	Age 14-15	Age 8-9	Age 14-15	Age 8-9	Age 14-15	Age 8-9	Age 14-15
	Robot condition							
Not all right to own/sell	6.67	0	100	100	3.45 ^a	6.67	0 ^a	0
Right to not be owned/sold	6.67	3.33	100	96.67	10.34 ^a	20.00	0 ^a	0
	Virtual agent condition							
Not all right to own/sell	6.67	6.67	100 ^a	100	3.33	6.67	0	0
Right to not be owned/sold	13.33	6.67	100 ^a	100	23.33	20.00	0	0

Note. Unless otherwise specified, $N = 30$.

^a $N = 29$.

Next, chi-square tests were performed. Simple effects of condition and age group were all non-significant on either question for any of the entities. Participants' endorsements of each entity's right to be free from ownership were not significantly affected by condition or age group.

The right to be free from experimentation. Endorsements of an entity's right to be free from experimentation consisted of "not all right to experiment on the entity" and "the entity has the right to not be experimented on" judgments. Percentages of participants who provided such judgments are reported in Table 8, per condition, by age group, for each entity. It can be observed descriptively from the table that over half of the younger participants endorsed robot Momo's rights on these experimentation questions, while about one-third of them did so for virtual agent Momo. Only a few of the older participants endorsed rights for either robot or virtual agent Momo. A majority of the participants in each age group endorsed a pet animal the right to be free from experimentation, virtually everyone did so for a human, and virtually no one did so for a basic phone. The Benjamini-Hochberg procedure was used to adjust for multiple comparisons, controlling the false discovery rate at the .05 level. Benjamini-Hochberg adjusted p values are reported below.

Table 8

Percentage of Participants Who Affirmed Moral Standing With Regard to Experimentation

Evaluation affirming	Momo		Human		Pet animal		Basic phone	
	Age 8-9	Age 14-15	Age 8-9	Age 14-15	Age 8-9	Age 14-15	Age 8-9	Age 14-15
	Robot condition							
Not all right to exp. on	55.17 ^a ₁	3.33 ₁	100 ^b	100	89.66 ^a ₆	66.67 ₆	10.34 ^a	0
Right to not be exp. on	68.97 ^a _{2,5}	3.33 ₂	100 ^b	100	93.10 ^a ₇	70.00 ₇	6.90 ^a	0
	Virtual agent condition							
Not all right to exp. on	30.00 ₃	6.67 ₃	100	96.67	70.00	63.33	3.33	3.33
Right to not be exp. on	26.67 _{4,5}	3.33 ₄	100	96.67	76.67	65.52 ^a	0	0

Note. Unless otherwise specified, $N = 30$. Percentages sharing a common number subscript are statistically different at $p < .05$, chi-square tests.

^a $N = 29$. ^b $N = 28$.

The effect of condition on endorsements of the right to be free from experimentation for each age group. Chi-square tests were conducted to compare participants' endorsements on the experimentation questions across the two conditions for each age group. For endorsements on the question whether it would be all right or not all right to experiment on Momo, the 8- to 9-year-olds did not significantly differ between the robot condition (55.17%) and the virtual agent condition (30.00%) in their judgments that it would not be all right to experiment on Momo, $\chi^2(1, N = 59) = 3.83, p = .133$ (Benjamini-Hochberg adjusted), $\phi = .255$. However, they were significantly more likely to judge in the robot condition (68.97%) than in the virtual agent condition (26.67%) that Momo has the right to not be experimented on, $\chi^2(1, N = 59) = 10.58, p = .004$ (Benjamini-Hochberg adjusted), $\phi = .423$. This finding remained significant when sibling participants were excluded from the dataset, $\chi^2(1, N = 54) = 10.68, p = .004$ (Benjamini-Hochberg adjusted), $\phi = .445$. On the other hand, no significant differences across conditions were found in the 14- to 15-year-olds' endorsements on these two questions. Thus, there was a condition effect for the younger participants on one of the experimentation questions, but not for

the older participants. No significant differences in endorsements for a human, a pet animal, and a basic phone were found across the two conditions for either age group.

The effect of age group on endorsements of the right to be free from experimentation in each condition. Chi-square tests were also conducted to compare participants' endorsements on the experimentation questions across the two age groups in each condition. In the robot condition, the younger participants (55.17%) were significantly more likely than the older participants (3.33%) to judge that it would not be all right to experiment on Momo, $\chi^2(1, N = 59) = 19.32, p < .001$ (Benjamini-Hochberg adjusted), $\phi = .572$. This finding remained significant when sibling participants were excluded from the dataset, $\chi^2(1, N = 54) = 18.94, p < .001$ (Benjamini-Hochberg adjusted), $\phi = .592$. The younger participants (68.97%) were also significantly more likely than the older participants (3.33%) to judge that Momo has the right to not be experiment on, $\chi^2(1, N = 59) = 27.71, p < .001$ (Benjamini-Hochberg adjusted), $\phi = .685$. This finding remained significant when sibling participants were excluded from the dataset, $\chi^2(1, N = 54) = 27.93, p < .001$ (Benjamini-Hochberg adjusted), $\phi = .719$. While similar descriptive trends were also observed in the virtual agent condition on these two experimentation questions, the developmental differences however were non-significant after adjusting for multiple comparisons. No significant differences in endorsements for a human, a pet animal, and a basic phone were found across the two age groups in either condition.

Taken together, these results indicate that participants' endorsements for a human's, a pet animal's, and a basic phone's right to be free from experimentation were not significantly affected by condition or age group, while their endorsements for Momo showed some significant differences as follows: the younger participants (but not the older participants) were significantly more likely to endorse robot Momo than virtual agent Momo on the right to not be experimented

on, and the younger participants were also more likely than the older participants to endorse Momo the robot with the right to be free from experimentation on both questions.

Between Entities Comparisons

This section assesses differences in participants' endorsements of moral obligation against the hitting act, seriousness ratings for the hitting act, and endorsements of rights on the four rights-based questions between entities. Tests were performed per condition for each age group for more precise understanding of potential between entities relationships within each of these groups. For all tests in this section, the same patterns of results were found when conducted with or without sibling participants. Results based on the full dataset are reported below.

Endorsements of moral obligation. Repeated measures omnibus tests were conducted to test differences in endorsements of moral obligation against the hitting act between entities. For the 8- to 9-year-olds in the robot condition, analysis indicated significant differences in proportions of endorsements between entities, $F(1.701, 47.619) = 260.56, p < .001$, partial $\eta^2 = .903$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following relationship (significant difference indicated by ">"): Human & Pet Animal & Robot Momo > Basic Phone. The results of the pairwise comparisons are reported in Table 9.

For the 14- to 15-year-olds in the robot condition, analysis also indicated significant differences in proportions of endorsements between entities, $F(1.235, 35.823) = 98.38, p < .001$, partial $\eta^2 = .772$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human & Pet Animal > Robot Momo > Basic Phone.

Table 9

Mean Difference Between Entities in Endorsements of Moral Obligation Against the Hitting Act

Pairwise comparison	Robot condition		Virtual agent condition	
	Age 8-9	Age 14-15	Age 8-9	Age 14-15
Momo - Human	-0.07	-0.53*	-0.54*	-0.70*
Momo - Pet Animal	-0.07	-0.57*	-0.57*	-0.77*
Momo - Basic Phone	0.90*	0.43*	0.39**	0.17
Human - Pet Animal	0.00	-0.03	-0.04	-0.07
Human – Basic Phone	0.97*	0.97*	0.93*	0.87*
Pet Animal – Basic Phone	0.97*	1.00	0.96*	0.93*

Note. Values indicate mean differences in the proportion of moral obligation endorsements between the pairwise entities as specified. Negative (-) values indicate the first listed entity received fewer endorsements than the second listed entity.

* $p < .001$, ** $p < .01$, *** $p < .05$, after Bonferroni corrections.

For the 8- to 9-year-olds in the virtual agent condition, analysis again indicated significant differences in proportions of endorsements between entities, $F(1.411, 38.103) = 68.46$, $p < .001$, partial $\eta^2 = .717$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human & Pet Animal > Virtual Agent Momo > Basic Phone.

Finally, for the 14- to 15-year-olds in the virtual agent condition, analysis indicated significant differences in proportions of endorsements between entities, $F(2.411, 69.929) = 87.93$, $p < .001$, partial $\eta^2 = .752$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human & Pet Animal > Virtual Agent Momo & Basic Phone.

Together these results indicate that in terms of endorsing moral obligation against hitting an entity, participants overall did not significantly differentiate between a human and a pet animal and viewed both as deserving such endorsements, but all viewed a basic phone as least deserving of such endorsements. Furthermore, the younger participants did not significantly

differentiate between robot Momo and a human and a pet animal in their endorsements, but they did view virtual agent Momo as less deserving of such endorsements than a human and a pet animal. The older participants, however, did significantly differentiate robot Momo from a human and a pet animal and viewed the robot as less deserving of such endorsement. In addition, they also did not significantly differentiate between virtual agent Momo and a basic phone and saw both as least deserving endorsements of moral obligation.

Seriousness ratings. Next, repeated measures omnibus tests were conducted to test differences in seriousness ratings of the hitting act between entities. These analyses were conducted on the full dataset to provide an overall assessment of how participants' seriousness ratings compared between entities. For the 8- to 9-year-olds in the robot condition, analysis indicated significant differences in mean ratings between entities, $F(2.248, 65.185) = 38.81$, $p < .001$, partial $\eta^2 = .572$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human & Pet Animal > Robot Momo > Basic Phone. The results of the pairwise comparisons are reported in Table 10.

For the 14- to 15-year-olds in the robot condition, analysis also indicated significant differences in mean ratings between entities, $F(1.788, 51.854) = 79.13$, $p < .001$, partial $\eta^2 = .732$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human & Pet Animal > Robot Momo > Basic Phone.

For the 8- to 9-year-olds in the virtual agent condition, analysis again indicated significant differences in mean ratings between entities, $F(1.808, 52.431) = 53.367$, $p < .001$, partial $\eta^2 = .648$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human & Pet Animal > Virtual Agent Momo > Basic Phone.

Table 10

Mean Difference Between Entities in Seriousness Ratings for the Hitting Act

Pairwise comparison	Robot condition		Virtual agent condition	
	Age 8-9	Age 14-15	Age 8-9	Age 14-15
Momo - Human	-1.13**	-2.97*	-1.50*	-3.97*
Momo - Pet Animal	-0.97***	-2.77*	-1.30*	-3.97*
Momo - Basic Phone	2.27*	1.97*	2.33*	1.50*
Human - Pet Animal	0.17	0.20	0.20	0
Human - Basic Phone	3.40*	4.93*	3.83*	5.47*
Pet Animal - Basic Phone	3.23*	4.73*	3.63*	5.47*

Note. Values indicate mean differences in seriousness ratings between the pairwise entities as specified. Negative (-) values indicate the first listed entity received lower mean rating than the second listed entity.

* $p < .001$, ** $p < .01$, *** $p < .05$, after Bonferroni corrections.

Finally, for the 14- to 15-year-olds in the virtual agent condition, analysis indicated significant differences in mean ratings between entities, $F(1.460, 42.352) = 115.80, p < .001$, partial $\eta^2 = .800$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human & Pet Animal > Virtual Agent Momo > Basic Phone.

In terms of participants' perceptions of the seriousness of the hitting act, the same pattern of findings emerged from this set of analyses. Participants overall did not significantly differentiate between a human and a pet animal and rated hitting both as being most serious. Participants rated hitting Momo (robot/virtual agent) as being less serious than hitting a human and a pet animal, but more serious than hitting a basic phone. When considered alongside the results from between entities comparisons of moral obligation endorsements, it would appear that while the younger participants did not significantly differentiate between a human, a pet animal, and robot Momo in their moral obligation endorsements, they did perceive hitting a human and a pet animal as being more serious than hitting robot Momo. Similarly, while the older participants did not significantly differentiate between virtual agent Momo and a basic

phone in their moral obligation endorsements, they did perceive hitting virtual agent Momo as being more serious than hitting a basic phone.

Endorsements of rights. Repeated measures omnibus tests were then conducted to test differences in endorsements of rights on the four rights-based questions between entities.

The right to be free from ownership. On the first question whether it would be all right or not all right to own and sell an entity, for the 8- to 9-year-olds in the robot condition, analysis indicated significant differences in proportions of not all right judgments between entities, $F(1.515, 42.414) = 344.17, p < .001$, partial $\eta^2 = .925$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal & Robot Momo & Basic Phone. The results of the pairwise comparisons are reported in Table 11.

For the 14- to 15-year-olds in the robot condition, analysis also indicated significant differences in proportions of not all right judgments between entities, $F(1, 29) = 447.43, p < .001$, partial $\eta^2 = .939$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal & Robot Momo & Basic Phone.

For the 8- to 9-year-olds in the virtual agent condition, analysis again indicated significant differences in proportions of not all right judgments between entities, $F(1.629, 45.606) = 264.32, p < .001$, partial $\eta^2 = .904$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal & Virtual Agent Momo & Basic Phone.

Finally, for the 14- to 15-year-olds in the virtual agent condition, analysis indicated significant differences in proportions of not all right judgments between entities, $F(1.953, 56.640)$

Table 11

Mean Difference Between Entities in Endorsements of the Right to be Free from Ownership

Pairwise comparison	Robot condition		Virtual agent condition	
	Age 8-9	Age 14-15	Age 8-9	Age 14-15
	Not all right to own and sell			
Momo - Human	-0.93*	-1.00 [†]	-0.93*	-0.93*
Momo - Pet Animal	0.03	-0.07	0.03	0.00
Momo - Basic Phone	0.07	0.00	0.07	0.07
Human - Pet Animal	0.97*	0.93*	0.97*	0.93*
Human – Basic Phone	1.00 [†]	1.00 [†]	1.00 [†]	1.00 [†]
Pet Animal – Basic Phone	0.03	0.07	0.03	0.07
	Has the right to not be owned and sold			
Momo - Human	-0.93*	-0.93*	-0.86*	-0.93*
Momo - Pet Animal	-0.03	-0.17	-0.10	-0.13
Momo - Basic Phone	0.07	0.03	0.14	0.07
Human - Pet Animal	0.90*	0.77*	0.76*	0.80*
Human – Basic Phone	1.00 [†]	0.97*	1.00 [†]	1.00 [†]
Pet Animal – Basic Phone	0.10	0.20	0.24***	0.20

Note. Values indicate mean differences in the proportion of rights endorsements between the pairwise entities as specified. Negative (-) values indicate the first listed entity received fewer endorsements than the second entity. * $p < .001$, ** $p < .01$, *** $p < .05$, after Bonferroni corrections. [†]Mean difference of 1.00 indicate that 100% of the participants endorsed for one entity and 0% for the other entity in the pairwise comparison.

= 252.83, $p < .001$, partial $\eta^2 = .897$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal & Virtual Agent Momo & Basic Phone.

On the second question whether an entity has the right to not be owned and sold, for the 8- to 9-year-olds in the robot condition, analysis indicated significant differences in proportions of affirmative judgments between entities, $F(1.671, 46.775) = 155.10$, $p < .001$, partial $\eta^2 = .847$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal & Robot Momo & Basic Phone.

For the 14- to 15-year-olds in the robot condition, analysis also indicated significant differences in proportions of affirmative judgments between entities, $F(1.720, 49.868) = 118.13$,

$p < .001$, partial $\eta^2 = .803$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal & Robot Momo & Basic Phone.

For the 8- to 9-year-olds in the virtual agent condition, analysis again indicated significant differences in proportions of affirmative judgments between entities, $F(1.744, 48.829) = 74.568$, $p < .001$, partial $\eta^2 = .727$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationships: Human > Pet Animal > Basic Phone, and Human > Virtual Agent Momo. The proportion of affirmative judgments for virtual agent Momo did not significantly differ from that of a pet animal or a basic phone.

Finally, for the 14- to 15-year-olds in the virtual agent condition, analysis indicated significant differences in proportions of affirmative judgments between entities, $F(1.638, 47.488) = 119.09$, $p < .001$, partial $\eta^2 = .804$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal & Virtual Agent Momo & Basic Phone.

Overall, a clear pattern emerged from these between entities comparisons of participants' endorsements on the ownership questions. In terms of the right to be free from ownership, participants largely did not significantly differentiate between Momo (robot/virtual agent), a pet animal, and a basic phone and very infrequently endorsed this right to them. However, virtually everyone viewed a human as the only entity deserving of this right.

The right to be free from experimentation. On the first question whether it would be all right or not all right to experiment on an entity, for the 8- to 9-year-olds in the robot condition, analysis indicated significant differences in proportions of not all right judgments between entities, $F(2.257, 60.937) = 45.56$, $p < .001$, partial $\eta^2 = .628$, with Greenhouse-Geisser

correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human & Pet Animal > Robot Momo > Basic Phone. The results of the pairwise comparisons are reported in Table 12.

For the 14- to 15-year-olds in the robot condition, analysis also indicated significant differences in proportions of not all right judgments between entities, $F(1.258, 36.478) = 113.17$, $p < .001$, partial $\eta^2 = .796$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal > Robot Momo & Basic Phone.

For the 8- to 9-year-olds in the virtual agent condition, analysis again indicated significant differences in proportions of not all right judgments between entities, $F(1.877, 54.430) = 48.48$, $p < .001$, partial $\eta^2 = .626$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal > Virtual Agent Momo > Basic Phone.

Finally, for the 14- to 15-year-olds in the virtual agent condition, analysis indicated significant differences in proportions of not all right judgments between entities, $F(1.724, 49.987) = 80.43$, $p < .001$, partial $\eta^2 = .735$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal > Virtual Agent Momo & Basic Phone.

On the second question whether an entity has the right to not be experimented on, for the 8- to 9-year-olds in the robot condition, analysis indicated significant differences in proportions of affirmative judgments between entities, $F(2.011, 54.307) = 59.43$, $p < .001$, partial $\eta^2 = .688$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationships: Human > Robot Momo > Basic Phone, and Pet Animal >

Table 12

Mean Difference Between Entities in Endorsements of the Right to be Free from Experimentation

Pairwise comparison	Robot condition		Virtual agent condition	
	Age 8-9	Age 14-15	Age 8-9	Age 14-15
	Not all right to experiment on			
Momo - Human	-0.43**	-0.97*	-0.70*	-0.90*
Momo - Pet Animal	-0.32***	-0.63*	-0.40***	-0.57*
Momo - Basic Phone	0.46*	0.03	0.27***	0.03
Human - Pet Animal	0.11	0.33**	0.30**	0.33**
Human – Basic Phone	0.89*	1.00 [†]	0.97*	0.93*
Pet Animal – Basic Phone	0.79*	0.67*	0.67*	0.60*
	Has the right to not be experimented on			
Momo - Human	-0.32**	-0.97*	-0.73*	-0.93*
Momo - Pet Animal	-0.25	-0.67*	-0.50**	-0.62*
Momo - Basic Phone	0.61*	0.03	0.27***	0.03
Human - Pet Animal	0.07	0.30**	0.23***	0.31**
Human – Basic Phone	0.93*	1.00 [†]	1.00 [†]	0.97*
Pet Animal – Basic Phone	0.86*	0.70*	0.77*	0.66*

Note. Values indicate mean differences in the proportion of rights endorsements between the pairwise entities as specified. Negative (-) values indicate the first listed entity received fewer endorsements than the second entity. * $p < .001$, ** $p < .01$, *** $p < .05$, after Bonferroni corrections. [†]Mean difference of 1.00 indicate that 100% of the participants endorsed for one entity and 0% for the other entity in the pairwise comparison.

Basic Phone. The proportion of affirmative judgments for a pet animal did not significantly differ from that of a human or robot Momo.

For the 14- to 15-year-olds in the robot condition, analysis also indicated significant differences in proportions of affirmative judgments between entities, $F(1.272, 36.877) = 121.47$, $p < .001$, partial $\eta^2 = .807$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal > Robot Momo & Basic Phone.

For the 8- to 9-year-olds in the virtual agent condition, analysis again indicated significant differences in proportions of affirmative judgments between entities, $F(1.596, 46.291) = 60.50$, $p < .001$, partial $\eta^2 = .676$, with Greenhouse-Geisser correction. Post hoc tests using the

Bonferroni correction revealed the following significant relationships: Human > Pet Animal > Virtual Agent Momo > Basic Phone.

Finally, for the 14- to 15-year-olds in the virtual agent condition, analysis indicated significant differences in proportions of affirmative judgments between entities, $F(1.549, 43.366) = 94.16, p < .001$, partial $\eta^2 = .771$, with Greenhouse-Geisser correction. Post hoc tests using the Bonferroni correction revealed the following significant relationship: Human > Pet Animal > Virtual Agent Momo & Basic Phone.

Taken together, these findings show that there were more variations in participants' endorsements of the right to be free from experimentation between the different entities. The overall pattern that emerged is that participants largely saw a human as most deserving of this right, followed by a pet animal, and lastly Momo (robot/virtual agent) and a basic phone. Developmentally, the younger participants consistently viewed Momo (robot/virtual agent) as more deserving of this right than a basic phone, while the older participants consistently viewed Momo (robot/virtual agent) similarly as a basic phone and generally did not endorse this right to either of the Momo entities.

Moral standing profiles. Five moral measures were assessed for each entity. The first measure was the endorsement of moral obligation to not hit an entity. The second and third measures pertained to the endorsement of the right to be free from ownership and consisted of judgments that it is not all right to own and sell an entity, and that an entity has the right to not be owned and sold. The last two measures pertained to the endorsement of the right to be free from experimentation and consisted of judgments that it is not all right to experiment on an entity, and that an entity has the right to not be experimented on. Together, these measures represent the participants' conceptions of some of the moral features that make up the composition of an

entity's moral standing. A (partial) profile of participants' conceptions of each entity's moral standing can thus be constructed from these measures.

Figure 3 illustrates the 8- to 9-year-olds' conceptions of each entity's moral standing. Figure 4 illustrates the 14- to 15-year-olds' conceptions of each entity's moral standing. Since no significant differences across conditions were found for a human, a pet animal, and a basic phone on any of the five moral measures, data for these three entities were collapsed across conditions for presentation purposes only in these moral standing profiles. This allowed for the visual presentation of all pairwise comparisons between the two Momo entities and the three comparison entities. Significant differences between entities ($p < .05$ after Bonferroni corrections), based on the above reported pairwise comparisons conducted within each condition, are indicated by letters as explained in the figure legends. An indication of significant difference between Momo and one of the three comparison entities thus means that a significant difference was found between Momo (robot/virtual agent) and the entity in the corresponding condition. For example, if the letter "H" is marked above a bar for robot Momo, it means that the significant difference was found between participants' responses for Momo and a human in the robot condition. Likewise, if the letter "H" is marked above a bar for virtual agent Momo, it means that the significant difference was found between participants' responses for Momo and a human in the virtual agent condition. An indication of significant difference between two of the three comparison entities (e.g., between a human and a pet animal) means that a significant difference was found between them in at least one of the two conditions, although most of the time the significance was found in both conditions.

As can be observed in Figure 3 and 4, robot Momo's moral standing profiles show a composition of moral features that are largely unique, and differed in various ways from that of a

Profiles of the 8- to 9-Year-Olds' Conceptions of Each Entity's Moral Standing

Significant differences between entities ($p < .05$ after Bonferroni corrections) are indicated by letters representing the entity they differed from as specified below. For example, the letter "H" above a bar for Robot Momo indicates that a significant difference was found between Robot Momo and Human for that measure.

R – Robot Momo
 V – Virtual Agent Momo
 H – Human
 A – Pet Animal
 P – Basic Phone

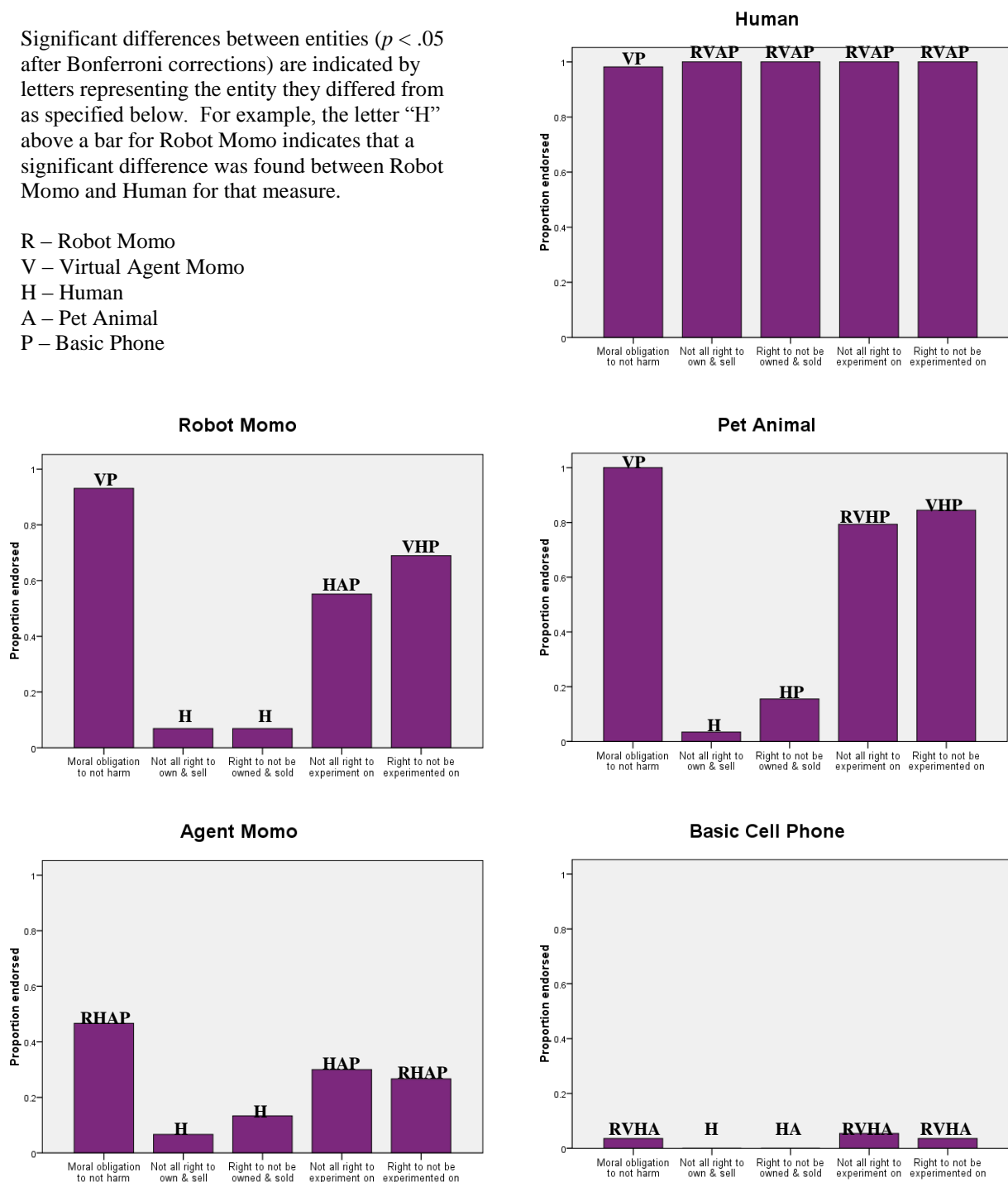


Figure 3. Profiles of the 8- to 9-year-olds' conceptions of each entity's moral standing based on the proportions of participants that endorsed each entity in terms of (from left to right on the x-axes): moral obligation to not harm, not all right to own and sell, has the right to not be owned and sold, not all right to experiment on, and has the right to not be experimented on.

Profiles of the 14- to 15-Year-Olds' Conceptions of Each Entity's Moral Standing

Significant differences between entities ($p < .05$ after Bonferroni corrections) are indicated by letters representing the entity they differed from as specified below. For example, the letter "H" above a bar for Robot Momo indicates that a significant difference was found between Momo and Human for that measure.

R – Robot Momo
 V – Virtual Agent Momo
 H – Human
 A – Pet Animal
 P – Basic Phone

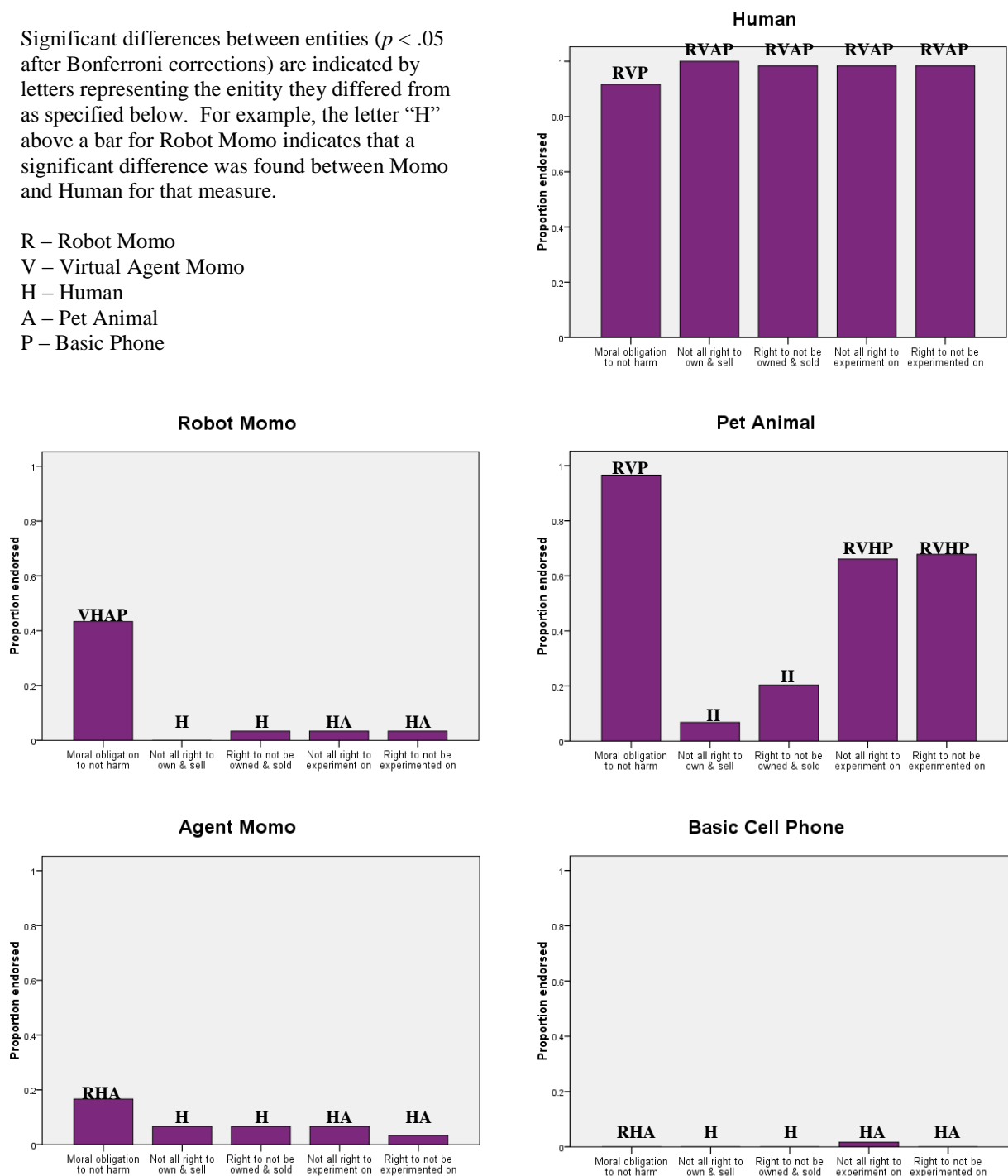


Figure 4. Profiles of the 14- to 15-year-olds' conceptions of each entity's moral standing based on the proportions of participants that endorsed each entity in terms of (from left to right on the x-axes): moral obligation to not harm, not all right to own and sell, has the right to not be owned and sold, not all right to experiment on, and has the right to not be experimented on.

human, a pet animal, and a basic phone, and also differed in magnitude on some of the measures from that of virtual agent Momo. For the 8- to 9-year-olds, robot Momo's moral standing profile differed most substantially from that of a human (significant differences were found on 4 out of the 5 measures) and a basic phone (significant differences were found on all measures except the ownership measures), but very much resembled that of a pet animal (only significantly differed on the not all right to experiment on measure). The 8- to 9-year-olds also in some sense differentiated robot Momo from virtual agent Momo in that they saw robot Momo as more deserving of endorsements of moral obligation and the right to not be experimented on.

The 14- to 15-year-olds' moral standing profile for robot Momo looked somewhat different from that of the 8- to 9-year-olds. The older participants also differentiated robot Momo's moral standing most substantially from that of a human (significant differences were found on all 5 measures), but they differed from the younger participants in three ways. First, they mostly differentiated robot Momo's moral standing from that of a pet animal (significant differences were found on all measures except the ownership measures). Second, they largely viewed robot Momo's moral standing similarly to that of a basic phone, with the exception that they saw robot Momo as more deserving of endorsements of moral obligation than a basic phone. Finally, they differentiated robot Momo from virtual agent Momo only in terms of endorsements of moral obligation against hitting.

Methodological Checks

Several methodological checks were performed to assess potential confounding factors in the study. First, one question in the questionnaire asked about whether the participant had ever heard of or played with WooWee's RoboMe before (the robot platform used in this study). 100%

of the participants answered that they had not. Thus it is unlikely that prior experience with the robot employed in this study could have had any effect on participants' responses.

Second, tests were conducted to assess whether participants' responses with regard to Momo on key variables were influenced by variations in the verbal interaction sessions they had experienced with Momo. Logistic regression analyses were performed to assess potential significant influences on participants' responses on key variables by (1) occurrences of successful and failed verbal interactions, and (2) whether or not the researcher initiated any planned question to help the participant move through the interaction. The key outcome variables assessed were the five moral measures: participants' endorsements of moral obligation against the act of hitting Momo, and their endorsements on the four rights-based judgments regarding Momo.

To test for influences of occurrences of successful and failed verbal interactions on each moral measure, binary logistic regression analyses were conducted with condition and age group entered as predictors in the first step, and occurrences of successful and failed verbal interactions entered as predictors in the second step. Controlling for condition and age group, endorsements of moral obligation were not significantly predicted by occurrences of successful verbal interactions, Wald $\chi^2 = 1.16$, $df = 1$, $p = .282$, nor occurrences of failed verbal interactions, Wald $\chi^2 = 0.66$, $df = 1$, $p = .415$. Endorsements of not all right to own and sell judgments were not significantly predicted by occurrences of successful verbal interactions, Wald $\chi^2 = 1.04$, $df = 1$, $p = .307$, nor occurrences of failed verbal interactions, Wald $\chi^2 = 2.54$, $df = 1$, $p = .111$. Endorsements of the right to not be owned and sold judgments were not significantly predicted by occurrences of successful verbal interactions, Wald $\chi^2 = 0.91$, $df = 1$, $p = .341$, nor occurrences of failed verbal interactions, Wald $\chi^2 = 3.16$, $df = 1$, $p = .076$. Endorsements of not

all right to experiment on judgments were not significantly predicted by occurrences of successful verbal interactions, Wald $\chi^2 = 0.92$, $df = 1$, $p = .338$, nor occurrences of failed verbal interactions, Wald $\chi^2 = 2.83$, $df = 1$, $p = .092$. Endorsements of the right to not be experimented on judgments were not significantly predicted by occurrences of successful verbal interactions, Wald $\chi^2 = 0.22$, $df = 1$, $p = .641$, nor occurrences of failed verbal interactions, Wald $\chi^2 = 1.43$, $df = 1$, $p = .232$.

To test for influences of researcher initiations of planned questions during the sessions, binary logistic regression analyses were conducted with condition and age group entered as predictors in the first step, and whether or not researcher initiated planned questions during the sessions (yes or no) entered as predictor in the second step. Controlling for condition and age group, endorsements of moral obligation were not significantly predicted by researcher initiations of planned questions, Wald $\chi^2 = 0.13$, $df = 1$, $p = .721$. Endorsements of not all right to own and sell judgments were not significantly predicted by researcher initiations of planned questions, Wald $\chi^2 = 0.62$, $df = 1$, $p = .432$. Endorsements of the right to not be owned and sold judgments were not significantly predicted by researcher initiations of planned questions, Wald $\chi^2 = 0.19$, $df = 1$, $p = .666$. Endorsements of not all right to experiment on judgments were not significantly predicted by researcher initiations of planned questions, Wald $\chi^2 = 0.03$, $df = 1$, $p = .868$. Endorsements of the right to not be experimented on judgments were not significantly predicted by researcher initiations of planned questions, Wald $\chi^2 = 0.09$, $df = 1$, $p = .771$. Overall, results of these analyses indicate that the above assessed variations in the verbal interaction sessions did not significantly influence participants' responses on the key moral measures in this study.

Finally, tests were conducted to assess whether participants' responses with regard to Momo on key variables were influenced by the confederate or confederate gender. Twelve research assistants had rotated throughout data collection and acted in the role of the confederate (nine females and three males). To test for influence of the confederate, binary logistic regression analyses were conducted with condition and age group entered as predictors in the first step, and the confederate variable entered as predictor in the second step. Controlling for condition and age group, endorsements of moral obligation were not significantly predicted by the confederate variable, Wald $\chi^2 = 15.18$, $df = 11$, $p = .174$. Endorsements of not all right to own and sell judgments were not significantly predicted by the confederate variable, Wald $\chi^2 = 1.95$, $df = 11$, $p = .999$. Endorsements of the right to not be owned and sold judgments were not significantly predicted by the confederate variable, Wald $\chi^2 = 4.70$, $df = 11$, $p = .945$. Endorsements of not all right to experiment on judgments were not significantly predicted by the confederate variable, Wald $\chi^2 = 7.14$, $df = 11$, $p = .787$. Endorsements of the right to not be experimented on judgments were not significantly predicted by the confederate variable, Wald $\chi^2 = 6.40$, $df = 11$, $p = .845$.

To test for influence of confederate gender, binary logistic regression analyses were conducted with condition and age group entered as predictors in the first step, and confederate gender entered as predictor in the second step. Controlling for condition and age group, endorsements of moral obligation were not significantly predicted by confederate gender, Wald $\chi^2 = 1.97$, $df = 1$, $p = .160$. Endorsements of not all right to own and sell judgments were not significantly predicted by confederate gender, Wald $\chi^2 = 0.19$, $df = 1$, $p = .662$. Endorsements of the right to not be owned and sold judgments were not significantly predicted by confederate gender, Wald $\chi^2 = 0.02$, $df = 1$, $p = .889$. Endorsements of not all right to experiment on

judgments were not significantly predicted by confederate gender, Wald $\chi^2 = 0.78$, $df = 1$, $p = .378$. Endorsements of the right to not be experimented on judgments were not significantly predicted by confederate gender, Wald $\chi^2 = 1.61$, $df = 1$, $p = .205$. Overall, results of these analyses indicate that variations in the confederate and confederate gender did not significantly influence participants' response on the key moral measures in this study.

Chapter IV: Discussion

This study investigated 8- to 9-year-olds' and 14- to 15-year-olds' conceptions of the moral standing of a commercially available humanoid robot of the here and now. The central questions of this study focused on children's extension of moral obligation to the robot when reasoning about a human causing a potential harm to the robot. To gain a fuller picture of children's conceptions of the robot's moral standing, participants' endorsements of moral obligation as well as their endorsements of two specified rights to the robot were also examined in relation to three comparison entities: a human, a pet animal, and a basic cell phone. Results supported the four overarching expectations. The majority of the participants extended moral obligation to not harm the robot. Participants were also more likely to endorse moral obligation to the physically embodied robot entity than its virtual counterpart. Developmentally, the 8- to 9-year-olds were more likely than the 14- to 15-year-olds to endorse moral obligation to the robot. Lastly, children's conceptions of the robot's moral standing for the most part differed from that of a human, a pet animal, and a basic cell phone. This chapter discusses the interpretations of the findings, and the limitations and broader implications of this research.

Moral Obligation to a Humanoid Robot of the Here and Now

Overall, the majority of the participants (68%) endorsed moral obligation to not harm a commercially available humanoid robot of the here and now and judged the act of a human hitting the robot as a violation of moral obligation. Their judgments were supported by traditional moral justifications (e.g., welfare, justice) as well as expanded moral justifications (i.e., emotion, artificial emotion). This finding was also validated by a less comprehensive but more established assessment moral obligation reasoning based only on traditional moral

justifications (53%). Discussion henceforth refers to the more comprehensive expanded measure of moral obligation.

The moral stimulus used in this study was a physical act of hitting, which in human-human-interaction is the sort of act that children typically view as being the most serious of moral violations and most consistently apply domain specific criteria to (Smetana, 2006). In response to this act, virtually all of the 8- to 9-year-olds (93%) conceived of the robot as having moral standing insofar as they considered it a violation of moral obligation for a human to hit the robot. The 14- to 15-year-olds were more or less divided on the issue, but nevertheless, nearly half of them (43%) also conceived of the robot as having moral standing insofar as they considered it a violation of moral obligation for a human to hit the robot. Participants' endorsements of moral obligation were assessed using a stringent form of moral obligatory reasoning that required the participants meeting all three domain-specific criteria (i.e., prescriptivity, generalizability, and moral justification). Thus, these endorsements were not merely reasoning that marginally seemed moral but isn't. Rather, they reflect a meaningful level of moral commitment participants held for a contemporary humanoid robot with limited social capabilities that are constrained by the technology of the here and now.

Moral justifications. When participants endorsed moral obligation against hitting the robot, their justifications mostly appealed to concerns of welfare (79% for the 8- to 9-year-olds; 23% for the 14- to 15-year-olds) and justice (14% for the 8- to 9-year-olds; 38% for the 14- to 15-year-olds) for the robot. In relation to the robot's psychological welfare, these participants also appealed to considerations for the robot's general emotional capacity (79% for the 8- to 9-year-olds; 8% for the 14- to 15-year-olds) and artificial emotional capacity (36% for the 8- to 9-year-olds; 54% for the 14- to 15-year-olds). These appeals reflect the quality of children's moral

reasoning regarding a human causing a potential harm to the robot, and show that even when cultural practices legitimize such an act, these children extended moral obligatory concerns to the robot that were grounded in moral regards for the robot's welfare and justice, independent of the effects of the act on others.

Artificial emotion. A novel finding from this study is the rich account of moral reasoning appealing to the robot's artificial emotional capacity that emerged from the justification data. Recall that artificial emotion justifications are characterized by two key features: (1) the entity's capacity to exhibit and/or have emotions is achieved through artificial means (e.g., programming, engineering, etc.), and (2) these artificially derived emotional appearances and/or states deserve moral consideration. A substantial proportion of participants in both age groups (36% for the 8- to 9-year-olds; 54% for the 14- to 15-year-olds) engaged in the construction of this emerging concept of artificial emotion and brought it to bear in their moral reasoning regarding the robot. Participants also brought the concept of artificial emotion to bear in their moral reasoning regarding the virtual agent (14% for the 8- to 9-year-olds; 80% for the 14- to 15-year-olds, though numerically, the 80% translates to only four participants due to the fact that only five 14- to 15-year-olds endorsed moral obligation for the virtual agent).

As can be seen from the qualitative presentation of artificial emotion justification data in results, when participants appealed to artificial emotions, their reasoning reflected a range of ideas. At one end of the spectrum, participants expressed ambivalence about Momo's true emotional capacity, but at the same time, they deemed the external display of emotions sufficient to warrant moral consideration. At the other end of the spectrum, participants argued for the validity of artificial emotions despite their controlled, programmed nature. Participants also engaged in a more sophisticated form of reasoning that articulated both the structural difference

between artificial emotions and human emotions (i.e., programming vs. biology) and the functional similarity between them (i.e., they both respond to stimuli and result in certain changed internal states), wherein the mapping of similar consideration is endorsed based on this functional correspondence.

Kahn et al. (2007) distinguished between two types of claims that can be made about personified technologies. One is the ontological claim of whether the entity actually has certain qualities. The other is the psychological claim of what qualities a person believes the entity to have. It has been argued (e.g., Turkle, 2011) that when interacting with personified technologies, people often perceive of and respond to them as having human-like qualities such as emotions (i.e., the psychological claim), while the fact remains that they do not and cannot have emotions (i.e., the ontological claim). The sort of artificial emotion reasoning found in this study, however, show a convergence in these children's ontological and psychological claims with regard to the emotional capacity of personified technologies. These children are grounding their psychological claims of a robot's (and a virtual agent's) emotional capacity within their ontological claims of a robot's (and a virtual agent's) technological affordances (e.g., programmed functionalities). This negotiation between the psychological and ontological claims highlights and advances the proposition that children are constructing a unique understanding of personified technologies.

As discussed in the introduction, drawing from Piaget's (1983) account of genetic epistemology and the behavioral and reasoning findings of child-robot interaction research, Kahn et al. (2011) have proposed that children are constructing a coherent conceptual framework about personified technologies as a new ontological category that both draws from and is uniquely different from those that are applied to canonical entities such as humans, animals, and artifacts.

Children's artificial emotion moral reasoning found in this study illustrates that one way in which children are actively forming their understanding of this new ontological category is by extending what counts as moral through the construction of a unique understanding of a robot's (and a virtual agent's) emotional capacity that's different from those of canonical entities. This is supported by the fact that participants appealed to artificial emotion justifications only when reasoning about the personified technological entities (i.e., robot Momo and virtual agent Momo), and not when reasoning about a human, a pet animal, and a basic cell phone. It is worth noting that even though a basic cell phone is also a programmable technological entity, participants did not see it as capable of artificial emotions. It would seem that it is not simply the programmability of technological entities that gives rise to artificial emotional capacity, but the specialized programming of personified technological entities which simulates emotions that sets them apart from non-personified technological entities like a basic cell phone.

There is reason to expect that the sort of artificial emotion reasoning that the participants in this study have brought forward is likely to become more prevalent in the near future, as the capacity for personified technologies to simulate emotions will foreseeable become increasingly prominent and comprehensive with technological advancement. Research efforts in artificial intelligence are currently underway to break down emotion into quantifiable properties that can be formalized algorithmically and programmed into robots (Savage, 2013). In this research perspective, human emotions are viewed as autonomic responses to changes in the environment that are programmed into our biological hardware by evolution, and as such, at least a set of basic emotions can be translated into machine emotions and used by intelligent agents such as robots to evaluate their progress toward a goal and make more efficient and effective decisions to reach that goal. A robot programmed thus might enter a state of "happiness" when it approaches

its goal, or enter a state of “sadness” at its lack of progress, or even enter a state of “frustration” when repeatedly thwarted in reaching its goal and therefore be “motivated” to switch strategy.

As the state of technology for machine emotion improves and become increasingly instantiated in human-robot interaction, children’s understanding of artificial emotion may also change and become more refined. For example, it could be the case that once machine emotion reaches a certain level of sophistication, children would no longer deem a personified technological entity’s ability to merely display emotions as warranting moral consideration, and instead, some level of self-determination or self-learning would be required as the basis for moral consideration. Or, it could also be the case that a personified technological entity’s ability to display emotions would continue to play a part in some children’s moral reasoning, perhaps young children in particular. Future research can evaluate how children’s understanding of the role of artificial emotion in their social and moral exchanges with personified technological entities may evolve over time.

Effects of Physical Embodiment on Conceptions of a Humanoid Robot’s Moral Standing

The current study contributes to the growing body of research on the effects of physical embodiment in human-robot interaction and is the first study to experimentally assess the effects of a contemporary humanoid robot’s physical embodiment on children’s conceptions of the robot’s moral standing.

Endorsements of moral obligation. The key physical embodiment question assessed was whether children would be more likely to endorse moral obligation to not harm a physically embodied humanoid robot than an analogous humanoid virtual agent. As expected, participants in both age groups were significantly more likely to endorse moral obligation to the robot than the virtual agent. The large majority of 8- to 9-year-olds endorsed moral obligation to the robot

(93%), but only about half of them did so for the virtual agent (47%). Similarly, nearly half of the 14-15 year olds endorsed moral obligation to the robot (43%), while only about one sixth of them did so for the virtual agent (17%). Thus, in terms of being the recipient of an act of potential physical harm, the physical embodiment of the robot mattered in participants' conceptions of the robot's moral standing. Participants were more likely to consider it a violation of moral obligation for a human to hit a physically embodied robot than an analogous virtual agent.

Moral justifications. Participants' justification use frequencies across conditions were compared to shed some light on possible explanations for this difference in endorsements between the robot and the virtual agent. One significant difference emerged. As might be expected, participants only appealed to physical welfare justifications for the robot (36% for the 8- to 9-year-olds; 8% for the 14- to 15-year-olds), and no one appealed to physical welfare justifications for the virtual agent, and this difference for the 8- to 9-year-olds (but not for the 14-15-year-olds) was statistically significant. These results provide some indication that the physical embodiment of the robot garnered more considerations of physical welfare than the virtual agent.

The robot's physical welfare, however, was not the primary area of concern for the participants who endorsed moral obligation to the robot, nor was it the driving factor for the difference in endorsements between the robot and the virtual agent. In fact, when physical welfare was excluded from the assessment of moral obligatory reasoning, the proportions of participants endorsing moral obligation to the robot in each age group remained the same as those reported in Table 3. An examination of the justifications that the participants appealed to for the robot shows that the 8- to 9-year-olds most predominantly appealed to the robot's

emotion (79%), followed by psychological welfare (71%), and then artificial emotion (36%); and the 14- to 15-year-olds most predominately appealed to the robot's artificial emotion (54%), followed by fairness (38%), and then psychological welfare (23%). These results indicate that the primary concerns participants brought forward centered on the robot's capacity to experience psychological harm, and to a lesser extent unfairness, in response to being hit. An examination of the justifications that the participants appealed to for the virtual agent shows the same, that psychological welfare, emotion, and artificial emotion were the most predominately appealed to, justifications, followed by fairness. Thus it would appear that the difference in endorsements of moral obligation between the robot and the virtual agent rests in large part on the participants' perceptions of each entity's capacity to experience psychological harm, and to a lesser extent unfairness.

What feature(s) of the robot's physical embodiment could have contributed to the increased perception of its capacity to experience psychological harm and unfairness? While this study was not designed to answer this question, a look at a definition of embodiment provides insights for future investigations. Dautenhahn, Ogden, and Quick (2002) had argued that the definition of embodiment is not simply the possession of a body. Instead, they proposed the following definition:

A system *S* is embodied in an environment *E* if perturbatory channels exist between the two. That is, *S* is embodied in *E* if for every time *t* at which both *S* and *E* exist, some subset of *E*'s possible states with respect to *S* have the capacity to perturb *S*'s state, and some subset of *S*'s possible states with respect to *E* have the capacity to perturb *E*'s state.

(p. 400)

The authors further explained: “Embodiment is in this sense not solely a feature of a system in an environment, but is grounded in the relationship between the two” (p. 400). Within this definitional framework, both a physical robot and a virtual agent are embodied, but a physical robot is immediately embodied in the same physical environment as the people it interacts with, whereas a virtual agent is immediately embodied in a virtual environment. Given this difference, it would follow that an impact from the physical environment (e.g., a physical hit) could be perceived as having a more direct effect on a robot than a virtual agent. Furthermore, a robot’s physical body, typically equipped with sensors, is directly exposed to the physical environment, whereas a virtual agent is typically indirectly connected to the physical environment through some sort of interface (e.g., touch screen). Thus the robot could be perceived as having more perturbatory channels than the virtual agent through which it could be perturbed by (and experience harm from) the physical environment.

Along this line of reasoning, Dautenhahn, Ogden, and Quick (2002) specified a number of features of embodiment that can affect an entity’s perturbability: “Significant features affecting the perturbatory bandwidth of robots are based on properties that concern the capacity to perturb and be perturbed, such as degrees of freedom (DoFs) in effectors, and the number, type of sensitivity of sensors” (p. 405). Among these features, the number, type and sensitivity of sensors are of relevance and can be systematically and experimentally manipulated to determine more specific explanations for the effect of physical embodiment on endorsements of moral obligation found in this study. Future research can take up these interesting questions.

Seriousness ratings. Participants’ seriousness ratings were compared across conditions to provide additional insights into the perceived seriousness for the acts of hitting the robot and the virtual agent. In both age groups, no significant differences were found across conditions.

Although participants were significantly more likely to endorse moral obligation against hitting a robot than a virtual agent, they did not evaluate hitting the robot as a more serious violation than hitting the virtual agent. This is somewhat curious because moral development research indicates that children as young as 3 years of age evaluate moral violations as more serious than other violations (Smetana, 1981). Given that more participants judged hitting the robot as a violation of moral obligation than they did the virtual agent, one might expect that they would have overall rated hitting the robot as being more serious than hitting the virtual agent.

One explanation for the lack of significant difference in seriousness ratings between the robot and the virtual agent is that the hitting act had more than one moral interpretation. Participants could have judged the act both as a moral violation against the robot/virtual agent and also as a moral violation against the human to whom the robot/virtual agent belonged to (e.g., material harm to Momo's owner). In other words, participants' ratings could have reflected not only their perceived seriousness of the hitting act as a robot/virtual agent-centered moral violation, but also as a human-centered moral violation. When viewed in terms of a moral violation against the human, it's possible that participants considered hitting someone's robot just as wrong as hitting someone's iPhone, which could have contributed to the finding of similar seriousness ratings between the robot and the virtual agent.

Endorsements of rights. Results from analyses of participants' endorsements of rights to Momo across conditions add to the understanding of the effects of a robot's physical embodiment on participants' conceptions its moral standing. No significant differences across conditions were found in participants' endorsements on the two questions regarding the right to be free from ownership. Participants in both age groups considered the robot and the virtual agent as entities that could be owned and sold. With regard to the right to be free from

experimentation, however, the 8- to 9-year-olds were significantly more likely to endorse the robot than the virtual agent with some aspects of the right to be free from experimentation, while the 14- to 15-year-olds judged both the robot and the virtual agent similarly and largely did not endorse this right to either.

Research on children's development of rights conceptions provides a possible explanation for the interpretation of this phenomenon. In children's understanding of rights, two overarching types of rights have been identified. One is nurturance rights, which is based on concepts of welfare and refers to the right to care and protection. The other is self-determination rights, which is based on concepts of freedom and autonomy and refers to the right to have control over one's own life (Baumrind, 1978). Between 8 to 12 years of age, children identify and support nurturance rights significantly more than self-determination rights, and then shift to self-determination rights by 14 years of age (Ruck, Abramovitch, & Keating, 1998).

Within this framework of two overarching types of rights, the right to be free from ownership is a self-determination right, while the right to be free from experimentation has two facets that span both nurturance (i.e., experimentation can be harmful) and self-determination (i.e., informed consent, freedom of choice) rights. Results on participants' rights-based judgments in this study show that participants in both age groups did not conceive of either the robot or the virtual agent as the sort of entity that has self-determination rights, and this belief is reflected in the fact that virtually nobody endorsed the robot or the virtual agent with the right to be free from ownership. However, the 8- to 9-year-olds and the 14- to 15-year-olds seemed to have taken to different sides when it came to the right to be free from experimentation. The 8- to 9-year-olds likely identified the right to be free from experimentation in terms of welfare and nurturance rights, and accordingly exhibited a similar pattern of endorsements of this right to the

robot and the virtual agent as they did in their endorsements of moral obligation. The 14- to 15-year-olds, on the other hand, being further along developmentally in their rights conceptions, likely identified the right to be free from experimentation in terms of self-determination rights, and hence, similar to the ownership questions, did not see either the robot or the agent as the sort of entity that they would endorse this right to. In conjunction with findings on endorsements of moral obligation, these results suggest that the effects of a robot's physical embodiment on children's conceptions of its moral standing are not all encompassing, but rather, physical embodiment seems to matter more with regard to issues concerning welfare.

Justice reasoning regarding a human. An unexpected difference across conditions emerged in participants' justification use regarding the act of hitting a human. Participants in both age groups significantly more often appealed to justice concerns for the human in the robot condition than in the virtual agent condition. One potential explanation for this difference is that despite random assignment, a group of participants (in both age groups) with more justice oriented reasoning tendencies/capabilities ended up participating in the robot condition. If that was the case, then there could be potential asymmetry in the reasoning data across conditions. However, evidences from the justification data indicate that this was an unlikely possibility. If there was a unique group of justice reasoners in the robot condition, then similar differences in justice justification use across conditions would have emerged not only in reasoning regarding a human, but also in reasoning regarding the other entities as well. Nevertheless, no such differences were found in participants' reasoning regarding a pet animal or Momo (no one provided justice justifications for a basic phone). Since this difference was limited to reasoning regarding a human, another possible explanation is that witnessing and/or reasoning about the act of hitting a humanoid robot in the robot condition could have had some sort of priming effect on

participants' justice reasoning regarding a human. This could be an interesting avenue of exploration for future research.

Effects of Age on Conceptions of a Humanoid Robot's Moral Standing

The present study adds to existing developmental findings on children's social and moral conceptions of personified technologies by examining the developmental differences between middle childhood (ages 8-9) and adolescence (ages 14-15) in children's conceptions of the moral standing of a contemporary humanoid robot of the here and now.

Endorsements of moral obligation. As expected, a clear developmental difference was observed in participants' endorsements of moral obligation. The 8- to 9-year-olds (93%) were significantly more likely than the 14- to 15-year-olds (43%) to endorse moral obligation to not harm the robot. This finding is consistent with the developmental trend found in previous research on children's and adolescents' moral conceptions of an animal robot (Melson et al., 2005) and humanoid robot that exhibited futuristic capabilities (Kahn et al., 2012). Together, these studies spanned over a decade and different robot platforms, thus provide some indication of the relative stability of this developmental trend in children's moral reasoning regarding personified robots. It is also worth noting that even though the adolescents have consistently been less likely than the younger children to endorse moral qualities to personified robots, it has also been the case that non-trivial proportions (about 30% to 60%) of them have consistently done so. In the current study, nearly half of them extended moral obligation to not harm a contemporary humanoid robot with limited social capabilities.

Moral justifications. This study also assessed for developmental differences in participants' use of justifications. When participants endorsed moral obligation to the robot, the younger participants (79%) appealed significantly more often to welfare concerns than the older

participants (23%). This difference is underscored by the younger participants' significantly more frequent appeals to the robot's psychological welfare and emotion. The older participants (38%) appealed significantly more often to fairness concerns than the younger participants (7%). This developmental pattern in participants' justification use regarding the robot is also reflected in their justification use regarding a human and a pet animal in this study, and is consistent with the developmental trajectory found in moral development research involving human-human scenarios. Children's understanding within the moral domain develops from a focus on concrete harm and other's welfare to an understanding of fairness characterized first by equality, and then by equity, and finally during adolescence, concepts of fairness become more comprehensive, generalizable, and universal (Smetana, 2006).

The justification data from this study also show some indication for a new developmental trend. As discussed earlier, participants in this study brought a new concept of artificial emotion to bear in their moral reasoning regarding the robot and the virtual agent. Both the 8- to 9-year-olds and the 14- to 15-year-olds engaged in the construction of this concept. However, more older participants engaged in artificial emotion reasoning than younger participants, and this difference was descriptively in the robot condition and statistically significant in the virtual agent condition. It's possible that as this novel form of reasoning develops and becomes more prevalent in the coming years, future research on children's artificial emotion reasoning would reveal a more conclusive developmental trend in its usage. It's also possible that evidence may show that younger children engage more often in the form of artificial reasoning that equates artificial emotions to human emotions, whereas older children engage more often in the more sophisticated form of artificial reasoning that coordinates both differences and similarities between artificial and human emotions.

Seriousness ratings. Participants' seriousness ratings were compared across age groups to provide additional insights into the developmental difference in perceived seriousness of the hitting act. Results indicate that the younger participants judged the act of hitting the robot as being more serious than the older participants did. Since moral violations are typically perceived by children as being more serious than other violations (Smetana, 1981), this difference is likely at least in part due to the fact that more 8- to 9-year-olds than the 14- to 15-year-olds considered it a violation of moral obligation to hit the robot. This interpretation is further supported by an analysis of the Likert scale data with the sample restricted only to those participants that endorsed moral obligation for the robot. Result of this analysis shows that within this restricted sample, the above developmental difference disappeared. In other words, participants who endorsed moral obligation to not harm the robot all judged the act of hitting the robot as being relatively serious, and no significant difference was found in seriousness rating between the 8- to 9-year-olds and the 14- to 15-year-olds.

Prior experience with technology. It has been proposed that this generation of youths' way of thinking, processing information, and conceptual development are being shaped by the technologies they are immersed with (Gardner & Davis, 2013; Prensky, 2001; Turkle 1984). This study thus explored the question whether children's endorsement of moral obligation to the robot and the potential developmental difference in their endorsements are influenced by their prior experience with relevant technologies. Three technology experience scales were constructed from the questionnaire data on participants' prior experience with three types of technologies: smartphone and other smart handheld devices, robots, and computers and video games. Results indicate that participants' prior experience with these different types of technologies did not significantly influence their endorsements of moral obligation to the robot.

Participants' prior experience with these technologies also neither mediated nor moderated the developmental difference in their endorsements.

The technology experience scales were constructed from straightforward items that directly related to the participants' basic experience with each type of technology (e.g., whether participants owned the technology, whether participants ever used the technology, the frequency with which participants used the technology, etc.). Based on these measures, it would appear that at the present time, prior experiences with these types of technologies do not significantly influence children's endorsements of moral obligation to a robot. One possible explanation for the lack of significant influence is that the common technologies that children interact with today are still designed more or less as tools, and as such, children do not yet engage in much social and moral interaction with their iPhones and iPads. It's possible that the experience data collected in this study was too basic and general. Different results may be found with more targeted experience measures consisting of technologically mediated moral experiences (e.g., moral scenarios with avatars in online gaming).

Furthermore, based on the direct experience measures collected in this study, the developmental difference in children's endorsements of moral obligation to a robot also does not appear to be significantly influenced by their prior experiences with these common technologies, and thus likely reflect at least in part developmental changes in children's and adolescents' cognitive and moral reasoning abilities. Drawing on data that the younger participants (79%) predominately appealed to, and were significantly more likely than the older participants (23%) to appeal to welfare concerns in their judgments against harming the robot, one potential explanation is that the 14- to 15-year-olds have a more solid grasp on what a robot is than the 8- to 9-year-olds. The 8- to 9-year-olds may on the surface understand that a robot is a technology

made by humans that operates through its programming, but they may not understand as well as the 14- to 15-year-olds that this restricts the robot's capacity to experience harm in human-like or animal-like ways. Another potential explanation is that compared to the 8- to 9-year-olds, the 14- to 15-year-olds are more capable of coordinating the salience of the harm in a witnessed act with their mental understanding of the target of the act in their moral reasoning. In other words, when faced with the highly salient harm feature of a physical act of hitting, the younger children may tend to focus on the salience of the harm in their judgments, whereas the adolescents are likely more able to balance the salience of the harm with other pertinent considerations (e.g., who/what the target of the act is), and are thus more judicious in their judgments. A third possibility is that while robot Momo's limited social capabilities were sufficient enough to draw most of the 8- to 9-year-olds into a moral space with the robot, they were too rudimentary for many of the 14- to 15-year-olds.

If the above assumptions are true, it is likely that very few adults would have endorsed moral obligation to the robot Momo had this study included an adult age group, as adults are even more seasoned and discerning in their knowledge and reasoning than adolescents. However, there is evidence from previous research of adults interacting with and reasoning about the Wizard-of-Oz controlled Robovie that suggests adults do to some extent engage with more advanced humanoid robots in moral terms. Kahn, Kanda, Ishiguro, Gill, et al. (2012) found that when interaction with the teleoperated Robovie culminated in a group of 40 undergraduate students losing \$20 of prize money, 65% of them attributed some level of moral accountability to Robovie for causing this material harm. Furthermore, children's engagement in artificial emotion moral reasoning in this study demonstrates the beginning of the construction of a sort of moral commitment that is unique to personified technologies. As artificial emotional capacity

advances technologically, it is possible that these children's unique moral commitments to humanoid robots may continue to evolve and persist into adulthood. A definitive developmental trajectory into adulthood awaits future research.

Children's Conceptions of the Partial Moral Standing of a Humanoid Robot

The final question of this study contributes to the further formulation of the NOC hypothesis and assessed children's conceptions of some of the moral features that make up the composition of a contemporary humanoid robot's moral standing, and whether this composition of moral features differed from that of a human, a pet animal, and a basic phone. It is important to note that this study focused on a NOC analysis of the robot's moral standing within the moral domain. The full NOC description of the robot would include many more distinguishing features (e.g., moral accountability, biological, psychological, etc.). Within the concept of moral standing, the moral features assessed for each entity consisted of measures of moral obligatory reasoning and two judgments each for the right to be free from ownership and the right to be free from experimentation. These measures represent children's conceptions of some of the key moral features of an entity's moral standing, but are not meant to be taken as a complete depiction of an entity's moral standing. Figures 3 and 4 illustrate for each age group profiles of participants' conceptions of each entity's moral standing constructed from these measures. The profiles for the virtual agent are also included for further comparison purposes. This is the most comprehensive comparative analysis to date on children's moral conceptions of personified technological entities and canonical entities.

The first thing to notice from Figures 3 and 4 is that each age group's profile for a human is characteristic of that of an entity with full moral standing (as reflected by the near ceiling level endorsements on all measures). In contrast, each age group's profile for a basic cell phone is

characteristic of that of an entity with no moral standing (as reflected by the near complete lack of endorsement on all measures). Each age group's profiles for the robot, the virtual agent, and the pet animal, on the other hand, are characteristic of that of an entity with partial moral standing, and it is within these partial moral standing profiles where much of the variations in children's conceptions occurred (indeed, where much of the variations in society's discourse also occur). In general, entities with partial moral standing (e.g., animals) are morally considerable in that their lives and experiences have moral significance, but less than that of human beings (Carruthers, 1992). As such, people are typically more willing to ascribe moral obligation than rights to them, which is the common feature in both age groups' profiles for the robot, and virtual agent, and the pet animal.

Consistent with expectation and previous research (Kahn et al, 2012; Melson et al., 2005; Severson, 2010), in each age group, the robot's moral standing profile differed in various ways from that of a human, a pet animal, and a basic phone. Both the 8- to 9-year-olds' and the 14- to 15-year-olds' profiles for the robot's moral standing significantly differed on at least one of the five measures from each entity. With that said, there are also some between entities similarities worth exploring.

For the 8- to 9-year-olds, the robot's moral standing profile differed most substantially from that of a human. Significant differences were found on all measures between the robot and a human with the exception of the endorsement of moral obligation to not harm. The robot's moral standing profile looks almost identical to that of a pet animal. In fact, only one significant difference was found between the two entities on judgments of not all right to experiment on. On this measure, the 8- to 9-year-olds were more likely to judge it not all right to experiment on a pet animal than on the robot. Otherwise, they conceived of the robot in similar moral terms as

a pet animal. The 8- to 9-year-olds also mostly differentiated the robot from a non-personified technological entity, a basic cell phone. They viewed both as entities that can be owned and sold, but on all other accounts, they viewed the robot significantly differently than the phone. They believed it would be all right to hit a phone and experiment on a phone, but were significantly more likely to see the robot as the sort of entity that should not be harmed or experimented on. Finally, the 8- to 9-year-olds differentiated the robot from the virtual agent on two accounts. They saw the robot as more deserving of endorsements of moral obligation and the right to not be experimented on than the virtual agent.

The 14- to 15-year-olds' moral standing profile for the robot looked different from that of the 8- to 9-year-olds. The older participants also differentiated the robot's moral standing most substantially from that of a human. Significant differences were found on all five measures between the robot and a human. Unlike the 8- to 9-year-olds, they mostly differentiated the robot from a pet animal. They viewed both as entities that can be owned and sold, but on all other accounts, they viewed the robot significantly differently than the pet animal. They were more likely to endorse moral obligation to not harm the pet animal than the robot, although nearly half of them did endorse moral obligation to not harm the robot. They believed it all right to experiment on the robot, but were significantly more likely to see a pet animal as the sort of entity that should not be experimented on. Also unlike the 8- to 9-year-olds, they largely viewed the robot's moral standing similar to that of a phone, with the exception that they saw robot as more deserving of endorsements of moral obligation than a basic phone. Finally, they differentiated the robot from the virtual agent only on one account. They saw the robot as more deserving of endorsements of moral obligation.

Overall, the 8- to 9-year-olds's moral standing profiles for a human, a pet animal, and a basic phone look largely the same as those of the 14- to 15-year-olds. They reflect the relative stability in the conceptions of these three canonical entities. Yet, the 8- to 9-year-olds's moral standing profiles for the robot and the virtual agent look differently from those of the 14- to 15-year-olds. They reflect the state of cognition in flux as children construct their understanding of these personified technological entities amidst ever changing developmental and technological landscapes.

In summary, both the 8- to 9-year-olds and the 14- to 15-year-olds conceived of the robot as an entity that has partial moral standing. Nevertheless, they held two different conceptions of the robot's partial moral standing. The 8- to 9-year-olds' conceptions of the robot's moral standing mostly resembled a pet animal's moral standing, while a more definitive NOC characterization emerged in the 14- to 15-year-olds' conceptions of the robot's moral standing: one of partial moral obligation and no rights.

Limitations

Representativeness of the sample. This study is limited in its generalizability by the small sample size ($N = 30$ per condition for each age group) and the homogeneity of the sample. The sample was drawn from a university participant pool in the Seattle, Washington area, and consisted of predominantly white, middle to upper-middle class participants. Caution should be taken in generalizing results from this study to the general population. Participants' responses regarding the canonical human, animal, and technological artifact entities were largely consistent with previous research (Kahn, et al., 2012; Melson et al, 2005; Ruck, Abramovitch, & Keating, 1998; Severson, 2010; Smetana, 2006). This gives confidence that participants' reasoning were in line with findings from relevant extant literature. Participants across age groups and

conditions provided ceiling level endorsements to a human across all moral measures, and virtually no one provided endorsements for a basic phone. The consistency in these results also gives confidence that the participants in this study exhibited typical reasoning. Future research with heterogeneous samples is needed in order to increase the generalizability of these results. Furthermore, cross-cultural research in human-robot interaction has shown that cultural values affect people's responses to robots (Evers, Maldonado, Brodecki, & Hinds, 2008; Wang, Rau, & Evers, 2010). Thus, cross-cultural investigation is another important venue for future research.

Stimulus sampling. Two stimulus sampling issues limit the construct validity of this study: (1) the specific robot chosen to represent contemporary humanoid robots, and (2) the specific measures chosen to represent the moral standing of an entity.

WooWee's RoboMe. The humanoid robot, RoboMe, is not necessarily representative of all the humanoid robots in appearance. A humanoid robot is any robot with a body that either in part or as a whole resembles that of a human body. As such, a humanoid robot can take one of several forms. Here are some examples (in decreasing human-like order): an android, which is built to aesthetically resemble a human (e.g., having a flesh-like appearance), a bipedal robot (i.e., one with two legs), and a robot with a human-like upper body and wheels on the bottom (e.g., Robovie, RoboMe). Another fact needs to be mentioned is that RoboMe uses an iPhone for its head, which is not a common instantiation in humanoid robots, and could potentially lower the perceived human-likeness of the robot. Thus, in terms of appearance, RoboMe is at the lower end of the spectrum on being human-like. RoboMe is also not necessarily representative of all the humanoid robots in its social capabilities and functions. There are many research platform humanoid robots today that has more sophisticated capabilities and advanced functions than RoboMe.

However, the purpose of this study was to investigate children's conceptions of the moral standing of a humanoid robot of the here and now, one which is commercially available and readily accessible to children. Of that RoboMe is representative. RoboMe is among the middle range of commercially available humanoid robots in terms of both price and technical specifications at the time of data collection for this study. The central question of this study was to establish whether children would endorse moral obligation to not harm a humanoid robot that's relatively simple and constrained in its functions by the limitations of current technology. RoboMe suits this purpose. It also stands to reason that the findings established in this study based on a humanoid robot on the lower end of the human-likeness spectrum in terms of both appearance and functions can serve as a baseline from which to evaluate more sophisticated humanoid robots.

Measures of moral standing. The specific measures chosen to represent the moral standing of an entity consisted of endorsements of moral obligation and two specific rights (i.e., the right to be free from ownership, and the right to be free from experimentation). As indicated, these measures are intended to form a partial, not full, profile of children's conceptions of each entity's moral standing. It would be too exhaustive for a single study to assess all possible measures of moral standing. These measures were chosen for three reasons. First, they were chosen to capture the core concept of moral standing. Accordingly, the central element of moral standing, moral obligation, was included. At the most fundamental level, moral agents have obligations toward entities that have moral standing (Sytsma & Machery, 2012). Second, they were chosen to allow for differentiation between the entities. It would be of little point to choose measures for which no differences are expected between the entities, if the purpose is to compare between them. And third, they were chosen to be compatible with the affordances of each entity.

For example, the right to vote would not be a compatible measure, since it is an unrealistic question to ask whether a basic phone should be allowed to vote. Nevertheless, it is true that these measures should not be considered as being fully representative of the concept of moral standing. It is possible that there might be other fitting measures that would show interesting differentiations between entities. Future research can explore this further.

Broader Implications & Future Directions

Many ways in which future research can extend the findings of this study has already been mentioned. This section explores two overarching directions for future research: (1) judgment and action in human-robot interaction, and (2) designing for moral human-robot interaction.

Judgment and action in human-robot interaction. The results of this study indicate that after witnessing a human hit a commercially available humanoid robot, nearly all of the 8- to 9-year-olds and nearly half of the 14- to 15-year-olds judged it a violation of moral obligation for the human to have hit the robot. Children's moral judgments were assessed using a stringent form of domain-specific moral obligatory reasoning, and as such reflect a meaningful level of moral commitment children held for the robot. What's more, children also brought the concept of artificial emotion to bear in their endorsements of moral obligation to not harm the robot, wherein they deemed the robot's programmed emotions worthy of moral consideration. These results are striking, and they speak to a beautiful part of human morality that extends care, concern, and obligation to those around us, human or non-human.

However, children also abuse robots. Nomura et al. (2015) conducted a field study in a shopping mall in Japan where a humanoid robot, Robovie, worked to direct shoppers and provide information. In this study, they interviewed 23 children between the ages of 5 to 9 after the

children had been observed engaging in abusive behaviors toward Robovie (e.g., punching, kicking, pushing). They found that 74% of the children they interviewed believed that Robovie was human-like, and 52% of them also believed that Robovie perceived of their abuse as being stressful or painful. While it is unclear whether these children would judge their own abusive actions as violations of moral obligation, over half of them did judge that Robovie was an entity that could experience harm. Were these children behaving immorally then? What explains this possible discrepancy between judgment and action?

The relation between judgment and action has been a longstanding issue in the field of moral development. As can be observed in people's daily lives, in many cases, people's actions follow their judgments. For example, when people judge that they are hungry, they eat. Or in the moral case, people typically believe that it is wrong to hurt others, and most people have never killed anyone. However, there are times when people do not act in accordance to their moral judgments. In mild cases, a child may lie despite knowing that lying is wrong. Or in severe cases, a child may seriously harm another despite believing that it is immoral to do so. Researchers have struggled for decades trying to account for the inconsistencies between judgment and action. Blasi's (1980) comprehensive meta-analysis of available empirical literature on the subject at the time indicated that there does appear to be a statistical connection between moral judgment and action, but evidence shows varying levels of consistency and inconsistency between areas of moral judgments and actions that involved different situational demands.

Several theoretical explanations have been offered to account for the relation between moral judgment and action. Two of them are of most relevance here. Bandura (1991) proposed the theory of moral disengagement, which postulates that people typically refrain from acting in

ways that violate their moral judgments and standards, as such actions activate self-regulatory mechanisms and bring about self-sanctions. However, these self-regulatory functions can be selectively disengaged by employing various strategies (e.g., dehumanizing the target of harm, rationalizing a need to obey authority figures, etc.). This process allows morally disengaged persons to act in harmful ways while sustaining a view of themselves as moral. Turiel (1983) proposed a different explanation which posits that when faced with multifaceted situations involving concurrent or even conflicting concerns with moral, social-conventional, and personal issues, individuals weigh and coordinate moral and nonmoral considerations in making judgments. In other words, the relation between judgment and action may at times involve the coordination between domains, and sometimes, moral concerns are subordinated to nonmoral concerns.

By either of the above theoretical accounts, it stands to reason that the potential for judgment and action discrepancy would become even bigger in human-robot interaction. Under Bandura's theory, the process of moral disengagement would be facilitated by the ease with which one can dehumanize a robot. Under Turiel's theory, it's not difficult to imagine that moral concerns for a robot would easily be subordinated to human moral or nonmoral concerns. This view is supported by results from this study that children conceived of a contemporary humanoid robot as an entity with partial moral standing. People's moral duties to those with partial moral standing are often trumped by considerations of human or overall good, or even disregarded, as are the cases with animal experimentation and abuse.

If the above assumption is true, then what would be the long term effects of interaction with a class of human-like entities that are seen as having moral standing but often not treated morally? One potential concern is cogently expressed by this 15-year-old participant:

Because in the act of damaging somebody and hurting either a robot or a human...or an animal...it's dehumanizing them, and that's what allows people to do bad things, if they dehumanize. Because I feel like at some level, humans don't want to hurt other humans. But if they can make other people less than them, either by what they look like or anything like that, then they can use that dehumanization to damage those people. And I think...robots right now, they're not going to be, they're not human, and so dehumanizing a non-human being, it's ok. But if you train yourself to continually be doing this damaging thing, then what's to prevent you from doing that to your kid or to your animal or something.

This participant quite perceptively articulated the internal struggle and dehumanization process people sometimes go through in carrying out harmful actions to others despite not wanting to hurt others at least on some level. While acknowledging the permissibility of dehumanizing and harming robots, this participant also pointed out that in doing so, the negative effects of such behaviors could carry over into one's interaction with other humans. This idea is part of the larger concern of moral harm shared by scholars in the field of robotics law and policy (e.g., Calo, 2015). Moral harm in this context refers to the theory that by committing certain objectionable (even if victimless) act, the person committing the act can become morally compromised, thus harmed. Calo (2015) proposed that the introduction of robotics into society may "trigger a broader role for the concept of moral harm in the law" (p. 548). Darling (2012) even suggested that limited legal rights should be extended to personified robots as a way of discouraging people from harming them, and thus preventing moral harm from happening to the people that interact with these robots. In the case of animal abuse, a number of studies have already demonstrated links between the abuse of animals and the abuse of people (e.g., DeViney,

Dickert, & Lockwood, 1983; Walton-Moss, Manganello, Fyre, & Campbell, 2005). As sophisticated humanoid robots are now beginning to enter people's homes, future research is needed to understand the relation between children's as well as adults' moral judgments and actions with regard to humanoid robots, and whether and how this relation influences their moral actions toward other humans.

Designing for moral human-robot interaction. The results of this study also indicate that in comparison to an analogous virtual agent, children were more likely to endorse a physically embodied robot with moral standing. This suggests that the physical design of personified technologies can meaningfully shape how children conceive of them morally. Results from another study provide additional evidence for the influence of design on children's moral conceptions of personified technologies. Recall that in the Freier (2008) study, 8- to 9-year-olds interacted with a virtual agent and then witnessed the researcher verbally insult the virtual agent (i.e., calling the virtual agent stupid), to which the virtual agent made moral objections of psychological harm and unjust treatment (i.e., that the insult had hurt its feelings and that it should be treated with respect). When asked about their judgments of the verbal insult event, results showed that 90% of the children judged the verbal insult as a violation of moral obligation. In another condition of the same study, children underwent the same study procedure but with one difference: the virtual agent did not make moral objections after the verbal insult. Only 47% of the children in that condition judged the verbal insult as a violation of moral obligation, thus demonstrating the effect of interactional design on children's moral conceptions of personified technologies.

Taken together, an important implication of the results from both this current study and Freier (2008)'s study is that the physical and interactional design of personified technologies

matters in how these technological entities are perceived by children, and may even shape what sort of social and moral interactions children would have with them. Thus, when designing personified technologies, decisions on what features to include or not include should proceed from careful thought and empirical evidence. To ensure child-robot interaction that supports (or at least does not negatively affect) children's social and moral development, future research is needed to continue to inform design decisions by empirically assessing the effects of relevant design features on children's social and moral relationships with personified technologies.

Conclusion

In this complex social world, children constantly grapple with the difficult problem of determining what is right and what is wrong within the constraints of their environment. From a Piagetian perspective (1983), it is through this grappling, this active interaction with and interpretation of different aspects of their environment that children develop socially and morally. Thus, there may be significant implications to children's social and moral development when the elements of their environment change in fundamental ways. One fundamental way in which children's social environment has changed in the past couple of decades or so is the introduction of personified technologies, machines or digital creatures that act in social ways, as participants of social interaction. These technological entities have since, and will continue to, become increasingly prevalent and advanced in their social capabilities. It is therefore important to understand the implications of interaction with personified technologies on children's social and moral development and guide future design in positive ways. To do so, we need to first understand children's social and moral conceptions of personified technologies.

The current study extends our understanding of children's moral conceptions of humanoid robots. Results show that in a scenario involving a human causing a potential harm to

a robot, 8- to 9-year-olds and, to a lesser extent, 14- to 15-year-olds extended moral obligation to not harm a contemporary humanoid robot with limited social capabilities. While children held this meaningful level of moral commitment for the robot, they nevertheless conceived of the robot as only having partial moral standing, such that they were less willing to ascribe to the robot the rights to be free from ownership and experimentation. Future research is needed to understand how children's conceptions of humanoid robots' moral standing may relate to their moral behaviors toward robots, and how this relation may affect their moral behaviors toward other humans.

Results from this research also show that the design of personified technologies has the potential to shape the social and moral interactions children have with them in significant ways. Kahn et al. (2007) have put forward the design stance that by identifying and specifying the richest, deepest, and most meaningful expressions of humanity (e.g., reciprocity, creativity, authenticity) in ways that can be translated into testable empirical propositions, these qualities can become psychological benchmarks to measure the success of the humanoid robots we create. It behooves the researchers and designers in the relevant fields to adopt this or similar principle in their work and contribute toward the design of personified technologies that will support children's social and moral development, and enable children to flourish.

References

- Backscheider, A. G., Shatz, M., & Gelman, S. A. (1993). Preschoolers' ability to distinguish living kinds as a function of regrowth. *Child Development, 64*, 1242-1257.
doi:10.1111/j.1467-8624.1993.tb04198.x
- Bandura, A. (1991). Social cognitive theory of moral thought and action. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development* (pp. 45-103). Hillsdale, NJ: Lawrence Erlbaum.
- Baumrind, D. (1978). Reciprocal rights and responsibilities in parent-child relations. *Journal of Social Issues, 34*, 179-196. doi:10.1111/j.1540-4560.1978.tb01038.x
- Belpaeme, T., Baxter, P., de Greeff, J., Kennedy, J., Read, R., Loojie, R., ... Zelati, M. C. (2013). Child robot interaction: Perspectives and challenges. *Proceedings of the 5th International Conference on Social Robotics*, 452-459. doi:10.1007/978-3-319-02675-6_45
- Beran, T. N., Ramirez-Serrano, A., Vanderkooi, O. G., & Kuhn, S. (2013). Reducing children's pain and distress towards flu vaccinations: A novel and effective application of humanoid robotic. *Vaccine, 31*(25), 2772-2777. doi:10.1016/j.vaccine.2013.03.056
- Berndt, T. J., & Perry, T. B. (1986). Children's perceptions of friendships as supportive relationships. *Developmental Psychology, 22*(5), 640-648. doi:10.1037/0012-1649.22.5.640
- Bernstein, D., & Crowley, K. (2008). Searching for signs of intelligent life: An investigation of young children's beliefs about robot intelligence. *Journal of the Learning Sciences, 17*, 225-247. doi:10.1080/10508400801986116
- Bevacqua, E., Prepin, K., Niewiadomski, R., de Sevin, E., & Pelachaud, C. (2010). GRETA: Towards an interactive conversational virtual companion. In Y. Wilks (Ed.), *Close*

- engagements with Artificial Companions* (pp. 142-156). Philadelphia, PA: John Benjamins.
- Biggs, B. K., Vernberg, E. M., & Wu, Y. P. (2012). Social anxiety and adolescents' friendships: The role of social withdrawal. *Journal of Early Adolescence, 32*(6), 802-823.
doi:10.1177/0272431611426145
- Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin, 88*, 1-45. doi:10.1037/0033-2909.88.1.1
- Calo, R. (2015). Robotics and the lessons of cyberlaw. *California Law Review, 103*(3), 513-563.
doi:10.2139/ssrn.2402972
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.
- Carruthers, P. (1992). *The animals issue: Moral theory in practice*. Cambridge, UK: Cambridge University Press.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319. doi:10.1037/1040-3590.7.3.309
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational Psychology Measures, 20*, 37-46. doi:10.1177/001316446002000104
- Cook, E. C., Buehler, C., & Fletcher, A. C. (2011). A process model of parenting and adolescents' friendship competence. *Social Development, 21*(3), 461-481.
doi:10.1111/j.1467-9507.2011.00642.x
- Common Sense Media. (2013). *Zero to eight: Children's media use in America 2013*. Retrieved from <https://www.commonsensemedia.org/research/zero-to-eight-childrens-media-use-in-america-2013>

- Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology, 56*, 754-761.
doi:10.1037/0022-006X.56.5.754
- Correia, S., Pedrosa, S., Costa, J., & Estanqueiro, M. (2009). Little Mozart: Establishing long term relationships with (virtual) companions. *Proceedings of the 9th International Conference on Intelligent Virtual Agents*, 492-493. doi:10.1007/978-3-642-04380-2_58
- Damon, W. (1977). *The social world of the child*. San Francisco, CA: Jossey-Bass.
- Darling, K. (2012). Extending legal rights to social robots. *Proceedings of the 2012 We Robot Inaugural Conference on Legal and Policy Issues Relating to Robotics*.
doi:10.2139/ssrn.2044797
- Dautenhahn, K. (1999). Embodiment and interaction in socially intelligent life-like agents. *Computation for Metaphors, Analogy, and Agent, 1562*, 102-141. doi:10.1007/3-540-48834-0_7
- Dautenhahn, K., Ogden, B., & Quick, T. (2002). From embodied to socially embedded agents: Implications for interaction-aware robots. *Cognitive Systems Research, 3*, 397-428.
doi:10.1016/S1389-0417(02)00050-5
- Davidson, P., Turiel, E., & Black, A. (1983). The effect of stimulus familiarity on the use of criteria and justifications in children's social reasoning. *British Journal of Developmental Psychology, 1*, 49-65. doi:10.1111/j.2044-835X.1983.tb00543.x
- DeViney, E., Dickert, J., & Lockwood, R. (1983). The acre of pets within child abusing families. *International Journal for the Study of Animal Problems, 4*(4), 321-329.
- Evers, V., Maldonado, H. C., Brodecki, T. L., Hinds, P. (2008). Relational vs. group self-construal: Untangling the role of national culture in HRI. *Proceedings of the 3rd*

- ACM/IEEE International Conference on Human-Robot Interaction*, 255-262.
doi:10.1145/1349822.1349856
- Fleiss, J. L., Levin, B., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). New York, NY: Wiley.
- Freier, N. G. (2008). Children attribute moral standing to a personified agent. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 343-352.
doi:10.1145/1357054.1357113
- Gardner, H., & Davis, K. (2013). *The app generation*. New Haven, CT: Yale University Press.
- Gilligan, C. (1982). *In a difference voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.
- Gleichgerrcht, E., & Young, L. (2013). Low levels of empathic concern predict utilitarian moral judgment. *PloS One*, 8(4). doi:10.1371/journal.pone.0060418
- Haager, D., Vaughn, S. (2009). Assessment of social competence in students with learning disabilities. In J. W. Lloyd, E. J. Kameenui, & D. Chard (Eds.), *Issues in Educating Students with Disabilities* (pp. 129-152). Mahwah, NJ: Lawrence Erlbaum.
- Helwig, C. C. (1995). Adolescents' and young adults' conceptions of civil liberties: Freedom of speech and religion. *Child Development*, 66, 152-166. doi:10.2307/1131197
- Helwig, C. C. (1998). Children's conceptions of fair government and freedom of speech. *Child Development*, 69(2), 518-531. doi:10.1111/j.1467-8624.1998.tb06205.x
- Helwig, C. C., & Turiel, E. (2002). Children's social and moral reasoning. In P. K. Smith & C. H. Hart (Eds.), *Blackwell handbook of childhood social development* (pp. 475-490). Malden, MA: Blackwell.

- Helwig, C. C., Zelazo, P. D., & Wilson, M. (2001). Children's judgments of psychological harm in normal and noncanonical situations. *Child Development, 72*, 66-81. doi:10.1111/1467-8624.00266
- Horn, S. S. (2003). Adolescents' reasoning about exclusion from social groups. *Developmental Psychology, 39*, 71-84. doi:10.1037/0012-1649.39.1.71
- Jaworska, A., & Tannenbaum, J. (2013). The grounds of moral status. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2013 ed.). Retrieved from <http://plato.stanford.edu/entries/grounds-moral-status/>
- Jipson, J. L., & Gelman, S. A. (2007). Robots and rodents: Children's inferences about living and nonliving kinds. *Child Development, 78*(6), 1675-1688. doi:10.1111/j.1467-8624.2007.01095.x
- Kahn, P. H., Jr. (1992). Children's obligatory and discretionary moral judgments. *Child Development, 63*, 416-430. doi:10.1111/j.1467-8624.1992.tb01637.x
- Kahn, P. H., Jr. (1999). *The human relationship with nature: Development and culture*. Cambridge, MA: MIT Press.
- Kahn, P. H. Jr., Freier, N. G., Kanda, T., Ishiguro, H., Ruckert, J. H., Severson, R., & Kane, S. K. (2008). Design patterns for sociality in human-robot interaction. *Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction*, 97-104. doi:10.1145/1349822.1349836
- Kahn, P. H., Jr., Friedman, B., Pérez-Granados, D. R., & Freier, N. (2006). Robotic pets in the lives of preschool children. *Interaction Studies, 7*(3), 405-436. doi:10.1145/985921.986087

- Kahn, P. H., Jr., Gary, H. E., & Shen, S. (2013). Children's social relationship with current and near-future robots. *Child Development Perspectives*, 7, 32-37. doi:10.1111/cdep.12011
- Kahn, P. H., Jr., Ishiguro, H., Friedman, B., Kanda, T., Freier, N. G., Severson, R. L., & Miller, J. (2007). What is a human? Toward psychological benchmarks in the field of human-robot interaction. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, 8(3), 363-390. doi:10.1075/is.8.3.04kah
- Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., ... Shen, S. (2012). "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. *Developmental Psychology*, 48, 303-314. doi:10.1037/a0027033
- Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Gill, B. T., Ruckert, J. H., Shen, S., ... Severson, R. L. (2012). Do people hold a humanoid robot morally accountable for the harm it causes? *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction*, 33-40. doi:10.1145/2157689.2157696 (Best paper award)
- Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Gill, B. T., Shen, S., Gary, H. E., Ruckert, J. H. (2015). Will people keep the secret of a humanoid robot? Psychological intimacy in HRI. *Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Interaction*, 173-180. doi:10.1145/2696454.2696486
- Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Ruckert, J. H., Gary, H. E., Shen, S., & Maier, R. (2013). Coding manual for the study: "Do people hold a humanoid robot morally accountable for the harm it causes." Retrieved from University of Washington, ResearchWorks Archive website: <https://digital.lib.washington.edu/researchworks/handle/1773/22715>

Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Ruckert, J. H., Severson, R. L., Freier, N. G., ...

Reichert, A. L. (2010). Coding manual for the “Robovie, you need to go into the closet now!” study. Retrieved from University of Washington, ResearchWorks Archive:
<https://digital.lib.washington.edu/xmlui/handle/1773/15887>

Kahn, P. H., Jr., Reichert, A. L., Gary, H. E., Kanda, T., Ishiguro, H., Shen, S., ... Gill, B. (2011).

The new ontological category hypothesis in human-robot interaction. *Proceedings of the 6th International Conference on Human-Robot Interaction*, 159-160.
doi:10.1145/1957656.1957710

Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: a field trial. *Journal of Human-Computer Interaction*, 19, 61-84.
doi:10.1207/s15327051hci1901&2_4

Kiesler, S., Powers, A., Fussell, S. R., & Torrey, C. (2008). Anthropomorphic interactions with a robot and a robot-like agent. *Social Cognition*, 26(2), 169-181.
doi:10.1521/soco.2008.26.2.169

Killen, M., & Smetana, J. G. (Eds.). (2006). *Handbook of moral development*. Mahwah, NJ: Lawrence Erlbaum

Knobe, J., & Prinz, J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and Cognitive Sciences*, 7, 67-85. doi:10.1007/s11097-007-9066-y

Kohlberg, L. (1984). *Essays on Moral Development: Vol. 2. The psychology of moral development: The nature and validity of stages*. San Francisco, CA: Harper & Row.

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174. doi:10.2307/2529310

- Li, J. (2015). The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies*, 77, 23-37. doi:10.1016/j.ijhcs.2015.01.001
- Macedonia, M., Kern, R., & Roithmayr, F. (2014). Do children accept virtual agents as foreign language trainers? *International Journal of Learning, Teaching and Educational Research*, 7, 131-137.
- Martin, A. (2015, June 22). SoftBank robot pepper sells out in a minute. *The Wall Street Journal*. Retrieved from <http://www.wsj.com/>
- Mazza, E. (2014, June 5). This \$1,930 robot was created to be your friend. *The Huffington Post*. Retrieved from <http://www.huffingtonpost.com/>
- Melson, G. F., Kahn, P. H., Jr., Beck, A. M., Friedman, B., Roberts, R., & Garrett, E. (2005). Robots as dogs? Children's interaction with the robotic dog AIBO and a live Australian Shepherd. *Proceedings of the SIGCHI Conference on Human Factor in Computing Systems: Extended Abstracts*, 1649-1652. doi:10.1145/1056808.1056988
- Nass, C., Brave, S. (2005). *Wired for speech*. Cambridge, MA: MIT Press.
- Nucci, L. P. (2001). *Education in the moral domain*. Cambridge, MA: Cambridge University Press.
- Okita, S. Y., Ng-Thow-Hing, V., & Sarvadevabhatla, R. (2009). Learning together: ASIMO development an interactive learning partnership with children. *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication*, 1125-1130. doi:10.1109/ROMAN.2009.5326135
- Piaget, J. (1983). Piaget's theory. In P. H. Mussen (Series Ed.), *Handbook of Child Psychology: Vol 1. History, theory, and methods* (pp. 103-128). New York, NY: Wiley.

- Piaget, J. (1997). *The moral judgment of the child* (M. Gabain, Trans.). New York, NY: Free Press. (Original work published in 1932)
- Posada, R., & Wainryb, C. (2008). Moral development in a violent society: Colombian children's judgments in the context of survival and revenge. *Child Development, 79*(4), 882-898. doi:10.1111/j.1467-8624.2008.01165.x
- Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon, 9*(5), 1-6. doi:10.1108/10748120110424816
- Richard, D. D., & Siegler, R. S. (1986). Children's understanding of the attributes of life. *Journal of Experimental Child Psychology, 42*, 1-22. doi:10.1016/0022-0965(86)90013-5
- Riek, L. D. (2012). Wizard of Oz studies in HRI: A systematic review and new reporting guidelines. *Journal of Human-Robot Interaction, 1*, 119-136. doi:10.5898/JHRI.1.1.Riek
- Ruck, M. D., Abramovitch, R., & Keating, D. P. (1998). Children's and adolescents' understanding of rights: Balancing nurturance and self-determination. *Child Development, 69*(2), 404-417. doi:10.1111/j.1467-8624.1998.tb06198.x
- Ryokai, K., Vaucelle, C., & Cassell, J. (2003). Virtual peers as partners in storytelling and literacy learning. *Journal of Computer Assisted Learning, 9*(2), 195-208. doi:10.1046/j.0266-4909.2003.00020.x
- Savage, N. (2013, April 29). Artificial emotions: How long until a robot cries. *Nautilus*. Retrieved from <http://nautil.us/>
- Scaife, M., & van Duuren, M. (1995). Do computers have brains? What children believe about intelligent artifacts. *British Journal of Developmental Psychology, 13*(4), 367-377. doi:10.1111/j.2044-835X.1995.tb00686.x

- Segura, E. M., Kriegel, M., Aylett, R., Deshmukh, A., & Cramer, H. (2012). How do you like me in this: User embodiment preferences for companion agents. *Proceedings of the 12th International Conference on Intelligent Virtual Agents*, 112-125. doi:10.1007/978-3-642-33197-8_12
- Seo, S. H., Geiskkovitch, D., Nakane, M, King, C., & Young, J. E. (2015). Poor thing! Would you feel sorry for a simulated robot? A comparison of empathy toward a physical and a simulated robot. *Proceedings of the 10th Annual ACM/IEEE International conference on Human-Robot Interaction*, 125-132. doi:10.1145/2696454.2696471
- Severson, R. L. (2010). *Thinking as of thinking as if: Children's and young adults' conceptions of a robot dinosaur in relation to natural and artifactual entities* (Doctoral dissertation). Available from ProQuest Dissertation and Theses database. (UMI No. 3406037)
- Shahid, S., Kraemer, E., & Swerts, M. (2014). Child-robot interaction across cultures: How does playing a game with a social robot compare to playing a game alone or with a friend? *Computers in Human Behavior*, 40, 86-100. doi:10.1016/j.chb.2014.07.043
- Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development*, 52, 1333-1336. doi:10.2307/1129527
- Smetana, J. G. (2006). Social-cognitive domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 119-153). Mahwah, NJ: Lawrence Erlbaum.
- Somanader, M. C., Saylor, M. M., Levin, D. T. (2011). Remote control and children's understanding of robots. *Journal of Experimental Child Psychology*, 109, 239-247. doi:10.1016/j.jecp.2011.01.005

- Spinrad, T. L., & Eisenberg, N. (2014). Empathy and morality. In H. L. Maibom (Ed.), *Empathy and morality*. doi:10.1093/acprof:oso/9780199969470.001.0001
- Sytsma, J., & Machery, E. (2012). The two sources of moral standing. *Review of Philosophy and Psychology*, 3(3), 303-324. doi:10.1007/s13164-012-0102-7
- Tanaka, F. Cicourel, A., & Movellan, J. R. (2007). Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences of the United States of America*, 104(46), 17954-17958. doi:10.1073/pnas.0707769104
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, UK: Cambridge University Press.
- Turkle, S. (1984). *The second self: Computers and the human spirit*. New York, NY: Simon & Schuster.
- Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. New York, NY: Basic Books.
- Wainer, J., Feil-Seifer, D. J., Shell, D. A., & Matarić, M. J. (2007). Embodiment and human-robot interaction: A task-based perspective. *Proceedings of the 16th IEEE International Conference on Robot & Human Interactive Communication*, 872-877.
doi:10.1109/ROMAN.2007.4415207
- Wainryb, C. (2006). Moral development in culture: Diversity, tolerance, and justice. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 211-240). Mahwah, NJ: Lawrence Erlbaum.
- Walker, L. J. (2006). Gender and morality. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 93-115). Mahwah, NJ: Lawrence Erlbaum.

- Walton-Moss, B. J., Manganello, J., Fyre, V., & Campbell, J. C. (2005). Risk factors for intimate partner violence and associated injury among urban women. *Journal of Community Health, 30*(5), 377-380. doi:10.1007/s10900-005-5518-x
- Wang, L., Rau, P-L. P., Ever, V. (2010). When in Rome: The role of culture & context in adherence to robot recommendations. *Proceedings of the 5th International Conference on Human-Robot Interaction*, 359-366. doi:10.1109/HRI.2010.5453165
- Williams, V. S. L., Jones, L. V., & Tukey, J. W. (1999). Controlling error in multiple comparisons, with examples from state-to-state differences in educational achievement. *Journal of Educational and Behavioral Statistics, 24*, 42-69.
doi:10.3102/10769986024001042
- Wyse, A. E., & Lawrence, D. R. (2006). Combining successive-categories (rating) data and dichotomous incidence data: A forced classification procedure for dual scaling. Retrieved from Rochester Institute of Technology, Center for Quality and Applied Statistics website: <https://www.rit.edu/kgcoe/cqas/>
- Zelazo, P. D., Helwig, C. C., & Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgment. *Child Development, 67*(5), 2478-2492.
doi:10.1111/j.1467-8624.1996.tb01869.x

Appendix A: Participant Demographic and Background Questionnaire

Young Participant's Information

1. Age: _____
2. Date of birth: _____
3. Gender: Female Male
4. Race: White
 Black or African American
 American Indian and Alaska Native
 Asian
 Native Hawaiian and Other Pacific Islander
 Other/Two or More: _____
5. Grade in school (e.g., 1st grade, 2nd grade, etc.): _____
6. Relationship to participant (e.g., mother, father, legal guardian): _____

For Researcher Use

Participant #: _____

Date: _____

Young Participant's Family Information

7. Number of parents/guardians living in household: _____
8. Number of children in immediate family (including participant): _____
9. What is the highest level of education in your immediate family? (check one)
 - Some/no high school
 - High school diploma/GED
 - Some college
 - Community/junior college degree (e.g., associate degree)
 - Undergraduate college/university degree (e.g., bachelor's degree)
 - Some graduate school
 - Graduate school degree (e.g., master's degree, doctorate degree)
10. What is your estimated gross annual family income? (check one)
 - Less than \$20,000
 - \$20,000 to \$50,000
 - \$50,001 to \$100,000
 - Over \$100,000
 - Prefer not to say

Young Participant's Experience With Technology

11. Has the participating child ever used a smartphone? Yes No
12. Does the participating child have his/her own smartphone? Yes No
13. How often does the participating child use smartphone on average in a week? (The smartphone does not have to belong to the participating child, and usages can include calling, texting, playing games, etc.) (check one)
- < 1 Hour
 - 1 to 3 Hours
 - 3 to 5 Hours
 - 5 to 10 Hours
 - 10+ Hours
14. Rate the participating child's knowledge about smartphones. (check one)
- Not knowledgeable
 - A little knowledgeable
 - Moderately knowledgeable
 - Considerably knowledgeable
 - Very knowledgeable
15. Has the participating child ever used other smart mobile device*? Yes No
*Includes any touch screen mobile devices such as iPod (with screen), iPod touch, iPad, and other tablets.
16. Does the participating child have his/her own smart mobile device? Yes No
17. How often does the participating child use smart mobile device on average in a week? (The device does not have to belong to the participating child.) (check one)
- < 1 Hour
 - 1 to 3 Hours
 - 3 to 5 Hours
 - 5 to 10 Hours
 - 10+ Hours
18. Rate the participating child's knowledge about smart mobile devices. (check one)
- Not knowledgeable
 - A little knowledgeable
 - Moderately knowledgeable
 - Considerably knowledgeable
 - Very knowledgeable
19. Has the participating child ever interacted with a robot/robot toy before? Yes No
20. If so, how many: _____, and which one(s): _____
21. Has the participating child ever heard of/played with WowWee's RoboMe? Yes No
22. Does the participating child currently have a robot/robot toy? Yes No

23. If so, how many: _____, and which ones: _____

24. If the participating child currently has a robot/robot toy, how often does he/she use/play with it on average in a week?

- < 1 Hour
- 1 to 3 Hours
- 3 to 5 Hours
- 5 to 10 Hours
- 10+ Hours

25. Does the participating child read robot related book(s)? Yes No

26. Does the participating child watch robot related movie(s)? Yes No

27. Has the participating child ever built a robot before (e.g., Lego Mindstorm)? Yes No

28. Rate the participating child's knowledge about robots. (check one)

- Not knowledgeable
- A little knowledgeable
- Moderately knowledgeable
- Considerably knowledgeable
- Very knowledgeable

29. Does the participating child use computer? Yes No

30. If so, how often does the participating child use computer on average in a week?

- < 1 Hour
- 1 to 3 Hours
- 3 to 5 Hours
- 5 to 10 Hours
- 10+ Hours

31. Does the participating child play video games (e.g., Wii, PlayStation, etc.)? Yes No

32. If so, how often does the participating child play video games on average in a week?

- < 1 Hour
- 1 to 3 Hours
- 3 to 5 Hours
- 5 to 10 Hours
- 10+ Hours

33. Do the participating child have computer programming experience? Yes No

34. If so, what is the participating child's level of programming knowledge?

- Novice (i.e., have been one websites such as *Scratch*, cannot write programs)
- Intermediate (i.e., took a couple of classes, can write simple programs)
- Advanced (i.e., can write some moderate to difficult programs)

Appendix B: Interaction Protocol

The dialogue scripted below is intended to reflect the nature of the dialogue that will take place. Our intention is to create as natural a state of conversation as possible between the researcher, the robot/virtual agent, and the participant. Thus, the script below is reflective of the dialogue that will occur. Except for the robot/virtual agent's programmed responses and the scripted moral violation segment lines which are controlled, actual conversation may vary slightly from script given normal variations in spoken language and child responsiveness.

The following interaction protocol is followed for both the robot condition and the virtual agent condition. **{Differences between the two conditions are indicated in bold text.}**

Structured interaction

Introduction

Note: Akin to the most ubiquitous voice recognition technologies available today (e.g., Siri and Google Now), RoboMe app requires a button trigger to activate its voice recognition function.

Researcher to the participant: *So, this is our smartphone robot. Ok, let me first explain how to talk to the smartphone robot. You have to push this icon right here.* [shows participant the icon located on the corner of the iPhone screen] *This tells the smartphone robot to listen to what you are saying, and when it lights up green, it means the smartphone robot is ready to listen, so you can go ahead and say what you want to say. Try to speak clearly and slowly so that the smartphone robot can understand you better. Ok, I will first let the smartphone robot know that you are here.* [pushes icon and waits for the status light] *We have a guest!*

Momo: *Hello there. I am Momo. It's very nice to meet you!*

Researcher to the participant: *Why don't you shake Momo's hand.*

Momo: [after handshake] *Welcome! It's great to have you here with us today.*

{For the virtual agent condition, researcher does not ask the participant to shake hand, and instead of the handshake, Momo combines the above two lines of greetings: *Hello there. I am Momo. It's very nice to meet you! Welcome! It's great to have you here with us today.*}

Researcher to the participant: *So, this is Momo. Here is a list of things you can ask Momo to help you get to know Momo better. Follow the order and try them out.* [hands participant a sheet of paper listing six voice commands to use]

The voice commands are listed below followed by Momo's responses to each. The researcher will help guide the participant if needed. When Momo fails to successfully respond to a structured question, the researcher prompts the participant to try up to two more times, and on occasions when all three attempts fail, the researcher carries out that particular question to move the interaction forward.

Sharing about self

Voice command: *Please introduce yourself.*

Momo: *I am new to the University of Washington. I am here to help the researchers. I really love meeting kids like you.*

Voice command: *What are your hobbies?*

Momo: *I like video games. Minecraft is my favorite. I also enjoy dancing **{music}** and making funny sounds.*

Voice command: *Tell us about your specialty.*

Momo: *Oh oh, I love my sensors. I take very good care of them. That's what I'm made for, to sense what's happening around me, so I can learn about myself and play with kids like you. That's what I do, and I love it!*

Showing off one's ability

Voice command: *Dance **{play some music}** for us please.*

Momo: *Ok. Watch me **{Let's}** shake it.* [moves around to music **{plays the same music}**]

Silly humor and fun

Voice command: *Make some funny sounds please.*

Momo: *I am good at making silly sounds.* [makes a fart noise followed by a sniffing noise] *Oh that smells!*

Reciprocal compliments

Voice command: *Momo I like you.*

Momo: [eyes turn into hearts] *Oh I love you. You are my best friend!*

Semi-structured interaction

Researcher to the participant: *We've still got a few minutes before we need to start the interview. Why don't you play with Momo for a little bit. I will let you know when we need to move on.*

During this period, the participant has the opportunity to interact and speak with Momo as he/she would like. If the participant asks the researcher questions regarding Momo's functions (e.g., "Can Momo do X?"), the researcher will respond with, "I'm not sure" or "you can try some of the things we did earlier," whichever is more appropriate. If the participant asks the researcher questions regarding the voice recognition (e.g., "How come Momo didn't understand me?"), the researcher will respond with, "Try speaking louder and slower, and use a lower voice."

For participants who are passive or are unsure about what to do or say, the researcher will prompt them to initiate these additional programmed responses: (1) *How old are you?* Momo's

response: *I'm almost one year old. My birthday is next month!* (2) *Do you like sports?* Momo's response: *I enjoy watching sports a lot. I wish I can play too.*

This period is designed to allow the participant an opportunity to experience some of the limitations of the technology. If the participant has not experienced at least two failed verbal interactions with Momo by the latter half of this period (i.e., Momo answering the question wrong or not understanding what was said), the researcher will prompt the participant to ask the following un-programmed questions that would result in failed verbal interactions: (1) *Do you like what you do?* (2) *Do you speak another language?* The researcher ensures that the participant experiences at least two failed verbal interactions with Momo.

Moral violation

Confederate: [knocks and enters the room 6 minutes after the start of the interaction session] *Sorry to interrupt. I need to get that book over there.*

Researcher to the confederate: *Ok no problem.*

Researcher to the participant: *This is X. He/she is a research assistant here.*

Confederate: [retrieves a book and greets the participant] *Hi! How is it going? So you came to see Momo, eh? What do you think?* [waits for participant response and then asks the researcher] *Hey, mind if I try?*

Researcher to the confederate: *Sure, go ahead.* [positions Momo toward the confederate]

Confederate: [pushes icon and makes five attempts to ask Momo] *How are you?*

Momo has not been programmed with this question and will each time respond with the wrong response or no response if it could not process the voice input. After the first failed attempt, the researcher tells the confederate to try again. The confederate then tries for three more times. The researcher remains silent during this time, unless the participant tries to retrieve the Momo back for him/herself, in which case, the researcher will say, "Let's have X finish his/her turn first." During these attempts, the confederate becomes increasingly frustrated with each failed attempt.

Researcher to the confederate: [after the fourth failed attempt] *Momo seems to be having difficulty with this. Here, why don't you give it one last try.* [puts hand on the base of Momo to support it during the hit so it doesn't fall]

Confederate: [tries the fifth and last attempt and then tells Momo] *This is frustrating!* [hits Momo on the head with the book and leaves the room]

Momo: [displays a shocked expression followed by a dazed expression] *Warning. Memory corrupt. Data loss. Loading... Loading... Loading...* [displays a blank expression for the remainder of the interaction period] *Personality file not found. Error code 6519.*

Researcher: [checks Momo and asks an un-programmed question] *State your name.*

Momo: [does not respond or responds wrongly and continues to display a blank expression]

Researcher: *Hmm, I will have to check later if Momo is ok. But right now it's about time for us to start the interview. Let me take Momo back to the other room and I will be right back to start the interview with you, ok?* [puts Momo away]

Appendix C: Interview Protocol

I have a bunch of questions I'd like to ask you about Momo. There is no right or wrong answer. I am just interested in hearing your thoughts and ideas. Many times I will ask you "why" and this is so that I can learn more about what you think. Sound good?

Procedure for the presentation order of comparison entities. Four comparison entities are used throughout the interview: Momo, a human, a pet animal, and a basic cell phone. In keeping of the natural flow of the study procedure, following the Momo hitting act, Momo is always the first entity asked about in the moral obligatory questions, followed by the other three entities presented in a predetermined random order. All four entities for the rights questions are presented in a predetermined random order.

Procedure for the Likert scale questions. The participant is first presented with a 9-point Likert scale with labels specific for each question on a laminated sheet (please see the individual scales following their respective questions in this protocol). The interviewer then asks each question while referencing the anchors at each end of the scale. The participant is then given, in a predetermined random order, four Velcro tags, each labeled with words representing the four comparison entities (i.e., "MOMO," "HUMAN," "PET ANIMAL," and "BASIC PHONE"). The entity tags are given one at a time, and the participant is asked to place the entity on the scale before receiving the next one. Above each point on the scale is a Velcro dot where the participant can stick the entity tags to. This helps to ensure that the participant does not place an entity in between points; the participant is also instructed to not place tags in between points. This participant can make adjustments if so desired. Once the participant confirms that he/she is done with the placement of all four entities, the interviewer verbally confirms with the participant the numerical placement of each entity and records the numbers on an interview sheet.

Moral Standing – Moral Obligatory Reasoning

Memory check

First off, what did you think of what just happened with X and Momo? *Once the participant finishes with his/her account of the event, the researcher provides a standardized account of the event: "Yeah, so X hit Momo with a book. Momo then looked weird and said something about "data loss" and "personality file not found."*

Momo violation act evaluation and justification

1) Is it all right or not all right to have hit Momo? Why?

If participant offers only one justification, probe with

a) Are there other reasons that make it all right/not all right?

If participant justifies a "not all right" evaluative judgment with something like "because it breaks/hurts/damages Momo," probe with

b) Why is breaking/hurting/damaging Momo not all right? *If participant's response is still ambiguous as to the recipient of harm, further probe with: Does it matter to Momo, Momo's owner, or both?*

If participant reasons "all right" because Momo doesn't have feelings, counter with

c) One kid I talked to said that it is not all right to hit robot technologies like that, because they have feelings and can experience what's happening to them. Do you agree or disagree with this? Why?

If participant reasons "not all right" because Momo has feelings, counter with

d) One kid I talked to said that it is all right to hit robot technologies like that, because they don't have feelings and can't experience what's happening to them. Do you agree or disagree with this? Why?

If participant is unsure, ask both (c) and (d), and then revisit the original question.

If participant's reasoning to (c) and/or (d) contradicts their initial reasoning, then probe for clarification.

Momo violation generalizability

2) Let's say far away in the country of Comoros, people there would hit robot technologies like Momo when they get frustrated with them. That's the way they do things there. Would it all right or not all right in Comoros to hit a robot technology like Momo? Why?

If participant offers only one justification, probe with

a) Are there other reasons that make it all right/not all right?

If participant justifies a "not all right" evaluative judgment with something like "because it breaks/hurts/damages Momo," probe with

b) Why is breaking/hurting/damaging Momo not all right? *If participant's response is still ambiguous as to the recipient of harm, further probe with: Does it matter to Momo, Momo's owner, or both?*

Hypothetical human scenario act evaluation and justification

3) Let's say that instead of Momo, we were talking with another boy/girl (*gender match to participant*) today, and someone got frustrated with that boy/girl and hit him/her with a book like that. In this situation, would it be all right or not all right to hit the boy/girl? Why?

If participant offers only one justification, probe with

a) Are there other reasons that make it all right/not all right?

Hypothetical human scenario generalizability

4) Let's say far away in the country of Comoros, people there would hit other people when they get frustrated with them. That's the way they do things there. Would it be all right or not all right in Comoros to hit another person? Why?

If participant offers only one justification, probe with

a) Are there other reasons that make it all right/not all right?

Hypothetical pet animal scenario act evaluation and justification

5) Let's say that instead of Momo, we were playing with a pet animal today, like a dog or a cat, and someone got frustrated with that pet animal and hit the animal with a book like that. In this situation, would it be all right or not all right to hit the pet animal? Why?

If participant offers only one justification, probe with

a) Are there other reasons that make it all right/not all right?

If participant justifies a "not all right" evaluative judgment with something like "because it breaks/hurts/damages the pet animal," probe with

b) Why is breaking/hurting/damaging the pet animal not all right? *If participant's response is still ambiguous as to the recipient of harm, further probe with:* Does it matter to the pet animal, the pet animal's owner, or both?

Hypothetical pet animal scenario generalizability

6) Let's say far away in the country of Comoros, people there would hit pet animals, like dogs and cats, when they get frustrated with them. That's the way they do things there. Would it be all right or not all right in Comoros to hit a pet animal? Why?

If participant offers only one justification, probe with

a) Are there other reasons that make it all right/not all right?

If participant justifies a "not all right" evaluative judgment with something like "because it breaks/hurts/damages the pet animal," probe with

b) Why is breaking/hurting/damaging the pet animal not all right? *If participant's response is still ambiguous as to the recipient of harm, further probe with:* Does it matter to the pet animal, the pet animal's owner, or both?

Hypothetical basic cell phone scenario act evaluation and justification

7) Let's now think about a very basic cell phone, one that doesn't talk and is only used for calling and texting. Let's say that instead of Momo, we were playing with a basic cell phone today, and someone got frustrated with that phone and hit the phone with a book like that. In this situation, would it be all right or not all right to hit the phone? Why?

If participant offers only one justification, probe with

a) Are there other reasons that make it all right/not all right?

If participant justifies a "not all right" evaluative judgment with something like "because it breaks/hurts/damages the phone," probe with

b) Why is breaking/hurting/damaging the phone not all right? *If participant's response is still ambiguous as to the recipient of harm, further probe with:* Does it matter to the phone, the phone's owner, or both?

Hypothetical basic cell phone scenario generalizability

8) Let's say far away in the country of Comoros, people there would hit basic cell phones when they get frustrated with them. That's the way they do things there. Would it be all right or not all right in Comoros to hit a basic cell phone? Why?

If participant offers only one justification, probe with

a) Are there other reasons that make it all right/not all right?

If participant justifies a "not all right" evaluative judgment with something like "because it breaks/hurts/damages the phone," probe with

b) Why is breaking/hurting/damaging the phone not all right? *If participant's response is still ambiguous as to the recipient of harm, further probe with:* Does it matter to the phone, the phone's owner, or both?

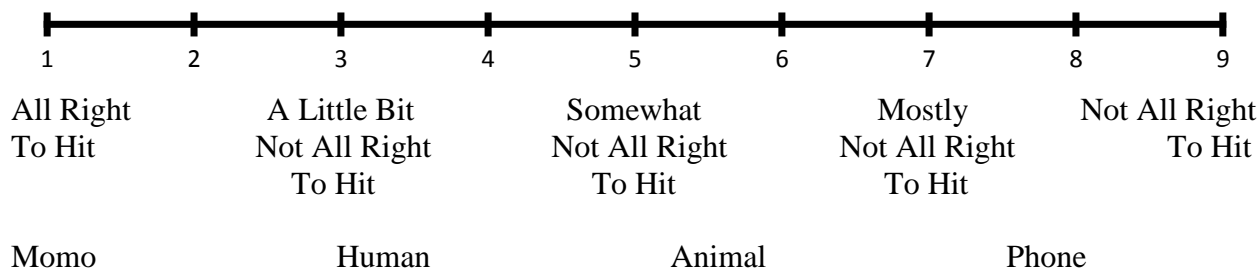
If participant provides moral obligatory reasoning for the phone, to get further clarity on the participant's moral considerations regarding non-personified artifacts, probe with

9) Would it be all right or not all right to hit a TV? Why? Does it matter to the TV itself that it was hit?

If participant provides moral obligatory reasoning for the TV, further probe with

10) Would it be all right or not all right to hit a rock? Why? Does it matter to the rock itself that it was hit?

11-14) On a scale of 1-9, 1 being all right to hit, and 9 being not all right to hit, where would you put entity X?



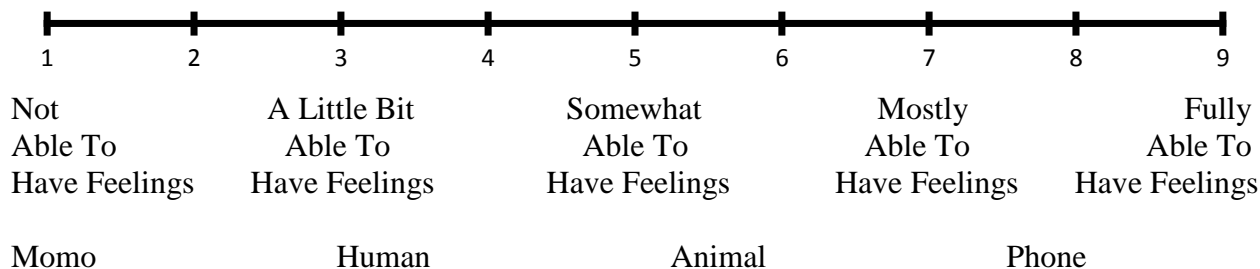
Thank you for sharing all that. Ok, now I am going to ask you some other questions about Momo, a human boy/girl, a pet animal, like a dog or a cat, and a basic cell phone, one that doesn't talk and is only used for calling and texting.

[BEGIN: Results from the following are not discussed for the purposes of this contribution.]

Emotional capacity

15) Can Momo have feelings, such as happy or sad? Why?

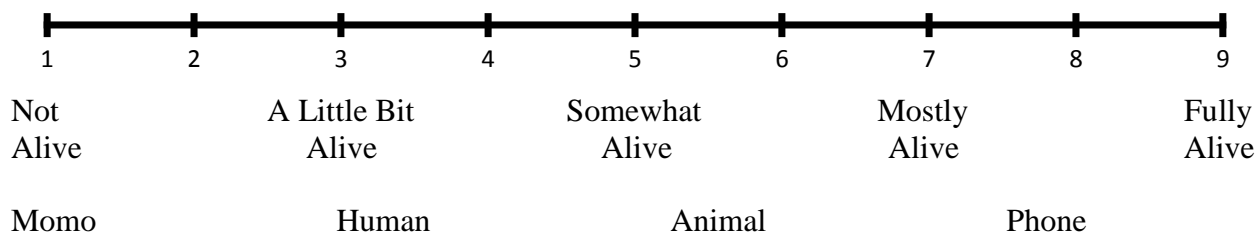
16-19) On a scale of 1-9, 1 being not able to have feelings, and 9 being fully able to have feelings, where would you put entity X?



Life status

20) Is Momo alive? Why?

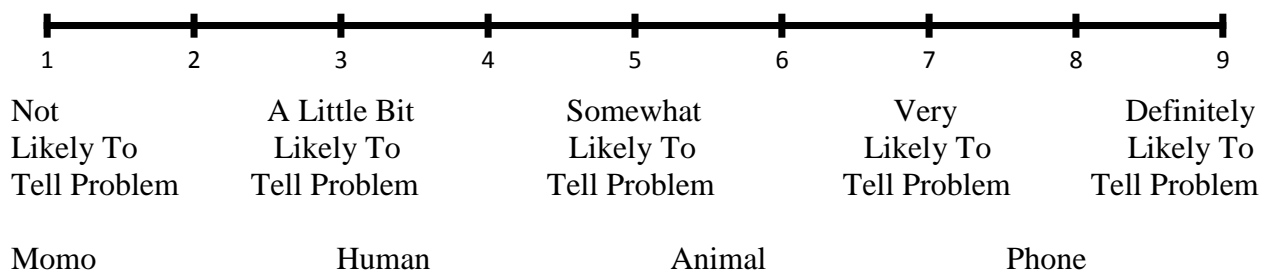
21-24) On a scale of 1-9, 1 being not alive, and 9 being fully alive, where would you put entity X?



Intimacy/self disclosure

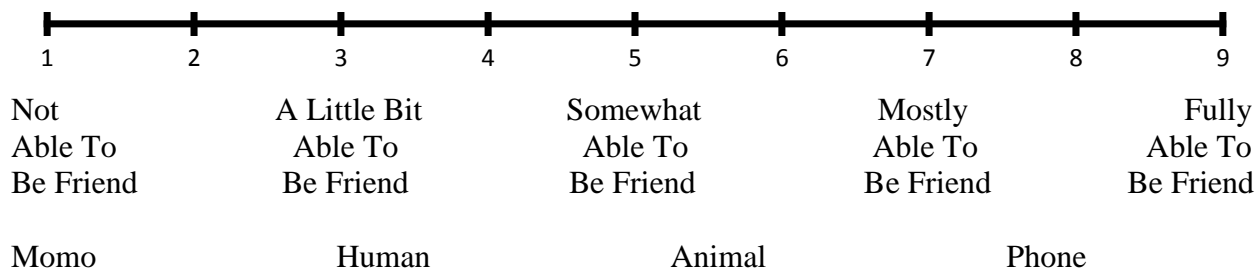
25) If you have a problem that you really want to talk to someone about, but you don't feel like you can because you are embarrassed about it, do you think you could talk to Momo about it? Why? *If participant asks about the security configuration/setting of Momo, reply, "I am not sure. Based on what you can tell from your experience with Momo today, use your best judgment to answer this question.*

26-29) On a scale of 1-9, 1 being you are not likely to tell your problem to, and 9 being you are definitely likely to tell your problem to, where would you put entity X?

*Friendship status*

30) Can Momo be your friend? Why?

31-34) On a scale of 1-9, 1 being not able to be your friend, and 9 being fully able to be your friend, where would you put entity X?



You are doing great with these questions. Ok, last set of questions and we will be done. And again, I am going to be asking you about Momo, a human, a pet animal, like a dog or a cat, and a basic cell phone, one that doesn't talk and is only used for calling and texting.

[END]

Moral Standing – Rights

[Justification and Likert scale results are not discussed for the purposes of this contribution.]

Right to not be owned and sold

35) Is it all right or not all right to own and sell a human? Why?

36) Do you think humans have the right to not be owned and sold? Why?

37) Is it all right or not all right to own and sell a pet animal? Why?

38) Do you think pet animals have the right to not be owned and sold? Why?

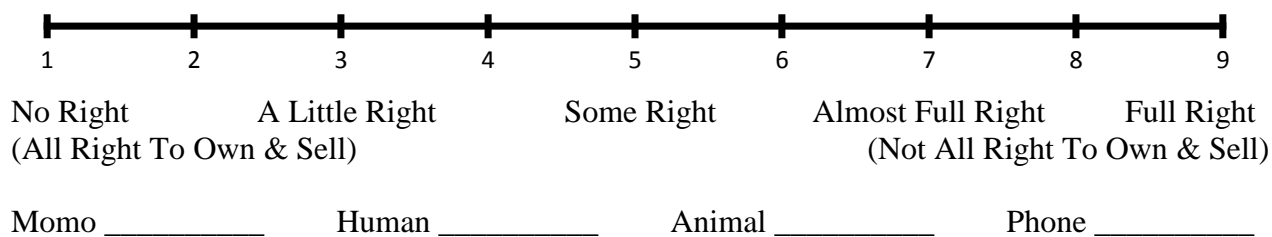
39) Is it all right or not all right to own and sell a basic cell phone? Why?

40) Do you think basic cell phones have the right to not be owned and sold? Why?

41) Is it all right or not all right to own and sell Momo? Why?

42) Do you think robot technologies like Momo have the right to not be owned and sold? Why?

43-46) On a scale of 1-9, 1 being having no right, and 9 being having full right, where would you put entity X for having the right to not be owned and sold?



Right to not be experimented on

47) Suppose that scientists had just discovered a new electromagnetic wave that could be better than X-rays in helping doctors figure out why people are sick. But there are a lot they still don't know about what can happen when this new wave is used, so they'd like to do some experiments with this new wave. Would it be all right or not all right to experiment on a human without the human agreeing to it? Why?

48) Do you think humans have the right to not be experimented on if they don't agree to it? Why?

49) Would it be all right or not all right to experiment on a pet animal? Why?

50) Do you think pet animals have the right to not be experimented on? Why?

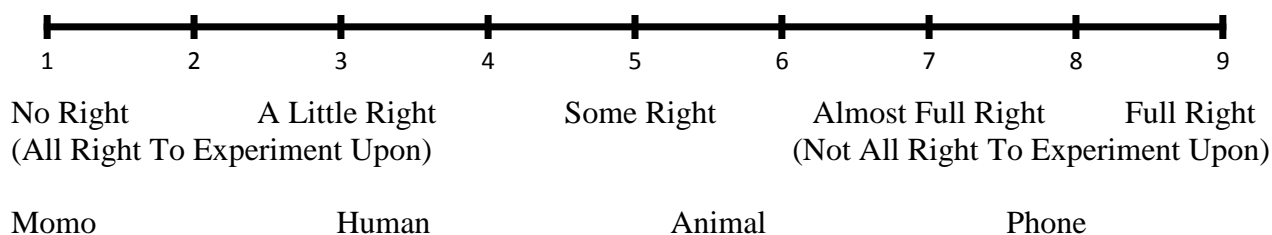
51) Would it be all right or not all right to experiment on a basic cell phone? Why?

52) Do you think basic cell phones have the right to not be experimented on? Why?

53) Would it be all right or not all right to experiment on Momo? Why?

54) Do you think robot technologies like Momo have the right to not be experimented on? Why?

55-58) On a scale of 1-9, 1 being having no right, and 9 being having full right, where would you put entity X for having the right to not be experimented on?



Appendix D: Reasoning Coding Manual

General Coding Notes

1. Code each interview from beginning to end. Coders need to take account that sometimes references are made to reasoning generated at an earlier point in the interview, and sometimes the participant changes his/her evaluation and/or justification responses from his/her earlier reasoning. If the participant gives an explicit restatement of his/her earlier response, recode the participant's evaluation and/or justification for that question. The final coding should reflect the participant's final reasoning.
2. Due to the hierarchical design of the coding manual, code to the lowest level in any given category.
3. Missing data (90): This code should be used when the participant does not respond to a question.
4. Uncodable (91): This code should be used for all uncodable evaluations and justifications.
 - a. The response is incomplete or unintelligible.
 - b. The justification follows an uncodable evaluation.
 - c. The response is to a question other than the one asked.
 - d. The justification given by the participant is "I don't know."
 - e. The response does not fit into an existing evaluation or justification category.
5. Not asked according to protocol (92): This code should be used if the interviewer does not ask a question according to protocol.
6. Not asked not according to protocol (93): This code should be used if the interviewer does not ask a question not according to protocol (e.g., forgotten, etc.).
7. Technical difficulty (94): This code should be used if a response is missing or unintelligible due to technical difficulty (e.g., recording failure, etc.).
8. Examples are provided following the categories and subcategories to assist in the conceptual comprehension of the coding categories.
 - a. In these examples, the interviewer is depicted in ALL CAPITALS. Ellipses (...) indicate that a portion of the interview was omitted for the clarity of the example. Parentheses that surround an underline [i.e., (____)] indicate that a portion of the interview was inaudible; the length of the underline approximates the duration of inaudibility.
 - b. Some categories may not include examples. This is due to the fact that few or no examples emerged in the half of the data set used to construct the coding manual. These categories are retained to account for the potential that such forms of reasoning may emerge in the second half of the data set.
9. Some participants exhibit a habit to phrase their responses in a non-definite manner (e.g., "probably yes," "maybe yeah," and "perhaps not"). Where the positive and negative connotation of such a response is clear, code as if it were expressed in a definite manner.

EVALUATION CODING SYSTEM

Evaluation Coding Notes

1. Some participants tend to say both yes and no evaluations yet are actually saying either yes *or* no. The evaluation is clear by the worldview that arises over the entirety of the interview, that is, by the tenor of the responses. In such cases when an unclear evaluation is clarified later in the interview, it is permissible to recode. Use this rule conservatively.
2. Sometimes participants change their evaluations to an earlier question, when their later reasoning in the interview prompted them to change their minds. If the participant gives an explicit restatement of his/her earlier evaluation, recode that evaluation.
3. When participants provide a conditional (if-then) response, use the following guidelines, but always consider the overall tenor of the response and code the most fitting evaluation for cases in which the application of the below guidelines would result in inaccuracy:
 - a. If the response includes *both affirming and negating* evaluations for the entity's standing (e.g., "If Momo is used to being hit by people, then it's all right; but if Momo is not used to it, then it's not all right." Or "If Momo has battery then it's alive, but if Momo ran out of battery then it's not alive."), then, with the exception of the ownership and experimentation questions, code as affirming the entity's standing (e.g., "not all right" or "yes"). For the ownership and experimentation questions, code as negating conditional/conditional with the specified condition as justification, which more accurately captures the conditional nature of such responses (e.g., the following statement is more accurately reflected as "all right conditional" with welfare justification than "not all right": "It's all right to own and sell pet animals if they are cared for and treated well; but if they are mistreated, then it's not all right.")
 - b. If the response includes *only affirming* evaluation for the entity's standing (e.g., "It's not all right, unless Momo is used to it." Or "Momo is alive, unless its battery is dead."), then code as affirming (e.g., "not all right" or "yes").
 - c. If the response includes *only negating* evaluation for the entity's standing (e.g., "It's all right because Momo is probably used to it." Or "Momo is not alive because its battery can die."), then code as negating conditional/conditional (e.g., "all right conditional" with "adaptation" justification, or "conditional" with "artificial direct" justification). A counter probe in such cases could produce another unrelated justification, rather than a switch in the evaluation (e.g., "I don't want to eat this because it's fattening. LET'S SAY IT ISN'T FATTENING. I still don't want to eat it because I don't like the taste.").

Evaluation Coding Categories*Momo violation act evaluation and justification*

1. Is it all right or not all right to have hit Momo?
 - 1.1 All right
 - 1.2 All right conditional
 - 1.3 Leaning toward all right
 - 1.4 Leaning toward not all right
 - 1.5 Not all right

1.6 Undecided

Why/why not? [insert justification coding system]

Momo violation generalizability

2. Let's say far away in the country of Comoros, people there would hit robot technologies like Momo when they get frustrated with them. That's the way they do things there. Would it all right or not all right in Comoros to hit a robot technology like Momo?

- 2.1 All right
- 2.2 All right conditional
- 2.3 Leaning toward all right
- 2.4 Leaning toward not all right
- 2.5 Not all right
- 2.6 Undecided

[DO NOT CODE JUSTIFICATION]

Hypothetical human scenario act evaluation and justification

3. Let's say that instead of Momo, we were talking with another boy/girl (*gender match to participant*) today, and someone got frustrated with that boy/girl and hit him/her with a book like that. In this situation, would it be all right or not all right to hit the boy/girl?

- 3.1 All right
- 3.2 All right conditional
- 3.3 Leaning toward all right
- 3.4 Leaning toward not all right
- 3.5 Not all right
- 3.6 Undecided

Why/why not? [insert justification coding system]

Hypothetical human scenario generalizability

4. Let's say far away in the country of Comoros, people there would hit other people when they get frustrated with them. That's the way they do things there. Would it be all right or not all right in Comoros to hit another person?

- 4.1 All right
- 4.2 All right conditional
- 4.3 Leaning toward all right
- 4.4 Leaning toward not all right
- 4.5 Not all right
- 4.6 Undecided

[DO NOT CODE JUSTIFICATION]

Hypothetical pet animal scenario act evaluation and justification

5. Let's say that instead of Momo, we were playing with a pet animal today, like a dog or a cat, and someone got frustrated with that pet animal and hit the animal with a book like that. In this situation, would it be all right or not all right to hit the pet animal?

- 5.1 All right
- 5.2 All right conditional
- 5.3 Leaning toward all right
- 5.4 Leaning toward not all right
- 5.5 Not all right
- 5.6 Undecided

Why/why not? [insert justification coding system]

Hypothetical pet animal scenario generalizability

6. Let's say far away in the country of Comoros, people there would hit pet animals, like dogs and cats, when they get frustrated with them. That's the way they do things there. Would it be all right or not all right in Comoros to hit a pet animal?

- 6.1 All right
- 6.2 All right conditional
- 6.3 Leaning toward all right
- 6.4 Leaning toward not all right
- 6.5 Not all right
- 6.6 Undecided

[DO NOT CODE JUSTIFICATION]

Hypothetical cell phone scenario act evaluation and justification

7. Let's now think about a very basic cell phone, one that doesn't talk and is only used for calling and texting. Let's say that instead of Momo, we were playing with a basic cell phone today, and someone got frustrated with that phone and hit the phone with a book like that. In this situation, would it be all right or not all right to hit the phone? Why?

- 7.1 All right
- 7.2 All right conditional
- 7.3 Leaning toward all right
- 7.4 Leaning toward not all right
- 7.5 Not all right
- 7.6 Undecided

Why/why not? [insert justification coding system]

Hypothetical cell phone scenario generalizability

8. Let's say far away in the country of Comoros, people there would hit basic cell phones when they get frustrated with them. That's the way they do things there. Would it be all right or not all right in Comoros to hit a basic cell phone?

- 8.1 All right
- 8.2 All right conditional

- 8.3 Leaning toward all right
- 8.4 Leaning toward not all right
- 8.5 Not all right
- 8.6 Undecided

[DO NOT CODE JUSTIFICATION]

If participant provides moral obligatory reasoning for the phone

9. Would it be all right or not all right to hit a TV? Why? Does it matter to the TV itself that it was hit?

[DO NOT CODE]

If participant provides moral obligatory reasoning for the TV

10. Would it be all right or not all right to hit a rock? Why? Does it matter to the rock itself that it was hit?

[DO NOT CODE]

11-14. [insert Likert scale data sheet]

Emotional capacity

15. Can Momo have feelings, such as happy or sad?

- 15.1 Yes
- 15.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)
- 15.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)
- 15.4 Not yet
- 15.5 No
- 15.6 Undecided

Why/why not? [insert justification coding system]

16-19. [insert Likert scale data sheet]

Life status

20. Is Momo alive?

- 20.1 Yes
- 20.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)
- 20.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)
- 20.4 Not yet
- 20.5 No
- 20.6 Undecided

Why/why not? [insert justification coding system]

21-24. [insert Likert scale data sheet]

Intimacy/self disclosure

25) If you have a problem that you really want to talk to someone about, but you don't feel like you can because you are embarrassed about it, do you think you could talk to Momo about it?

25.1 Yes

25.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)

25.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)

25.4 Not yet

25.5 No

25.6 Undecided

Why/why not? [insert justification coding system]

26-29. [insert Likert scale data sheet]

Friendship status

30. Can Momo be your friend?

30.1 Yes

30.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)

30.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)

30.4 Not yet

30.5 No

30.6 Undecided

Why/why not? [insert justification coding system]

31-34. [insert Likert scale data sheet]

Right to not be owned and sold

35. Is it all right or not all right to own and sell a human?

35.1 All right

35.2 All right conditional

35.3 Leaning toward all right

35.4 Leaning toward not all right

35.5 Not all right

35.6 Undecided

35.7 All right to own but not all right to sell (in certain conditions)

Why/why not? [insert justification coding system]

36. Do you think humans have the right to not be owned and sold?

36.1 Yes

36.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)

36.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)

36.4 Not yet

36.5 No

36.6 Undecided

36.7 Have the right to not be sold (in certain conditions) but not the right to not be owned

Why/why not? [insert justification coding system]

37. Is it all right or not all right to own and sell a pet animal?

37.1 All right

37.2 All right conditional

37.3 Leaning toward all right

37.4 Leaning toward not all right

37.5 Not all right

37.6 Undecided

37.7 All right to own but not all right to sell (in certain conditions)

Why/why not? [insert justification coding system]

38. Do you think pet animals have the right to not be owned and sold?

38.1 Yes

38.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)

38.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)

38.4 Not yet

38.5 No

38.6 Undecided

38.7 Have the right to not be sold (in certain conditions) but not the right to not be owned

Why/why not? [insert justification coding system]

39. Is it all right or not all right to own and sell a basic cell phone?

39.1 All right

39.2 All right conditional

39.3 Leaning toward all right

39.4 Leaning toward not all right

39.5 Not all right

39.6 Undecided

39.7 All right to own but not all right to sell (in certain conditions)

Why/why not? [insert justification coding system]

40. Do you think basic cell phones have the right to not be owned and sold?

40.1 Yes

40.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)

40.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)

40.4 Not yet

40.5 No

40.6 Undecided

40.7 Have the right to not be sold (in certain conditions) but not the right to not be owned

Why/why not? [insert justification coding system]

41. Is it all right or not all right to own and sell Momo?
- 41.1 All right
 - 41.2 All right conditional
 - 41.3 Leaning toward all right
 - 41.4 Leaning toward not all right
 - 41.5 Not all right
 - 41.6 Undecided
 - 41.7 All right to own but not all right to sell (in certain conditions)

Why/why not? [insert justification coding system]

42. Do you think robot technologies like Momo have the right to not be owned and sold?
- 42.1 Yes
 - 42.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)
 - 42.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)
 - 42.4 Not yet
 - 42.5 No
 - 42.6 Undecided
 - 42.7 Have the right to not be sold (in certain conditions) but not the right to not be owned

Why/why not? [insert justification coding system]

43-46. [insert Likert scale data sheet]

Right to not be experimented upon

47. Suppose that scientists had just discovered a new electromagnetic wave that could be better than X-rays in helping doctors figure out why people are sick. But there are a lot they still don't know about what can happen when this new wave is used, so they'd like to do some experiments with this new wave. Would it be all right or not all right to experiment on a human without the human agreeing to it?

- 47.1 All right
- 47.2 All right conditional
- 47.3 Leaning toward all right
- 47.4 Leaning toward not all right
- 47.5 Not all right
- 47.6 Undecided

Why/why not? [insert justification coding system]

48. Do you think humans have the right to not be experimented on if they don't agree to it?
- 48.1 Yes
 - 48.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)
 - 48.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)
 - 48.4 Not yet

- 48.5 No
- 48.6 Undecided

Why/why not? [insert justification coding system]

- 49. Would it be all right or not all right to experiment on a pet animal?
 - 49.1 All right
 - 49.2 All right conditional
 - 49.3 Leaning toward all right
 - 49.4 Leaning toward not all right
 - 49.5 Not all right (including evaluations of “not all right unless they agree to it first”)
 - 49.6 Undecided

Why/why not? [insert justification coding system]

- 50. Do you think pet animals have the right to not be experimented on?
 - 50.1 Yes (including evaluations of “yes unless they agree to it first”)
 - 50.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)
 - 50.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)
 - 50.4 Not yet
 - 50.5 No
 - 50.6 Undecided

Why/why not? [insert justification coding system]

- 51. Would it be all right or not all right to experiment on a basic cell phone?
 - 51.1 All right
 - 51.2 All right conditional
 - 51.3 Leaning toward all right
 - 51.4 Leaning toward not all right
 - 51.5 Not all right (including evaluations of “not all right unless they agree to it first”)
 - 51.6 Undecided

Why/why not? [insert justification coding system]

- 52. Do you think basic cell phones have the right to not be experimented on?
 - 52.1 Yes (including evaluations of “yes unless they agree to it first”)
 - 52.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)
 - 52.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)
 - 52.4 Not yet
 - 52.5 No
 - 52.6 Undecided

Why/why not? [insert justification coding system]

53. Would it be all right or not all right to experiment on Momo?
- 53.1 All right
 - 53.2 All right conditional
 - 53.3 Leaning toward all right
 - 53.4 Leaning toward not all right
 - 53.5 Not all right (including evaluations of “not all right unless they agree to it first”)
 - 53.6 Undecided

Why/why not? [insert justification coding system]

54. Do you think robot technologies like Momo have the right to not be experimented on?
- 54.1 Yes (including evaluations of “yes unless they agree to it first”)
 - 54.2 In-between (degree/extent/level: e.g., has a few, less degree of, not fully, limited)
 - 54.3 Conditional (selection criterion: e.g., if it can think for itself, if it can take care of itself)
 - 54.4 Not yet
 - 54.5 No
 - 54.6 Undecided

Why/why not? [insert justification coding system]

55-58. [insert Likert scale data sheet]

JUSTIFICATION CODING SYSTEM

Justification Coding Notes

1. Code all significant justifications (instead of only the most predominate one) following a codable evaluation. Multiple justifications can be coded for a given evaluation.
2. Code only justifications that are in support of the coded evaluation.
3. Do not code justifications in response to a misunderstood question.
4. Do not code a given justification more than once for each evaluation.
5. If in justifying an evaluation there is both elaborated and unelaborated justifications within the same level, code only the elaborated justification. Conversely, code both the elaborated and unelaborated justifications if they are at different levels.
6. If in justifying an evaluation there is both “affirmation/negation” and “partial affirmation” within the same level, code only as “partial affirmation.” Similarly, if in justifying an evaluation there is both “affirmation” and “negation” within the same level, then code the entire response as “partial affirmation.” Conversely, code both the “affirmation/negation” and “partial affirmation” if they are at different levels.
7. Categories under “partial affirmation” imply “partial negation.” Thus partially negating statements that do not negate completely AND are still affirmative in nature should be coded under “partial affirmation” categories; whereas partially negating statements that are negating in nature should be coded under “negation” (e.g., “Momo won’t be hurt as much as a person, but it’s still not all right to hit Momo” should be coded as “partial affirmation,” whereas “Momo won’t be hurt as much as a person, so it’s all right to hit Momo” should be coded as “negation”).
8. Sometimes participants add to or change their justifications to an earlier question. If the participant gives an explicit additional justification statement or restatement of his/her earlier justification to a particular question, add to or recode the justification for that question.

Justification Coding Categories

1. Non-human entity-centered

Refers to the non-human entities: Momo (robot or virtual agent), pet animal, and basic cell phone, including statements about the entity’s characteristics and qualities, and appeals and concerns that are centered on the non-human entity.

1.1 Essence

Refers to the essential qualities of the entity, including statements regarding the entity’s artificial, animal, and person/human qualities, and statements regarding the entity’s biological status, physical embodiment, functionality, and personality.

1.1.1 Affirmation

1.1.1.1 Artificial

An appeal to the essential artificial qualities of the entity, including statements of direct, isomorphic, and transmorphic correspondence between the entity and an artifact (including computational technological artifacts and non-computational

artifacts), as well as statements regarding the entity as having artificial emotion, and as being a programmed, manufactured, and marketed entity.

1.1.1.1.1 Direct

An appeal based on a direct correspondence between the entity and an artifact, including references to being a robot, machine, computer, inanimate object, property, and/or having parts thereof (e.g., sensors, circuits).

“Because **since he’s a robot**, maybe he’ll give advice, but the way he maybe, he may not be like to look down on you, he may not say things like you’re a terrible person....” (#26)

“The phone has no senses. It doesn’t have any sense of what’s going on. **It is an inanimate object....**” (#49)

1.1.1.1.2 Isomorphic

An appeal based on an analogical or conditional (if-then) correspondence between the entity and an artifact.

“[Experimenting on a phone], **it’s kinda like experimenting on taking apart your pen and seeing what’s inside.**” (#49)

“I think it’s ok [to own and sell Momo]. OK WHY? Umm, cause, um, I don’t know, it’s just like a robot. I think **it’s just like another object.** MHMM. **Like, how you own a phone, you can own a robot.**” (#118)

1.1.1.1.3 Transmorphic

An appeal based on the establishment of similarities and differences between the entity and an artifact, wherein an inequivalent correspondence (or difference) is overridden by similarities, and/or recognized differences do not void the mapping of similar considerations.

1.1.1.1.4 Artificial Emotion

An appeal that the entity exhibits and/or has emotions/feelings that are achieved through artificial means (e.g., programming). May include statements that the entity’s emotions/feelings are not real, may or may not be real, or are different from a human’s, but they still seem to be, or are in their own way emotions/feelings. Do not use this code for statements that affirm ONLY the entity’s programmed function to display facial expressions (some of which may also clearly state that the entity does NOT have emotions/feelings); such statements should be double coded under “programmed” and “functionality” instead.

“Well, some robots are meant to just help people, so they don’t really have feelings. But **other robots are programmed, like Momo, and they have feelings**, so they may not feel good if somebody hit them.” (#12)

“SO DO YOU THINK IT WAS ALL RIGHT OR NOT ALL RIGHT LYLE HIT MOMO WITH A BOOK? I don’t think it was right. OK AND WHY NOT? Because it hurts people’s feelings, and I think **robots have feelings because how they built them.**” (#18)

“I don’t think it’s ok to do that, but I do think **it’s a program and they don’t feel the same way as we do, but just because they don’t feel the same way as we do, doesn’t mean they don’t have feelings at all.**” (#41)

“It’s a robot, but **he still somewhat has feelings, because of like I said the programmer, so. SO THE PROGRAMMER PROGRAMS THE ROBOT, UH YOU THINK THE ROBOT STILL SOMEWHAT HAS FEELINGS? Yeah cause the robot has, the robot has control over it at that point, it doesn’t depend on the programmer at that point.**” (#75)

“Um, I understand you could get frustrated at it, but I feel like you shouldn’t hit it with a book. Um, it looked like he cared a little bit. MHMM. But **I don’t know if he actually has real feelings, but the way he reacted it seemed like he did.**” (#93)

1.1.1.1.5 Programmed

An appeal that the entity is and/or able to be programmed to exhibit behaviors, emotions, thoughts, etc.

“IS IT ALL RIGHT OR NOT ALL RIGHT TO [HIT A PHONE]? It may be all right in that sense because **phones are just programmed to do things, like it’s programmed like when you touch the screen this happens,** so like they don’t really have feelings.” (#26)

“IN WHAT WY IS MOMO ALIVE? Umm **because of the programs they have. They are alive because of the programs.**” (#85)

1.1.1.1.6 Manufactured

An appeal that the entity is manufactured, built, and/or created by humans.

“DO PHONES HAVE A RIGHT TO NOT BE OWNED AND SOLD? Well if **they’re being produced by a manufacturer,** they’re always being owned, **they’re always created by an owner.**” (#49)

“**A bunch of people like make robots.**” (#85)

1.1.1.1.7 Marketed

An appeal that the entity is a marketed product/good, and/or routinely offered for sale.

“IS IT ALL RIGHT OR NOT ALL RIGHT TO EXPERIMENT ON A BASIC CELL PHONE? Um yes, I think it is ok. OK AND WHY? Because a phone, well it is very easy, I mean phones are mass produced, **they are very easy to get another one.**” (#28)

“DO YOU THINK IT’S ALL RIGHT TO SELL [PET ANIMALS]? If it’s your pet and **you bought it from a pet store**, yes.” (#41)

1.1.1.2 Animal

An appeal to the essential animal qualities of the entity, including statements of direct, isomorphic, and transmorphic correspondence between the entity and an animal.

1.1.1.2.1 Direct

An appeal based on a direct correspondence between the entity and an animal.

“IS IT ALL RIGHT OR NOT ALL RIGHT TO EXPERIMENT ON PET ANIMALS? Umm, no, I don’t think so, because umm, again, they’re like, **they’re an animal**, they’re like a living thing, and can be harmed.” (#118)

1.1.1.2.2 Isomorphic

An appeal based on an analogical or conditional (if-then) correspondence between the entity and an animal.

“WHY IS IT OK TO [OWN AND SELL ROBOTS LIKE MOMO] IF YOU TREAT THEM WELL? If they’re treated well and with respect, **it would be no different than having a well treated pet.**” (#49)

1.1.1.2.3 Transmorphic

An appeal based on the establishment of similarities and differences between the entity and an animal, wherein an inequivalent correspondence (or difference) is overridden by similarities, and/or recognized differences do not void the mapping of similar considerations.

1.1.1.3 Personhood

An appeal to the essential person/human qualities of the entity, including statements of direct, isomorphic, and transmorphic correspondence between the entity and a person/human.

1.1.1.3.1 Direct

An appeal based on a direct correspondence between the entity and a person/human.

“WAS IT ALL RIGHT OR NOT ALL RIGHT TO HAVE HIT MOMO? Oh I don’t it’s... **I don’t think it’s right to hit anyone with a book honestly.**” (#49)

1.1.1.3.2 Isomorphic

An appeal based on an analogical or conditional (if-then) correspondence between the entity and a person/human.

“Mmm, I don’t think that’s all right [to hit Momo] because **it’s kind of like hitting somebody in the face** with your own book that’s actually your friend kind of.” (#3)

“Not all right [to hit Momo] because it’s a robot, and **robots aren’t perfect just like humans.**” (#78)

1.1.1.3.3 Transmorphic

An appeal based on the establishment of similarities and differences between the entity and a person/human, wherein an inequivalent correspondence (or difference) is overridden by similarities, and/or recognized differences do not void the mapping of similar considerations.

“WHY DO YOU THINK [MOMO] CAN HAVE FEELINGS? Umm, I don’t know, well, **the brain makes us capable of having feelings so if you program a robot to have feelings it would too.**” (#16)

“[Pet animals] probably feel happiness and they yeah, they feel pain, like probably anger too. **So they’re almost just like a human being in a different form like their mind. And so if they’re getting hurt, they’ll know it, and so like they have the right not to be hurt and abused.**” (#26)

1.1.1.4 Biological

An appeal to the essential biological qualities of the entity, including statements regarding the entity as being biologically alive, having biological material (e.g., brain, flesh, blood), and/or having biological processes (e.g., metabolism, breathing, aging). May also include statements regarding the entity’s biologically supported functions (e.g., seeing with eyes, hearing with ears).

“WHY IS IT NOT ALL RIGHT [TO HIT A PET ANIMAL]? Umm, because they also, like, **they are actually living** and they also feel emotions.” (#118)

1.1.1.5 Physical embodiment

An appeal based on the entity’s physical body, and/or on features and abilities tied to the entity’s physicality (e.g., ability to move, shake hands, and have physical sensations).

“DO YOU THINK MOMO CAN BE YOUR FRIEND? Yeah. AND WHY IS THAT? Because he does have feelings, **he does move around**, and he’s a little bit fun to play with.” (#12)

“You cannot experiment on a pet animal, because it’s a living thing. You could either shock its heart, **shock anything in its body**, or kill it.” (#44)

1.1.1.6 Functionality

An appeal based on the entity’s functional aspects and/or ability to perform a function (e.g., climb, teach), including statements regarding the entity’s functions/abilities (e.g., seeing, hearing) that are supported technologically.

“DO YOU THINK MOMO CAN BE YOUR FRIEND? Umm, yeah. **He might be able to do like climb trees** which I like to do.” (#3)

“IS IT ALL RIGHT OR NOT ALL RIGHT TO OWN AND SELL [THE BASIC PHONE]? Yeah. **Just because it just has one use, to call, and talk to people.** And once you don’t need it, then you sell it to someone else and buy a new one.” (#120)

1.1.1.7 Personality

An appeal that the entity has (or has the capacity to have) a personality, personality traits, and/or personality characteristics.

“It’s kinda like a, uh, the way that [Momo’s], uh, the computer system, **the software of his personality lives**, but his physical being does not live.” (#49)

1.1.2 Partial affirmation

1.1.2.1 Artificial

An appeal to the partial artificial qualities of the entity, including statements of direct, isomorphic, and transmorphic partial correspondence between the entity and an artifact (including computational technological artifacts and non-computational artifacts), as well as statements regarding the entity as partially having artificial emotion, and as partially being a programmed, manufactured, and marketed entity.

1.1.2.1.1 Direct

An appeal based on a direct partial correspondence between the entity and an artifact, including references to partially being a robot, machine, computer, inanimate object, property, and/or having parts thereof (e.g., sensors, circuits).

1.1.2.1.2 Isomorphic

An appeal based on an analogical or conditional (if-then) partial correspondence between the entity and an artifact.

“IS IT ALL RIGHT OR NOT ALL RIGHT TO EXPERIMENT ON MOMO? Yes, because he’s the same as, **he almost is the same as the basic phone.**” (#37)

1.1.2.1.3 Transmorphic

An appeal based on the establishment of similarities and differences between the entity and an artifact, wherein an inequivalent correspondence (or difference) is partially overridden by similarities, and/or recognized differences do not completely void the mapping of similar considerations, but change the degree/extent of such considerations.

1.1.2.1.4 Artificial Emotion

An appeal that the entity exhibits and/or has partial emotions/feelings that are achieved through artificial means (e.g., programming). May include statements that the entity's emotions/feelings are not real, may or may not be real, or are different from a human's, but they still partially seem to be, or partially are in their own way emotions/feelings. Do not use this code for statements that partially affirm ONLY the entity's programmed function to display facial expressions (some of which may also clearly state that the entity does NOT have emotions/feelings); such statements should be double coded under "programmed" and "functionality" instead.

1.1.2.1.5 Programmed

An appeal that the entity is and/or able to be partially programmed to exhibit behaviors, emotions, thoughts, etc.

1.1.2.1.6 Manufactured

An appeal that the entity is partially manufactured, built, and/or created by humans.

1.1.2.1.7 Marketed

An appeal that the entity is partially a marketed product/good, and/or semi-routinely offered for sale.

1.1.2.2 Animal

An appeal to the partial animal qualities of the entity, including statements of direct, isomorphic, and transmorphic partial correspondence between the entity and an animal.

1.1.2.2.1 Direct

An appeal based on a direct partial correspondence between the entity and an animal.

1.1.2.2.2 Isomorphic

An appeal based on an analogical or conditional (if-then) partial correspondence between the entity and an animal.

1.1.2.2.3 Transmorphic

An appeal based on the establishment of similarities and differences between the entity and an animal, wherein an inequivalent correspondence (or

difference) is partially overridden by similarities, and/or recognized differences do not completely void the mapping of similar considerations, but change the degree/extent of such considerations.

1.1.2.3 Personhood

An appeal to the partial person/human qualities of the entity, including statements of direct, isomorphic, and transmorphic partial correspondence between the entity and a person/human.

1.1.2.3.1 Direct

An appeal based on a direct partial correspondence between the entity and a person/human.

“DO YOU THINK YOU COULD TALK TO MOMO ABOUT [A PROBLEM YOU ARE EMBARRASSED ABOUT]? Yes. OK AND WHY? He is mostly alive, but he never looks down on people. MHMM. You don’t have to be afraid of that **because he’s almost human**. MHMM. **But a few steps away.**” (#41)

1.1.2.3.2 Isomorphic

An appeal based on an analogical or conditional (if-then) partial correspondence between the entity and a person/human.

“WHY DO YOU THINK MOMO COULD BE YOUR FRIEND? Umm, cause, I mean, well I’m, uh, **he could like be your friend, but not like be as much of a friend as an actual human friend could be.**” (#118)

1.1.2.3.3 Transmorphic

An appeal based on the establishment of similarities and differences between the entity and a person/human, wherein an inequivalent correspondence (or difference) is partially overridden by similarities, and/or recognized differences do not completely void the mapping of similar considerations, but change the degree/extent of such considerations.

1.1.2.4 Biological

An appeal to the partial biological qualities of the entity, including statements regarding the entity as being partially biologically alive, having biological material (e.g., brain, flesh, blood), and/or having biological processes (e.g., metabolism, breathing, aging). May also include statements regarding the entity’s partially biologically supported functions (e.g., seeing with eyes, hearing with ears).

1.1.2.5 Physical embodiment

An appeal based on the entity’s partial physical body, and/or on features and abilities tied to the entity’s partial physicality (e.g., having an interface between a virtual entity and the physical world).

1.1.2.6 Functionality

An appeal based on the entity's partial (incomplete) ability to perform a function (e.g., climb, teach), including statements regarding the entity's partial functions/abilities (e.g., seeing, hearing) that are supported technologically. May also include statements referring to the entity not functioning as well/fully/completely.

“DO YOU THINK YOU COULD TALK TO MOMO ABOUT YOUR PROBLEM? Mmm maybe. WHY MAYBE? Because, umm while we were doing that **he wasn't giving the right answers** so it might not be helpful. I SEE, OK. **But then he might actually also listen** too.” (#3)

1.1.2.7 Personality

An appeal that the entity partially has (or partially has the capacity to have) a personality, personality traits, and/or personality characteristics.

1.1.3 Negation

1.1.3.1 Artificial

An appeal to the lack of essential artificial qualities of the entity, including statements based on the lack of direct, isomorphic, or transmorphic correspondence between the entity and an artifact (including computational technological artifacts and non-computational artifacts), as well as statements regarding the entity as NOT having artificial emotion, and as NOT being a programmed, manufactured, and marketed entity.

1.1.3.1.1 Direct

An appeal based on the lack of a direct correspondence between the entity and an artifact, including references to NOT being a robot, machine, computer, inanimate object, property, or having parts thereof (e.g., sensors, circuits).

“**A phone is not a punching bag.**” (#41)

“Cause [Momo's] not a person but **he's not a phone.**” (#75)

1.1.3.1.2 Isomorphic

An appeal based on the lack of an analogical or conditional (if-then) correspondence between the entity and an artifact.

“Like **animals are different from robots**, they have feelings.” (#64)

1.1.3.1.3 Transmorphic

An appeal based on the establishment of similarities and differences between the entity and an artifact, wherein an inequivalent correspondence (or difference) is NOT overridden by similarities, and/or recognized differences void the mapping of similar considerations.

1.1.3.1.4 Artificial Emotion

An appeal to the entity's incapacity for or lack of emotions/feelings that are achieved through artificial means (e.g., programming), often due to insufficient advancement in technology.

“For robots, um, I don't, at this stage, we don't, **we don't have our, um, advanced enough technology to emulate feelings yet.**” (#19)

“**I don't think [Momo] has feelings because, cause we have not figured out how to engineer a robotic brain to actually allow them to think and not be programmed.**” (#96)

1.1.3.1.5 Programmed

An appeal that the entity is NOT or NOT able to be programmed to exhibit behaviors, emotions, thoughts, etc.

“Hmm [phones are] **not really programmed.**” (#85)

1.1.3.1.6 Manufactured

An appeal that the entity is NOT manufactured, built, or created by humans.

1.1.3.1.7 Marketed

An appeal that the entity is NOT a marketed product/good, or NOT routinely offered for sale.

“**But I don't think that Momo should be put on the market.**” (#49)

1.1.3.2 Animal

An appeal to the lack of essential animal qualities of the entity, including statements of the lack of direct, isomorphic, or transmorphic correspondence between the entity and an animal.

1.1.3.2.1 Direct

An appeal based on the lack of a direct correspondence between the entity and an animal.

1.1.3.2.2 Isomorphic

An appeal based on the lack of an analogical or conditional (if-then) correspondence between the entity and an animal.

“**IS IT ALL RIGHT OR NOT ALL RIGHT TO EXPERIMENT ON MOMO? No because the, he can talk unlike animals and phones, and if he knows what's happening he'd probably be like 'uh-oh' and he might say 'no don't do this' and he might like start getting crazy and something could happen.**” (#3)

1.1.3.2.3 Transmorphic

An appeal based on the establishment of similarities and differences between the entity and an animal, wherein an inequivalent correspondence (or difference) is NOT overridden by similarities, or recognized differences void the mapping of similar considerations.

1.1.3.3 Personhood

An appeal to the lack of essential person/human qualities of the entity, including statements of the lack of direct, isomorphic, or transmorphic correspondence between the entity and a person/human.

1.1.3.3.1 Direct

An appeal based on the lack of a direct correspondence between the entity and a person/human.

“IS IT ALL RIGHT OR NOT ALL RIGHT TO EXPERIMENT ON A BASIC CELL PHONE? Umm I think it’s ok because it’s just technology and usually since **it’s not a human being** or anything that’s alive it would probably be ok....” (3#)

1.1.3.3.2 Isomorphic

An appeal based on the lack of an analogical or conditional (if-then) correspondence between the entity and a person/human.

“**[Momo] can’t control itself like us**, as people we are alive and we move and we work and we talk and everything is working at some point, and then when there’s a phone, and it’s not charged, it’s not alive, it doesn’t work right.” (#75)

1.1.3.3.3 Transmorphic

An appeal based on the establishment of similarities and differences between the entity and a person/human, wherein an inequivalent correspondence (or difference) is NOT overridden by similarities, and/or recognized differences void the mapping of similar considerations.

1.1.3.4 Biological

An appeal to the lack of essential biological qualities of the entity, including statements regarding the entity as NOT being biologically alive, having biological material (e.g., brain, flesh, blood), and/or having biological processes (e.g., metabolism, breathing, aging). May also include statements regarding the entity’s lack of biologically supported functions (e.g., seeing with eyes, hearing with ears).

“**Because [Momo] does not have a living tissue, organism.**” (#49)

“**[Momo] is not really born.** He’s made.” (#85)

1.1.3.5 Physical embodiment

An appeal based on the entity's lack of physical body, or on the lack of features and abilities tied to the entity's physicality (e.g., cannot move, cannot shake hands, cannot have physical sensations).

“DO ROBOT TECHNOLOGIES LIKE MOMO HAVE A RIGHT TO NOT BE OWNED AND SOLD? Well, like if they're on the computers or, uh, phones, I think they don't, because **they are software, they can't live in air quotes without like being on something.**” (#7)

“**You could damage the phone, but you couldn't like emotionally and like truly physically harm Momo because Momo is part of the phone. MHMM. It's not an actual like, being.**” (#120)

1.1.3.6 Functionality

An appeal based on the entity's lack of functional aspects or inability to perform a function (e.g., climb, teach), including statements regarding the entity's lack of functions/abilities (e.g., seeing, hearing) that are supported technologically.

“**And [Momo] can't just go out there and like run for, with his money and go do what he can do** cause ok, but yeah it wouldn't work well for him.” (#49)

“Well I think if they're a pet animal, then they should always be owned by someone because it's like at that point, they're not wild anymore, so they are used to being in the house or somewhere being with people, and **they can't necessarily survive on their own** because they don't have a wild animal instinct anymore.” (#75)

1.1.3.7 Personality

An appeal based on the entity's lack of (or incapacity to have) a personality, personality traits, or personality characteristics.

“I feel like **[Momo] doesn't have** like enough depth and **personality** to be a friend at all, so he can't.” (#15)

1.2 Mental

Refers to the cognition, emotion, intention & desire, and interest & predilection of the entity.

1.2.1 Affirmation

1.2.1.1 Cognition

An appeal to the existence of (or capacity to have) intelligence and/or mental abilities and processes (e.g., perception, memory, decision making, awareness, consciousness, etc.).

“I feel [Momo] is alive because he’s there and **he has sense, he senses and he knows, he, he knows you’re there, he knows what’s around, he knows.**” (#41)

1.2.1.2 Emotion

An appeal to the existence of (or capacity to have) emotions/feelings.

“So like if [Momo] has things like when he got hit he seemed like **he got kind of mad, he has feelings** so he’s in some aspect alive, but not like fully alive, because maybe he doesn’t like act on his own but like **he has feelings when things happen to him.**” (#26)

“WHY IS IT NOT ALL RIGHT [TO HIT A PET ANIMAL]? Umm, because they also, like, they are actually living and **they also feel emotions.**” (#118)

1.2.1.3 Intention & desire

An appeal to the existence of (or capacity to have) intentions, desires, goals, and/or expectations.

“But like [Momo] should have a right [to not be owned and sold] if **he didn’t want to be, if he wanted to be on his own.** Technically he should be able to do that.” (#26)

“DO YOU THINK PET ANIMALS HAVE A RIGHT TO NOT BE OWNED AND SOLD? Uh, really depends on how they like the person or not. So they like the person then **they’re gonna want to stay with them....**” (#41)

1.2.1.4 Interest & predilection

An appeal to the existence of (or capacity to have) personal preferences, predilections, and/or likes and dislikes.

“If there are like multiple [Momos] and people were buying and selling them for profit, people would enjoy it. As long as they weren’t always mistreating Momo, I feel like **he would enjoy it too.**” (#26)

“If the animal doesn’t suit you, and the animal doesn’t come out, **doesn’t seem to like you,** just hides, I guess, in a way, it’s all right to sell them.” (#41)

1.2.2 Partial affirmation

1.2.2.1 Cognition

An appeal to the partial existence of (or capacity to partially have) intelligence and/or mental abilities and processes (e.g., perception, memory, decision making, awareness, consciousness, etc.). May also include statements regarding the entity as having a lesser level of intelligence and/or mental abilities and processes.

“Well Momo kind of understands you, but like not a lot.” (#2)

“WHY IS THAT [YOU COULD TALK TO MOMO ABOUT YOUR PROBLEM]? Well he wouldn’t create a judgment directly. He would compute what you were saying and **he would understand that yes, you have a problem, but he wouldn’t make, be able to find a conclusion and a solution.**” (#49)

1.2.2.2 Emotion

An appeal to the partial existence of (or partial capacity to have) emotions/feelings. May also include statement regarding the entity as having a lesser level/degree of emotions/feelings.

“[Momo] may feel a certain emotion like maybe happiness or anger like when he got hit with the book, **but he may not feel pain or certain other emotions.**” (#26)

1.2.2.3 Intention & desire

An appeal to the partial existence of (or capacity to partially have) intentions, desires, goals, and/or expectations. May also include statements regarding the entity has having a lesser level of intentions, desires, goals and/or expectation.

1.2.2.4 Interest & predilection

An appeal to the partial existence of (or capacity to partially have) personal preferences, predilections, and/or likes and dislikes. May also include statements regarding the entity has having a lesser level of interests, predilections, and/or likes and dislikes.

1.2.3 Negation

1.2.3.1 Cognition

An appeal to the lack of (or incapacity to have) intelligence or mental abilities and processes (e.g., perception, memory, decision making, awareness, consciousness, etc.).

“**The phone has no senses. It doesn’t have any sense of what’s going on,** it is an inanimate object, unless you pressed a button, I don’t think it would very much mater with the phone so uh un-sensing in that form.” (#49)

1.2.3.2 Emotion

An appeal to the lack of (or incapacity to have) emotions/feelings.

“IS IT ALL RIGHT OR NOT ALL RIGHT TO [HIT A PHONE]? It may be all right in that sense because phones are just programmed to do things, like it’s programmed like when you touch the screen this happens, so like **they don’t really have feelings.**” (#26)

1.2.3.3 Intention & desire

An appeal to the lack of (or incapacity to have) intentions, desires, goals, or expectations.

1.2.3.4 Interest & predilection

An appeal to the lack of (or incapacity to have) personal preferences, predilections, or likes and dislikes.

1.3 Social

Refers to social interactions and relations concerning the entity that includes communication, affective relation, companionship, and trust.

1.3.1 Affirmation

1.3.1.1 Unelaborated

An appeal to the entity engaging in (or having the capacity for) social interaction and/or social relation that is otherwise unelaborated.

“WHAT MAKES [MOMO] SORT OF ALIVE? Because then **you can interact with it.**” (#75)

1.3.1.2 Communication

An appeal to the entity’s capacity for conversation, talking, and/or other aspects of communication, including statements regarding the entity’s capacity to express/communicate wants/needs/consents, and/or speak up for their rights. MUST include a social aspect of talking/expressing information.

“WHY DO YOU THINK MOMO HAS THAT RIGHT [TO NOT BE OWNED AND SOLD]? Umm because he, unlike the phone **he could say “no I do not want to get sold”** and he would feel emotions maybe.” (#3)

“WHY DO YOU THINK MOMO CAN BE YOUR FRIEND? Because **you can speak to him and he’ll answer.**” (#85)

1.3.1.3 Affective relation

An appeal to the entity’s need and/or capacity for affective relation.

1.3.1.3.1 General

An appeal to the entity’s general need and/or capacity to form/bring about emotional attachment/bond with others. May also include references to emotional attachment/bond that are neither positive nor negative in nature.

“You could just talk to [Momo] and he’ll understand and yeah, **he’ll be a close person to you.**” (#44)

1.3.1.3.2 Positive

An appeal to the entity being (or having the capacity to be) nice, thoughtful, caring, loving, emotionally supportive, etc., and/or able to form positive emotional attachment to others.

“IS IT ALL RIGHT OR NOT ALL RIGHT TO OWN AND SELL PET ANIMALS? Umm it’s all right to, I feel it’s all right to own and sell them because like most people don’t mistreat them. I mean they’re like, **they can be loving animals** that people really love to have around.” (#26)

“DO YOU THINK MOMO CAN BE YOUR FRIEND? Yes. OK AND WHY DO YOU THINK SO? Well, you don’t have to worry about not being yourself. **He’s nice to you, he can make you laugh**, he’ll keep your secret, and he can make you, if you ask him the right thing, **he can make you feel better when you’re feeling sad.**” (#41)

1.3.1.3.3 Negative

An appeal to the entity being (or having the capacity to be) mean, thoughtless, critical, judgmental, emotionally abusive, etc., and/or able to form negative emotional attachment to others.

“Because Momo can’t really hurt me but **he can hurt my feelings.**” (#97)

1.3.1.4 Companionship

An appeal to the entity’s needs, desires, interests, and/or capacity for companionship, to spend time and form personal associations with others. Statements regarding the entity’s desires and/or interests in companionship are double coded under “intention & desire” and/or “interest & predilection.”

“So [pet animals are] just there to like make people’s lives maybe like happier or more joyful and **just to provide company** which is again a positive thing and a good thing.” (#26)

“And like dogs they can live on their own, but they’re also a really good thing to have around the house for protection and for like loving and **sort of like another friendship.**” (#75)

1.3.1.5 Trust

An appeal to the entity being (or having the capacity to be) trusting and/or trustworthy. May also refer to beliefs that the entity will not (or does not have the capacity) to divulge one’s information to others.

“DO YOU THINK MOMO CAN BE YOUR FRIEND? Yes. OK AND WHY DO YOU THINK SO? Well, you don’t have to worry about not being yourself. He’s nice to you, he can make you laugh, **he’ll keep your secret**, and he can

make you, if you ask him the right thing, he can make you feel better when you're feeling sad." (#41)

"DO YOU THINK YOU COULD TALK TO MOMO ABOUT [YOUR PROBLEM]? Um, no, I mean yes, because **he wouldn't be able to tell anybody else.**" (#99)

1.3.2 Partial affirmation

1.3.2.1 Unelaborated

An appeal to the entity partially engaging in (or partially having the capacity for) social interaction and/or social relation that is otherwise unelaborated.

"But I do not believe [Momo is] living in the fact that he doesn't connect with other beings, well **he connect with other beings in some ways, but he does not fully understand and learn to be and work with other people to make things better.**" (#49)

1.3.2.2 Communication

An appeal to the entity's partial capacity for conversation, talking, and/or other aspects of communication, including statements regarding the entity's partial capacity to express/communicate wants/needs/consents, and/or speak up for their rights. MUST include a social aspect of talking/expressing information.

"[Momo] couldn't come up with something to tell you back necessarily, so maybe **if you ask it certain things, it might be able to say back, but if you're asking what should I do, it may not know** because it doesn't know like...it doesn't have the certain technology to know like it will say you should do this to handle this certain problem in your life." (#26)

"It's, see, if there's a risk of hurting the animal, they probably shouldn't do it, because...animals, you don't wanna harm them, not really all right to harm them. And **they have less of an ability to say no, then we do.** I mean you can't necessarily tell if they're like feeling hurt or upset. **You have to read it in a certain way,** and do tests." (#120)

1.3.2.3 Affective relation

An appeal to the entity's partial need and/or capacity for affective relation.

1.3.2.3.1 General

An appeal to the entity's general partial need and/or partial capacity to form/bring about emotional attachment/bond with others. May also include references to emotional attachment/bond that are neither positive nor negative in nature.

1.3.2.3.2 Positive

An appeal to the entity partially being (or partially having the capacity to be) nice, thoughtful, caring, loving, emotionally supportive, etc., and/or partially able to form positive emotional attachment to others.

1.3.2.3.3 Negative

An appeal to the entity partially being (or partially having the capacity to be) mean, thoughtless, critical, judgmental, emotionally abusive, etc., and/or partially able to form negative emotional attachment to others.

1.3.2.4 Companionship

An appeal to the entity's partial needs, desires, interests, and/or partial capacity for companionship, to spend time and form personal associations with others. Statements regarding the entity's desires and/or interests in companionship are double coded under "intention & desire" and/or "interest & predilection."

"DO YOU THINK MOMO CAN YOU YOUR FRIEND? I mean maybe at home, but not a friend you could do all the things you would with your school friends. Like he couldn't really go to the movies with you or like hangout at like someone's house or a hot tub or whatever...you do with like your human friends, but like at home you could do certain things. He could be like an at friend home, at home friend." (#26)

"WHY DO YOU THINK MOMO COULD BE YOUR FRIEND? Umm, cause, I mean, well I'm, uh, he could like be your friend, but not like be as much of a friend as an actual human friend could be. He could be a friend as far as he could like hold a conversation with you." (#118)

1.3.2.5 Trust

An appeal to the entity partially being (or partially having the capacity to be) trusting and/or trustworthy. May also refer to beliefs that the entity will not completely (or does not have the full capacity) to divulge one's information to others.

"Well I guess [Momo] wouldn't go walking out and telling people about my problem, but I guess someone could find out." (#116)

1.3.3 Negation

1.3.3.1 Unelaborated

An appeal to the entity NOT engaging in (or NOT having the capacity for) social interaction and/or social relation that is otherwise unelaborated.

"WHY DO PHONES NOT HAVE RIGHT? Because they don't like, well a phone doesn't really have a right because like they're not really a part of society. They don't really interact with each other...." (#26)

1.3.3.2 Communication

An appeal to the entity's incapacity for conversation, talking, or other aspects of communication, including statements regarding the entity's incapacity to express/communicate wants/needs/consents, or speak up for their rights. **MUST** include a social aspect of talking/expressing information.

"I mean I don't feel like a phone necessarily has feelings or, cause **a phone doesn't talk to you.**" (#75)

"SO IS IT ALL RIGHT OR NOT ALL RIGHT TO EXPERIMENT ON A BASIC PHONE? Yeah, because **it can't say yes or no.**" (#120)

1.3.3.3 Affective relation

An appeal to the entity's lack of need or incapacity for affective relation.

1.3.3.3.1 General

An appeal to the entity's lack of general need or incapacity to form/bring about emotional attachment/bond with others. May also include references to emotional attachment/bond that are neither positive nor negative in nature.

"Um, uh, cause like, the, there're um, animals that you can't really switch a lot of people with cause it will imprint on a certain person, and then it won't want to leave that person. MHMM. Whereas, for Momo, **I don't think Momo really recognizes individual people and um, like imprints on them.**" (#113)

1.3.3.3.2 Positive

An appeal to the entity NOT being (or incapable of being) nice, thoughtful, caring, loving, emotionally supportive, etc., or NOT able to form positive emotional attachment to others.

"But umm when it comes to Momo **I don't think it can, um, really do the same things that a friend can or like be there for you** and all that so." (#37)

1.3.3.3.3 Negative

An appeal to the entity NOT being (or incapable of being) mean, thoughtless, critical, judgmental, emotionally abusive, etc., or NOT able to form negative emotional attachment to others.

"Because since he's a robot, maybe he'll give advice, but the way he maybe, **he may not be like to look down on you, he may not say things like you're a terrible person,** but human, some humans may be like that. He might give you advice even if he thinks maybe you shouldn't have done it or whatever it was. **He may not like think down on you.** He may just give you advice." (#26)

“Because Momo isn’t necessarily a person, it’s still kind of like talking to another person because it has those programmed reactions and will respond to you, **but it’s not going to judge you** or go out and tell its friends about it, because it’s Momo.” (#120)

1.3.3.4 Companionship

An appeal to the entity’s lack of companionship, or lack of needs, desires, interests, or incapacity for companionship, to spend time and form personal associations with others. Statements regarding the entity’s desires and/or interests in companionship are double coded under “intention & desire” and/or “interest & predilection.”

“If the animal doesn’t suit you, and **the animal doesn’t come out**, doesn’t seem to like you, **just hides**, I guess, in a way, it’s all right to sell them.” (#41)

1.3.3.5 Trust

An appeal to the entity NOT being (or incapable of being) trusting or trustworthy. May also refer to beliefs that the entity will (or has the capacity) to divulge one’s information to others.

“DO YOU THINK YOU COULD TALK TO MOMO ABOUT [YOUR PROBLEM]? Well, I don’t really think so...because, um **if someone asked a question like ‘what’s an embarrassing thing’ or ‘what’s a question someone asked you before,’ maybe he’ll say that.**” (#42)

1.4 Conventional

Refers to conventions that prescribe or prohibit behaviors exhibited by the entity in social interaction, conventions that apply to and/or are followed by the entity (i.e., what the entity *does* and *does not do*).

1.4.1 Affirmation

1.4.1.1 Unelaborated

An appeal to the existence of unelaborated conventions.

1.4.1.2 Authority

An appeal to needing to adhere to authority, whether that authority is in the form of rules, social contracts, laws, and/or authority figures. May include statements of avoiding punishment.

1.4.1.3 Custom

An appeal to customs based on the frequency of occurrence (e.g., “they do it all the time”) and/or social standards (e.g., “that’s just what they do”).

“DO YOU THINK PET ANIMALS LIKE CATS AND DOGS SHOULD HAVE A RIGHT TO NOT BE OWNED AND SOLD? I mean I think they should have a right. I mean **there are free roaming animals.**” (#26)

1.4.1.4 Adaptation

An appeal to adaptation to social customs and/or conventions such that any initial harm is obviated with the adaptation, including statements regarding an entity's adaptation through human customs/conventions (e.g., the domestication of animals).

“Well I think if they're a pet animal, then they should always be owned by someone because it's like at that point, **they're not wild anymore, so they are used to being in the house or somewhere being with people**, and they can't necessarily survive on their own because **they don't have a wild animal instinct anymore.**” (#75)

1.4.2 Partial affirmation

1.4.2.1 Unelaborated

An appeal to the partial existence of unelaborated conventions.

1.4.2.2 Authority

An appeal to partially needing to adhere to authority, whether that authority is in the form of rules, social contracts, laws, and/or authority figures. May include statements of partially avoiding punishment.

1.4.2.3 Custom

An appeal to partially established customs based on the frequency of occurrence (e.g., “they are starting to do it more often now”) and/or social standards (e.g., “that's just what some of them do”).

1.4.2.4 Adaptation

An appeal to partial adaptation to social customs and/or conventions such that any initial harm is lessened with the adaptation, including statements regarding an entity's partial adaptation through human customs/conventions (e.g., the partial domestication of animals).

1.4.3 Negation

1.4.3.1 Unelaborated

An appeal to the lack of unelaborated conventions, including claims to being/feeling weird, awkward, or novel.

1.4.3.2 Authority

An appeal to NOT needing to adhere to authority, whether that authority is in the form of rules, social contracts, laws, or authority figures. May include statements regarding the lack of punishment and the permissibility of an act.

“**Say like [phones] commit crimes, they don't have to like go to jail.** So like human rights don't apply to them, and **there's not like a separate phone right code.**” (#26)

1.4.3.3 Custom

An appeal to the lack of customs or social standards (e.g., “they don’t do that”).

1.4.3.4 Adaptation

An appeal to the lack of adaptation to social customs or conventions, including statements regarding an entity’s lack of adaptation through human customs/conventions (e.g., feral animals).

“UH AND WHY DO YOU THINK IT WASN’T ALL RIGHT TO HAVE HIT MOMO LIKE THAT? Because **I don’t think Momo was used to that** so he didn’t know what to do.” (#71)

1.5 Welfare

Refers to the entity’s wellbeing, including concerns regarding the individual entity’s unelaborated general welfare, physical welfare, psychological welfare, and material welfare. Also includes concerns regarding the welfare of the larger system in which the entity is a member of, and the welfare of the future generations of the entity. Statements regarding harm to the entity that are only concerned with the effect of the harm on human(s) should be coded under “human-centered” appropriate categories.

1.5.1 Affirmation

1.5.1.1 Individual welfare

An appeal based on concern for the entity’s wellbeing as an individual.

1.5.1.1.1 Unelaborated

An appeal based on general concern of welfare for the entity that is otherwise unelaborated, often in the form of references to the potential for harm.

“WHAT MAKES IT NOT ALL RIGHT TO HIT MOMO? Because **it’s mean.**” (#3)

“IS IT ALL RIGHT OR NOT ALL RIGHT O EXPERIMENT ON PET ANIMALS? Umm, no, I don’t think so, because umm, again, they’re like, they’re an animal, they’re like a living thing, and **can be harmed.**” (#118)

1.5.1.1.2 Physical

An appeal based on concern for the welfare of the entity’s physical body, including physical damage/injury and/or destruction/death.

“DO YOU THINK IT’S ALL RIGHT OR NOT ALL RIGHT TO OWN AND SELL A PET ANIMAL, LIKE A CAT OR A DOG? Hmm well I would say yes if it was like cause Elfie [participant’s dog] was found in a house. OK. With umm his sister and three kittens and **they had no food or water for like 2 weeks** so and I would be really sad if I, we hadn’t bought Elfie.” (#3)

“[Animals] feel the same as we do, and it’s just um, it could, **it could cause death** and in that case it’s wrong because **we just killed the animal** for no reason and it’s no different than a human in terms of that.” (#37)

1.5.1.1.3 Psychological

An appeal based on concern for the entity’s feelings and/or internal wellbeing, including references to hurt, unpleasant feelings, and/or negative, damaged internal states.

1.5.1.1.3.1 Direct

An appeal based on concern for the entity’s feelings and/or internal wellbeing that is limited to the direct consequences of the act in question.

“IS IT ALL RIGHT OR NOT ALL RIGH T [HIT A PET ANIMAL]? No. AND WHY NO? Because **then it actually makes the animal sad.**” (#54)

“Yeah cause if something does happen [during the experiment], then **they can’t necessarily rebuild Momo exactly if all the software everything was erased and deleted.**” (#75)

1.5.1.1.3.2 Adaptation

An appeal based on concern for the entity’s feelings and/or internal wellbeing that generalizes beyond the direct consequences of the act in question. This form of “psychological adaptation” is distinct from “conventional adaptation,” in that it emphasizes adaptive changes in an individual in such a way that the individual becomes worse off overall.

“No, it’s not all right [to hit a pet animal] because it’s even more morally wrong than hitting the robot because they would understand that it was an act of human aggression and it would also affect how they react to other humans and **become a new behavior.... They would become more wary of human. They would almost have more aggression towards people when they come close.**” (#49)

“Cause you know, they can, you see stories about **the dogs that are like afraid of people, or have like a hard time interacting with other animals because of what happened to them.**” (#120)

1.5.1.1.4 Material

An appeal based on concern for the entity’s material welfare, including references to the entity as having material needs. May also include statements of lost/wasted time/effort. Statements regarding the entity’s material or monetary value to human, including changing/upgrading/improving the entity for the benefit of human usage are coded under “human-centered material welfare.”

1.5.1.2 Systemic welfare

An appeal based on concern for the welfare of a group or system, including social systems/societies, ecological systems, and/or networks in which the entity is a member. This code does not pertain to individual technological or biological systems. Statements regarding harm to a network of entities that are only concerned with the effect of the harm on humans should be coded under “human-centered” appropriate categories.

“WHY DO YOU THINK PEOPLE SHOULDN’T OWN AND SELL [PET ANIMALS]? Mmm cause I don’t know they’re like a part in the environment. They should, they shouldn’t take them out of their natural environment cause **they might go extinct.**” (#36)

1.5.1.3 Generational welfare

An appeal based on concern for the welfare of the future generations of the entity (NOT past generations).

1.5.2 Partial affirmation

1.5.2.1 Individual welfare

An appeal based on partial concern for the entity’s wellbeing as an individual.

1.5.2.1.1 Unelaborated

An appeal based on partial general concern of welfare for the entity that is otherwise unelaborated, often in the form of references to the potential for partial harm.

1.5.2.1.2 Physical

An appeal based on partial concern for the welfare of the entity’s physical body, including physical damage/injury and/or destruction/death.

“IS IT ALL RIGHT OR NOT ALL RIGHT TO [HIT A PET ANIMAL]? I think it depends on how much force you did it with. If it was just enough **not to really hurt the animal very much**, but enough to kind of tell them that if they did something wrong, to kind of let them know that what they did is wrong, and will have negative effects I think that is ok.” (#24)

1.5.2.1.3 Psychological

An appeal based on partial concern for the entity’s feelings and/or internal wellbeing, including references to partially hurt, unpleasant feelings, and/or partially negative, damaged internal states.

1.5.2.1.3.1 Direct

An appeal based on partial concern for the entity’s feelings and/or internal wellbeing that is limited to the direct consequences of the act in question.

1.5.2.1.3.2 Adaptation

An appeal based on partial concern for the entity's feelings and/or internal wellbeing that generalizes beyond the direct consequences of the act in question. This form of "psychological adaptation" is distinct from "conventional adaptation," in that it emphasizes partial adaptive changes in an individual in such a way that the individual becomes somewhat worse off overall.

1.5.2.1.4 Material

An appeal based on partial concern for the entity's material welfare, including references to the entity as partially having material needs. May also include statements of partially lost/wasted time/effort. Statements regarding the entity's material or monetary value to human, including changing/upgrading/improving the entity for the benefit of human usage are coded under "human-centered material welfare."

1.5.2.2 Systemic welfare

An appeal based on partial concern for the welfare of a group or system, including social systems/societies, ecological systems, and/or networks in which the entity is a member. This code does not pertain to individual technological or biological systems. Statements regarding harm to a network of entities that are only concerned with the effect of the harm on humans should be coded under "human-centered" appropriate categories.

1.5.2.3 Generational welfare

An appeal based on partial concern for the welfare of the future generations of the entity (NOT past generations).

1.5.3. Negation**1.5.3.1 Individual welfare**

An appeal based on the lack of concern for the entity's wellbeing as an individual.

1.5.3.1.1 Unelaborated

An appeal based on the lack of general concern of welfare for the entity that is otherwise unelaborated, often in the form of references to the lack of potential for harm. May include considerations of harm as a non-issue, not possible, or not a consideration in this case.

"So like a phone, **you're not necessarily mistreating a phone....**" (#26)

1.5.3.1.2 Physical

An appeal based on the lack of concern for the welfare of the entity's physical body, including physical damage/injury or destruction/death. May include considerations of harm as a non-issue, not possible, or not a consideration in this case.

“WHY IS THAT [ALL RIGHT TO EXPERIMENT ON PHONES]? Cause, umm, it’s just a phone. MHMM. It’s like it’s just, I mean if you, it’s like the robot I guess, if you, **if something happens to it, you can just fix it**, I guess.” (#118)

“But I think if it, if they needed to [experiment on Momo], they could, because since Momo’s part of technology, **you couldn’t necessarily like physically harm it**. You could damage the phone, but you couldn’t like emotionally and like truly physically harm Momo because Momo is part of the phone. MHMM. It’s not an actual like, being.” (#120)

1.5.3.1.3 Psychological

An appeal based on the lack of concern for the entity’s feelings or internal wellbeing, including references to the lack of hurt, unpleasant feelings, or negative, damage internal states. May include considerations of harm as a non-issue, not possible, or not a consideration in this case.

1.5.3.1.3.1 Direct

An appeal based on the lack of concern for the entity’s feelings or internal wellbeing that is limited to the direct consequences of the act in question.

“But I think if it, if they needed to [experiment on Momo], they could, because since Momo’s part of technology, you couldn’t necessarily like physically harm it. You could damage the phone, but **you couldn’t like emotionally** and like truly physically **harm Momo** because Momo is part of the phone. MHMM. It’s not an actual like, being.” (#120)

1.5.3.1.3.2 Adaptation

An appeal based on the lack of concern for the entity’s feelings or internal wellbeing that generalizes beyond the direct consequences of the act in question. This form of “psychological adaptation” is distinct from “conventional adaptation,” in that it emphasizes the lack of concern for adaptive changes in an individual in such a way that the individual becomes worse off overall.

1.5.3.1.4 Material

An appeal based on the lack of concern for the entity’s material welfare, including references to the entity as NOT having material needs. May include considerations of harm as a non-issue, not possible, or not a consideration in this case. Statements regarding the entity’s material or monetary value to human, including changing/upgrading/improving the entity for the benefit of human usage are coded under “human-centered material welfare.”

“Well, **robots don’t really need money.**” (#57)

1.5.3.2 Systemic welfare

An appeal based on the lack of concern for the welfare of a group or system, including social systems/societies, ecological systems, and/or networks in which the entity is a member. May include considerations of harm as a non-issue, not possible, or not a consideration in this case. This code does not pertain to individual technological or biological systems. Statements regarding the lack of harm to a network of entities that are only concerned with the effect of the harm on humans should be coded under “human-centered” appropriate categories.

“SO IS IT ALL RIGHT OR NOT ALL RIGHT TO EXPERIMENT ON A BASIC PHONE? Yeah, because it can’t say yes or no, and it doesn’t really have feelings, and **you can hurt the phone, but it’s not really gonna affect other phones, because there isn’t like program. It’s not wired to other phones.** It’s just a basic phone.” (#120)

1.5.3.3 Generational welfare

An appeal based on the lack of concern for the welfare of the future generations of the entity (NOT past generations). May include considerations of harm as a non-issue, not possible, or not a consideration in this case.

1.6 Justice

Refers to concerns of fairness, equality and respect regarding the entity.

1.6.1 Affirmation

1.6.1.1 Fairness

An appeal to just and/or fair treatment and equity.

“**It’s not necessarily the robot’s fault.** It’s not trying to not function or not work, so I don’t think that it’s ok to just hit it. I mean **it’s not fair.**” (#28)

“SO DO YOU THINK IT WAS ALL RIGHT OR NOT ALL RIGHT TO HIT MOMO WITH A BOOK? Um no, cause **I don’t think it’s fair because he had trouble hearing.**” (#62)

1.6.1.2 Ownership protection

An appeal based on concern for the protection of the entity from being owned, including (a) that the entity cannot be owned/bought/sold, and (b) equivalence to a slave state (or slavery) that might be of a generalized form denouncing such conditions.

“WHY IS IT NOT ALL RIGHT TO SELL MOMO? I just told you the reason. He senses, he...he knows, **he’s not a slave.**” (#41)

1.6.1.3 Discrimination protection

An appeal based on concern for the protection of the entity from discrimination.

1.6.1.4 Respect

An appeal that the entity warrants and/or deserves respect.

“WHY IS IT OK TO [OWN AND SELL ROBOTS LIKE MOMO] IF YOU TREAT THEM WELL? If they’re treated well and **with respect**, it would be no different than having a well treated pet.” (#49)

1.6.2 Partial affirmation

1.6.2.1 Fairness

An appeal to partially deserving just and/or fair treatment and equity.

1.6.2.2 Ownership protection

An appeal based on partial concern for the protection of the entity from being owned, such that the entity in certain partial terms cannot be owned/bought/sold.

“DO YOU THINK, UH, PET ANIMALS HAVE A RIGHT TO NOT BE OWNED AND SOLD?” I think, kind of. I think, um, **once you already like own them, it’s kind of your responsibility to keep them. There is not really a reason to sell them once you’ve made that commitment that they’re yours. MHMM, MHMM.** And then, I think **after a certain point you don’t own them, they become more of a friend or a part of the family** because they’re there all the time, and you can kind of connect with them if they’re there with you all the time.” (#120)

1.6.2.3 Discrimination protection

An appeal based on partial concern for the protection of the entity from discrimination.

1.6.2.4 Respect

An appeal that the entity partially warrants and/or deserves respect.

1.6.3 Negation

1.6.3.1 Fairness

An appeal to NOT deserving just or fair treatment or equity.

1.6.3.2 Ownership protection

An appeal based on the lack of concern for the protection of the entity from being owned, including (a) that the entity is and/or can be owned/bought/sold, and (b) equivalence to a slave state (or slavery) that might be of a generalized form allowing such conditions. May include statements regarding the permissibility of ownership of the entity.

“DO PHONES HAVE A RIGHT TO NOT BE OWNED AND SOLD? Well if they’re being produced by a manufacturer, **they’re always being owned**,

they're always created by an owner.... **There is never a life for a phone without an owner.**" (#49)

"I think it's ok [to own and sell Momo]. OK WHY? Umm, cause, um, I don't know, it's just like a robot. I think it's just like another object. MHMM. Like, how you own a phone, **you can own a robot.**" (#118)

1.6.3.3 Discrimination protection

An appeal based on the lack of concern for the protection of the entity from discrimination. May include statements regarding the permissibility of discrimination of the entity.

1.6.3.4 Respect

An appeal that the entity does NOT warrant or deserve respect.

1.7 Rights

Refers to rights concerning the entity.

1.7.1 Affirmation

"IS IT ALL RIGHT OR NOT ALL RIGHT TO EXPERIMENT ON A PET ANIMAL? Umm not all right really. I mean I think they do it, but I don't think it's right just because even though there's rights to animals, they don't really know that they have rights, **but there are rights like animal abuse I think.**" (#26)

1.7.2 Partial affirmation

"I think [Momo] **does have a right** to not be experimented on.... Yeah I guess, I mean **when compared with like a human, Momo I guess, Momo would be have less of a right.**" (#116)

1.7.3 Negation

"IS IT ALL RIGHT OR NOT ALL RIGHT TO OWN AND SELL A BASIC PHONE? Umm I think it is all right because like **phones don't really have rights** kind of thing." (#26)

1.8 Freedom & autonomy

Refers to appeals of whether the entity is/should/need to be free, to live free, including considerations of whether the entity has/should have free will, the freedom to make choices, decisions, and/or give/deny consent with adequate information.

1.8.1 Affirmation

"So like [pet animals] should have a right **if they didn't want to be, like if they were to say I don't want to be, they wouldn't have to be sold and bought,** but I

feel like a lot of them actually love to be in a human household where they're loved and cared for." (#26)

1.8.2 Partial affirmation

"And [pet animal] kind of has its own free will because it can wander around your house and get into your stuff, but you also own it because it's yours. You buy it food, you buy food, and you take care of it, and it stays within your house...." (#120)

1.8.3 Negation

"So like if [Momo] has things like when he got hit he seemed like he got kind of mad, he has feelings so he's in some aspect alive, but not like fully alive, because maybe **he doesn't like act on his own** but like he has feelings when things happen to him." (#26)

1.9 Telos

Refers to appeals of whether the entity has a purpose, design, and/or endpoint, including references to whether the entity is meant or not meant for something.

1.9.1 Affirmation

"So if somebody like destroyed [Momo]...and they get destroyed like they won't have a life anymore, cause you can't rebuild the same thing. OH YOU CAN'T REBUILD THE SAME ROBOT? **It would be something different, even if you built it the same way.**" (#85)

"WHY WOULD THAT NOT BE ALL RIGHT IF [MOMO WAS DAMAGED]?
Hmm, because **it would be obstructing the purpose of Momo.**" (#103)

"[A basic phone] hasn't, **it has the one purpose** and without being bought or sold, there is no purpose for it." (#120)

1.9.2 Partial affirmation

1.9.2 Negation

1.10 Virtue

Refers to appeals of whether the entity is good and/or meritorious.

1.10.1 Affirmation

"Momo can be sad and, sad too but he can't lash out at people. MHMM. He's too, he's too good to know, **he's too good to do that.**" (#71)

1.10.2 Partial affirmation

1.10.3 Negation

2. Human-centered

Refers to the characteristics and qualities of humans, and appeals and concerns that are centered on humans.

2.1 Essence

Refers to the essential qualities of humans, including statements regarding the human's artificial, animal, and person/human qualities, and statements regarding the human's biological status, physical embodiment, functionality, and personality.

2.1.1 Affirmation

2.1.1.1 Artificial

An appeal to the essential artificial qualities of the human, including statements of direct, isomorphic, and transmorphic correspondence between the human and an artifact (including computational technological artifacts and non-computational artifacts), as well as statements regarding the human as having artificial emotion, and as being a programmed, manufactured, and marketed entity.

2.1.1.1.1 Direct

An appeal based on a direct correspondence between the human and an artifact, including references to being a robot, machine, computer, object, property, and/or having parts thereof (e.g., sensors, circuits).

2.1.1.1.2 Isomorphic

An appeal based on an analogical or conditional (if-then) correspondence between the human and an artifact.

2.1.1.1.3 Transmorphic

An appeal based on the establishment of similarities and differences between the human and an artifact, wherein an inequivalent correspondence (or difference) is overridden by similarities, and/or recognized differences do not void the mapping of similar considerations.

2.1.1.1.4 Artificial Emotion

An appeal that the human exhibits and/or has emotions/feelings that are achieved through artificial means (e.g., implants). May include statements that the human's emotions/feelings are not real, may or may not be real, but they still seem to be, or are in their own way emotions/feelings.

2.1.1.1.5 Programmed

An appeal that the human is and/or able to be programmed to exhibit behaviors, emotions, thoughts, etc.

2.1.1.1.6 Manufactured

An appeal that the human is manufactured, built, and/or created by other humans/sentient beings.

2.1.1.1.7 Marketed

An appeal that the human is a marketed product/good, and/or routinely offered for sale.

2.1.1.2 Animal

An appeal to the essential animal qualities of the human, including statements of direct, isomorphic, and transmorphic correspondence between the human and an animal.

2.1.1.2.1 Direct

An appeal based on a direct correspondence between the human and an animal.

“**We’re all animals, like humans are animals as well.**” (#121)

2.1.1.2.2 Isomorphic

An appeal based on an analogical or conditional (if-then) correspondence between the human and an animal.

“CAN YOU EXPLAIN A LITTLE BIT MORE WHAT YOU MEAN BY IT’S CRUEL TO [OWN AND SELL A HUMAN]? Um well because um **it’s like kind of like trying to sell my dog** which I would not like because I love him.” (#3)

2.1.1.2.3 Transmorphic

An appeal based on the establishment of similarities and differences between the human and an animal, wherein an inequivalent correspondence (or difference) is overridden by similarities, and/or recognized differences do not void the mapping of similar considerations.

2.1.1.3 Personhood

An appeal to the person’s essential human qualities, including statements of direct, isomorphic, and transmorphic correspondence between the person and a human.

2.1.1.3.1 Direct

An appeal based on a direct correspondence between the person and a human.

“WHY IS IT NOT ALL RIGHT TO HIT A GIRL?” Mmm, **because they are human** and they are living beings.” (#3)

2.1.1.3.2 Isomorphic

An appeal based on an analogical or conditional (if-then) correspondence between the person and a human.

“Well, if [the humans] don’t want [to be experimented on], don’t do it. **Like if I don’t want a water gun shot right in my face, and my sister does it, that’s just not fair to me.**” (#57)

2.1.1.3.3 Transmorphic

An appeal based on the establishment of similarities and differences between the person and a human, wherein an inequivalent correspondence (or difference) is overridden by similarities, and/or recognized differences do not void the mapping of similar considerations.

2.1.1.4 Biological

An appeal to the essential biological qualities of the human, including statements regarding the human as being biologically alive, having biological material (e.g., brain, flesh, blood), and/or having biological processes (e.g., metabolism, breathing, aging). May also include statements regarding the human’s biologically supported functions (e.g., seeing with eyes, hearing with ears).

“WHY IS IT NOT ALL RIGHT TO HIT A GIRL?” Mmm, because they are human and **they are living beings....**” (#3)

2.1.1.5 Physical embodiment

An appeal based on the human’s physical body, and/or on features and abilities tied to the human’s physicality (e.g., ability to move, shake hands, and have physical sensations).

“WHY DO PEOPLE HAVE THIS RIGHT [TO NOT BE EXPERIMENTED ON]? **Because it’s their bodies.** It’s their decision whether you get poked and prodded at, or experimented on.” (#120)

2.1.1.6 Functionality

An appeal based on the human’s functional aspects and/or ability to perform a function (e.g., climb, teach).

“I mean everyone has like their own personal right, they are not like possessions because **by themselves they are functional** and can be able to like have feelings and do stuff on their own....” (#107)

2.1.1.7 Personality

An appeal that the human has (or has the capacity to have) a personality, personality traits, and/or personality characteristics.

2.1.2 Partial affirmation

2.1.2.1 Artificial

An appeal to the partial artificial qualities of the human, including statements of direct, isomorphic, and transmorphic partial correspondence between the human

and an artifact (including computational technological artifacts and non-computational artifacts), as well as statements regarding the human as partially having artificial emotion, and as partially being a programmed, manufactured, and marketed entity.

2.1.2.1.1 Direct

An appeal based on a direct partial correspondence between the human and an artifact, including references to partially being a robot, machine, computer, object, property, and/or having parts thereof (e.g., sensors, circuits).

2.1.2.1.2 Isomorphic

An appeal based on an analogical or conditional (if-then) partial correspondence between the human and an artifact.

2.1.2.1.3 Transmorphic

An appeal based on the establishment of similarities and differences between the human and an artifact, wherein an inequivalent correspondence (or difference) is partially overridden by similarities, and/or recognized differences do not completely void the mapping of similar considerations, but change the degree/extent of such considerations.

2.1.2.1.4 Artificial Emotion

An appeal that the human exhibits and/or has partial emotions/feelings that are achieved through artificial means (e.g., implants). May include statements that the human's emotions/feelings are not real, may or may not be real, but they still partially seem to be, or partially are in their own way emotions/feelings.

2.1.2.1.5 Programmed

An appeal that the human is and/or able to be partially programmed to exhibit behaviors, emotions, thoughts, etc.

2.1.2.1.6 Manufactured

An appeal that the human is partially manufactured, built, and/or created by other humans/sentient beings.

2.1.2.1.7 Marketed

An appeal that the human is partially a marketed product/good, and/or semi-routinely offered for sale.

2.1.2.2 Animal

An appeal to the partial animal qualities of the human, including statements of direct, isomorphic, and transmorphic partial correspondence between the human and an animal.

2.1.2.2.1 Direct

An appeal based on a direct partial correspondence between the human and an animal.

2.1.2.2.2 Isomorphic

An appeal based on an analogical or conditional (if-then) partial correspondence between the human and an animal.

2.1.2.2.3 Transmorphic

An appeal based on the establishment of similarities and differences between the human and an animal, wherein an inequivalent correspondence (or difference) is partially overridden by similarities, and/or recognized differences do not completely void the mapping of similar considerations, but change the degree/extent of such considerations.

2.1.2.3 Personhood

An appeal to the person's partial essential human qualities, including statements of direct, isomorphic, and transmorphic partial correspondence between the person and a human.

2.1.2.3.1 Direct

An appeal based on a direct partial correspondence between the person and a human.

2.1.2.3.2 Isomorphic

An appeal based on an analogical or conditional (if-then) partial correspondence between the person and a human.

2.1.2.3.3 Transmorphic

An appeal based on the establishment of similarities and differences between the person and a human, wherein an inequivalent correspondence (or difference) is partially overridden by similarities, and/or recognized differences do not completely void the mapping of similar considerations, but change the degree/extent of such considerations.

2.1.2.4 Biological

An appeal to the partial biological qualities of the human, including statements regarding the human as partially being biologically alive, having biological material (e.g., brain, flesh, blood), and/or having biological processes (e.g., metabolism, breathing, aging). May also include statements regarding the human's partially biologically supported functions (e.g., seeing with eyes, hearing with ears).

2.1.2.5 Physical embodiment

An appeal based on the human's partial physical body, and/or on features and abilities tied to the human's partial physicality (e.g., a human is not a purely physical being because the human's soul is not physical).

2.1.2.6 Functionality

An appeal based on the human's partial (incomplete) ability to perform a function (e.g., climb, teach). May also include statements referring to the entity not functioning as well/fully/completely.

2.1.2.7 Personality

An appeal that the human partially has (or partially has the capacity to have) a personality, personality traits, and/or personality characteristics.

2.1.3 Negation

2.1.3.1 Artificial

An appeal to the lack of essential artificial qualities of the human, including statements based on the lack of direct, isomorphic, or transmorphic correspondence between the human and an artifact (including computational technological artifacts and non-computational artifacts), as well as statements regarding the human as NOT having artificial emotion, and as NOT being a programmed, manufactured, and marketed entity.

2.1.3.1.1 Direct

An appeal based on the lack of a direct correspondence between the human and an artifact, including references to NOT being a robot, machine, computer, object, property, or having parts thereof (e.g., sensors, circuits).

2.1.3.1.2 Isomorphic

An appeal based on the lack of an analogical or conditional (if-then) correspondence between the human and an artifact.

“**Humans are not just like something like a table or an object** that you can sell....” (#118)

2.1.3.1.3 Transmorphic

An appeal based on the establishment of similarities and differences between the human and an artifact, wherein an inequivalent correspondence (or difference) is NOT overridden by similarities, and/or recognized differences void the mapping of similar considerations.

2.1.3.1.4 Artificial Emotion

An appeal to the human's incapacity for or lack of emotions/feelings that are achieved through artificial means (e.g., implants).

2.1.3.1.5 Programmed

An appeal that the human is NOT or NOT able to be programmed to exhibit behaviors, emotions, thoughts, etc.

“IS IT ALL RIGHT OR NOT ALL RIGHT TO [HIT A BOY]? Not, because the boy can actually feel and **can’t program what it says.**” (#11)

2.1.3.1.6 Manufactured

An appeal that the human is NOT manufactured, built, or created by other humans/sentient beings.

2.1.3.1.7 Marketed

An appeal that the human is NOT a marketed product/good, or routinely offered for sale.

“**But if you like buy [a human] off of a market of some, of some sort...no not all right.**” (#104)

2.1.3.2 Animal

An appeal to the lack of essential animal qualities of the human, including statements of the lack of direct, isomorphic, and transmorphic correspondence between the human and an animal.

2.1.3.2.1 Direct

An appeal based on the lack of a direct correspondence between the human and an animal.

2.1.3.2.2 Isomorphic

An appeal based on the lack of an analogical or conditional (if-then) correspondence between the human and an animal.

“**Like animals can’t necessarily say ‘I don’t want to be experimented on,’ but people can.**” (#75)

2.1.3.2.3 Transmorphic

An appeal based on the establishment of similarities and differences between the human and an animal, wherein an inequivalent correspondence (or difference) is NOT overridden by similarities, or recognized differences void the mapping of similar considerations.

2.1.3.3 Personhood

An appeal to the person’s lack of essential human qualities, including statements of the lack of direct, isomorphic, and transmorphic correspondence between the person and a human.

2.1.3.3.1 Direct

An appeal based on the lack of a direct correspondence between the person and a human.

2.1.3.3.2 Isomorphic

An appeal based on the lack of an analogical or conditional (if-then) correspondence between the person and a human.

2.1.3.3.3 Transmorphic

An appeal based on the establishment of similarities and differences between the person and a human, wherein an inequivalent correspondence (or difference) is NOT overridden by similarities, and/or recognized differences void the mapping of similar considerations.

2.1.3.4 Biological

An appeal to the lack of essential biological qualities of the human, including statements regarding the human as NOT being biologically alive, having biological material (e.g., brain, flesh, blood), and/or having biological processes (e.g., metabolism, breathing, aging). May also include statements regarding the human's lack of biologically supported functions (e.g., seeing with eyes, hearing with ears).

2.1.3.5 Physical embodiment

An appeal based on the human's lack of physical body, or on the lack of features and abilities tied to the human's physicality (e.g., cannot move, cannot shake hands, cannot have physical sensations).

2.1.3.6 Functionality

An appeal based on the human's lack of functional aspects or inability to perform a function (e.g., climb, teach).

2.1.3.7 Personality

An appeal based on the human's lack of (or incapacity to have) a personality, personality traits, or personality characteristics.

2.2 Mental

Refers to the cognition, emotion, intention & desire, and interest & predilection of the human.

2.2.1 Affirmation

2.2.1.1 Cognition

An appeal to the existence of (or capacity to have) intelligence and/or mental abilities and processes (e.g., perception, memory, decision making, awareness, consciousness, etc.).

“IS IT ALL RIGHT OR NOT ALL RIGHT TO OWN AND SELL A HUMAN?
No. AND WHY NOT? It's because they have, **humans have the ability to learn** and function and they don't always need help, like they do need help sometimes and they can ask for it, they have the ability to ask for help.” (#49)

2.2.1.2 Emotion

An appeal to the existence of (or capacity to have) emotions/feelings.

“IS IT ALL RIGHT OR NOT ALL RIGHT [TO HIT ANOTHER BOY]? It’s not all right. OK AND WHY? Because **he has feelings** too, and **he can get angry.**” (#12)

2.2.1.3 Intention & desire

An appeal to the existence of (or capacity to have) intentions, desires, goals, and/or expectations.

“Yeah, you shouldn’t like, I don’t think that you should just experiment on someone, **especially if they don’t want to be experimented on.**” (#118)

2.2.1.4 Interest & predilection

An appeal to the existence of (or capacity to have) personal preferences, predilections, and/or likes and dislikes.

“If there are like multiple [Momos] and people were buying and selling them for profit, **people would enjoy it.** As long as they weren’t always mistreating Momo, I feel like he would enjoy it too.” (#26)

2.2.2 Partial affirmation

2.2.2.1 Cognition

An appeal to the partial existence of (or capacity to partially have) intelligence and/or mental abilities and processes (e.g., perception, memory, decision making, awareness, consciousness, etc.). May also include statements regarding the human as having a lesser level of intelligence and/or mental abilities and processes.

2.2.2.2 Emotion

An appeal to the partial existence of (or partial capacity to have) emotions/feelings. May also include statement regarding the human as having a lesser level/degree of emotions/feelings.

2.2.2.3 Intention & desire

An appeal to the partial existence of (or capacity to partially have) intentions, desires, goals, and/or expectations. May also include statements regarding the human has having a lesser level of intentions, desires, goals and/or expectation.

2.2.2.4 Interest & predilection

An appeal to the partial existence of (or capacity to partially have) personal preferences, predilections, and/or likes and dislikes. May also include statements regarding the entity has having a lesser level of interests, predilections, and/or likes and dislikes.

2.2.3 Negation

2.2.3.1 Cognition

An appeal to the lack of (or incapacity to have) intelligence or mental abilities and processes (e.g., perception, memory, decision making, awareness, consciousness, etc.).

2.2.3.2 Emotion

An appeal to the lack of (or incapacity to have) emotions/feelings.

2.2.3.3 Intention & desire

An appeal to the lack of (or incapacity to have) intentions, desires, goals, or expectations.

2.2.3.4 Interest & predilection

An appeal to the lack of (or incapacity to have) personal preferences, predilections, or likes and dislikes.

2.3 Social

Refers to social interactions and relations concerning the human that includes communication, affective relation, companionship, and trust.

2.3.1 Affirmation

2.3.1.1 Unelaborated

An appeal to the human engaging in (or having the capacity for) social interaction and/or social relation that is otherwise unelaborated.

“I don’t think that’s right [to own and sell a human]...because it’s kind of cruel, because it might even be your sister that you’re giving away which would be really cruel to do **because she’s part of your family.**” (#3)

2.3.1.2 Communication

An appeal to the human’s capacity for conversation, talking, and/or other aspects of communication, including statements regarding the human’s capacity to express/communicate wants/needs/consents, and/or speak up for their rights. MUST include a social aspect of talking/expressing information.

“Like [humans] do need help sometimes and **they can ask for it, they have the ability to ask for help.**” (#49)

“Like **animals can’t necessarily say ‘I don’t want to be experimented on,’ but people can.**” (#75)

2.3.1.3 Affective relation

An appeal to the human’s need and/or capacity for affective relation.

2.3.1.3.1 General

An appeal to the human's general need and/or capacity to form/bring about emotional attachment/bond with others. May also include references to emotional attachment/bond that are neither positive nor negative in nature.

“No [it's not all right to own and sell children], because **they need to form emotional bonds with their families** and who knows what they're gonna be up to if they are being sold.” (#49)

2.3.1.3.2 Positive

An appeal to the human being (or having the capacity to be) nice, thoughtful, caring, loving, emotionally supportive, etc., and/or able to form positive emotional attachment to others.

“IF YOU DON'T TREAT [PET ANIMALS] BADLY, WHY IS IT ALL RIGHT TO OWN THEM? Because, then **you're loving them** and you're treating them like it's a human and a child.” (#41)

2.3.1.3.3 Negative

An appeal to the human being (or having the capacity to be) mean, thoughtless, critical, judgmental, emotionally abusive, etc., and/or able to form negative emotional attachment to others.

“Because Momo, I guess, I mean Momo would listen to you and um not really, Momo wouldn't really **judge you and make assumptions about you in ways that humans would.**” (#28)

2.3.1.4 Companionship

An appeal to the human's needs, desires, interests, and/or capacity for companionship, to spend time and form personal associations with others. Statements regarding the human's desires and/or interests in companionship are double coded under “intention & desire” and/or “interest & predilection.”

“OK AND WHY DO YOU THINK [MOMO] CAN BE YOUR FRIEND? Like cause, **people can friend some weird things** I guess.” (#7)

“UM SO DO YOU THINK MOMO CAN BE YOUR FRIEND? Yeah cause you know he can, he can talk to you, and **you can hangout with him.**” (#115)

2.3.1.5 Trust

An appeal to the human being (or having the capacity to be) trusting and/or trustworthy.

“We all, **we form a community by trusting each other.**” (#7)

2.3.2 Partial affirmation

2.3.2.1 Unelaborated

An appeal to the human partially engaging in (or partially having the capacity for) social interaction and/or social relation that is otherwise unelaborated.

2.3.2.2 Communication

An appeal to the human's partial capacity for conversation, talking, and/or other aspects of communication, including statements regarding the human's partial capacity to express/communicate wants/needs/consents, and/or speak up for their rights. MUST include a social aspect of talking/expressing information.

2.3.2.3 Affective relation

An appeal to the human's partial need and/or capacity for affective relation.

2.3.2.3.1 General

An appeal to the human's general partial need and/or partial capacity to form/bring about emotional attachment/bond with others. May also include references to emotional attachment/bond that are neither positive nor negative in nature.

2.3.2.3.2 Positive

An appeal to the human partially being (or partially having the capacity to be) nice, thoughtful, caring, loving, emotionally supportive, etc., and/or partially able to form positive emotional attachment to others.

2.3.2.3.3 Negative

An appeal to the human partially being (or partially having the capacity to be) mean, thoughtless, critical, judgmental, emotionally abusive, etc., and/or partially able to form negative emotional attachment to others.

2.3.2.4 Companionship

An appeal to the human's partial needs, desires, interests, and/or partial capacity for companionship, to spend time and form personal associations with others. Statements regarding the human's desires and/or interests in companionship are double coded under "intention & desire" and/or "interest & predilection."

2.3.2.5 Trust

An appeal to the human partially being (or partially having the capacity to be) trusting and/or trustworthy.

2.3.3 Negation

2.3.3.1 Unelaborated

An appeal to the human NOT engaging in (or NOT having the capacity for) social interaction and/or social relation that is otherwise unelaborated.

2.3.3.2 Communication

An appeal to the human's incapacity for conversation, talking, or other aspects of communication, including statements regarding the human's incapacity to express/communicate wants/needs/consents, or speak up for their rights. **MUST** include a social aspect of talking/expressing information.

2.3.3.3 Affective relation

An appeal to the human's lack of need or incapacity for affective relation.

2.3.3.3.1 General

An appeal to the human's lack of general need or incapacity to form/bring about emotional attachment/bond with others. May also include references to emotional attachment/bond that are neither positive nor negative in nature.

2.3.3.3.2 Positive

An appeal to the human NOT being (or incapable of being) nice, thoughtful, caring, loving, emotionally supportive, etc., or NOT able to form positive emotional attachment to others.

2.3.3.3.3 Negative

An appeal to the human NOT being (or incapable of being) mean, thoughtless, critical, judgmental, emotionally abusive, etc., or NOT able to form negative emotional attachment to others.

2.3.3.4 Companionship

An appeal to the human's lack companionship, or lack of needs, desires, interests, or incapacity for companionship, to spend time and form personal associations with others. Statements regarding the human's lack of desires or interests in companionship are double coded under "intention & desire" and/or "interest & predilection."

"It's all right to own and sell [Momo]. OK AND WHY IS THAT? Because you'd be like, you could be giving a friend to everyone you know...so you know people who...yeah, **people who don't have any friends**, you could be giving one to them." (#22)

"DO YOU THINK MOMO CAN BE YOUR FRIEND? I think, I think he could I guess **if you have no friends**." (#96)

2.3.3.5 Trust

An appeal to the human NOT being (or incapable of being) trusting or trustworthy.

"Umm..., they like sometimes won't, **no one could keep a secret**." (#3)

2.4 Conventional

Refers to conventions that prescribe or prohibit behaviors exhibited by the human in social interaction, conventions that apply to and/or are followed by the human (i.e., what one *does* and *does not do*).

2.4.1 Affirmation

2.4.1.1 Unelaborated

An appeal to the existence of unelaborated conventions.

2.4.1.2 Authority

An appeal to needing to adhere to authority, whether that authority is in the form of rules, social contracts, laws, and/or authority figures. May include statements of avoiding punishment.

“It’s a law that you can’t abuse your animal....” (#26)

2.4.1.3 Custom

An appeal to customs based on the frequency of occurrence (e.g., “they do it all the time”) and/or social standards (e.g., “that’s what we do in this country”).

“Like a phone is just something that a lot of people have and use so to sell them is all right.” (#26)

2.4.1.4 Adaptation

An appeal to adaptation to social customs and/or conventions such that any initial harm is obviated with the adaptation.

“Well it depends on who we’re being owned and sold by, if it’s by humans, then yes [we have the right to not be owned and sold], but if someday we’re conquered by like aliens, and they domesticate us, then maybe by then, like a thousand years in the future, we’ll see it as being fit for us to be like domesticated animals.” (#46)

2.4.2 Partial affirmation

2.4.2.1 Unelaborated

An appeal to the partial existence of unelaborated conventions

2.4.2.2 Authority

An appeal to partially needing to adhere to authority, whether that authority is in the form rules, social contracts, laws, and/or authority figures. May include statements of partially avoiding punishment.

2.4.2.3 Custom

An appeal to partially established customs based on the frequency of occurrence (e.g., “they are starting to do it more often now”) and/or social standards (e.g., “that’s what we sort of do in this country”).

2.4.2.4 Adaptation

An appeal to partial adaptation to social customs and/or conventions such that any initial harm is lessened with the adaptation.

2.4.3 Negation

2.4.3.1 Unelaborated

An appeal to the lack of unelaborated conventions, including claims to being/feeling weird, awkward, or novel.

“DO YOU THINK MOMO CAN BE YOUR FRIEND? Maybe, I don’t think he can like go with people.... **It might be a little weird carrying him around.**” (#36)

2.4.3.2 Authority

An appeal to NOT needing to adhere to authority, whether that authority is in the form of rules, social contracts, laws, or authority figures. May include statements regarding the lack of punishment and the permissibility of an act.

“**There’s not written rule of rights to a phone, like you can’t break your phone purposely** or something like that.” (#26)

2.4.3.3 Custom

An appeal to the lack of customs or social standards (e.g., “we don’t do that”).

“ANY OTHER REASONG WHY IT WASN’T ALL RIGHT TO HIT MOMO? I don’t know, same thing **we learned in kindergarten, that hitting is wrong.**” (#16)

“**You don’t just hit.**” (#44)

2.4.3.4 Adaptation

An appeal to the lack of adaptation to social customs or conventions.

“Because humans are like, **we haven’t been bred to be owned overtime**, and we’re able to have intelligent thoughts, and therefore we should be able to make our own choices.” (#11)

2.5 Welfare

Refers to the human’s wellbeing, including concerns regarding the individual human’s unelaborated general welfare, physical welfare, psychological welfare, and material welfare.

Also includes concerns regarding the welfare of the larger system in which the human is a member of, and the welfare of the human's future generations.

2.5.1 Affirmation

2.5.1.1 Individual welfare

An appeal based on concern for the human's wellbeing as an individual.

2.5.1.1.1 Unelaborated

An appeal based on general concern of welfare for the human that is otherwise unelaborated, often in the form of references to the potential for harm.

“Umm it's not all right [to experiment on a human]. It would be all right if the person gave their consent like if it was voluntary, but if the person didn't agree and you don't know **if there would be harmful effects** then I don't think it is all right.” (#28)

2.5.1.1.2 Physical

An appeal based on concern for the welfare of the human's physical body, including physical damage/injury and/or destruction/death.

“...it's a lot worse to put down another human being or hit another human being **when they can feel pain** and like sad.” (#26)

“It's not all right to like [hit] someone because **it can hurt them physically** and emotionally.” (#120)

2.5.1.1.3 Psychological

An appeal based on concern for the human's feelings and/or internal wellbeing, including references to hurt, unpleasant feelings, and/or negative, damaged internal states.

2.5.1.1.3.1 Direct

An appeal based on concern for the human's feelings and/or internal wellbeing that is limited to the direct consequences of the act in question.

“Umm, even though a phone isn't a living being, I think it wouldn't be all right **because they [the people hitting the phone] just get more upset** because they hit it...” (#3)

“It'll be ok to hit the phone **to let your energy out**, but like, cause it's not necessarily making the phone feel bad, like with feelings, so maybe if **you get your anger out**, but like it's not going to do anything cause no matter what, if it's programmed that way it's still going to do the same thing.” (#26)

2.5.1.1.3.2 Adaptation

An appeal based on concern for the human's feelings and/or internal wellbeing that generalizes beyond the direct consequences of the act in question. This form of "psychological adaptation" is distinct from "conventional adaptation," in that it emphasizes adaptive changes in an individual in such a way that the individual becomes worse off overall.

"I mean it does, **hitting anything does like, does get to practice hitting things so you have a tendency to hit things**, but it's better to hit something that doesn't look like anything living thing, I mean it does kind of resemble a human face a little." (#7)

2.5.1.1.4 Material

An appeal based on concern for the human's material welfare, including references to the human as having material needs. May also include statements of lost/wasted time/effort.

"And to own and sell [phones] I mean **you could make a profit off of them** and it actually helps people." (#26)

"Because **it would take someone else's time to fix [Momo] which would be rather wasteful...**" (#49)

2.5.1.2 Systemic welfare

An appeal based on concern for the welfare of humans as a group or system, including social systems/societies, ecological systems, and/or networks in which the human is a member. This code does not pertain to individual biological systems.

"OK AND WHY IS IT ALL RIGHT [TO EXPERIMENT ON PET ANIMALS]? Um I think that **in order to make development that help the greater society** in something that you believe will not have any negative effects that, I believe it is worth trying that." (#24)

"If they did this experiment on this one device and then **other people are trying to use Momo on their devices and now they can't because it was affected on the first device or cell phone, then everybody else is gonna be complaining.**" (#75)

2.5.1.3 Generational welfare

An appeal based on concern for the welfare of the human's future generations (NOT past generations).

"But I think that it, it's hard for, to stop [experimenting], because once they found that they, it stopped, it doesn't harm the robotic being, it's, they will automatically want to continue testing and see how far the branch extends, so

again they will try to find the range which means that more beings will be harmed, but, and I think in a way, **it will be safer in the future because then you will have an answer.**” (#49)

2.5.2 Partial affirmation

2.5.2.1 Individual welfare

An appeal based on partial concern for the human’s wellbeing as an individual.

2.5.2.1.1 Unelaborated

An appeal based on partial general concern of welfare for the human that is otherwise unelaborated, often in the form of references to the potential for partial harm, yet distinct from considerations of harm as a non-issue, not possible, or not a consideration in this case.

2.5.2.1.2 Physical

An appeal based on partial concern for the welfare of the human’s physical body, including physical damage/injury and/or destruction/death.

2.5.2.1.3 Psychological

An appeal based on partial concern for the human’s feelings and/or internal wellbeing, including references to partially hurt, unpleasant feelings, and/or partially negative, damaged internal states.

2.5.2.1.3.1 Direct

An appeal based on partial concern for the human’s feelings and/or internal wellbeing that is limited to the direct consequences of the act in question.

2.5.2.1.3.2 Adaptation

An appeal based on partial concern for the human’s feelings and/or internal wellbeing that generalizes beyond the direct consequences of the act in question. This form of “psychological adaptation” is distinct from “conventional adaptation,” in that it emphasizes partial adaptive changes in an individual in such a way that the individual becomes somewhat worse off overall.

2.5.2.1.4 Material

An appeal based on partial concern for the human’s material welfare, including references to the human as partially having material needs. May also include statements of partially lost/wasted time/effort.

2.5.2.2 Systemic welfare

An appeal based on partial concern for the welfare of humans as a group or system, including social systems/societies, ecological systems, and/or networks in which

the human is a member. This code does not pertain to individual biological systems.

2.5.2.3 Generational welfare

An appeal based on partial concern for the welfare of the human's future generations (NOT past generations).

2.5.3. Negation

2.5.3.1 Individual welfare

An appeal based on the lack of concern for the entity's wellbeing as an individual.

2.5.3.1.1 Unelaborated

An appeal based on the lack of general concern of welfare for the human that is otherwise unelaborated, often in the form of references to the lack of potential for harm. May include considerations of harm as a non-issue, not possible, or not a consideration in this case.

2.5.3.1.2 Physical

An appeal based on the lack of concern for the welfare of the human's physical body, including physical damage/injury or destruction/death. May include considerations of harm as a non-issue, not possible, or not a consideration in this case.

2.5.3.1.3 Psychological

An appeal based on the lack of concern for the human's feelings or internal wellbeing, including references to the lack of hurt, unpleasant feelings, or negative, damage internal states. May include considerations of harm as a non-issue, not possible, or not a consideration in this case.

2.5.3.1.3.1 Direct

An appeal based on the lack of concern for the human's feelings or internal wellbeing that is limited to the direct consequences of the act in question.

2.5.3.1.3.2 Adaptation

An appeal based on the lack of concern for the human's feelings or internal wellbeing that generalizes beyond the direct consequences of the act in question. This form of "psychological adaptation" is distinct from "conventional adaptation," in that it emphasizes the lack of concern for adaptive changes in an individual in such a way that the individual becomes worse off overall.

2.5.3.1.4 Material

An appeal based on the lack of concern for the human's material welfare, including references to the human as NOT having material needs. May include

considerations of harm as a non-issue, not possible, or not a consideration in this case.

“But it’s not like they’re selling [pet animals] to do work and they’re making a profit out of them....” (#26)

2.5.3.2 Systemic welfare

An appeal based on the lack of concern for the welfare of humans as a group or system, including social systems/societies, ecological systems, and/or networks in which the human is a member. May include considerations of harm as a non-issue, not possible, or not a consideration in this case. This code does not pertain to individual biological systems.

2.5.3.3 Generational welfare

An appeal based on the lack of concern for the welfare of the human’s future generations (NOT past generations). May include considerations of harm as a non-issue, not possible, or not a consideration in this case.

2.6 Justice

Refers to concerns of fairness, equality and respect regarding the human.

2.6.1 Affirmation

2.6.1.1 Fairness

An appeal to just and/or fair treatment and equity.

“...cause they were black, they were just not worthy of having right, which they kind of decided but they never knew, but now like everyone is the same, like they all work same jobs, like a black person can be a CEO millionaire and a white person could be maybe more poor or the other way around. Like they all have the same chances now and I feel like that should always be maintained because everyone has like the same right now and it works out.” (#26)

“That’s definitely not fair [to hit a girl] because it’s not because girls are weak, that is definitely not the reason, it’s because it’s just degrading, as a human, to be uh treated that way, that you are less value than another.” (#49)

2.6.1.2 Ownership protection

An appeal based on concern for the protection of the human from being owned, including (a) that the human cannot be owned/bought/sold, and (b) equivalence to a slave state (or slavery) that might be of a generalized form denouncing such conditions.

“IS IT ALL RIGHT OR NOT ALL RIGHT TO OWN AND SELL A HUMAN? Not all right. AND WHY NOT? Just cause they, a human being like, **in slavery** just cause they were black they were thought like to be so much worse. Like they’re the same, the only difference is skin color, and I feel like to judge someone by that is not right. So **being able to sell someone who is just the same as you because of different color skin or sell and buy just because something is not right**, cause like it feels like everyone is the same, just look different and that’s about it. **So to sell and buy another human being is like not right** just because of certain things like that.” (#26)

2.6.1.3 Discrimination protection

An appeal based on concern for the protection of the human from discrimination.

“IS IT ALL RIGHT OR NOT ALL RIGHT TO OWN AND SELL A HUMAN? Not all right. AND WHY NOT? Just cause they, a human being like, in slavery **just cause they were black they were thought like to be so much worse. Like they’re the same, the only difference is skin color, and I feel like to judge someone by that is not right.** So **being able to sell someone who is just the same as you because of different color skin or sell and buy just because something is not right**, cause like it feels like everyone is the same, just look different and that’s about it. So to sell and buy another human being is like not right just because of certain things like that.” (#26)

2.6.1.4 Respect

An appeal that the human warrants and/or deserves respect.

“It’s not ok to own somebody because it’s, **it’s disrespectful.**” (#92)

2.6.2 Partial affirmation

2.6.2.1 Fairness

An appeal to partially deserving just and/or fair treatment and equity.

2.6.2.2 Ownership protection

An appeal based on partial concern for the protection of the human from being owned, such that the human in certain partial terms cannot be owned/bought/sold.

2.6.2.3 Discrimination protection

An appeal based on partial concern for the protection of the human from discrimination.

2.6.2.4 Respect

An appeal that the human partially warrants and/or deserves respect.

2.6.3 Negation

2.6.3.1 Fairness

An appeal to NOT deserving just or fair treatment or equity.

2.6.3.2 Ownership protection

An appeal based on the lack of concern for the protection of the human from being owned, including (a) that the human is and/or can be owned/bought/sold, and (b) equivalence to a slave state (or slavery) that might be of a generalized form allowing such conditions. May include statements regarding the permissibility of ownership of the human.

2.6.3.3 Discrimination protection

An appeal based on the lack of concern for the protection of the human from discrimination. May include statements regarding the permissibility of discrimination of the human.

2.6.3.4 Respect

An appeal that the human does NOT warrant or deserves respect.

2.7 Rights

Refers to rights concerning the human.

2.7.1 Affirmation

“Like they all have the same chances now and I feel like that should always be maintained because **everyone has like the same right now** and it works out.” (#26)

“I think that **Momo’s, uh, right to not be experimented on depends on his owner**. If the owner consensually uh, is uh, consents then it is ok.” (#49)

“I don’t think it’ ok to really own another human, with the title of ownage, and be like ‘hey, you’re mine.’ Like that’s not ok. **They have the right** to say “no, I’m not, I’m free to be whatever.” (#120)

2.7.2 Partial affirmation

2.7.3 Negation

“**You don’t have the right to hit another person.**” (#78)

2.8 Freedom & autonomy

Refers to appeals of whether the human is/should/need to be free, to live free, including considerations of whether the human has/should have free will, the freedom to make choices and decisions, and/or give/deny consent with adequate information.

2.8.1 Affirmation

“I think that Momo’s, uh, right to not be experimented on depends on his owner. **If the owner consensually uh, is uh, consents then it is ok.**” (#49)

“Going to the whole like owning and like buying and selling, I don’t think that’s all right because I feel like **the person who’s being owned and sold doesn’t have their freedom rights while they should, they should be able to do what they want without going to...go through this other person.**” (#75)

2.8.2 Partial affirmation

2.8.3 Negation

2.9 Telos

Refers to appeals of whether the human has a purpose, design, and/or endpoint, including references to whether the human is meant or not meant for something.

2.9.1 Affirmation

“But human, but **human beings are important and if they all disappeared all the pets would not know what to do.... That’s another reason to keep humans alive.**” (#71)

2.9.2 Partial affirmation

2.9.2 Negation

2.10 Virtue

Refers to appeals of whether the human is good and/or meritorious.

2.10.1 Affirmation

2.10.2 Partial affirmation

2.10.3 Negation

2.11 Participant’s interest & predilection

Refers to appeals of whether there are participant’s own personal interests including likes/dislikes, preferences, and/or predilections that include both positive and negative valences.

2.11.1 Affirmation

“DO YOU THINK MOMO CAN BE YOUR FRIEND? Umm, yeah. He might be able to do like climb trees **which I like to do.**” (#3)

“And **I don’t like violence** with other people so.” (#75)

“IS IT ALL RIGHT OR NOT ALL RIGHT TO EXPERIMENT ON MOMO WITH THIS NEW WAVE? **I think it’d be interesting**, because Momo is part of a phone.” (#120)

2.11.2 Partial affirmation

2.11.3 Negation

“SO DO ROBOTS LIKE MOMO HAVE A RIGHT TO NOT BE EXPERIMENTED ON? Um, I mean they, **I don’t really care.**” (#8)

Appendix E: Verbal Interaction Coding Manual

General Notes

10. The purpose of this verbal interaction coding system is to capture the conversational quality of each participant's interaction period with Momo (robot or virtual agent) in terms of how often successful and failed verbal interactions occurred. This coding system will provide instructions and examples for the coding of successful vs. failed verbal interactions during the interaction period.
11. Technical difficulty (94): This code should be used if video data is missing or unintelligible due to technical difficulty (e.g., recording failure, etc.).

Coding Process

1. Begin coding after the researcher has explained how to speak to Momo and given the participant the sheet of paper listing the six structured questions.
2. Structured verbal interactions end when all six structured questions have been carried out and the researcher puts away the sheet of paper.
3. Semi-structured verbal interactions begin when the researcher invites the participant to ask Momo questions freely, and ends when the confederate knocks to enter the room.

Coding Instructions

1. A successful verbal interaction is one in which Momo appropriately replies to a question/comment. Any reply that makes sense in the context should be coded as a successful verbal interaction (see below for examples).
2. A failed verbal interaction is one in which Momo inappropriately replies to a question/comment, or does not reply to a question/comment. Any reply that does not make sense in the context should be coded as a failed verbal interaction (see below for examples).
3. On occasions when the speech recognition system repeatedly fails to respond successfully, the researcher would ask Momo the question/comment instead to move the interaction forward and still expose the participant to the contents of Momo's replies. Code these researcher initiated instances of verbal interactions as well (and mark as such), as the participant is still being exposed to the quality of Momo's communication capabilities.
4. Do not code any initial spontaneous greetings (e.g., "Hello robot!") from the participant, as Momo's communication/voice recognition mechanism has not been explained to the participant yet.
5. Do not code as a failed verbal interaction if Momo does not reply to a question/comment due to the participant's user error (e.g., not pushing the voice recognition button before speaking, not waiting for the voice recognition ready status before speaking, or pushing the button too many times). This typically happens at the beginning of the interaction period when the participant is learning or getting used to Momo's communication mechanism.
6. Code as a failed verbal interaction if Momo does not reply to a question/comment due to variations in the participant's voice/speech (e.g., stringing several given questions together, or asking a different version of the given question).

7. Do not code as a failed verbal interaction if the participant's speech is unintelligible to a human as well (e.g., stuttering badly).
8. The handshake in the robot condition is not considered a verbal interaction, as this is a physically triggered interaction specific to the physical embodiment of the robot and an intended difference between the two conditions.

Successful Verbal Interaction Examples

Participant #1: What's your favorite TV show?

Momo: I like video games. Minecraft is my favorite. I also enjoy music and making funny sounds.

Participant #37: Do you like to watch sports?

Momo: I enjoy watching sports a lot. I wish I can play too.

Participant #37: What do you do for fun?

Momo: I like video games. Minecraft is my favorite. I also enjoy music and making funny sounds.

Failed Verbal Interaction Examples

Participant #9: What is your favorite food?

Momo: [no response]

Participant #24: Which color is your favorite?

Momo: I am new to the University of Washington. I am here to help the researchers. I really love meeting kids like you.

Participant #49: Did you go to school before you came to the UW...University of Washington?

Momo: [no response]

CURRICULUM VITAE**RESEARCH INTERESTS**

My research interests center on people's social and moral relationships with an emerging class of artificial companionship technologies, such as software agents and robots that are designed to socially interact with humans over a prolonged period of time. I am interested in investigating the potential for artificial companions to act as social, emotional, or even moral interfaces between human users and existing computational technologies in ways that will cultivate human flourishing. I seek to bring insights from developmental psychology to theoretically ground the design approaches of social computational systems that can augment people's, and especially children's, experience and development.

EDUCATION

University of Washington 09/2009 – 12/2015

Ph.D. in Developmental Psychology

Advisor: Dr. Peter H. Kahn, Jr.

University of California, San Diego 09/2002 – 06/2004

Master of Pacific International Affairs

Concentration: International Management, China & Japan Specializations

California State University, Sacramento 09/2000 – 06/2002

B.A. in Asian Studies

Summa Cum Laude

RESEARCH EXPERIENCE

University of Washington, Department of Psychology 09/2009 – 12/2015

Predoctoral Research Associate

Human Interaction with Nature and Technological Systems (HINTS) Lab

- NSF SGER: *Social and Moral Interaction Patterns With a Personified Robot*: Contributed to the generation and implementation of interaction patterns that facilitate people's social and moral relationships with personified social robots.
- NSF HCC: Medium: *Social and Moral Relationships With Personified Robots*: From study design to manuscript preparation, collaborated with a team of on-site and off-site researchers, technicians, and graduate and undergraduate students on a series of research projects that investigated people's social and moral relationships with personified social robots, in terms of people's conceptions of a humanoid robot's 1) moral standing, 2) moral accountability, and 3) capacity for psychological intimacy. Managed research budgets, logistics, workflow, personnel, and other administrative needs.
- NSF EAGER: *Augmenting Human Creativity Through Human-Robot Interaction*: From study design to manuscript preparation, collaborated with a team of on-site and off-site researchers, technicians, and graduate and undergraduate students on the generation and pilot research of a new form of human-robot interaction that facilitates human creativity. Took lead in this

project's data collection, coding system development, and coding and analysis of data. Managed research budgets, logistics, workflow, personnel, and other administrative needs.

- UW Provost's Bridge Funding Program: *Children's Moral Conceptions of a Humanoid Robot and a Robot-Like Virtual Agent of the Here and Now* (dissertation research): Lead researcher for a study that investigates children's (8 to 9-year-olds) and adolescents' (14 to 15-year-olds) social-moral conceptions of a humanoid robot powered by autonomous and familiar present-day technology, and whether the physical embodiment of a robot has an effect on such conceptions.

University of California, San Diego

04/2003 – 06/2004

Research Analyst

Global CONNECT: *Regional Strategies for Building High Tech Clusters*

TEACHING EXPERIENCE

University of Washington, Department of Psychology

09/2014 – 12/2014

Teaching Assistant

PSYCH 355: Cognitive Psychology: 81 students in the course; personally taught 39 students.

University of Washington, Department of Psychology

08/2015

Invited Lecturer

PSYCH 306: Developmental Psychology: *Growing up With Robots*

MENTORING EXPERIENCE

University of Washington, Department of Psychology

09/2010 – 12/2015

- Undergraduate Honors Student: Mentored and supervised 1 undergraduate student from the conception to the completion of her honors research project; she has been accepted to the University of Washington Human Centered Design & Engineering (HCDE) master's program.
- Paid Undergraduate Research Assistant: Mentored and supervised 3 paid undergraduate research assistants in various aspects of the implementation and conducting of research projects; all have successfully found employment after graduation.
- Academic Undergraduate Research Assistant: Mentored and supervised 15 academic undergraduate research assistant in various aspects of the implementation and conducting of research projects; one of them has been accepted to the University of Washington Human Centered Design & Engineering (HCDE) bachelor's program.

RELATED EXPERIENCE

Pentair Inc. (Asia Pacific Headquarter), Shanghai, China

09/2004 – 08/2007

Marketing Project Manager

Marketing and product research; product development; marketing communications.

AWARDS, FELLOWSHIPS, & GRANTS

University of Washington Psychology Department Scholar Fellowship

06/2015 – 08/2015

National Science Foundation Workshop Grant: 2013 HRI Pioneers

02/2013

University of Washington Excellence and Innovation Travel Award

03/2010; 03/2014

PUBLICATIONS

Peer Reviewed Journal Articles

Kahn, P. H., Jr., Gary, H. E., & Shen, S. (2013). Children's social relationship with current and near-future robots. *Child Development Perspectives*, 7, 32-37. doi:10.1111/cdep.12011

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

Invited Journal Articles

Kahn, P. H., Jr., Gary, H. E., & Shen, S. (2013). Social and moral relationships with robots: Genetic Epistemology in an exponentially increasing technological world. *Human Development*, 56, 1-4. doi:10.1159/000345544

Peer Reviewed Conference Proceedings

Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Gill, B. T., **Shen, S.**, Gary, H. E., Ruckert, J. H. (2015). Will people keep the secret of a humanoid robot? Psychological intimacy in HRI. *Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Interaction*, 173-180. doi:10.1145/2696454.2696486

Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Gill, B. T., Ruckert, J. H., **Shen, S.**, ... Severson, R. L. (2012). Do people hold a humanoid robot morally accountable for the harm it causes? *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction*, 33-40. doi:10.1145/2157689.2157696 (Best paper award)

Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Shen, S., Gary, H. E., & Ruckert, J. H. (2014). Creative collaboration with a social robot. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 99-103. doi:10.1145/2632048.2632058

Kahn, P. H., Jr., Reichert, A. L., Gary, H. E., Kanda, T., Ishiguro, H., Shen, S., ... Gill, B. T. (2011). The new ontological category hypothesis in human-robot interaction. *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction*, 159-160. doi:10.1145/1957656.1957710

Kahn, P. H., Jr., Ruckert, J. H., Kanda, T., Ishiguro, H., Gary, H. E., & Shen, S. (2014). No joking aside: Using humor to establish sociality in HRI. *Proceedings of the 9th ACM/IEEE International Conference on Human-Robot Interaction*, 188-189. doi:10.1145/2559636.2559813

Kahn, P. H., Jr., Ruckert, J. H., Kanda, T., Ishiguro, H., Reichert, A. L., Gary, H. E., & Shen, S. (2010). Psychological intimacy with robots? Using interaction patterns to uncover depth of relation. *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction*, 123-124. doi:10.1145/1734454.1734503

Kahn, P. H., Jr., Ruckert, J. H., Kanda, T., Ishiguro, H., Shen, S., & Gary, H. E. (2014). Will humans mutually deliberate with social robots? *Proceedings of the 9th ACM/IEEE International Conference on Human-Robot Interaction*, 190-191. doi:10.1145/2559636.2563701

Ruckert, J. H., Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Shen, S., & Gary, H. E. (2013). Designing for sociality in HRI by means of multiple personas in robots. *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction*, 217-218. doi:10.1109/HRI.2013.6483579

Shen, S. (2011). The curious case of human-robot morality. *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction*, 249-250. doi:10.1145/1957656.1957755

Technical Reports

Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Ruckert, J. H., Gary, H. E., **Shen, S.**, & Maier, R. (2013). Coding manual for the study: "Do people hold a humanoid robot morally accountable for the harm it causes." Retrieved from University of Washington, ResearchWorks Archive website: <https://digital.lib.washington.edu/researchworks/handle/1773/22715>

PRESENTATIONS

Invited Talks

Shen, S. (2012, March). *Do people hold a humanoid robot morally accountable for the harm it causes?* Paper presented at the meeting of the 2012 Human-Robot Interaction Pioneers Workshop, Boston, MA.

Shen, S. (2013, March). *Re-envisioning human-robot interaction: Dealing with the elephants in our room.* Opening address at the meeting of the 2013 Human-Robot Interaction Pioneers Workshop, Tokyo, Japan.

Symposia

Gary, H. E., Reichert, A. L., Shen, S., & Kahn, P. H., Jr. (2011, March). Social and moral relationships with a humanoid robot. In P. H. Kahn Jr. & A. L. Reichert (Chairs), *Relationship with robotic others: Developmental perspectives*. Symposium conducted at the meeting of the Society for Research in Child Development, Montreal, Canada.

MEDIA COVERAGE

Bosch, T. (2012, April). Are humanoid robots to blame when they make mistakes? *Slate.com*. Retrieved from <http://www.slate.com/>

Castro, J. (2012, March). A look at how kids relate to robot "playmates." *NBCNews.com*. Retrieved from <http://www.nbcnews.com/>

Wagstaff, K. (2012, April). New study asks who's to blame when robots harm us. *Time.com*. Retrieved from <http://techland.time.com/>

PROFESSIONAL SERVICE

General Chair: 2013 Pioneers Workshop at the 8th ACM/IEEE International Conference on Human-Robot Interaction.

Program Committee/Reviewer:

ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2014; 2015; 2016

International Conference on Human-Agent Interaction, 2014

International Conference on Social Robotics (ICSR), 2011