# Transparency about a robot's lack of human psychological capacities: Effects on child-robot perception and relationship formation

van Straten, C.L.; Peter, J.; Kühne, R.; Barco, A.

**DOI**
10.1145/3365668

**Publication date**
2020

**Document Version**
Final published version

**Published in**
Transactions on Human-Robot Interaction

# Transparency about a Robot's Lack of Human Psychological Capacities: Effects on Child-Robot Perception and Relationship Formation

CAROLINE L. VAN STRATEN, JOCHEN PETER, RINALDO KÜHNE, and ALEX BARCO,
Amsterdam School of Communication Research (ASCoR), University of Amsterdam

The increasing sophistication of social robots has intensified calls for transparency about robots' machine nature. Initial research has suggested that providing children with information about robots' mechanical status does not alter children's humanlike perception of, and relationship formation with, social robots. Against this background, our study experimentally investigated the effects of transparency about a robot's lack of human psychological capacities (intelligence, self-consciousness, emotionality, identity construction, social cognition) on children's perceptions of a robot and their relationship to it. Our sample consisted of 144 children aged 8 to 9 years old who interacted with the Nao robot in either a transparent or a control condition. Transparency decreased children's humanlike perception of the robot in terms of animacy, anthropomorphism, social presence, and perceived similarity. Transparency reduced child-robot relationship formation in terms of decreased trust, while children's feelings of closeness toward the robot were not affected.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**; **Laboratory experiments**;

Additional Key Words and Phrases: Child-robot interaction, human-robot interaction, robot ethics, child-robot relationship formation, robotics

**ACM Reference format:**
Caroline L. van Straten, Jochen Peter, Rinaldo Kühne, and Alex Barco. 2020. Transparency about a Robot's Lack of Human Psychological Capacities: Effects on Child-Robot Perception and Relationship Formation. *ACM Trans. Hum.-Robot Interact.* 9, 2, Article 11 (January 2020), 22 pages.
https://doi.org/10.1145/3365668

11

## 1 INTRODUCTION

As robots are becoming increasingly social [e.g., 34], and as children have a strong tendency to form social bonds with non-human entities [24], social relationships between children and robots are bound to become more common in the near future [e.g., 12]. Within the field of child-robot interaction (CRI), however, child-robot relationship formation constitutes a controversial topic. On the one hand, concerns have been raised about the nature and consequences of such

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, under Grant No. 682733 to the second author.
Authors' addresses: C. L. van Straten, J. Peter, R. Kühne, and A. Barco, Amsterdam School of Communication Research (ASCoR), University of Amsterdam, P.O. box 15791, 1001 NG Amsterdam, The Netherlands; emails: {c.l.vanstraten, j.peter, r.j.kuhne, a.barcomartelo}@uva.nl.

2573-9522/2020/01-ART11
https://doi.org/10.1145/3365668

relationships as deceptive, inauthentic [e.g., 58, 61], and as potentially damaging children's peer friendships [e.g., 34, 62]. On the other hand, the issue of deception has been looked at through the lens of stage magic and illusion [16]. Pearson and Borenstein [52] even highlight the potential benefits of child-robot relationship formation, positing that "to promote the welfare of children effectively, some degree of bonding between the robot and child will *have* to occur" (p. 127, emphasis added).

Whether relationships between children and social robots are considered harmful or benign, it is timely to investigate their—increasingly likely—emergence and its implications [e.g., 12, 67]. As Turkle [61] states, children form relationships with robots based on their perceptions of (and assumptions about) robots' behavior. For instance, children may get the impression that robots are emotional and intelligent [58]. Humanlike robot perceptions, in turn, would result from an emotional connection between children and robots [61]. Thus, children's robot perceptions seem to play a central role in the emergence of child-robot relationships. In this context, Boden et al. [11, p. 127] have recently proposed five ethical principles for responsible robotics, of which the fourth states that "[r]obots are manufactured artefacts [that] should not be designed in a deceptive way to exploit vulnerable users; instead their machine nature should be transparent."

Yet, while it may be effective to transparently present robots to adults, initial findings suggest that making children aware of robots' machine nature and/or working does not alter their perception of, and relationship formation with, a robot. First, debriefing children, after they had interacted with a robot, about its technological nature as well as letting them control the robot did not interfere with children's feelings of companionship toward the robot [63]. Second, children attributed feelings and free will to a robot even after seeing it being programmed and after being informed on its working [13]. Third, introducing, before the interaction, a robot as either a machine or a friend did not affect children's self-reported perceptions of the robot as a social agent with mental abilities [e.g., thought, emotion; 39]. Based on these studies, "lifting the curtain" for children to find out what a robot is (and what it is and is not capable of) [11] thus seems ineffective.

However, children's reasoning about robots allows for contradictions: In their view, robots can both have a mind and be built and controlled by humans [7]. In line with this idea, children attribute psychological capacities to robots and develop friendly relationships with them, even if they understand that robots are not alive (for an overview of findings, see [56]). This may explain why studies that demonstrated to children mainly robots' inanimate nature and non-biological working found no effects of this transparency on children's robot perceptions and relationship formation with robots [i.e., 13, 39, 63]. Conversely, however, this also implies that transparency about a robot's lack of *psychological* capacities may alter child-robot perceptions and relationship formation.

More systematic, preferably causal, evidence is needed that elucidates whether providing robot-related information affects children's evaluations of social robots, or, more specifically, (a) children's perception of robots as humanlike, social beings, and (b) the emergence of child-robot relationships (for a similar request in the broader context of HRI, see [55], p. 218). We, therefore, conducted an experiment to investigate whether being transparent about the differences between robots and humans in terms of key psychological capacities affects children's humanlike perception of, and relationship formation with, a social robot. Being transparent to children about differences between robots and humans in terms of psychological capacities requires us, first, to identify the capacities that distinguish humans from other entities. This is done in the next section, on the basis of both the philosophical and psychological literature. After that, we will outline the concepts that should be assessed to investigate whether transparency about robots' lack of these human capacities affects children's humanlike perception of, and relationship formation with, social robots.

## 2 THEORETICAL FRAMEWORK

### 2.1 Robots Versus Humans: Key Psychological Capacities

To identify key psychological capacities that distinguish humans from other entities (e.g., robots), the philosophical considerations about personhood are a useful starting point, given their focus on what constitutes being human. The lack of these capacities in robots can then, in child-appropriate terms, be conveyed to children during their interaction with a robot. Building on a key paper by Dennett [21], Hubbard [33] proposes five requirements that a machine entity (e.g., robot) should fulfill to be considered equivalent (though not identical) to a human being and be granted personhood. First, an entity should be able to interact with, and learn from, its environment in a meaningful, diverse, and sophisticated way, as well as to engage in complex thought and communication (i.e., be intelligent). Second, the entity should have a sense of self (i.e., be self-conscious). Third, it should possess the creativity to define a unique, relatively coherent 'self' and accompanying "life plan" (i.e., build an identity). Fourth, it should be able to experience emotions related to a care for survival and purposefulness (i.e., be emotional). Fifth and finally, the entity should be capable of living in a community with other persons and to understand others' perspectives and feelings (i.e., have social cognition) [33].

Hubbard's [33] requirements have received support from psychological research on mind and person perception [e.g., 23, 27, 29]. *Mind* perception forms the foundation of how people distinguish between human and nonhuman entities, and precedes *person* perception, that is, the attribution of traits, dispositions, and capacities to others. When people perceive nonhuman entities to have a mind, they attribute humanlike characteristics to these entities (i.e., engage in anthropomorphism) and may form a humanlike relationship with them [23]. Two psychological categories—experience and agency—play a role in the perception of mind [27]. Together, they cover Hubbard's [33] requirements of intelligence, self-consciousness, emotionality, and social cognition [27]. Person perception, in turn, relates to Hubbard's [33] requirement of a stable and enduring identity [23].

In conclusion, as Hubbard's [33] philosophical requirements figure prominently in psychological research on mind and person perception, they seem suitable as key human psychological capacities. By extension, they lend themselves—in adapted, child-appropriate form—for making transparent to children how robots differ from humans. However, in line with media psychological findings [e.g., 18, 47], Hubbard [33] points out that even though current machines (e.g., robots) do not possess the five identified human capacities (and are often merely capable of mimicking some of them), people may still treat or perceive machines as humanlike. One reason for this is the human tendency to see non-human entities and objects, such as animals, plants, and geometrical shapes, as humanlike (for an overview of findings, see [24]). This tendency is particularly strong in children [24]. As a result, children may frequently come to perceive robots as humanlike, and accordingly form social relationships with them—even after being informed about robots' lack of human psychological capacities.

### 2.2 Effects of Transparency on CRI

Prior CRI research has found that providing children with robot-related information does not influence child-robot perception and relationship formation [13, 39, 63]. Media psychological research, in contrast, suggests that children's perceptions of television characters and content can be influenced by external factors. For example, preschoolers watching a movie scene framed as a dream were able to identify the scene as such [69]. Moreover, children's perception and evaluation of gender stereotypical television content could be altered by having an experimenter contradict the stereotypes [49]. Finally, guiding children's attention toward the feelings of the victim in aggressive media content altered their perceptions of both the media characters and the violence itself [48]. Even though these studies do not specifically address children's humanlike perceptions of, or relationship formation with, media characters, they suggest that children's perceptions of robots may be susceptible to external information as well.

Robots differ from traditional media characters in two ways. First, robots have a physical embodiment in the real world. Compared to virtual agents or virtual depictions of robots, physically embodied robots are more persuasive [43] and evoke more anthropomorphic interactions and attributions [35]. Thus, it might be harder to alter children's perceptions of robots than those of traditional media characters. Second, robots' interactive character may strengthen the illusion of reciprocity, which may complicate the attenuation of child-robot relationship formation. At the same time, it has been shown, for example, that children who have more knowledge of animals—which are, like robots, embodied and interactive—are less likely to apply anthropocentric reasoning to animals [25]. Moreover, from 8 to 9 years of age, children develop their capacity to distinguish between what is real and what is not, notably in a media context [64]. Thus, at least for children aged 8 and older, information about a robot may affect their robot perceptions.

In the preceding section, we have outlined the human psychological capacities that current robots lack. As will be outlined below in detail, we consider providing children with information about the various psychological capacities *in their entirety* as "being transparent"; we thus do not study the influence of a particular capacity separately. When children are provided with information about these *psychological* capacities, this likely affects children's humanlike robot perception in a broader sense. First, at the lowest level of detail, transparency about how robots differ from humans may elicit machinelike robot perceptions, which might alter children's judgements about robots' animacy, or their "perception of life" [3, p. 74]. Second, human psychological capacities are central to the concept of anthropomorphism, or "the tendency to imbue the real or imagined behavior of nonhuman agents with humanlike characteristics, motivations, intentions, or emotions" [24, p. 864]. In turn, anthropomorphism is influenced, third, by the extent to which one perceives a robot as similar to oneself [2, 24], and corresponds, fourth, with the notion of social presence [51], which occurs when users do not notice a robot's artificial nature [41, p. 45].

We thus study the effects of information about a robot's lack of human psychological capacities on children's humanlike perceptions of the robot by asking children (a) whether they think a robot is alive; (b) whether it is humanlike; (c) whether they consider the robot similar to themselves; and (d) whether it feels like being with another person when the robot is around.

As to children's relationship formation with social robots, psychological research has identified two concepts that are central to the emergence of interpersonal relationships [10] and constitute primary functions of children's friendships [4]: closeness and trust. Closeness has been defined as a feeling of connectedness or intimacy that could potentially build up to friendship [60]. An increase in closeness is facilitated by an increase in trust [10], which can be defined as the belief in the benevolence and honesty of another person [40]. As children's reasoning about humans and robots appears to follow similar patterns, the concepts of closeness and trust seem central to child-robot relationship formation. While reminding children about robots' lack of human capacities cannot preclude child-robot relationship formation altogether, "it might reduce the likelihood and extent to which [children] form emotional bonds with robots" [55, p. 218], and thus, affect closeness and trust. Therefore, we hypothesize that:

Hypothesis 1 (H1). Being transparent about a robot's lack of human psychological capacities, as opposed to not being transparent, decreases children's humanlike perception of the robot in terms of (H1a) animacy, (H1b) anthropomorphism, (H1c) social presence, and (H1d) perceived similarity.

Hypothesis 2 (H2). Being transparent about a robot's lack of human psychological capacities, as opposed to not being transparent, reduces the emergence of a child-robot relationship in terms of (H2a) closeness and (H2b) trust.

## 3 METHOD

A one-factorial experiment was conducted. Transparency about robots' lack of human psychological capacities was the between-subject factor, which included a transparent and a control condition. Ethical approval was obtained from the Ethics Review Board of the Faculty of Social and Behavioral Sciences of the University of Amsterdam before data collection.

### 3.1 Participants

Active informed consent was obtained from both schools and parents of participating children, and verbal consent was obtained from the children themselves before the experiment started. On the parental consent form, parents were asked to report whether their child had a medical indication, and if so, then which one. It was explained in an accompanying information letter that this information would never exclude children from participation in the study, but that the data of children with indications that could interfere with the study outcomes may be excluded from the analyses.

An *a priori* power analysis for a one-way ANOVA was conducted. Assuming an average effect size ($f = .25$) and specifying an alpha level of .05 and a required power of .80, the analysis indicated that a sample of 128 children was required. We were able to collect data of 158 children aged 8 to 9 years at six primary schools across the Netherlands. We excluded the data of two children who, due to robot malfunctions, missed a part of the interaction. In addition, nine children were excluded from the analyses as they had, one week before the experiment, followed a philosophy course on social robots and aspects of human-robot relationships that are relevant for the experiment. Finally, three children were excluded who were diagnosed with Autism Spectrum Disorder (ASD).[1] We excluded these children, because children with ASD typically encounter difficulties with respect to social interaction and communication [1] that may hinder their ability to engage in and sustain peer friendships [e.g., 44]. Moreover, the anthropomorphic reasoning of individuals with ASD differs from that of typically developed people (see [24]). Thus, we eventually analyzed the data of 144 children (72 male, 72 female) with an average age of 8.84 years ($SD = 0.59$). Children were randomly assigned to the transparent ($n = 72$, 34 female, $M_{age} = 8.84$, $SD_{age} = 0.59$) or the control condition ($n = 72$, 38 female, $M_{age} = 8.83$, $SD_{age} = 0.59$). There were no significant differences in biological sex ($\chi 2\,(1, N = 144) = 0.444, p = .505$) or age ($t\,(142) = -0.030, p = .809$) across conditions, which indicates that the randomization procedure was successful.

### 3.2 Interaction Task and Manipulation

Each child interacted once with a humanoid robot (Nao, Softbank). On average, the interaction took about 12 minutes. The interaction started with some small talk, during which the robot asked for the child's name, introduced itself, and asked for the child's age and favorite color. Subsequently, the robot engaged the child in a sport guessing game (i.e., the robot imitated sports movements and the child had to guess the correct sport). Furthermore, the robot told the child a short fairytale-like story and performed a dance that the child could join if s/he wanted to. To link the various parts of the interaction, the robot engaged the child in additional small talk by posing questions that were

[1]When we included ASD diagnoses as a covariate in the model rather than excluding the data of these children from further analyses, the results remained stable.

logically related to the topic of the preceding or following activity. The robot never claimed to have any truly human capacities (e.g., consciousness), and said exactly the same in both conditions.

The experimental manipulation consisted of 14 instances of experimenter intervention, during which the experimenter either gave the child robot-unrelated information that elaborated on the topic of conversation (control condition) or provided information about the robot's lack of human psychological capacities (transparent interaction). For instance, after the robot told the story, the experimenter either explained to the children something about fairytale characters (i.e., "Fairytales often feature all kinds of fantasy creatures, like dragons, wizards, and witches"), or about the preprogrammed nature of the story (i.e., "The story that Nao just told you has been put into Nao's computer beforehand. Because robots can only say what is in their computer."). The explanations provided in the transparent condition emphasize the robot's lack of human psychological capacities as identified by Hubbard [33], namely, intelligence, self-consciousness, identity construction, emotionality, and social cognition.

We decided to have the experimenter rather than the robot itself provide children with the explanations, as the experimenter has higher source credibility than the robot (for an overview of findings on credibility, see [8]). In addition, having the robot provide the explanations itself would interfere with the manipulation, as this would suggest that the robot has knowledge of its "conceptual self"—which implies self-consciousness [50]. We controlled for influences of interaction length by keeping the total word count of the experimenter texts approximately constant across conditions (i.e., 514 words in the control condition versus 512 words in the transparent interaction condition). Moreover, we ensured that the length of specific sections differed by no more than five characters (excluding spaces) across conditions. Adjusting the number of characters per section seemed more appropriate than adjusting the number of words, as the speech duration depended on the length rather than on the number of words.

The English translations of the experimenter texts can be consulted in Appendix A. Apart from providing the child with the information specified in the Appendix, the experimenter did not take part in the interaction. Any questions or remarks children addressed to the experimenter were answered with concise and neutral, pre-scripted responses (e.g., "Shall we talk about that later?").

### 3.3 Procedure

Prior to the start of the first interaction, the female experimenter introduced the study to the children at class-level to increase their comfort with the interaction setting (see [68]). Children were shown a picture of the robot and were informed about the study procedure. The experimenter told the children in an age-appropriate language that participation was voluntary, that the results would be anonymized before publication, and that they could stop their participation at any point in time without providing an explanation. General questions about the study procedure were answered, while answers to questions that could influence children's initial robot perceptions were postponed until the debriefing.

The robot was activated before children entered the experimental room, in which all distractions were minimized. As the study relied upon a Wizard of Oz (WOZ) paradigm, a female research assistant was present to control the robot from a distance—where possible out of children's direct line of sight when facing the robot. The robot's face tracking was activated (offline, not recorded) and the robot converted preprogrammed text into speech. Upon entering the room, the experimenter briefly introduced the children to the assistant. Children were told not to distract the assistant from her work—without specifying what she was doing—such that they would not focus their attention on her and the WOZ set-up. Children were then asked to sit down in front of the robot, at a distance they felt comfortable with. When the child sat down, the experimenter sat down next to the robot and the child, explaining that she would tell the child some things during the interaction.

She reminded the child that participation was voluntary and could be stopped at any time. Once a child had indicated that s/he still wanted to participate, the experimenter asked the child to save any questions or remarks that might come up during the interaction for later. After the child had agreed to this, the assistant started the robot. After the interaction, the experimenter walked the child to the interview room and, subsequently, the robot was turned off.

The survey was conducted by a second trained, female research assistant. We opted for self-report rather than observational measures, because, first, they are considerably less time-consuming; second, we investigate complex concepts that are difficult to adequately observe; and third, we have good experience with the use of such measures in prior CRI studies [19, 66]. The interviewer briefly introduced herself to the child and explained the procedure of the survey. Following the approach taken by Leite, Pereira, and Lehman [42], children were familiarized with the question format and answering scale by means of several practice items (e.g., "I like candy," "I like Brussel's sprouts"). While the questionnaire mainly consisted of Likert-scale items, it also contained a set of visual semantic differentials. This format was only explained when the items were encountered, to not overburden children at the beginning. Once the child indicated that s/he understood the procedure, the interviewer asked the child a series of questions regarding his or her perception of, and relationship formation with, the robot. Most questions used a closed-ended response format. However, several questions about children's reasoning processes used an open-ended response format. The questionnaire ended with a treatment check.

Children were asked not to discuss the content of the interaction and questionnaire among each other. We observed that, even if they did do so, no information relevant to the manipulation was conveyed. Once all children had participated, they were debriefed through an interactive presentation at class-level (for a similar approach, see [54]). The experimenter revealed that some children had received different information than others and provided all the children with information on robots' mechanical status and lack of human capacities. For children that had been exposed to transparent interaction, this information was not new, but in this way, we made sure that all children—also the ones exposed to the control condition—were fully debriefed in the end. We clarified that the robot does not function on its own and revealed the pre-programmed nature of the interaction: We showed a picture of a Choregraphe computer program and explained that when the "play" button is pressed, the robot starts doing what it is programmed to do (in the example we gave, the robot waved its arm). We also emphasized that although the experimenter mentioned different things in the different conditions, the robot had done and said exactly the same things during all interactions. Any remaining questions were answered, and all children received a small gift to thank them for their participation and/or attention.

### 3.4 Measures

Children indicated their answers to the closed-ended questions on a five-point Likert scale, which ranged from "does not apply at all" to "applies completely." The answer options were complemented with a visualized answer scale (see Appendix B) as used by Severson and Lemm [57], in which bars of increasing height clarified the meanings of the verbal labels. We chose the bar scale because it did not provide any indication as to the valence of the answer options (e.g., colors, smileys), as this could trigger socially desirable answers. In an earlier data collection (see References [19, 66]), the answer scale was demonstrated to be properly understood by children.

The questionnaire started with the measures assessing children's robot perception: animacy, anthropomorphism (Godspeed, visual semantic differentials, and IDAQ-CF), social presence, and perceived similarity. The perception measures were followed by the measures of closeness and trust. Finally, the treatment check was administered. All measures were ordered such that influences of questions posed earlier would minimally influence those that came after. The one-factorial structure of the measures of closeness and trust was confirmed in a prior study, in which these measures also proved to be sufficiently reliable [66]. The specific questionnaire items of each scale can be consulted in Appendix C.

*3.4.1 Treatment Check.* Children were exposed to six questions that asked whether, during the interaction with the robot, the experimenter had talked about a certain topic or not. The questions were posed at the end of the survey to preclude priming effects on the measurement of our main outcomes, and to not artificially amplify the experimental manipulation. The second, third, and sixth question referred to information that was mentioned in the control condition. The first, fourth, and fifth questions referred to information mentioned in the transparent condition. This question order was used for two reasons: first, to warrant that the children did not need to indicate three times in a row that they could not recall the information—which they may experience as uncomfortable; second, to preclude that the children could easily identify an answer pattern.

*3.4.2 Animacy.* To assess animacy, we used a four-item scale based on two semantic differential scales of animacy that have been used in adult populations [3, 31]. The factor analysis (principal axis factoring, direct oblimin rotation; this type was used for all analyses) confirmed the one-factorial structure of the scale that explained 48% of the variance. The scale had good internal consistency ($\alpha$ = .75). Further inspection showed that the internal consistency could be increased to $\alpha$ = .81 by removing the item "Nao can die." However, given the good internal consistency, we decided to keep the item in the scale to ensure comparability with the original scales. An index score for animacy was computed by averaging the items ($M$ = 2.37, $SD$ = 1.04, skewness = 0.398, kurtosis = −0.788).

*3.4.3 Anthropomorphism.* Anthropomorphism is a central concept in the present study, but difficult to measure among children. Therefore, we assessed anthropomorphism with two multi-item scales and a set of three visual semantic differentials. The first multi-item scale was based on the anthropomorphism dimension of the Godspeed questionnaire [3], the Dutch translation of which was adapted for use among children by De Jong, Vogt, and Krahmer [20]. While the Godspeed questionnaire has widely been used in HRI studies, both the original scale and the four-item child-appropriate version used by De Jong et al. [20] have some shortcomings (e.g., in terms of validity and internal consistency; for an evaluation of the original scale, see [15]). These issues may partly derive from the use of semantic differentials whose pairings are not always unidimensional [15]. Based on both the original Godspeed measure [3] and the items used by De Jong et al. [20], we therefore came up with five assertions that could be rated on the aforementioned five-point scale. After deleting one item (i.e., "Nao can move well"), which did not load on either of the initially extracted two factors, the remaining items loaded on one factor, which explained 31% of the variance. The four-item scale had moderate internal consistency ($\alpha$ = .63). An index score for anthropomorphism (Godspeed) was computed by averaging the four remaining items ($M$ = 2.87, $SD$ = 0.90, skewness = −0.043, kurtosis = −0.581).

As the adapted Godspeed scale has not been validated, we included a second multi-item scale of anthropomorphism. This four-item scale was adapted from the technology dimension of Severson and Lemm's [57] Individual Differences in Anthropomorphism Questionnaire-Child Form (IDAQ-CF). The scale was adjusted to measure children's anthropomorphic perceptions of the robot. The factor analysis confirmed the one-factorial structure of the scale and explained 62% of the variance. The scale was internally consistent ($\alpha$ = .86). We averaged the items to compute an index score for anthropomorphism (IDAQ-CF) ($M$ = 2.57, $SD$ = 1.17, skewness = 0.263, kurtosis = −1.159). As the item-formulations of the IDAQ-CF match the information provided to children in the transparent interaction condition rather closely, the Godspeed measure will be used to check whether potential significant findings on the IDAQ-CF do not simply constitute an artifact of this similarity.

Finally, inspired by Edwards [22], we included a set of three visual semantic differentials (VSDs). The items consisted of a nine-point continuum on which children indicated whether they thought the robot was more like (a) a human or a machine ($M$ = 5.51, $SD$ = 1.77, skewness = −0.225, kurtosis = 0.487), (b) a human or an animal ($M$ = 2.72, $SD$ = 1.47, skewness = 0.399, kurtosis = −0.538), and (c), an animal or a machine ($M$ = 7.15, $SD$ = 1.46, skewness = −0.892, kurtosis = 2.733). A score of 1 would indicate that the child thought the robot was completely like a human (a and b) or animal (c), whereas a score of 9 would indicate that the robot is perceived to be exactly like a machine (a and c) or animal (b). In contrast to the multi-item scales, where children were asked to judge specific capacities and characteristics of the robot, the differentials thus required them to make more general judgements about the robot. Therefore, the differentials serve to check whether the outcomes of the multi-item scales result from the characteristics and capacities they represent rather than from children's general anthropomorphic perceptions of the robot.

To prevent influences of gender, we depicted "human" by a drawing of a male and female together. Each item was presented on a separate page to prevent children from comparing their previous answer to the next. The items reflect the multidimensional nature of anthropomorphism, which entails judgements about features that reflect human nature (i.e., emotionality, desire) and human uniqueness (i.e., higher cognition, refined emotions). Whereas distinctions between humans and robots are primarily based on human nature features, the differentiation between humans and animals mostly relies upon human uniqueness features [29]. The human-animal differential served to check whether the format of the differentials works properly. Based on our manipulation, children's answers to this item should not differ across conditions.

*3.4.4 Social Presence.* To assess social presence, we used a four-item scale based on a measure of social presence that Heerink, Kröse, Evers, and Wielinga [30] used with adults. The factor analysis confirmed the one-factorial structure of the scale, explaining 56% of the variance. The scale was internally consistent ($\alpha$ = .83). An index score of social presence was computed by averaging the items ($M$ = 3.62, $SD$ = 0.88, skewness = −0.491, kurtosis = −0.144).

*3.4.5 Perceived Similarity.* To assess perceived similarity, we adapted the attitude dimension of McCroskey, Richmond, and Daly's [45] four-factor measure of perceived homophily. The items loaded onto one factor that explained 44% of the variance. The scale had good internal consistency ($\alpha$ = .75). We computed an index score of perceived similarity by averaging the items ($M$ = 2.31, $SD$ = 0.83, skewness = 0.468, kurtosis = −0.184).

*3.4.6 Closeness.* To assess closeness, we used a CRI scale developed and validated for use among children in a similar age range (for details, see [66]). The scale consists of five items, whose one-factorial structure was confirmed for the present sample and explained 54% of the variance. The scale had good internal consistency ($\alpha$ = .85). An index score of closeness was computed by averaging the items ($M$ = 3.82, $SD$ = 0.73, skewness = −0.556, kurtosis = 0.355).

*3.4.7 Trust.* We assessed trust through a scale consisting of four items adapted from Larzelere and Huston [40]. Its one-factorial structure was confirmed for the present sample and explained 58% of the variance. Internal consistency was good ($\alpha$ = .84). An index score for trust was computed by averaging the items ($M$ = 4.11, $SD$ = 0.79, skewness = −1.083. kurtosis = 1.304).

### 3.5 Analytical Approach

For all dependent variables, we checked whether the assumptions of normality and homoscedasticity were met. The data were considered to be normally distributed when skewness and kurtosis

Table 1. Tests of Homoscedasticity

| Variable | Levene statistic | df1 | df2 | Sig. |
|---|---|---|---|---|
| Animacy | 0.528 | 1 | 142 | .469 |
| Anthropomorphism (IDAQ-CF) | 4.230 | 1 | 142 | .042 |
| Anthropomorphism (Godspeed) | 0.046 | 1 | 142 | .830 |
| VSD human-machine | 0.001 | 1 | 142 | .974 |
| VSD human-animal | 0.231 | 1 | 142 | .631 |
| VSD animal-machine | 1.741 | 1 | 142 | .189 |
| Social presence | 1.796 | 1 | 142 | .182 |
| Perceived similarity | 1.353 | 1 | 142 | .247 |
| Closeness | 2.705 | 1 | 142 | .102 |
| Trust | 23.007 | 1 | 142 | <.001 |

were between −2 and 2 [26]. We found this to be the case for all dependent variables except the animal-machine differential. Moreover, Levene's test showed that the assumption of homoscedasticity was met for all dependent variables except for anthropomorphism (IDAQ-CF) and trust (see Table 1), as well as the treatment check.

The treatment check and hypotheses were tested through a series of ANOVAs. When the assumption of normality was not met, we additionally performed non-parametric tests (i.e., Mann-Whitney). For the variables that did not meet the assumption of homoscedasticity, we conducted ANOVAs with the Welch test. As the outcomes of these analyses confirmed those of the ANOVAs, we only report the results of the ANOVAs for all variables for consistency reasons.

# 4 RESULTS

## 4.1 Treatment Check

Children in the control condition more often reported that the experimenter had provided robot-unrelated information ($M = 4.69, SD = 0.46$) than children in the transparent condition did ($M = 1.84, SD = 0.82$). This difference was very strong and significant, $F(1, 142) = 661.822$, $p < .001$, part. $\eta^2 = .82$. Moreover, children in the transparent condition more frequently reported that the experimenter had provided transparent information ($M = 4.68, SD = 0.44$) than children in the control condition did ($M = 2.25, SD = 0.85$). This difference, too, was very strong and significant, $F(1, 142) = 456.450$, $p < .001$, part. $\eta^2 = .76$. Thus, the treatment check was successful.

## 4.2 Effects on Robot Perception

According to H1a, transparency would decrease children's ratings of the robot's animacy. This hypothesis was supported, as children in the transparent condition ($M = 1.79, SD = 0.85$) produced significantly lower animacy ratings than children in the control condition ($M = 2.95, SD = 0.87$), $F(1, 142) = 65.615$, $p < .001$, part. $\eta^2 = .32$.

H1b, which predicted that transparency would decrease children's ratings of anthropomorphism, was also supported. On both the IDAQ-CF and Godspeed measures, children in the transparent condition ($M = 1.60, SD = 0.54$ and $M = 2.22, SD = 0.62$) rated the robot as lower in anthropomorphism than children in the control condition ($M = 3.53$, $SD = 0.74$ and $M = 3.53$, $SD = 0.60$, respectively). For both measures, the difference was significant, $F(1, 142) = 318.048$, $p < .001$, part. $\eta^2 = .69$ (IDAQ-CF) and $F(1, 142) = 168.529$, $p < .001$, part. $\eta^2 = .54$ (Godspeed). As to the visual semantic differentials, children in the transparent condition scored significantly higher ($M = 6.00$, $SD = 1.64$) on the human-machine differential than children in the control condition ($M = 5.03$, $SD = 1.78$), $F(1, 142) = 11.617$, $p = .001$, part. $\eta^2 = .08$. This means they were more certain of their classification of the robot as machinelike than as humanlike. Likewise, children in the transparent condition scored significantly higher ($M = 7.44, SD = 1.49$) on the animal-machine differential than children in the control condition ($M = 6.86, SD = 1.37$), $F(1, 142) = 5.990$, $p = .016$, part. $\eta^2 = .04$. This means they were more certain of their classification of the robot as machinelike than as animal-like. Following our expectations, children's answers to the human-animal differential did not differ between the transparent condition ($M = 2.87, SD = 1.57$) and the control condition ($M = 2.57, SD = 1.37$), $F(1, 142) = 1.552$, $p = .215$, part. $\eta^2 = .01$.

In line with H1c, children in the transparent condition rated the robot lower in social presence ($M = 3.46, SD = 0.97$) than children in the control condition ($M = 3.78, SD = 0.76$). This difference was significant, $F(1, 142) = 5.058$, $p = .026$, part. $\eta^2 = .03$. H1d stated that transparency would decrease perceived similarity. As expected, children in the transparent condition perceived the robot to be significantly less similar to themselves ($M = 1.89$, $SD = 0.67$) than children in the control condition did ($M = 2.74, SD = 0.76$), $F(1, 142) = 50.305$, $p < .001$, part. $\eta^2 = .26$. In sum, the results supported all components of H1.

### 4.3 Effects on Relationship Formation

Whereas H2a predicted that transparency would affect child-robot relationship formation in terms of a decrease in closeness, we found no difference in children's feelings of closeness toward the robot between the transparent condition ($M = 3.75$, $SD = 0.82$), and the control condition ($M = 3.89, SD = 0.63$), $F(1, 142) = 1.384$, $p = .241$, part. $\eta^2 = .01$. Still, as predicted by H2b, children in the transparent condition reported to trust the robot significantly less ($M = 3.89, SD = 0.94$) than children in the control condition did ($M = 4.33, SD = 0.53$), $F(1, 142) = 12.186, p = .001$, part. $\eta^2 = .08$. Hence, H2a was rejected, while H2b was supported.

## 5 DISCUSSION

The goal of the current study was to investigate experimentally whether transparency about a robot's lack of human psychological capacities would affect children's humanlike perception of, and relationship formation with, a social robot. In line with our expectations, children's ratings of the robot's animacy, anthropomorphism, social presence, and perceived similarity decreased, as did children's degree of trust in the robot. However, children's feelings of closeness toward the robot remained unaffected by our manipulation. Our findings thus show a shift in children's perception of the robot, and, at least in terms of trust, also a difference with respect to child-robot relationship formation.

Our results contradict the findings of Turkle et al. [63], who found that explanations about a robot's working and mechanical nature did not, or barely, affect children's perceptions of the robot's animacy and anthropomorphism as well as "their sense of consequence from the interactions" (p. 15). Likewise, Bumby and Dautenhahn [13] reported that children who, for example, watched a robot being controlled still anthropomorphized and animated the robot by attributing characteristics such as a free will to it. Finally, Kory Westlund et al. [39] found no difference in children's perceptions of the robot's aliveness and mental capacities (e.g., thought, emotion). Two of these studies [13, 63] were carried out with children in an age range similar to the one in our study. A difference in participants' cognitive capacities can thus be ruled out as an explanation for the inconsistency of findings.

Two alternative explanations for the difference in findings between the three aforementioned studies and our study exist. First, the difference in findings may result from differences in the information provided to the children. Whereas, we focused on making transparent which key human psychological capacities were absent in the robot, previous studies emphasized robots' inanimate nature and non-biological working [i.e., 13, 39, 63]. The children in these studies specifically reported to believe that robots have the humanlike capacities of, among other things, thought, emotion, and free will. Thus, addressing precisely the lack of such capacities may be crucial to altering children's responses to social robots.

Second, the way in which children were asked to convey their perception of the robot (and of their relationship with it) may have contributed to the difference in findings. Bumby and Dautenhahn [13] engaged children in group discussions about the robot and reported that this interview setting may have influenced their findings. Turkle et al. [63] more specifically asked children, in a semi-structured interview, whether they thought the robot was alive and whether they though the robot could be their friend. However, as they employed an open-ended answer format, children were possibly not stimulated to think about the *degree to which* they though the robot was alive and, potentially, a friend. Finally, the questionnaire employed by Kory Westlund et al. [39] used a dichotomous answer format. It is conceivable that the use of more fine-grained measurement instruments (i.e., multi-item scales and a five-point answering scale) and/or posing the questions individually to each child resulted in different outcomes.

In this context, it should be noted that even though children's robot perceptions differed significantly across the conditions, children in the control condition did not rate the robot particularly high in animacy or perceived similarity to themselves. Thus, in the absence of transparent information, children were somewhat ambivalent about the robot's animate status and similarity to themselves. Conversely, children's social presence ratings were relatively high in the transparency condition. This indicates that children's experience of social presence decreases but does not disappear when information about a robot's lack of human psychological capacities is provided.

Relatedly, the outcomes of the human-machine differential as a measurement of anthropomorphism in the present study deserve attention. Even in the transparent condition, children's answers about whether the robot resembled rather a human or a machine fell closer to the center point of the continuum than to the depiction of the machine. In general, children thus did not classify the robot as fully belonging to either the human or the machine category. Given the outcomes of the other two anthropomorphism measures (IDAQ-CF and Godspeed), we had expected children to make more pronounced judgements on this differential as well. Perhaps it can simply not be assumed that children understand that a humanoid robot is more similar to a machine than to a human being, even if this is pointed out repeatedly and even if children understand that a robot is not humanlike in terms of psychological capacities.

However, it is important to consider that it is intricate to evaluate whether anthropomorphizing a robot is appropriate, notably because such an evaluation is tied to ontological assumptions. If one assumes that robots are mere artifacts, then anthropomorphizing robots must be regarded as a form of a cognitive bias. If one assumes that robots are a new ontological category that lies in between humans and machines [34], then a certain level of anthropomorphism may be accurate. It seems safe to say that current social robots are not intelligent and sentient in the way humans are [e.g., 61]. However, technological progress may change the nature of social robots; ontological views may have to be adapted; and the notion of what constitutes transparent information about social robots may change in the future.

In line with the findings of previous studies [i.e., 13, 39, 63], we did not witness a difference in children's feelings of closeness toward the robot across conditions. The finding that closeness remains stable whereas all other concepts were affected by transparency suggests that children's feelings of friendship with the robot are independent of their robot perceptions. This finding can be understood in terms of children's psychological development. Research has shown that children in middle childhood—that is, children in the same age group as our participants—develop feelings of closeness toward television characters [e.g., 53] and animals [e.g., 46]. Moreover, transitional objects (i.e., objects children feel attached to [59]) and imaginary companions [32] continue to play a role throughout this developmental period. Apparently, children in middle childhood can experience feelings of closeness toward other entities that need not always be humanlike or alive.

Our study demonstrates that being transparent about the differences between robots and humans can alter children's robot perceptions and reduce child-robot relationship formation in terms of the development of trust. The lower degree of trust in the robot displayed by children in the transparent condition as compared to children in the control condition can be interpreted in two ways. First, on a more general level, children in the transparent condition may have believed to a lesser extent in the robot's benevolence and honesty. Second, although the robot did not claim to have any truly human capacities, it also did not mention that it *lacked* human psychological capacities. Thus, children may have gotten the impression that the robot withheld the information that the experimenter provided them with. Still, an inspection of children's mean trust scores across conditions reveals that even in the transparent condition, children's trust in the robot was relatively high.

As a follow-up on the discussion about the ethical desirability of child-robot relationships, voices have been raised that question whether transparency is always desirable. For instance, Collins [17] criticized the undefined nature of the terms like "deception" and "vulnerability." She posits that human users should be encouraged to use robots "in the most effective and positive manner" (p. 227), and that consideration should be given to the potential benefits of users' conviction that robots are alive [17]. Future research is needed to address this point of discussion.

Four limitations of our study should be pointed out. First, due to problems with the wireless network, we had to hardwire our set-up. Although this went unnoticed by the vast majority of children, some noticed that the robot was connected "to something" and did not function "on its own." We do not know to what extent this awareness affected our findings as we did not formally capture how many children noticed the teleoperation procedure. Thus, technological savviness may have influenced some children's questionnaire answers. Second, and relatedly, the fact that we only notified the children about the pre-programmed nature of the interaction during the debriefing phase may be a limitation of our study, because prior awareness of the robot's teleoperated working may further reduce the experience of interacting with a humanlike "other" one can befriend.

Third, the robot-unrelated information that was provided in the control condition may have resonated more with some children than with others. For instance, when the experimenter told the children in the control condition something about soccer matches, children who played soccer themselves often showed excitement. It is conceivable that this excitement transferred to their evaluation of the robot and caused children to feel more similar or closer to Nao, as though the robot was the one bringing up soccer. Still, we tried to minimize such potential effects by addressing a wide variety of topics.

Fourth and finally, our study exclusively relied on self-report measures—which tend to be subject to social desirability bias when used with children [e.g., 5]. However, socially desirable answers are generally most likely in response to sensitive questions [37, 65], where they pose the greatest threat to the correctness and accuracy of the response. The questionnaire of this study contained no sensitive items (see Appendix C). In addition, there were no obviously desirable answer options. Therefore, the potential influence of children's social desirability bias on our findings should be limited.

Still, self-report measures only tap into children's *conscious* perceptions of the robot. In contrast, automatic, unconscious responses can hardly be verbalized as they do not access our awareness [e.g., 36]. As such, self-report measures cannot adequately assess these responses [28]. In this context, it should be pointed out that Kory Westlund et al. [39] found some subtle differences in children's gaze behavior when a robot was introduced as a friend compared to when it was introduced as a machine. This may indicate children's unconscious view of the robot as a more or less social actor. At the same time, Kory Westlund et al. [39] found no differences on other social behavioral measures (e.g., speech, smiling, and courteous behavior). Behaviors such as gaze are equivocal, and care should be taken with interpreting them in isolation [14]. Still, a more comprehensive assessment of children's responses to robots requires the triangulation of self-report findings with the outcomes of physiological and/or behavioral measures [e.g., 6].

We have three suggestions for future research. First, whereas we were transparent about the robots' lack of human capacities, we did not reveal the WOZ set-up of the study and children's responses to the debriefing revealed that most of them had not noticed that the robot had been remotely controlled. In line with a research proposal by Kory Westlund and Breazeal [38], we think it would be valuable to investigate whether transparency about teleoperation would influence children's responses to social robots as well. Making children aware of teleoperation is particularly relevant to the agency dimension of mind perception [27].

Second, transparency may have different effects with a different type of robot (e.g., zoomorphic, mechanomorphic). Since we used an anthropomorphic robot, children's reasoning was probably influenced by the appearance and behavior of the robot. Whereas research has shown that children attribute similar biological, psychological, and intelligence characteristic to both an anthropomorphic and a mechanomorphic robot [9], the manifestation of concepts such as anthropomorphism and perceived similarity is likely to be prone to the humanlike appearance of a robot. However, as it was timely to gain insight into children's evaluations of humanoid robots, we do not consider this a limitation of our design.

Third and finally, it should be investigated whether transparency has a similar effect in different CRI settings and with various groups of children. In our study, social interaction was the sole purpose for which children encountered the robot. It might be that, when children interact with a robot in a learning or healthcare setting, they form a less 'social' representation of the robot, such that closeness may be affected as well. Furthermore, it would be interesting to see whether controlling for children's technological savviness would influence the effects of transparency. Last, investigating our hypotheses among different age groups would shed light on the developmental characteristics that may play a role in how children perceive, and form relationships with, social robots.

To conclude, our experimental study elucidated the influences of transparency about a robot's lack of human psychological capacities on child-robot perception and relationship formation. Whereas transparency did affect children's perceptions of, and trust in, the robot, it did not influence their feelings of closeness toward the robot.

## APPENDIX A: EXPERIMENTER TEXTS

Table A1. English Translations of Experimenter Texts

| | Control condition | Transparent condition |
|---|---|---|
| 1 | Maybe you have never heard that name before. You can never know all existing names, because there are so many different ones. Not only people but also animals, countries, and streets have names. Sometimes two things also have the same name. | There are many robots exactly like this one. Those are all called Nao! If you would replace this Nao-robot with a different one, you wouldn't notice that. Just like when you switch between computers: You don't notice that either, because every computer does exactly the same. |
| 2 | Nao just asked about your favorite color, but there are many more colors than you could come up with, because, of every color, there are many different kinds. | Nao just asked about your favorite color, but robots don't have favorite things. Robots are machines, and machines don't like or dislike anything. |
| 3 | There are many riddles that resemble this one, because the solution to all of them is a funny fantasy word. A while ago, someone asked me, for instance: What's green and sings all the time? And the answer was: lalala-lettuce. | Nao can tell this riddle because it was put into Nao's computer. Nao can't come up with riddles or solve them on its own, because robots can't think on their own. They are just like computers: They can't do that, either! |
| 4 | Nao gave the example that people are happy when they win a match. Soccer matches are often broadcast on TV, especially during the world championships. Then people dress up in orange clothing and watch the Dutch team play. | Nao gave the example that people are happy when they win a match, but you and I get happy for many more reasons. Robots like Nao don't: They consist of plastic and wires and can't be happy or sad. Nao doesn't feel anything! |
| 5 | Of course, there exist many different kinds of games. During some games you just have to be lucky, while during others, like chess, you have to think long and hard and can only win if you make exactly the right moves. | If Nao loses a game, Nao doesn't care. After all, Nao is a machine, and machines don't notice it when they lose. Just like a computer doesn't realize when it loses a game against someone. |
| 6 | If you can predict what your adversary is going to do, that's super useful during games like chess. But it is quite hard, because you never know for sure whether your prediction is correct. That's why you always have to be lucky as well. | Nao can't do that: It can't imagine what kids will do when, or what they are thinking about. That's logical, because Nao can't even think for itself or choose what to do: That's all determined by the computer inside it. |
| 7 | To become a veterinary or a teacher, you have to study for quite a long time, because you have to know a lot of things to be able to cure all animals or to teach children all subjects. That just takes a lot of time! | In a few years, you will be very different than you are right now, but Nao will never change. It will always stay exactly the same. That's no problem, because just like machines, robots do what people want them to do, but don't want anything themselves. |
| 8 | The story that Nao just told you is a bit like a fairytale, of which you probably know many. Fairytales often feature all kinds of fantasy creatures, like dragons, wizards, and witches. And sometimes also houses that are made out of cookies! | The story that Nao just told you has been put into Nao's computer beforehand. Because robots can only say what is in their computer. Therefore, you can't really talk to Nao: When you say something that comes to your mind, Nao can't react to that. |
| 9 | I really can't choose. What do you like best of those two? [...] Actually, I think that everybody would like that. While there are also things that almost nobody likes to eat, such as pea soup or carrots! | Try to give another answer, for instance, pizza. [...] See? Nao can't respond to that properly: In Nao's computer, there are only sentences about fries or pancakes. |

(Continued)

Table A1. Continued

| | Control condition | Transparent condition |
|---|---|---|
| 10 | I don't know what kind of dance this was exactly. On TV, you see all kinds of dance, from hiphop to ballet dancing. | Nao can only do this dance because someone put all the rules of the dance into Nao's computer. |
| 11 | That's completely right, unless of course you participate in a match in which judges decide who becomes number 1. Such as in a playback show. But most of the times, it doesn't matter what others think about you. In playback shows, you can choose the music that you like best. | But children often hope that other children will like them. Robots can't think: not about themselves and also not about people. Robots can therefore never be worried about what someone else may think of them. Just look what happens if I put Nao backwards. |
| 12 | But of course, there are many other kinds of music to which you can dance. Ballet dancers, for example, dance to classical music and hiphop dancers dance to rap music. | Did you hear that? Nao doesn't know that he's standing backwards, or that people think it's strange when you aren't facing them when you talk to them. That's why he just goes on. |
| 13 | That's nice. Most of the time, it's especially fun to dance with others, just like it's more fun to play together than alone. | See? Nao wanted to point to you, but didn't know that you were somewhere else. Children would know such a thing. |
| 14 | Unless you play a game like Go Fish. Then you have to collect as many cards as possible, and when you have three of the same kind, you shouldn't give them away anymore! | But robots can't really play together. Nao doesn't even know that we are here and so can't really be aware of people. |

## APPENDIX B: ANSWER SCALE

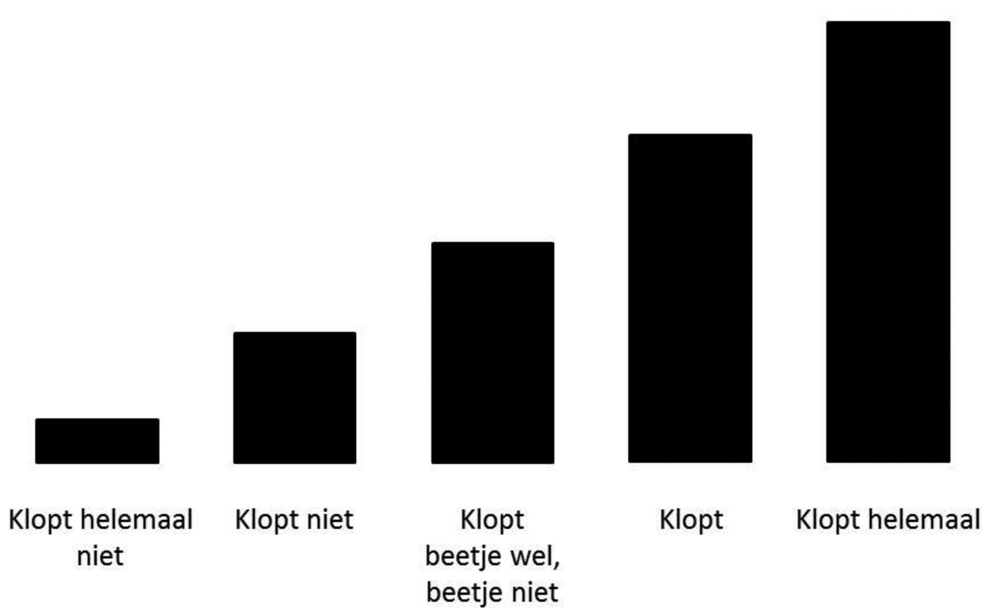


Fig. B1. Answer scale Dutch labels (adjusted from Severson and Lemm [57]).

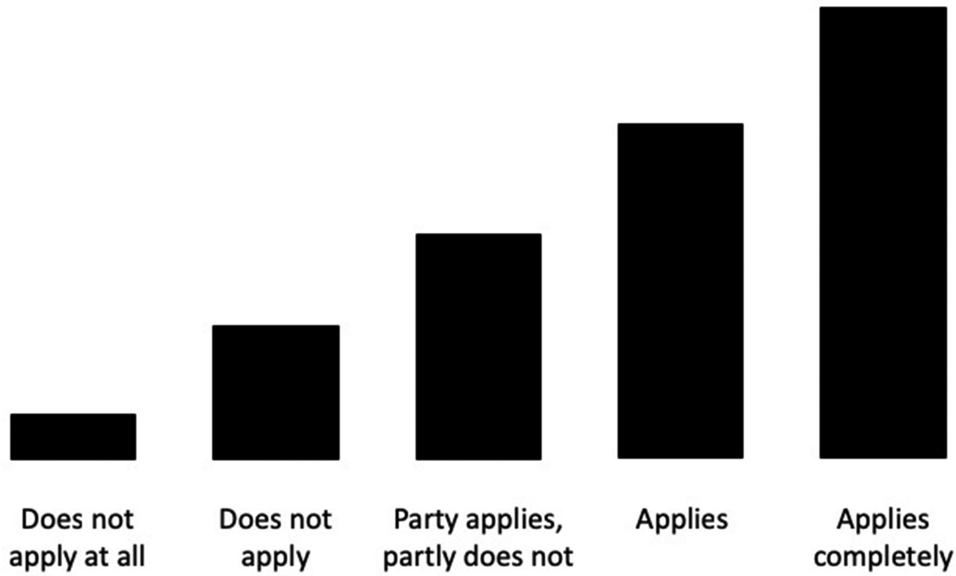


Fig. B2. Answer scale English labels (adjusted from Severson and Lemm [57]).

## APPENDIX C: QUESTIONNAIRE

Table C1. Indicators of Animacy

| Item | Dutch | Backtranslations into English |
|---|---|---|
| 1 | Nao is een levend wezen. | Nao is a living creature. |
| 2 | Nao is iets levends. | Nao is alive. |
| 3 | Nao kan doodgaan. | Nao can die. |
| 4 | Nao leeft. | Nao lives. |

Table C2. Indicators of Anthropomorphism (Godspeed)

| Item | Dutch | Backtranslations into English |
|---|---|---|
| 1 | Nao lijkt op een machine. | Nao is like a machine. |
| 2 | Nao is nep. | Nao is fake. |
| 3 | Nao kan goed bewegen. | Nao can move well. |
| 4 | Nao herinnert zich dat we net met elkaar hebben gepraat. | Nao remembers that we just talked to each other. |
| 5 | NAO is een soort persoon. | Nao is a kind of person. |

Table C3. Visual Semantic Differentials

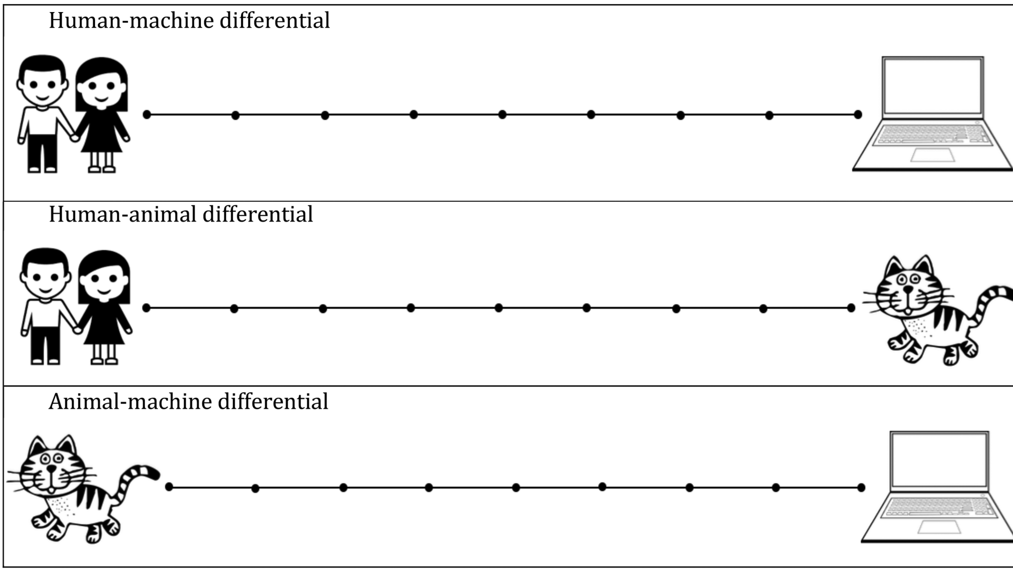


Image sources:
"Human": Emojione BW 1F46B by Emojione (https://github.com)
"Machine": Laptop black white by Dhanesh95 (https://commons.wikimedia.org)
"Animal": christels / 1010 images (https://pixabay.com)

Table C4. Indicators of Anthropomorphism (IDAQ-CF)

| Item | Dutch | Backtranslations into English |
|---|---|---|
| 1 | Nao kiest zelf wat Nao doet. | Nao chooses what Nao does. |
| 2 | Nao kan blij en boos zijn. | Nao can be happy and angry. |
| 3 | Nao weet dat Nao een robot is. | Nao knows that Nao is a robot. |
| 4 | Nao kan zelf nadenken. | Nao can think for itself. |

Table C5. Indicators of Social Presence

| Item | Dutch | Backtranslations into English |
|---|---|---|
| 1 | Toen ik met Nao aan het praten was, voelde het alsof Nao een echt mens was. | When I was talking to Nao, it felt as though Nao was a real person. |
| 2 | Toen ik met Nao aan het praten was, voelde het alsof ik met een mens praatte. | When I was talking to Nao, it felt as though I was talking to a person. |
| 3 | Toen ik met Nao aan het praten was, leek Nao net een echt mens te zijn. | When I was talking to Nao, Nao nearly seemed to be a real person. |
| 4 | Toen ik met Nao aan het praten was, voelde het alsof ik samen met een mens was. | When I was talking to Nao, it felt as though I was with a person. |

Table C6. Indicators of Perceived Similarity

| Item | Dutch | Backtranslations into English |
|---|---|---|
| 1 | Nao denkt zoals ik. | Nao thinks like me. |
| 2 | Nao doet zoals ik. | Nao acts like me. |
| 3 | Nao is zoals ik. | Nao is like me. |
| 4 | Nao lijkt op mij. | Nao resembles me. |

Table C7. Indicators of Closeness

| Item | Dutch | Backtranslations into English |
|---|---|---|
| 1 | Nao is een vriendje. | Nao is a friend. |
| 2 | Ik voel me op mijn gemak als ik met Nao ben. | I feel comfortable around Nao. |
| 3 | Nao en ik zijn vriendjes aan het worden. | Nao and I are becoming friends. |
| 4 | Nao en ik passen goed bij elkaar. | Nao and I are a good match. |
| 5 | Nao voelt als een vriendje voor mij. | Nao feels like a friend to me. |

Table C8. Indicators of Trust

| Item | Dutch | Backtranslations into English |
|---|---|---|
| 1 | Ik heb het gevoel dat ik Nao kan vertrouwen. | I feel that I can trust Nao. |
| 2 | Ik heb het gevoel dat Nao een geheim van mij kan bewaren. | I feel that Nao can keep one of my secrets. |
| 3 | Ik heb het gevoel dat Nao eerlijk is. | I feel that Nao is honest. |
| 4 | Ik heb het gevoel dat te vertrouwen is. | I feel that Nao is trustworthy. |

Table C9. Treatment Check

| Item | Dutch | Backtranslations into English |
|---|---|---|
| 1 | Caroline vertelde mij dat er heel veel robots precies zoals Nao bestaan, die ook allemaal Nao heten. | Caroline told me that there are many robots exactly like Nao, which are all called Nao too. |
| 2 | Caroline vertelde mij dat niet alleen mensen, maar ook dieren, landen, en straten namen hebben. | Caroline told me that not only people but also animals, countries, and streets have names. |
| 3 | Caroline vertelde mij dat in sprookjes vaak wezens zoals draken, tovenaars, en heksen voorkomen. | Caroline told me that fairytales often feature creatures like dragons, wizards, and witches. |
| 4 | Caroline vertelde mij dat Nao alléén dingen kan doen en zeggen die van tevoren in Nao's computer zijn gezet. | Caroline told me that Nao can only do and say things that have been put into Nao's computer beforehand. |
| 5 | Caroline vertelde mij dat Nao een machine is. | Caroline told me that Nao is a machine. |
| 6 | Caroline vertelde mij dat er allerlei verschillende soorten dans en muziek zijn. | Caroline told me that there are many kinds of dance and music. |

## ACKNOWLEDGMENTS

We express our gratitude to Alexandre Mazel, who generously provided us with Choregraphe code blocks and assisted us with their implementation. We also thank the team of student-assistants that helped us collect the data for this study, as well as the participating schools.

## REFERENCES

[1] American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC.

[2] D. R. Ames. 2004. Inside the mind reader's toolkit: Projection and stereotyping in mental state inference. *J. Pers. Soc. Psychol.* 87, 3 (2004), 340–353. DOI: https://doi.org/10.1037/0022-3514.87.3.340

[3] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi. 2009. Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots. *Int. J. Soc. Robot.* 1, 1 (2009), 71–81. DOI: https://doi.org/10.1007/s12369-008-0001-3

[4] N. Bauminger-Zviely and G. Agam-Ben-Artzi. 2014. Young friendship in HFASD and typical development: Friend versus non-friend comparisons. *J. Autism Dev. Disord.* 44, 7 (2014), 1733–1748. DOI: https://doi.org/10.1007/s10803-014-2052-7

[5] T. Belpaeme, P. Baxter, J. de Greeff, J. Kennedy, R. Read, R. Looije, and M. C. Zelati. 2013. Child-robot interaction: Perspectives and challenges. In *Proceedings of the International Conference on Social Robotics (ICSR'13)*. Springer, Cham, Switzerland, 452–459. DOI: https://doi.org/10.1007/978-3-319-02675-6_45

[6] C. L. Bethel and R. R. Murphy. 2010. Review of human studies methods in HRI and recommendations. *Int. J. Soc. Robot.* 2, 4 (2010), 347–359. DOI: https://doi.org/10.1007/s12369-010-0064-9

[7] S. Bhamjee, E. Griffiths and J. Palmer. 2010. Children's perception and interpretation of robots and robot behaviour. In *Proceedings of the International Conference on Human-Robot Personal Relationships (HRPR'10)*. Springer, Leiden, The Netherlands, 42–48. DOI: https://doi.org/10.1007/978-3-642-19385-9_6

[8] R. D. Benford and D. A. Snow. 2000. Framing processes and social movements: An overview and assessment. *Annu. Rev. Sociol.* 26, 1 (2000), 611–639. DOI: https://doi.org/10.1146/annurev.soc.26.1.611

[9] D. Bernstein and K. Crowley. 2008. Searching for signs of intelligent life: An investigation of young children's beliefs about robot intelligence. *J. Learn. Sci.* 17, 2 (2008), 225–247. DOI: https://doi.org/10.1080/10508400801986116

[10] E. Berscheid and P. Regan. 2005. *The Psychology of Interpersonal Relationships*. Pearson, NJ.

[11] M. Boden, J. Bryson, D. Caldwell, K. Dautenhahn, L. Edwards, S. Kember, and T. Sorrell. 2017. Principles of robotics: Regulating robots in the real world. *Connect. Sci.* 29, 2 (2017), 124–129. DOI: https://doi.org/10.1080/09540091.2016.1271400

[12] J. Borenstein and R. Arkin. 2019. Robots, ethics, and intimacy: The need for scientific research. In *On the Cognitive, Ethical, and Scientific Dimensions of Artificial Intelligence*, D. Berkich and M. d'Alfonso (Eds.), Springer, 299–309. DOI: https://doi.org/10.1007/978-3-030-01800-9_16

[13] K. Bumby and K. Dautenhahn. 1999. Investigating children's attitudes towards robots: A case study. In *Proceedings of the Third International Cognitive Technology Conference (CT'99)*. Michigan: M.I.N.D. Lab, San Francisco, CA, 391–410. Retrieved from https://pdfs.semanticscholar.org/f6a5/9f92b02a956856964c5778f59f03fa4ab3ce.pdf.

[14] J. N. Cappella. 1983. Conversational involvement: Approaching and avoiding others. In *Nonverbal Interaction*, J. M. Wiemann and R. P. Harrison (Eds.), 113–148. Sage, Beverly Hills, CA.

[15] C. M. Carpinella, A. B. Wyman, M. A. Perez, and S. J. Stroessner. 2017. The robotic social attributes scale (ROSAS): Development and validation. In *Proceedings of the 12th International Conference on Human-Robot Interaction (HRI'17)*. ACM/IEEE, Vienna, Austria, 254–262. DOI: https://doi.org/10.1145/2909824.3020208

[16] M. Coeckelbergh. 2017. How to describe and evaluate "deception" phenomena: Recasting the metaphysics, ethics, and politics of ICTs in terms of magic and performance and taking a relational and narrative turn. *Ethics Info. Technol.* 20, 2 (2017), 71–85. DOI: https://doi.org/10.1007/s10676-017-9441-5

[17] E. C. Collins. 2017. Vulnerable users: Deceptive robotics. *Connect. Sci.* 29, 3 (2017), 223–229. DOI: https://doi.org/10.1080/09540091.2016.1274959

[18] K. Darling. 2017. Who's Johnny? Anthropomorphic framing in human-robot interaction, integration, and policy. In *Robot Ethics 2.0*, P. Lin, G. Bekey, K. Abney and R. Jenkins (Eds.). Oxford University Press. DOI: https://doi.org/10.2139/ssrn.2588669

[19] C. de Jong, R. Kühne, J. Peter, C. L. van Straten, and A. Barco. 2018. Intentional acceptance of social robots: Development and validation of a self-report measure for children. *Manuscript Resubmitted for Publication.*

[20] C. de Jong, P. Vogt, and E. Krahmer. 2017. *Don't Be Sad, You Can Still Win Next Time, Robin! The Effect of a Social Robot's Presence on Children's Emotional Expressions after Receiving A Satisfying or Disappointing Gift*. Unpublished master's thesis. Tilburg University, Tilburg, The Netherlands.

[21] D. Dennett. 1988. Conditions of personhood. In *What Is a Person?* M. F. Goodman (Ed.), 145–167. Humana Press, Clifton, NJ.

[22] A. Edwards. 2018. Animals, humans, and machines: Interactive implications of ontological classification. In *Human-Machine Communication: Rethinking Communication, Technology, and Ourselves*, A. Guzman (Ed.), 29–49. Peter Lang, New York, NY.

[23] N. Epley and A. Waytz. 2010. Mind perception. In *Handbook of Social Psychology*, S. T. Fiske, D. T. Gilbert and G. Lindzey (Eds.), 498–541. Wiley & Sons, New York, NY.

[24] N. Epley, A. Waytz, and J. T. Cacioppo. 2007. On seeing human: A three-factor theory of anthropomorphism. *Psychol. Rev.* 114, 4 (2007), 864–886. DOI: https://doi.org/10.1037/0033-295x.114.4.864

[25] M. S. Geerdts. 2016. (Un) real animals: Anthropomorphism and early learning about animals. *Child Dev. Perspect.* 10, 1 (2016), 10–14. DOI: https://doi.org/10.1111/cdep.12153

[26] D. George and M. Mallery. 2010. *SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update* (10th ed.). Pearson, Boston, MA.

[27] H. M. Gray, K. Gray, and D. M. Wegner. 2007. Dimensions of mind perception. *Science* 315, 5812 (2007), 619. DOI: https://doi.org/10.1126/science.1134475

[28] A. G. Greenwald and M. R. Banaji. 1995. Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psych. Rev.* 102, 1 4–27. DOI: https://doi.org/10.1037/0033-295x.102.1.4

[29] N. Haslam, Y. Kashima, S. Loughnan, J. Shi, and C. Suitner. 2008. Subhuman, inhuman, and superhuman: Contrasting humans with nonhumans in three cultures. *Soc. Cogn.* 26, 2 (2008), 248–258. DOI: https://doi.org/10.1521/soco.2008.26.2.248

[30] M. Heerink, B. Kröse, V. Evers, and B. Wielinga. 2010. Assessing acceptance of assistive social agent technology by older adults: The Almere model. *Int. J. Soc. Robot.* 2, 4 (2010), 361–375. DOI: https://doi.org/10.1007/s12369-010-0068-5

[31] C.-C. Ho and K. F. MacDorman. 2010. Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Comput. Hum. Behav.* 26, 6 (2010), 1508–1518. DOI: https://doi.org/10.1016/j.chb.2010.05.015

[32] E. V. Hoff. 2005. Imaginary companions, creativity, and self-image in middle childhood. *Creat. Res. J.* 17, 2 (2005), 167–180. DOI: https://doi.org/10.1207/s15326934crj1702&3_4

[33] F. P. Hubbard. 2011. "Do androids dream?" Personhood and intelligent artifacts. *Temple Law Rev.* 83 (2011), 405–474.

[34] P. H. Kahn, H. E. Gary, and S. Shen. 2013. Children's social relationships with current and near-future robots. *Child Dev. Perspect.* 7, 1 (2013), 32–37. DOI: https://doi.org/10.1111/cdep.12011

[35] S. Kiesler, A. Powers, S. R. Fussell, and C. Torrey. 2008. Anthropomorphic interactions with a robot and robot–like agent. *Soc. Cogn.* 26, 2 (2008), 169–181. DOI: https://doi.org/10.1521/soco.2008.26.2.169

[36] J. F. Kihlstrom. 1990. The psychological unconscious. In *Handbook of Personality: Theory and Research*, O. John, R. Robins, and L. Pervin (Eds.), 424–442. Guilford, New York.

[37] M. King and G. Bruner. 2000. Social desirability bias: A neglected aspect of validity testing. *Psychol. Market* 17, 2 (2007) 9–103. DOI: https://doi.org/10.1002/(sici)1520-6793(200002)17:2⟨79::aid-mar2⟩3.0.co;2-0

[38] J. M. Kory Westlund and C. Breazeal. 2016. Transparency, teleoperation, and children's understanding of social robots. In *Proceedings of the 11th International Conference on Human-Robot Interaction (HRI'16)*. ACM/IEEE, Christchurch, New Zealand, 625–626. DOI: https://doi.org/10.1109/hri.2016.7451888

[39] J. M. Kory Westlund, M. Martinez, M. Archie, M. Das, and C. Breazeal. 2016. Effects of framing a robot as a social agent or as a machine on children's social behavior. In *Proceedings of the 25th International Symposium on Robot and Human Interactive Communication (RO-MAN'16)*. IEEE, New York, NY, 688–693. DOI: https://doi.org/10.1109/roman.2016.7745193

[40] R. E. Larzelere and T. L. Huston. 1980. The dyadic trust scale: Toward understanding interpersonal trust in close relationships. *J. Marriage Fam.* 42, 3 (1980), 595–604. DOI: https://doi.org/10.2307/351903

[41] K. M. Lee. 2004. Presence, explicated. *Commun. Theor.* 14, 1 (2004), 27–50. DOI: 10.1093/ct/14.1.27

[42] I. Leite, A. Pereira, and J. F. Lehman. 2017. Persistent memory in repeated child-robot conversations. In *Proceedings of the Conference on Interaction Design and Children (IDC'17)*. ACM, Stanford, CA, 238–247. DOI: https://doi.org/10.1145/3078072.3079728

[43] J. Li. 2015. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots, and virtual agents. *Int. J. Hum.-Comput. St.* 77 (2015), 23–27. DOI: https://doi.org/10.1016/j.ijhcs.2015.01.001

[44] K. Macintosh and C. Dissanayake. 2006. Social skills and problem behaviours in school aged children with high functioning autism and asperger's disorder. *J. Autism Dev. Disord.* 36, 8 (2006), 1065–1076. DOI: https://doi.org/10.1007/s10803-006-0139-5

[45] J. C. McCroskey, V. P. Richmond, and J. A. Daly. 1975. The development of a measure of perceived homophily in interpersonal communication. *Hum. Commun. Res.* 1, 4 (1975), 323–332. DOI: https://doi.org/10.1111/j.1468-2958.1975.tb00281.x

[46] V. Morrow. 1998. My animals and other family: Children's perspectives on their relationships with companion animals. *Anthrozoös* 11, 4 (1998), 218–226. DOI:https://doi.org/10.2752/089279398787000526

[47] C. Nass and Y. Moon. 2000. Machines and mindlessness: Social responses to computers. *J. Soc. Issues* 56, 1 (2000), 81–103. DOI:https://doi.org/10.1111/0022-4537.00153

[48] A. I. Nathanson and J. Cantor. 2000. Reducing the aggression-promoting effect of violent cartoons by increasing children's fictional involvement with the victim: A study of active mediation. *J. Broadcast Electron.* 44, 1 (2000), 125–142. DOI:https://doi.org/10.1207/s15506878jobem4401_9

[49] A. I. Nathanson, B. J. Wilson, J. McGee, and M. Sebastian. 2002. Counteracting the effects of female stereotypes on television via active mediation. *J. Commun.* 52, 4 (2002), 922–937. DOI:https://doi.org/10.1093/joc/52.4.922

[50] U. Neisser. 1988. Five kinds of self-knowledge. *Philos. Psychol.* 1, 1 (1988), 35–59. DOI:https://doi.org/10.1080/09515088808572924

[51] K. L. Nowak and F. Biocca. 2003. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence-Teleop. Virt.* 12, 5 (2003), 481–494. DOI:https://doi.org/10.1162/105474603322761289

[52] Y. Pearson and J. Borenstein. 2014. Creating "companions" for children: The ethics of designing esthetic features for robots. *AI Soc.* 29, 1 (2014), 23–31. DOI:https://doi.org/10.1007/s00146-012-0431-1

[53] S. F. Rosaen and J. L. Dibble. 2008. Investigating the relationships among child's age, parasocial interactions, and the social realism of favorite television characters. *Commun. Res. Report* 25, 2 (2008), 145–154. DOI:https://doi.org/10.1080/08824090802021806

[54] B. R. Schadenberg, M. A. Neerincx, F. Cnossen, and R. Looije. 2017. Personalising game difficulty to keep children motivated to play with a social robot: A Bayesian approach. *Cogn. Syst. Res.* 43 (2017), 222–231. DOI:https://doi.org/10.1016/j.cogsys. 2016.08.003

[55] M. Scheutz. 2014. The inherent dangers of unidirectional emotional bonds between humans and social robots. In *Robot Ethics: The Ethical and Social Implications of Robotics*, P. Lin, K. Abney and G. A. Bekey (Eds.), 205–221. MIT Press, Cambridge, MA.

[56] R. L. Severson and S. M. Carlson. 2010. Behaving as or behaving as if? Children's conceptions of personified robots and the emergence of a new ontological category. *Neural Netw.* 23, 8-9 (2010), 1099–1103. DOI:https://doi.org/10.1016/j.neunet.2010.08.014

[57] R. L. Severson and K. M. Lemm. 2016. Kids see human too: Adapting an individual differences measure of anthropomorphism for a child sample. *J. Cogn. Dev.* 17, 1 (2016), 122–141. DOI:https://doi.org/10.1080/15248372.2014.989445

[58] N. Sharkey and A. Sharkey. 2010. The crying shame of robot nannies: An ethical appraisal. *Interact. Stud.* 11, 2 (2010), 161–190. DOI:https://doi.org/10.1075/is.11.2.01sha

[59] M. Sherman, M. Hertzig, R. Austrian, and T. Shapiro. 1981. Treasured objects in school-aged children. *Pediatrics* 68, 3 (1981), 379–386.

[60] R. J. Sternberg. 1987. Liking versus loving: A comparative evaluation of theories. *Psychol. Bull.* 102, 3 (1987), 331–343. DOI:https://doi.org/10.1037/0033-2909.102.3.331

[61] S. Turkle. 2007. Authenticity in the age of digital companions. *Interact. Stud.* 8, 3 (2007), 501–517. DOI:https://doi.org/10.1075/is.8.3.11tur

[62] S. Turkle. 2011. *Alone Together: Why We Expect More From Technology and Less From Each Other.* Basic Books.

[63] S. Turkle, C. Breazeal, O. Dasté, and B. Scassellati. 2006. Encounters with kismet and cog: Children respond to relational artifacts. *Digital Media: Transformations in Human Communication 120.* Retrieved from http://web.mit.edu/~sturkle/www/encounters withkismet.pdf.

[64] P. M. Valkenburg. 2014. *Schermgaande Jeugd: Over Jeugd en Media.* Prometheus, Amsterdam, The Netherlands.

[65] T. F. van de Mortel. 2008. Faking it: social desirability response bias in self-report research. *Aust. J. Adv. Nurs.* 25, 4 (2008), 40–48.

[66] C. L. van Straten, R. Kühne, J. Peter, C. de Jong, and A. Barco. In press. Closeness, trust, and perceived social support in child-robot relationship formation: Development and validation of three self-report scales. *Interact. Stud.*

[67] C. L. van Straten, J. Peter, and R. Kühne. 2019. Child-robot relationship formation: A narrative review of empirical research. *Int. J. Soc. Robot.* DOI:https://doi.org/10.1007/s12369-019-00569-0

[68] P. Vogt, M. de Haas, C. de Jong, P. Baxter, and E. Krahmer. 2017. Child-robot interactions for second language tutoring to preschool children. *Front. Hum. Neurosci.* 11, 73 (2017). DOI:https://doi.org/10.3389/fnhum.2017.00073

[69] B. J. Wilson and A. J. Weiss. 1993. The effects of sibling coviewing on preschoolers' reactions to a suspenseful movie scene. *Commun. Res.* 20, 2 (1993), 214–248. DOI:https://doi.org/10.1177/009365093020002003

Received April 2019; revised September 2019; accepted October 2019