



Comparing the Effects of a Different Social Partner (Social Robot vs. Human) on Children's Social Referencing in Interaction

Nils F. Tolksdorf^{1*}, Camilla E. Crawshaw² and Katharina J. Rohlfing¹

¹ Faculty of Arts and Humanities, Psycholinguistics, Paderborn University, Paderborn, Germany, ² Faculty of Rehabilitation Sciences, Language and Communication, TU Dortmund University, Dortmund, Germany

OPEN ACCESS

Edited by:

Adriana Bus,
University of Stavanger, Norway

Reviewed by:

Xin Zhao,
East China Normal University, China
Glenn Smith,
University of South Florida,
United States

*Correspondence:

Nils F. Tolksdorf
nils.tolksdorf@upb.de

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Education

Received: 04 June 2020

Accepted: 14 December 2020

Published: 14 January 2021

Citation:

Tolksdorf NF, Crawshaw CE and
Rohlfing KJ (2021) Comparing the
Effects of a Different Social Partner
(Social Robot vs. Human) on
Children's Social Referencing in
Interaction. *Front. Educ.* 5:569615.
doi: 10.3389/feduc.2020.569615

Social robots have emerged as a new digital technology that is increasingly being implemented in the educational landscape. While social robots could be deployed to assist young children with their learning in a variety of different ways, the typical approach in educational practices is to supplement the learning process rather than to replace the human caregiver, e.g., the teacher, parent, educator or therapist. When functioning in the role of an educational assistant, social robots will likely constitute a part of a triadic interaction with the child and the human caregiver. Surprisingly, there is little research that systematically investigates the role of the caregiver by examining the ways in which children involve or check in with them during their interaction with another partner—a phenomenon that is known as social referencing. In the present study, we investigated social referencing in the context of a dyadic child–robot interaction. Over the course of four sessions within our longitudinal language-learning study, we observed how 20 pre-school children aged 4–5 years checked in with their accompanying caregivers who were not actively involved in the language-learning procedure. The children participating in the study were randomly assigned to either an interaction with a social robot or a human partner. Our results revealed that all children across both conditions utilized social referencing behaviors to address their caregiver. However, we found that the children who interacted with the social robot did so significantly more frequently in each of the four sessions than those who interacted with the human partner. Further analyses showed that no significant change in their behavior over the course of the sessions could be observed. Findings are discussed with regard to the caregiver's role during children's interactions with social robots and the implications for future interaction design.

Keywords: child-robot interaction, early childhood education, social referencing, children, social robots, humanoid robots

INTRODUCTION

In recent years, there has been a growth in studies examining young children's interactions with social robots as partners in learning environments (Belpaeme et al., 2018). Benefiting from the presence of an embodied social agent and the ability to use various social signals, social robots could offer wide-ranging opportunities to support and expand early childhood education by providing

new ways to engage children in social interaction. The areas of application range from language learning (Vogt et al., 2019) to the promotion of children's growth mindsets (Park et al., 2017) or supporting children's development of computational thinking skills (Ioannou and Makridou, 2018). Social robots can also be used in therapy-centered activities by delivering interventions to children on the autistic spectrum to support their social and emotional communication abilities (Boccanfuso et al., 2017; Cao et al., 2019). Whereas, social robots may assist young children in their learning across these various fields, the common purpose of approaches incorporating this technology in educational practices is typically to supplement the educational process rather than to replace the human caregiver, e.g., the teacher, parent, educator or therapist. More precisely, for both ethical and present technological reasons, a social robot is seen as a tool which can play a potential supportive role within an interaction in an educational setting but not as the sole interaction partner of the child (Coeckelbergh et al., 2016; Kennedy et al., 2016; Tolksdorf et al., 2020b).

Given this preferred configuration of a triadic interaction between the child, caregiver, and a social robot, there is surprisingly little research paying explicit attention to the role of the caregiver or systematically examining the ways in which children involve or check in with them during their interaction with a robot. From current studies, it seems clear that children accept social robots as informants (Breazeal et al., 2016; Oranç and Küntay, 2020) that need, however, to be introduced by a caregiver to establish a good learning environment (Vogt et al., 2017). More specifically, relating to the phenomenon of social referencing, Rohlfling et al. (2020a) argue that young children react with uncertainty when facing a robot for the first time; in such a situation, children typically refer visually to their caregivers to regulate their emotions and to gauge the situation (Hornik et al., 1987). Although Rohlfling et al. (2020a) noticed that some children at the age of 5 years verbally referred to their caregiver during an educational child-robot interaction, little is known about how often children non-verbally request support from their caregivers during such interactions. To investigate this phenomenon, non-verbal behavior has to be considered. Recently, Tolksdorf and Mertens (2020) showed that in child-robot interactions, children make use of non-verbal signals to a large extent, especially when engaging in a complex communicative task, such as retelling an event or retrieving a newly acquired word from memory. It is thus reasonable to focus the investigation of the role of the caregiver in child-robot interaction on the systematic similarities and differences between children's multimodal interaction behavior with robots vs. human partners. Furthermore, some effects upon children's behavior appear to occur in the first exposure and disappear when children familiarize with the situation (Feinman et al., 1992). In this respect, the literature lacks a perspective that considers children's social referencing during a long-term interaction occurring across multiple points in time. Our study addresses this research gap: We explored how pre-school children involved their caregiver over the course of a long-term language learning study, comparing children's behavior when interacting with a social robot or a human

interaction partner within two conditions that were designed to be directly comparable.

Our motivation for the direct comparison stems from research providing wide evidence that an adult can be a helpful resource for a child experiencing an unfamiliar situation, providing them with guidance on how to interpret and navigate it (Feinman et al., 1992). However, within the area of child-robot interaction, very few attempts have been made to consider the caregiver's role during triadic interactions between the child, the robot, and the caregiver. From a more implicit angle, Vogt et al. reported that at the beginning of language lessons with a social robot, certain pre-school children needed assistance and had to be encouraged by the caregiver to interact with the robot in order to respond to it (Vogt et al., 2019). This observation strongly suggests that a caregiver plays a crucial role when a novel partner is first introduced. This is corroborated by a study which showed that even very young children at the age of 1 year often extended their dyadic interaction with a Keepon robot into triadic interactions with their human caregivers when they were trying to share pleasure and surprise within the interaction (Kozima and Nakagawa, 2006). Whereas, children obviously benefit from caregivers' involvement at the beginning of a novel situation through aligning with their emotional interpretation, they also seem to receive specific cues for how to communicatively manage the interaction. A study by Serholt (2018) thoroughly analyzed dialogical breakdowns during an educational child-robot interaction at a primary school in Sweden. They demonstrated that although the children were able to solve some dialogical problems on their own, in most of the cases they were dependent on the human caregiver, especially when technological problems occurred (Serholt, 2018). In a more recent work, Rohlfling et al. (2020a) systematically investigated the caregivers' role during a single learning situation within a child-robot interaction and focused on the verbal ways in which the caregiver provided support to their child. The results revealed that a caregiver did not have to adopt an active role during the interaction but provided valuable instructions on how to repair the interaction when dialogical breakdowns occurred (Rohlfling et al., 2020a). We need to emphasize, however, that in this particular study, the analysis of children's social referencing behavior was limited to verbal turns toward the caregiver, such as requesting help explicitly.

Together, these few studies indicate that triadic interactions often emerge within child-robot interactions when a caregiver is present and that children particularly rely on the caregiver at certain stages. However, because prior work has not directly compared child-robot interactions to ones with a human partner, research can only speculate on what is typical of social referencing behavior during interaction with a robotic partner. Current insights are limited to single, one-off interactions or are based on observations made in the context of dialogical or technological problems occurring with the robotic system. Our study extends previous work by systematically investigating children's visual social referencing over a long-term interaction within parallel learning situations across two different conditions: interaction with either a social robot or a human partner. We assumed that children in both groups would involve their caregivers because

they were all faced with a novel and unfamiliar situation. Our main research goals were to explore the extent to which a child would involve their caregiver over the course of a long-term interaction as well as to examine the similarities and differences in children's social referencing between a child-robot interaction and a human-human interaction within an educational setting. The following hypotheses were addressed in our study:

(H1) We expect that an interaction with a social robot will lead to more social referencing in children compared to an interaction with a human partner. This hypothesis is grounded in research demonstrating that children rely on their caregivers during novel situations, including encounters with social robots (Kozima and Nakagawa, 2006; Serholt, 2018; Rohlfing et al., 2020a). A robot, which differs in its behavior from a human, might be perceived as less familiar by the children, meaning children's greater familiarity with human partners in contrast to social robots will lead to fewer attempts to involve the caregiver during a human-human interaction.

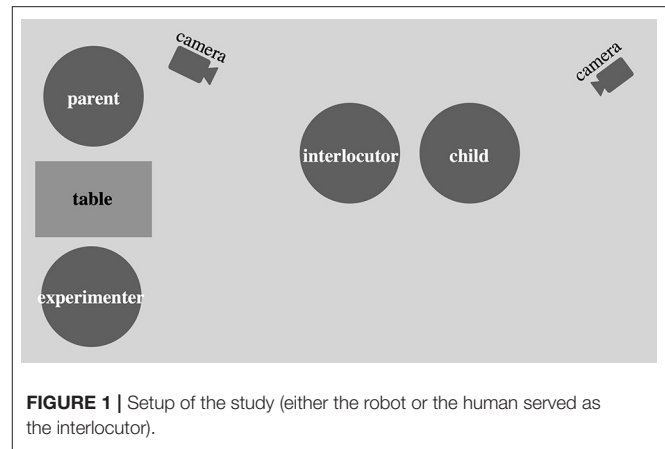
(H2) It is also expected that instances of children's social referencing will decrease over the course of a long-term interaction. This is anticipated because prior research has shown that children's social referencing varies in relation to their familiarity with a situation (Walden and Baxter, 1989). Moreover, with increasing repetition within an interaction, the interactional demands become more predictable for the child (Bruner, 1983; Rohlfing et al., 2016). This might result in children becoming less dependent on guidance from their caregivers.

MATERIALS AND METHODS

This study is part of a broader ongoing study in which we investigate how children learn a specific linguistic structure within a recurrent interaction. In this study, our main goal was to investigate children's social referencing, and how this behavior may differ when children interact with either a novel social robot or an unfamiliar human interaction partner.

Participants

Originally, 21 pre-school children participated in our explorative long-term study. Data from one child had to be excluded because they did not attend all sessions. For this reason, 20 children (6 females), ranging in age from 4;0 to 5;8 [years;months] (mean age = 5;0, $SD = 0.6$) were included in the final analysis, and their behavior was assessed across four sessions. The children were recruited in local kindergartens, libraries, via newspapers, and through our database of families willing to participate in research studies. The children and their parents were recruited from the wider areas of the Paderborn region (North Rhine-Westphalia, Germany). Parents were present during all interactions but were not actively involved in the interaction. In compliance with university ethics procedures for research with children, parents provided written consent prior to their children's participation. Children also provided verbal assent prior to taking part in the interaction and the interaction could be discontinued at any time



at no disadvantage to the child. Each child received stickers and a toy to thank them for their participation.

General Procedure

The children and their parents were invited to come to the laboratory at Paderborn University for four sessions within a 2-week period. Each session lasted ~20–35 min and all sessions were video recorded. The course of the sessions and the learning situation was explained to the parents by the experimenter; it was also communicated to the children in a child-oriented way. In our study we used a between-subjects experimental design with two conditions. The children were randomly assigned to a parallel learning situation with either (1) the social robot or (2) the human interlocutor. The final distribution consisted of 11 children (four females) interacting with the social robot and nine children (two females) interacting with the human interlocutor. During the experiment, every child taking part was accompanied by one parent; as illustrated in **Figure 1**, the child sat next to the interlocutor (either the robot or the human) at a 90-degree angle. The parent sat to the left of the child while the experimenter sat behind the child and operated the robot in the condition with the robot or sat in that position and avoided interaction with the parent or child in the condition with the human interlocutor. Parents were additionally instructed to avoid talking to their children during the experimental part of the children's interaction with the robot or the human interlocutor.

To design the learning situation, we were guided by existing theoretical concepts of learning, emphasizing that communication is jointly organized by the interaction partners in a multimodal way and toward a goal (Rohlfing et al., 2016). The resulting design of the learning setting therefore involved activities with which pre-school children are familiar. More specifically, a story was told by the robot or the human that had been created to frame the word learning situation. The story contained the plot of the interlocutor's trip to our university and the things they had seen on their journey. This narrative served as a context in which the novel words were provided as input over the course of the interaction. This setting was selected because previous work has shown the context of a story to be particularly facilitative for children's word learning (e.g.,



FIGURE 2 | Teaching of the novel word either by (A) the human or (B) the robot interloctor.



FIGURE 3 | Test situation with either the human (A) or the robot interloctor (B).

Horst, 2013; Nachtigäller et al., 2013). The referents of the target words were presented as pictures hanging on the wall. They were covered by a small cloth and the child had to uncover each one to see the target referent at the request of the interloctor (see **Figure 2**).

After the exposure to the training situation, children from both groups were tested for retention of the target words. To do so, we used a routinized activity for children and embedded the test procedure within a shared picture book reading situation (Grimminger and Rohlfing, 2017). In this test the child was asked to turn the pages while the interloctor talked about the pictures with the child and elicited the trained words (see **Figure 3**).

In the following sections, we present the details of the experimental procedures for each condition (firstly the robot condition and then the human condition). We further detail our endeavors to make the learning settings as similar as possible.

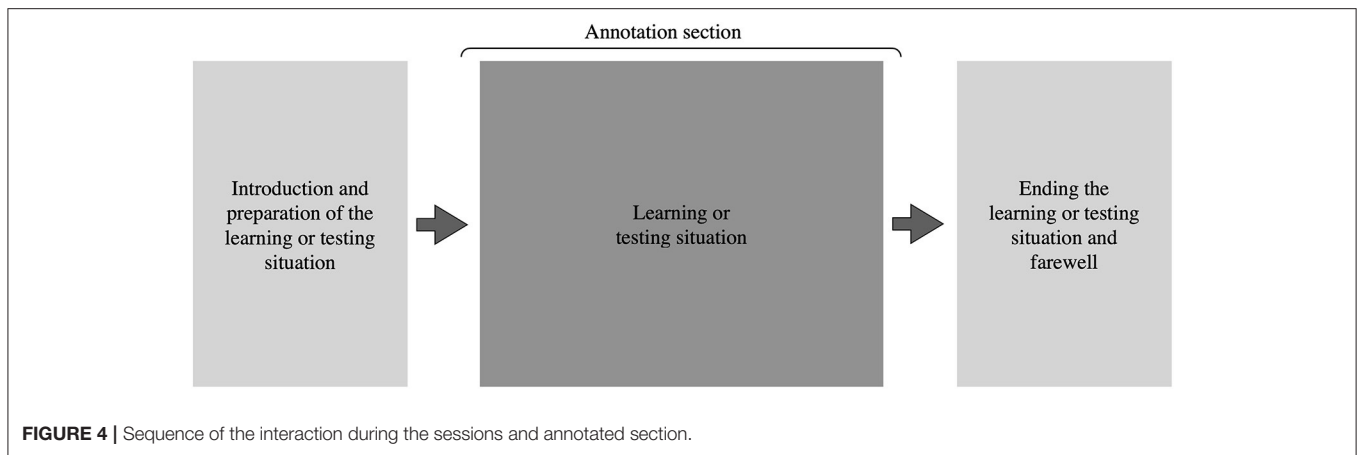
Procedure With the Robot Interloctor

In the first of the four sessions, there was a short warm-up phase in order to decrease the novelty effect of the robot (Kanero et al., 2018). The script for the learning situation with the robot was then launched, during which the robot introduced itself and shared the story containing the new words with the child. To make the robot's interaction behavior responsive to the child's communicative actions (Tolksdorf and Mertens, 2020) and in accordance with recent research postulating an important role of multimodal joint activities (Rohlfing et al., 2016, 2019), the robot performed a series of actions. First, the robot accompanied the novel words with pointing gestures in front of its upper

body to coordinate the child's attention and to establish a shared reference. It also coordinated its gaze between the child and the referents of the target words. After naming the first four target words, the robot then walked with the child to the two remaining target referents in order to make the situation more natural and to take advantage of the physical presence of the robot (van den Berghe et al., 2019). Once the robot had finished the story, it thanked the child and said goodbye. In the second session, the learning situation was repeated and the robot told the same story but adapted its greeting and farewell. In the third session, a similar learning situation took place. Afterwards, the retention task was administered, in which the child and the robot were engaged in the shared picture-book-reading situation. In the fourth session, the child was tested again on retention of the target words.

Procedure With the Human Interloctor

As in the robot condition, the experimental procedure in the human condition consisted of four sessions between the child and a human interloctor. The human interloctor was the same person in all of the sessions across the condition and a non-native speaker of German (one of the authors). This design decision was made to address the novelty of the robot as an unfamiliar interaction partner and to render the word-learning story told to the child as plausible. Like in the robot condition, a caregiver was present as well as another experimenter further to the one acting as the interloctor (see **Figure 2**). In the human condition, the interloctor also introduced herself in the first session with a short backstory, whereupon the learning situation began. Following the introduction, the same story and stimuli from the robot condition were used during the learning situation and only minor edits were made to make the story more representative of human experience (e.g., "born in..." rather than "built in..."). The human interloctor used an equal amount of deictic pointing gestures at matching points in the story to the robot condition, carrying these out within the same upper-body area and holding the gestures for an equivalent duration of time. Parallel to the robot condition, the interloctor further coordinated her gaze between the child and the referents of the target words. For purposes of achieving a fair comparison, we also wanted to keep the verbal and linguistic input in the human condition as close to the robot condition as possible and the human interloctor tried to copy these to a degree that would appear most natural. These elements included not only the language used but also factors such as emphasis and speed. As within the robot condition, the child was asked to uncover the pictures in a randomized order. After naming the first four target words, the human interloctor also suggested moving over again so that they were in range of the remaining two items. Once the human interloctor had finished telling the story, she thanked the child and said goodbye. In the second and third session, another learning situation took place and the story was told again, but with an appropriately adapted greeting and farewell by the human interloctor, just as it did in the condition with the robot. Following the learning situation in the third session, the retention task took place and involved the same shared picture book reading interaction as in the robot condition, during which the child was asked to turn the pages



and the human interlocutor asked them about the trained words. In the final (fourth) session, the retention task was conducted again after which the human interlocutor thanked the child and said goodbye.

Stimuli

The robot used in our study was the Nao robot from Softbank Robotics, which is a small, toy-like, humanoid robot used widely in child-robot interaction studies (Belpaeme et al., 2018). The Nao is 58 cm high with 25 degrees of freedom of motion with its body. Teleoperation was employed to enable the robot to act contingently (Kennedy et al., 2017). We implemented the behaviors in the NAO robot by using the *Choregraphe* Software and used the integrated text-to-speech production of the robot, with German language enabled and speech reduced to 85% speed to achieve a more natural pronunciation. The target words imparted by the robot were spoken at a speed of 75% speed in order to emphasize them verbally. The target items consisted of six morphologically complex German words (noun-adjective compounds such as “quince yellow [quittengelb]”) that represented different colors as features of different objects. Each item was presented as a picture on paper measuring 14.8×21 cm.

Coding of the Children's Behavior

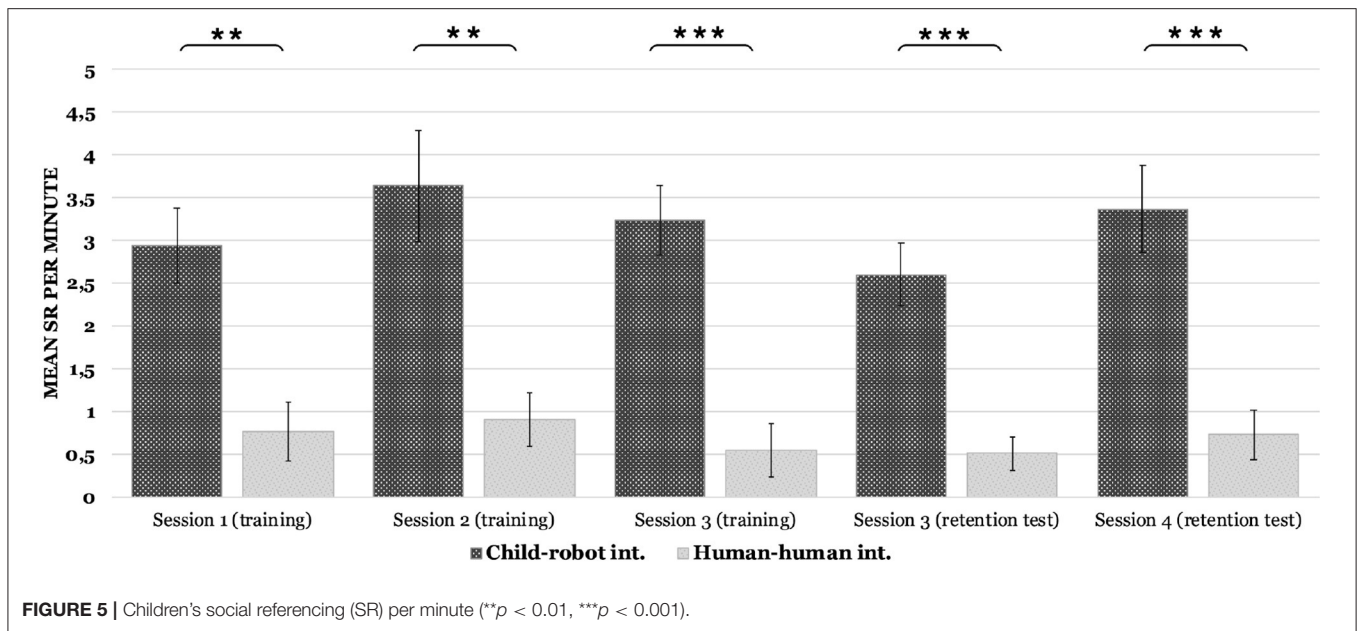
In our analysis, we were interested in children's social referencing to their caregivers during the learning and testing situations with either the social robot or the human interlocutor across the different sessions. In typical social referencing paradigms, children are confronted with unfamiliar situations, e.g., unknown interaction partners or novel toys, and then the children's non-verbal behavior in terms of their looking is measured along with the (visual or vocal) cues of their caregiver (Baldwin and Moses, 1996). In the present study, we did not analyze the kinds of cues that the caregiver provides to his or her child but rather focused on the instances in which the child independently sought to visually check in with their caregiver. Our focus on children's non-verbal behavior was also motivated by the fact that the values of the children's verbal initiation to their caregivers tended toward zero across almost all learning and testing situations. For this reason and with reference to Vaish and Striano, we coded

children's looking behaviors in the direction of their caregivers (Vaish and Striano, 2004). We measured this non-verbal social behavior across the period of time in the interaction during which the robot or the human interlocutor shared the story and taught the new words (see **Figure 4**).

We chose to analyze this sequence in each session because at this stage, all children had already achieved a certain familiarity with the novel interaction partner. This also represented the main part of the interaction while a welcome or farewell situation would represent a different social situation with its own contextually appropriate social behaviors (Vaughn and La Greca, 1992). Examining the selected sequence is particularly relevant because it provides an opportunity to understand how children involve a caregiver during a learning situation with a social robot in comparison to a human interaction partner. We additionally coded children's looking during the testing of retention of the target words (the situation in which the interlocutor asked the child about the trained words). As the duration of the interactions varies slightly between children, the children's looking to their parents were expressed in proportion per minute. To evaluate coding reliability, two coders independently coded 15% of the data. We used Cohen's kappa to measure intercoder agreement for children's looking ($\kappa = 0.954$).

RESULTS

Our data shows that children in both groups demonstrated social referencing behaviors during all their interactions with the novel interlocutor (robot or human) and attempted to involve or check in with their caregiver. At first, in order to investigate the effect of the different conditions and sessions, we performed an ANOVA type statistic (ATS), with children's social referencing as the dependent variable, condition as the between-subjects independent variable, and time as the within-subjects independent variable. Due to non-normally distributed data and small sample size, the ATS was used which represents a non-parametric equivalent of a mixed ANOVA (Akritas et al., 1997) and exactly meets the α -level while being conservative. It is robust in studying small sample sizes and longitudinal data considering its progression over time (Noguchi et al., 2012)



and has been applied in developmental approaches (Viertel, 2019). Contrary to our hypothesis, there was no main effect of time, $F_{(3,\infty)} = 0.638$, $p = 0.166$, and no significant interaction between experimental condition and time, $F_{(3,\infty)} = 0.427$, $p = 0.133$, indicating that no significant changes in children's social referencing behavior were found in either group over the entire course of the sessions, including all learning and test situations. However, there was a highly significant main effect of condition $F_{(1, 16.99)} = 49.08$, $p < 0.001$, demonstrating that children in the human condition displayed social referencing significantly less often than their peers interacting with the robotic partner.

In a second step, pairwise *post-hoc* analysis using Bonferroni correction was conducted to determine the differences between the groups in each training and testing situation. As depicted in **Figure 5**, we found significant differences between the groups in children's social referencing (SR) in the first session ($p < 0.01$) $Z = -2.929$, with a large effect size of $r = 0.65$; accordingly, children in the robot condition showed significantly more social referencing per minute ($M = 2.93$; $SD = 1.45$) than the children interacting with the human ($M = 0.75$; $SD = 1.03$).

In session 2, a significant difference with a large effect size could be observed again ($p < 0.01$, $Z = -3.154$, $r = 0.71$). Children interacting with the robot demonstrated more occurrences of social referencing ($M = 3.63$; $SD = 2.14$) than those children interacting with the human interlocutor ($M = 0.9$; $SD = 0.93$).

During the third session, in the last training situation, the groups again differed significantly ($p < 0.001$, $Z = -3.58$, $r = 0.80$). Among the children interacting with the robot, the frequency of attempts to involve their caregiver was higher ($M = 3.23$; $SD = 1.34$) than among the children interacting with the human interlocutor ($M = 0.54$; $SD = 0.93$). In the following retention task (also in session 3), the level of social referencing

among the children in the robot condition was again greater ($M = 2.6$; $SD = 1.2$) compared to the human interlocutor condition ($M = 0.50$; $SD = 0.58$), and the groups again differed significantly ($p < 0.001$, $Z = -3.459$, $r = 0.77$).

In the last session, during which the retention task was repeated, a significant difference could again be seen between the two groups ($p < 0.001$, $Z = -3.304$, $r = 0.74$). The children who interacted with the robot utilized more social referencing toward their parents ($M = 3.36$; $SD = 1.67$), while the group of children interacting with the human interlocutor were clearly less likely to involve their parents ($M = 0.72$; $SD = 0.82$).

Finally, we took a closer look at the nature of children's social referencing during the interactions with the robotic or the human partner and qualitatively analyzed at which particular stages the children involved their caregivers and sought guidance from them. **Table 1** presents an overview of the different interactional contexts in which children's social referencing was situated during the long-term interaction. We identified four interactional main contexts and three additional contexts specifically occurring during the retention tests in which the children involved their caregivers.

Certain occurrences of social referencing appeared exclusively in the interaction with the robot, such as an involvement of the caregiver after a delay in the dialogue occurred and the robot required too much time to provide an adequate utterance. Additionally, children who interacted with the robot checked in with their parents before manipulating some elements of the setting, for example, when uncovering the pictures at the request of the robot. An explanation for this type of social referencing could be that although children socially conform with what a social robot suggests (Vollmer et al., 2018), social conformity is even higher in interaction with a human interaction partner and children are less dependent on additional reassurance from their caregiver to follow an instruction. Another context in

TABLE 1 | Interactional contexts of children's social referencing (SR).

Contexts of SR	Robot condition	Human condition	Examples
Delays or interruptions in the dialogue	✓	-	The child expects an utterance from the robot
Reassurance before a manipulation of the setting	✓	-	Child is requested by the robot to uncover the pictures
New interactional event	✓	✓	A new book page is opened, the interlocutor introduces a new referent, the interlocutor performs a gesture
Naming of the unknown target words	✓	✓	The interlocutor names a target word
Test specific contexts of SR			
Reassurance before the production of a target word	✓	✓	Before producing the requested word, the child refers to the parent
No retrieval of the target word	✓	✓	The child fails to retrieve the target word and refers to the parent
Successful retrieval of a target word	✓	✓	The child is able to retrieve the target word and refers to the parent

which the children turned to their caregiver was when the target words were named while the story was being told. This type of social referencing occurred in both conditions, but was more pronounced in interaction with the robot. Since the target words were unknown to the children, it is reasonable that the children addressed their parents to obtain more information in order to dissolve the ambiguous situation. More specifically, this observation ties in with accounts suggesting that children seem to need reassurance from familiar interaction partners (such as their parents) to trust information provided by less familiar partners (Ehli et al., 2020). The last context in which social referencing occurred concerned the occurrence of a new event during the interaction. This included situations in the dialogue such as when a picture was uncovered, a new page was turned in the book during the retention test, the interlocutor began to walk to the position for the final two target referents, or when the interlocutor displayed a specific communicative means such as a gesture. In such situations, the children in the robot condition particularly extended the dyadic situation into a triadic interaction with their caregiver as they were clearly more unfamiliar with the robot's interaction behavior. For example, they often shared their positive surprise when the robot performed a pointing gesture, which on the other hand was possibly more ambiguous for the child compared to a human gesture. Thus, children's social referencing in these situations can be explained in two ways: on the one hand by their preference for sharing affective experiences with their caregivers, since they have only limited skills for downregulating their arousal on their own (Ainsworth et al., 2015). On the other hand, some of the situations could also have been unclear to a child, which led them to turn to their caregiver to disambiguate the ongoing situation (a strategy for information seeking).

During the retention tests, we observed additional contexts in which social referencing occurred: The children additionally referred to their caregivers before they produced a target word to the interlocutor, when they failed to retrieve a target word, and in cases in which they successfully retrieved a target word. The first two contextual types can be attributed to children's attempt to gather information or reassurance within the ambiguous situation as well as receiving emotional support from their familiar social partner when confronted with the situation of

not being able to produce the requested word. Children's social referencing occurring in the context of a successful production of a target word mostly reflected the child's intention to share its positive affect with their caregiver and its joy in successfully contributing to the communicative task.

To summarize, our analysis of children's social referencing during the interactions revealed that each child involved their caregiver and used them as a resource during their interaction with a novel interlocutor. In this vein, we identified several contextual factors in which children consistently involved their caregivers, indicating that the familiar social partner fulfilled diverse functions during the entirety of the long-term interaction. However, in accordance with our first hypothesis, the children who interacted with the robot were significantly more likely to approach their caregiver across all sessions (during both the language learning situations as well as the test situations) than the children who interacted with the human interlocutor. Regarding our second hypothesis that children's social referencing would decrease over the course of the long-term interaction, our results appear to demonstrate the reverse, indicating that social referencing occurred consistently during the interactions, especially during the interaction with the robot.

DISCUSSION

Our study involved a long-term interaction in an educational setting through which we explored how children interacting with an unfamiliar partner—either a social robot or a human—involved their caregiver by displaying behavior known as social referencing. The motivation for our approach was informed by recent research suggesting that a caregiver might serve as a helpful resource in an interaction between a child and a social robot (Serholt, 2018; Rohlfsing et al., 2020a). In terms of novel aspects, the present study has focused on children's non-verbal means of social referencing and analyzed children's looking behavior within an educational setting of language learning over multiple sessions. Overall, our results show that not only do children socially refer to their caregiver in a novel learning situation with a social robot and that they do so significantly more often than during an interaction with an unfamiliar human interlocutor, but that this behavior also persists long-term: Across

four sessions, we found continuous social referencing to the caregiver that was significantly more pronounced in the group of children interacting with a social robot than in the group interacting with an unfamiliar human partner. Contrary to our prior assumption, we could not observe a significant decrease in children's social referencing in both groups despite the repetition of the interaction and increasing familiarity with the situation. Whereas, there appeared to be a slight decreasing tendency from the second to the third learning situation in each group, this trend may have been slowed down by the subsequent novel situation of the retention task, which again increased children's reliance on the caregiver despite increasing familiarity with the interaction partner. This could be supported by the observation that some types of social referencing occur specifically in the test situations. Since children's social referencing is sensitive to contextual variation (Feinman et al., 1992), we had expected the difference between the groups to change during a test situation (an interaction where the imparted knowledge was now being assessed). Surprisingly, the group effect persisted and we could not observe a significant change between the situations in terms of children's involvement of their caregivers. Thus, children interacting with a social robot turned to their caregivers more often than children interacting with an unfamiliar human partner. A reason for the lack of increase in the test situation could be that children became familiar with the interlocutor as well as the interactional demands at that point in time, since three interactions had already taken place before the first retention test was administered.

The large difference in children's social referencing behavior between an interaction with the human vs. robotic partner is striking. One explanation for our findings is that a human partner naturally responds to various social cues (Kahle and Argyle, 2014) from the child in ways that social robots are not yet capable of, given their present technological limitations. More specifically, current social robots are not yet able to adapt to their social partner by rapidly processing and responding to this on-line non-verbal communicative information (Belpaeme et al., 2018). In contrast, in more familiar human-human interaction, a child might expect to share and exchange perceptual information and emotional attitudes with an interaction partner such as the caregiver (Butterworth and Jarrett, 1991; Tomasello, 1999). These exchanges not only enable the child to temporally synchronize with the interaction partner but also to monitor the state of the interaction partner and to establish a mutual understanding of the unfolding situation (Kozima et al., 2004). In fact, despite practice to reduce it, the human partner was observed to naturally use many subtle backchanneling signals in response to the children. This might mean that the robot does not meet certain expectations about the ongoing social interaction or that it is not possible to establish a mutual monitoring in the same way with a robotic partner. In other words, because the robot does not pick up on the child's non-verbal cues, the "flow of conversation" (Wrede et al., 2010) is disrupted resulting in children looking for reassurance or other solutions to repair. It appears that the process of mutual monitoring during an unfolding interaction represents "a key mechanism of a social interaction enabling

partners to align and jointly act" (Rohlfing et al., 2020b: 4). According to Clark and Krych (2004), a successful dialogue requires the speaker to monitor both their own actions and the understanding of the addressee during the interaction and, if necessary, adapt their actions to the addressee, who in turn continuously provides the speaker with information about their current level of understanding. For example, research has demonstrated that if the interactants are unable to monitor each other, they make eight times more errors than if they benefit from mutual monitoring (Clark and Krych, 2004). This is even more pertinent to settings where the goal is to convey knowledge to a learner. Moreover, this joint bilateral process is multimodal, involving verbal and non-verbal means of communication (Vollmer et al., 2010). Although the learning and testing procedures in the present study were kept consistent between the conditions and we endeavored to design the settings as similarly as possible regarding the verbal and non-verbal input, it is conceivable that the children interacting with the human partner could take advantage of some social adaptation or confirmation (e.g., mutual gaze or nodding) that the children interacting with the robot could not. Consequently, those children interacting with the robot were more frequently dependent on their caregiver as an additional resource for interpreting the ongoing situation and overcoming the disruption in the flow of conversation.

Along these lines, we further observed that the timing of the social robot's interactional behavior (the fluency of its turn-taking) might also have contributed to the greater proportion of social referencing among the children interacting with it. In multimodal reciprocal interaction between humans, children are used to rapidly exchanging utterances via multiple channels such as speech, gaze or gesture (Rohlfing et al., 2019). This fluent exchange is organized in a way in which the interactants' contributions can minimally overlap (Levinson, 2016). Although we employed teleoperation in order to enable the robot to act contingently with an appropriate timing and manner, the higher latency of the robot's responses compared to those of the human interaction partner was unavoidable. Due to this increased latency, it is possible that the children were confused, which might have resulted in them seeking guidance from their caregiver. In the study by Rohlfing et al. (2020a), it was shown that caregivers' suggestions during communicative breakdowns were about what to expect from a differently acting partner and how to cope with it: Children were advised to speak louder, to repeat themselves or to wait. Clearly, children thus expect and rely upon concrete coping strategies solicited from their caregivers for how to repair the flow of a conversation. In contrast, children interacting with the human interlocutor might have benefited from a more fluid interaction with a minimum of interruptions or delays, which could have decreased children's need for or reliance on guidance from their caregiver in such situations. In fact, in the present study, social referencing during delays in the dialogue was only observed within the robot condition. The duration of the interaction with the robot further supports this interpretation as in many cases, it lasted longer in comparison to the human-human interaction, although the

human interlocutor deliberately slowed down the speed of her utterances.

A considerable strength of our study relates to the close parallels between the designs of the two conditions, allowing us to make finely grained comparisons of their outcomes. This also played out in the positioning of both the robot and human interaction partners as peer-learners or informal tutors for the child. The backstory during the familiarization phase in the robot condition was designed to reinforce this interpretation and this was mirrored in the human condition. As the human interaction partner was not a native speaker of German, it was credible that they had learnt these new words and concepts and that the child (as a young native speaker of German) could be a peer learner. We actually observed one of the participants making reference to this backstory during a later experimental session when they asked: *"Did you just learn 'blue' when you came to Germany?"* In this case, the human interaction partner replied: *"Yes, in my language we have a different word for it,"* in order to answer the child in as authentically a way as possible.

With respect to the methodology applied in this study, we acknowledge the fact that other possibilities exist for the assessment of children's social referencing. Whereas, we decided to operationalize social referencing in line with typical paradigms and measured the cases in which the child visually checked in with their caregiver, one could also think of verbal behavior as a form of explicitly turning to the caregiver. In our study, however, hardly any child used this form of social referencing suggesting that our sample may have been too young to use this form. Another decision that we made concerned the amount of social referencing that we considered on average for each child. An alternative would be to look at how social referencing as a behavior develops (increases or decreases) individually over the course of one session. However, such an analysis of the development within a session would have depended heavily on the individual and his or her temperamental characteristics such as shyness (de Rosnay et al., 2006). Thus, we argue, when focusing on changes within a session, an assessment is only valuable if it is considered in relation to individual personality traits (Tolksdorf et al., 2020a), which were not examined in the work presented here.

We would also like to point to the possibility that the study design and procedure could have impacted our results. Adapting the design of the interaction from the robot experimental setting to be suitably comparable when taking place with a human interaction partner required us to make certain decisions. These pertained to verbal, non-verbal, and pragmatic aspects which relate to children's expectations about social roles and behaviors. In fact, this process of designing a comparable setting highlighted to us the difficulty inherent in controlling for potential behavioral differences between interactions and in realizing an appropriate introduction of the interaction partner so as to avoid biasing the data. The pilot sessions were therefore a crucial part of the design process. We reviewed these videotaped sessions as a group, noting and making revisions for the human interaction partner's verbal input and delivery, including non-verbal signals. As a consequence, the human interaction partner attempted to limit their visual checking as

well as the coordination of their gaze between the child and the target word referents in order to be comparable with the robot. As already mentioned above, they further endeavored to inhibit their nodding and other confirmatory signals such as corrective feedback and praise. During the interaction, the human interaction partner also observed that many of the children sought eye contact with her beyond the points in the learning situation specifically scripted for it, to which it was challenging not to respond instinctively. Although implicit, non-verbal, subconscious behaviors like nodding and eye contact were therefore very difficult to consistently control for, the human interaction partner was more successful at constraining explicit verbal feedback.

LIMITATIONS

As already suggested above, there are some limitations to our study regarding the generalizability of our results. First, our obtained results were likely influenced by the specific setting employed. The specific social robot used and the human interlocutor and their social behavior during the learning situation may have affected the child's behavior. A different social robot such as a pet-like or semi-humanoid social robot (Neumann, 2020) as well as another human interlocutor may have led to other results. Second, it is also important to emphasize that this study is limited by a relatively small sample size. We have to point out, however, that by conducting a long-term study over multiple sessions, the repeated measurement of the variable of interest over time strengthens the replicability and robustness of our findings (Smith and Little, 2018), while also allowing us to provide a particularly nuanced view of the development of children's behavior. Even though the sample size employed is in accordance with prior studies using a similar paradigm (Mcgregor et al., 2009) and we found clear differences between the groups in each session, each with a large effect size, further studies are necessary to validate our findings.

Yet another possible limitation concerns the representativeness of our results because the two groups were not balanced in gender. In this respect, past research suggests that children's social referencing to their parents is generalizable across gender (Feinman et al., 1992; Aktar et al., 2013). This finding indicates that similar behavior can be expected across gender groups, which is supported by the fact that gender-specific effects are traditionally low in the literature of social referencing (de Rosnay et al., 2006). Third, despite our best attempts to minimize differences between conditions, it is possible that some non-verbal and subconscious processes of familiarization and socialization could have already played out before the learning-situation part of the interaction began. For example, the child could already have been subconsciously aware that their caregiver had accepted the human interaction partner as a non-threat simply due to their presence and social cues during the conversation about the ethics procedures with the experimenter. However, this level of contact is arguably comparable to the familiarization and warm-up stages carried out in the robot condition, considering that the human

interaction partner was also unknown to the child and further represented some novelty factor to them as a person from a different country. Finally, there were a number of other issues that we had to consider involving the course of the interactions. One question concerned whether to tightly control the length of the possible interaction between the child and human interaction partner, preventing any further interaction beyond that of the learning situation. This would have been more comparable to the robot condition (switching it on and off) but less natural and could have been disruptive for the child. As a result, when the children and their caregiver were saying goodbye and leaving, there was slightly more interaction with the human interaction partner than they would have had with the robot. This was almost unavoidable because of the need to appropriately meet their social expectations. We contemplated inventing a story for the human interaction partner that would require them to leave suddenly to “go to a meeting” or “catch a train,” allowing them to cut short the interaction with the child. However, we eventually decided that this was ethically an unnecessary level of deception and that it could cause the child to develop negative ideas about the reliability of the interaction partner, regarding time-keeping for example, potentially influencing the interaction.

CONCLUSIONS AND PERSPECTIVES

In conclusion, the findings presented here have important implications for both carrying out research with social robots and implementing them within educational practice. Our results indicate that a caregiver serves as an important resource in children's interactions with digital learning tools such as social robots (at the current technological stage) and show that social referencing emerges as an important phenomenon in child–robot interaction. In this vein, our study not only revealed that children frequently initiate non-verbal exchanges between themselves and their caregivers at certain stages during an interaction with a robot, but also do so consistently over the long-term. In contrast, children who interacted in the same learning and testing situations with an unfamiliar human interlocutor addressed their caregiver to a considerably reduced amount.

At this point, we would like to be clear in our objective: When designing technology, the solution should not be to reduce social referencing within a child–robot interaction; instead, it is important to focus on where the child needs additional support in those interactions with the goal of developing more child-oriented technologies. Thus, referring to our observations above, we can postulate some crucial aspects when designing child-oriented social robots: On the one hand, educational social robots require means of monitoring the child's engagement and understanding in the interaction while simultaneously enabling the child to monitor the state of the robot. For this reason, a social robot should utilize multiple social signals to allow the child to better interpret the ongoing interaction and to allow socioemotional perception toward the robot. This ties in with recent research highlighting that emotional expressions by a robot are expected to occur at specific stages within an interaction and can be beneficial (Fischer et al., 2019). However, many of

the current social robots, including the Nao robot used here, are highly restricted in terms of their capabilities for affective expression (Song and Yamada, 2017). On the other hand, to further minimize interruptions and irritations in a child–robot dialogue, future implementations of social robots in educational contexts need a better multimodal turn-taking model in terms of the timing of the reciprocal interaction with a child (Baxter et al., 2013; Tolksdorf and Mertens, 2020). This would include, for example, a better child-specific speech recognition (Kennedy et al., 2017) to allow for a more contingent interaction and to reduce the latency of the responses of the robotic system. Clearly, the integration of these technologies will require further technical advances across a range of processes within artificial intelligence and robotics (Belpaeme et al., 2018). To conclude, addressing children's behavior and recognizing their emotional states within these interactions can inform future digital technologies and better enable their integration into the educational landscape. In future work, it would also be of interest to explore the role of a child's individual temperament in their social behavior and learning within the interaction with a social robot. This would further shed light on how suitable learning environments for children can be created in the digital world in the future.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Review Board of Bielefeld University (EUB 2014-043). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

NT, CC, and KR conceived, designed, and piloted the study and drafted the manuscript. NT recruited participants and analyzed the data. NT and CC conducted the data collection. All authors commented on, edited, and revised the manuscript prior to submission.

FUNDING

This research was supported by the Digital Society research program funded by the Ministry of Culture and Science of the German State of North Rhine-Westphalia.

ACKNOWLEDGMENTS

We gratefully thank all the parents and children who participated in this study.

REFERENCES

- Ainsworth, M. D., Blehar, M. C., Waters, E., and Wall, S. N. (2015). *Patterns of Attachment. A Psychological Study of the Strange Situation Classic*. New York, NY: Psychology Press.
- Akritis, M. G., Arnold, S. F., and Brunner, E. (1997). Nonparametric hypotheses and rank statistics for unbalanced factorial designs. *J. Am. Stat. Assoc.* 92, 258–265. doi: 10.1080/01621459.1997.10473623
- Aktar, E., Majdandžić, M., de Vente, W., and Bögels, S. M. (2013). The interplay between expressed parental anxiety and infant behavioural inhibition predicts infant avoidance in a social referencing paradigm: expressed parental anxiety and infant behavioural inhibition. *J. Child Psychol. Psychiatry* 54, 144–156. doi: 10.1111/j.1469-7610.2012.02601.x
- Baldwin, D., and Moses, L. J. (1996). The ontogeny of social information gathering. *Child Dev.* 67, 1915–1935. doi: 10.2307/1131601
- Baxter, P., Kennedy, J., Belpaeme, T., Wood, R., Baroni, I., and Nalin, M. (2013). “Emergence of turn-taking in unstructured child-robot social interactions,” in *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Tokyo: IEEE), 77–78. doi: 10.1109/HRI.2013.6483509
- Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., and Tanaka, F. (2018). Social robots for education: a review. *Sci. Rob.* 3:eaat5954. doi: 10.1126/scirobotics.aat5954
- Boccanfuso, L., Scarborough, S., Abramson, R. K., Hall, A. V., Wright, H. H., and O’Kane, J. M. (2017). A low-cost socially assistive robot and robot-assisted intervention for children with autism spectrum disorder: field trials and lessons learned. *Auton. Robot.* 41, 637–655. doi: 10.1007/s10514-016-9554-4
- Breazeal, C., Harris, P. L., DeSteno, D., Kory Westlund, J. M., Dickens, L., and Jeong, S. (2016). Young children treat robots as informants. *Top. Cogn. Sci.* 8, 481–491. doi: 10.1111/tops.12192
- Bruner, J. S. (1983). *Child’s Talk: Learning to Use Language*. 1st ed. New York, NY: W.W. Norton.
- Butterworth, G., and Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *Br. J. Dev. Psychol.* 9, 55–72. doi: 10.1111/j.2044-835X.1991.tb00862.x
- Cao, H.-L., Esteban, P. G., Bartlett, M., Baxter, P., Belpaeme, T., Billing, E., et al. (2019). Robot-enhanced therapy: development and validation of supervised autonomous robotic system for autism spectrum disorders therapy. *IEEE Robot. Automat. Mag.* 26, 49–58. doi: 10.1109/MRA.2019.2904121
- Clark, H. H., and Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *J. Mem. Lang.* 50, 62–81. doi: 10.1016/j.jml.2003.08.004
- Coeckelbergh, M., Pop, C., Simut, R., Peca, A., Pintea, S., David, D., et al. (2016). A survey of expectations about the role of robots in robot-assisted therapy for children with ASD: Ethical acceptability, trust, sociability, appearance, and attachment. *Sci. Eng. Ethics* 22, 47–65. doi: 10.1007/s11948-015-9649-x
- de Rosnay, M., Cooper, P. J., Tsigaras, N., and Murray, L. (2006). Transmission of social anxiety from mother to infant: an experimental study using a social referencing paradigm. *Behav. Res. Ther.* 44, 1165–1175. doi: 10.1016/j.brat.2005.09.003
- Ehli, S., Wolf, J., Newen, A., Schneider, S., and Voigt, B. (2020). Determining the function of social referencing: the role of familiarity and situational threat. *Front. Psychol.* 11:538228. doi: 10.3389/fpsyg.2020.538228
- Feinman, S., Roberts, D., Hsieh, K.-F., Sawyer, D., and Swanson, D. (1992). “A critical review of social referencing in infancy,” in *Social Referencing and the Social Construction of Reality in Infancy*, ed S. Feinman (Boston, MA: Springer US), 15–54. doi: 10.1007/978-1-4899-2462-9_2
- Fischer, K., Jung, M., Jensen, L. C., and aus der Wieschen, M. V. (2019). “Emotion expression in HRI – when and why,” in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Daegu: IEEE), 29–38. doi: 10.1109/HRI.2019.8673078
- Grimminger, A., and Rohlfing, K. J. (2017). “Can you teach me?: children teaching new words to a robot in a book reading scenario,” in *Proceedings WOCCI 2017: 6th International Workshop on Child Computer Interaction* (Glasgow: ISCA), 28–33. doi: 10.21437/WOCCI.2017-5
- Hornik, R., Risenhoover, N., and Gunnar, M. (1987). The effects of maternal positive, neutral, and negative affective communications on infant responses to new toys. *Child Dev.* 58:937. doi: 10.2307/1130534
- Horst, J. S. (2013). Context and repetition in word learning. *Front. Psychol.* 4:149. doi: 10.3389/fpsyg.2013.00149
- Ioannou, A., and Makridou, E. (2018). Exploring the potentials of educational robotics in the development of computational thinking: a summary of current research and practical proposal for future work. *Educ. Inf. Technol.* 23, 2531–2544. doi: 10.1007/s10639-018-9729-z
- Kahle, L. R., and Argyle, M. (2014). *Attitudes and Social Adaptation: A Person-Situation Interaction Approach*. Kent: Elsevier Science.
- Kanero, J., Geçkin, V., Oranç, C., Mamus, E., Küntay, A. C., and Gökşun, T. (2018). Social robots for early language learning: current evidence and future directions. *Child Dev. Perspect.* 12, 146–151. doi: 10.1111/cdep.12277
- Kennedy, J., Baxter, P., Senft, E., and Belpaeme, T. (2016). “Heart vs hard drive: children learn more from a human tutor than a social robot,” in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Christchurch: IEEE), 451–452. doi: 10.1109/HRI.2016.7451801
- Kennedy, J., Lemaignan, S., Montassier, C., Lavalade, P., Irfan, B., Papadopoulos, F., et al. (2017). “Child speech recognition in human-robot interaction: evaluations and recommendations,” in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (Vienna: ACM), 82–90. doi: 10.1145/2909824.3020229
- Kozima, H., and Nakagawa, C. (2006). “Social robots for children: practice in communication-care,” in *9th IEEE International Workshop on Advanced Motion Control, 2006* (Istanbul: IEEE), 768–773. doi: 10.1109/AMC.2006.1631756
- Kozima, H., Nakagawa, C., and Yano, H. (2004). Can a robot empathize with people? *Artif. Life Rob.* 8, 83–88. doi: 10.1007/s10015-004-0293-9
- Levinson, S. C. (2016). Turn-taking in human communication—origins and implications for language processing. *Trends Cogn. Sci.* 20, 6–14. doi: 10.1016/j.tics.2015.10.010
- Mcgregor, K. K., Rohlfing, K. J., Bean, A., and Marschner, E. (2009). Gesture as a support for word learning: the case of under. *J. Child Lang.* 36, 807–828. doi: 10.1017/S0305000908009173
- Nachtigäller, K., Rohlfing, K. J., and Mcgregor, K. K. (2013). A story about a word: does narrative presentation promote learning of a spatial preposition in German two-year-olds? *J. Child Lang.* 40, 900–917. doi: 10.1017/S0305000912000311
- Neumann, M. M. (2020). Social robots and young children’s early language and literacy learning. *Early Childhood Educ. J.* 48, 157–170. doi: 10.1007/s10643-019-00997-7
- Noguchi, K., Gel, Y. R., Brunner, E., and Konietzschke, F. (2012). nparLD: an R software package for the nonparametric analysis of longitudinal data in factorial experiments. *J. Stat. Softw.* 50, 1–23. doi: 10.18637/jss.v050.i12
- Oranç, C., and Küntay, A. C. (2020). Children’s perception of social robots as a source of information across different domains of knowledge. *Cogn. Dev.* 54:100875. doi: 10.1016/j.cogdev.2020.100875
- Park, H. W., Rosenberg-Kima, R., Rosenberg, M., Gordon, G., and Breazeal, C. (2017). “Growing growth mindset with a social robot peer,” in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (Vienna: ACM), 137–145. doi: 10.1145/2909824.3020213
- Rohlfing, K. J., G., Leonardi, I., Nomikou, J., Raczaszek-Leonardi, and Hüllermeier, E. (2019). Multimodal turn-taking: motivations, methodological challenges, and novel approaches. *IEEE Trans. Cogn. Dev. Syst.* 12, 260–271. doi: 10.1109/TCDS.2019.2892991
- Rohlfing, K. J., Grimminger, A., and Wrede, B. (2020a). “The caregiver’s role in keeping a child-robot interaction going,” in *International Perspectives on Digital Media and Early Literacy: The Impact of Digital Devices on Learning, Language Acquisition and Social Interaction*, eds K. J. Rohlfing and C. Müller-Brauers (London: Routledge), 73–89. doi: 10.4324/9780429321399
- Rohlfing, K. J., Cimiano, P., Scharlau, I., Matzner, T., Buhl, H. M., Buschmeier, H., et al. (2020b). Explanation as a social practice: Toward a conceptual framework for the social design of AI systems. *IEEE Trans. Cogn. Dev. Syst.* 1–12. doi: 10.1109/TCDS.2020.3044366
- Rohlfing, K. J., Wrede, B., Vollmer, A.-L., and Oudeyer, P.-Y. (2016). An alternative to mapping a word onto a concept in language acquisition: pragmatic frames. *Front. Psychol.* 7:470. doi: 10.3389/fpsyg.2016.00470
- Serholt, S. (2018). Breakdowns in children’s interactions with a robotic tutor: a longitudinal study. *Comput. Hum. Behav.* 81, 250–264. doi: 10.1016/j.chb.2017.12.030

- Smith, P. L., and Little, D. R. (2018). Small is beautiful: in defense of the small-N design. *Psychon. Bull. Rev.* 25, 2083–2101. doi: 10.3758/s13423-018-1451-8
- Song, S., and Yamada, S. (2017). “Expressing emotions through color, sound, and vibration with an appearance-constrained social robot,” in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (Vienna: ACM), 2–11. doi: 10.1145/2909824.3020239
- Tolksdorf, N. F., and Mertens, U. (2020). “Beyond words: children's multimodal responses during word learning with a robot,” in *International Perspectives on Digital Media and Early Literacy: The Impact of Digital Devices on Learning, Language Acquisition and Social Interaction*, eds K. J. Rohlfing and C. Müller-Brauers (London: Routledge), 90–102.
- Tolksdorf, N. F., Siebert, S., Zorn, I., Horwath, I., and Rohlfing, K. J. (2020a). Ethical considerations of applying robots in kindergarten settings: towards an approach from a macroperspective. *Int. J. Soc. Rob.* doi: 10.1007/s12369-020-00622-3. [Epub ahead of print].
- Tolksdorf, N. F., Viertel, F., and Rohlfing, K. J. (2020b). “Do shy children behave differently than non-shy children in a long-term child-robot interaction?: an analysis of positive and negative expressions of shyness in kindergarten children,” in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (Cambridge: ACM), 488–490. doi: 10.1145/3371382.3378367
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Vaish, A., and Striano, T. (2004). Is visual reference necessary? Contributions of facial versus vocal cues in 12-month-olds' social referencing behavior. *Dev. Sci.* 7, 261–269. doi: 10.1111/j.1467-7687.2004.00344.x
- van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., van der Ven, S., and Leseman, P. (2019). Social robots for language learning: a review. *Rev. Educ. Res.* 89, 259–295. doi: 10.3102/0034654318821286
- Vaughn, S., and La Greca, A. (1992). “Beyond greetings and making friends: social skills from a broader perspective,” in *Contemporary Intervention Research in Learning Disabilities*, ed B. Y. L. Wong (New York, NY: Springer New York), 96–114. doi: 10.1007/978-1-4612-2786-1_6
- Viertel, F. E. (2019). *Training of Visual Perspective-taking (level 1) by Means of Role Reversal Imitation in 18-20-month-olds With Particular Regard to the Temperamental Trait Shyness*. Paderborn: Paderborn University Library.
- Vogt, P., de Haas, M., de Jong, C., Baxter, P., and Kraemer, E. (2017). Child-robot interactions for second language tutoring to preschool children. *Front. Hum. Neurosci.* 11, 1–6. doi: 10.3389/fnhum.2017.00073
- Vogt, P., van den Berghe, R., de Haas, M., Hoffman, L., Kanero, J., Mamus, E., et al. (2019). “Second language tutoring using social robots: a large-scale study,” in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Daegu: IEEE), 497–505. doi: 10.1109/HRI.2019.8673077
- Vollmer, A.-L., Pitsch, K., Lohan, K. S., Fritsch, J., Rohlfing, K. J., and Wrede, B. (2010). “Developing feedback: how children of different age contribute to a tutoring interaction with adults,” in *2010 IEEE 9th International Conference on Development and Learning* (Ann Arbor, MI: IEEE), 76–81. doi: 10.1109/DEVLRN.2010.5578863
- Vollmer, A.-L., Read, R., Trippas, D., and Belpaeme, T. (2018). Children conform, adults resist: A robot group induced peer pressure on normative social conformity. *Sci. Rob.* 3:eaat7111. doi: 10.1126/scirobotics.aat7111
- Walden, T. A., and Baxter, A. (1989). The effect of context and age on social referencing. *Child Dev.* 60, 1511–1518. doi: 10.2307/1130939
- Wrede, B., Kopp, S., Rohlfing, K., Lohse, M., and Muhl, C. (2010). Appropriate feedback in asymmetric interactions. *J. Pragmat.* 42, 2369–2384. doi: 10.1016/j.pragma.2010.01.003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Tolksdorf, Crawshaw and Rohlfing. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.