Taylor & Francis
Taylor & Francis Group

International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: www.tandfonline.com/journals/hihc20

The Influence of Robot Social Behaviors on Second Language Learning in Preschoolers

Jun Yin, Wenjiao Guo, Wei Zheng, Mingyu Ren, Simin Wang & Yingdong Jiang

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
2024, VOL. 40, NO. 7, 1600–1608
https://doi.org/10.1080/10447318.2022.2144828

Taylor & Francis
Taylor & Francis Group

Check for updates

The Influence of Robot Social Behaviors on Second Language Learning in Preschoolers

Jun Yin[a,b], Wenjiao Guo[a,b], Wei Zheng[a,b], Mingyu Ren[a,b], Simin Wang[a,b], and Yingdong Jiang[a,b]

[a]Department of Psychology, Ningbo University, Ningbo, China; [b]Center of Group Behavior and Social Psychological Service, Ningbo University, Ningbo, China

ABSTRACT

The use of social robots has flourished in the educational field. Using them in education appropriately entails considering whether robot social behaviors influence language learning outcomes and have additive influences on the outcomes of language learning. In this study, we used a social robot (i.e., Alpha Mini) to teach Chinese preschool children ($N = 83$, recruited from a preschool; aged 6–7 years) new English words; it had the pitch of a human boy, and the words' referents were displayed on a tablet. During one learning session, the robot demonstrated one of several types of social behavior: verbal behavior (i.e., asking children to look at this robot), gestural behavior (i.e., pointing to the word's referent), both verbal and gestural behaviors, and control (i.e., purely saying target words to the children). Children were randomly assigned to one of four conditions. Their vocabulary learning outcomes were measured by pre- and posttests of the children's vocabulary of the target words. Our results show that the learning outcome in each type of social behavior was greater than zero, suggesting that any type of robot social behavior benefits word learning. Importantly, when robot social behavior included verbal demonstrations, the preschoolers learned more vocabulary than when the robot displayed other social behaviors (i.e., control and gestural behavior); however, gestural behavior showed no greater learning outcomes than the control condition. Hence, the efficacy of robot tutoring in preschoolers' second language learning depends upon a robot's social behavior, and social verbal behavior is the most effective behavior in increasing learning gains. Therefore, to achieve better learning outcomes, social verbal behavior should be critically considered to enhance robots' tutoring function.

1. Introduction

Robots are not limited to industrial manufacturing but have also progressively stepped into our daily lives and socially interact with humans (i.e., social robots; Broadbent, 2017). The use of social robots has flourished in the educational field, increasing the need for technical assistance in teaching and learning (Belpaeme et al., 2018; Toh et al., 2016). More than passively showing knowledge like other tools, social robots are able to be physically present, actively interact with learners and provide one-to-one tutoring support, similar to human teaching behaviors. As teaching is inherently a social process, with the expectation of achieving positive outcomes for learners, social robots are developed with the social characteristics of humans (i.e., social fidelity), such as humanlike appearance, emotional expression, verbal communication, and gestural demonstration (Belpaeme et al., 2018; Sinatra et al., 2021; Toh et al., 2016). To use them in education appropriately and maximize their benefits, how effective they are at achieving learning outcomes and the effects of their social attributes on learning performance must be considered.

Tutoring is a common role for social robots in education (Belpaeme et al., 2018). Due to the need for teacher-like support in education and technical limitations, robots act as knowledgeable teachers to supervise learners and provide direct support for them (Belpaeme et al., 2018; Toh et al., 2016). Additionally, the robot acting as a peer cooperates with students and learns with them (Chen et al., 2020), or the student takes on the role of an instructor and teaches the robot (Tanaka & Matsuzoe, 2012). Since the accumulation of knowledge in childhood largely relies on the teacher's assistance in the modern educational system, the tutor robot has been mainly applied to children, as confirmed by a meta-analysis showing that most studies (58%) focused on children (Belpaeme et al., 2018). Importantly, using robots in education has been proven to be effective in increasing learners' cognitive outcomes, including knowledge and comprehension (Alemi & Haeri, 2020; Belpaeme et al., 2018; Lee et al., 2011; Movellan et al., 2009). To obtain learning gains, the robot was set up with a humanlike external appearance as an avenue to emulate social fidelity (Belpaeme et al., 2018; Sinatra et al., 2021). Although few studies have directly manipulated the robot's appearance, the existence of

CONTACT Jun Yin yinjun1@nbu.edu.cn Department of Psychology, Ningbo University, No. 616 Fenghua Rd, Ningbo 315211, China
Supplemental data for this article is available online at https://doi.org/10.1080/10447318.2022.2144828.

the uncanny valley implies that robots should not be designed with an overly humanlike appearance, as when robots look almost but not perfectly human, they arouse uncanny feelings among humans, leading people to avoid them (MacDorman & Chattopadhyay, 2016; Mori, 1970; Yin et al., 2021). Nevertheless, when further removed from the uncanny valley, a humanlike appearance increases learning outcomes, as shown by a meta-analysis (Belpaeme et al., 2018). Specifically, it was found that the effect size of learning gain when using a Keepon robot ($d = 0.56$) with a yellow exterior without arms or legs was smaller than that when using a Nao robot ($d = 0.76$) with a more humanlike appearance with arms, legs, and a head. Further meta-analysis revealed that an anthropomorphic appearance has no effect on general behavioral performance (Roesler et al., 2021). Therefore, when considering promoting robot teaching outcomes, appearance should be carefully designed.

Social robots are also developed with social behavior functions, such as emotional expression, verbal communication, and gestural demonstration, to copy human tutoring (Belpaeme et al., 2018; Sinatra et al., 2021), given that accompanying teaching with social behaviors is beneficial for learning outcomes in humans. Many studies have shown that almost any social behavior exhibited by robots has a positive impact on general learning outcomes. For example, young children treated robots with social interaction ability as knowledgeable and informative interlocutors, and such a tendency was greater for robots with contingent responsiveness than for those lacking this quality (Breazeal et al., 2016). Even when robots socially interacted with children and taught them category knowledge in a sorting task, children can achieve learning outcomes with both truly present robots and virtually present robots (Kennedy et al., 2015a).

Moreover, robots with different social behaviors have been largely used to teach a first or second language, as these robots can perform various actions and enable learners to interact with their real-life physical environments, both of which are important for language learning (Kanero et al., 2018; van den Berghe et al., 2019). Importantly, the cognitive gains (e.g., pronunciation, vocabulary) from robot teaching can be extended to robot-assisted language lessons (Alemi & Haeri, 2020; Kanero et al., 2018; Lee et al., 2011; van den Berghe et al., 2019). When robots taught children second language words accompanied by social behaviors, such as playing together on a Microsoft Surface Pro 4, the learning effects lasted for a week (Rintjema et al., 2018), and even robots with minimal social behaviors, such as responding to children's choices using physical actuators while teaching words, affected vocabulary scores (Movellan et al., 2009). Moreover, when robots were set to teach children new words by telling stories, verbal behaviors with questioning and encouraging utterances significantly increased the vocabulary learned (Chen et al., 2020), and if the robots' verbal demonstrations matched the children's language ability, the children were more likely to learn the target words (Kory-Westlund & Breazeal, 2019a). Interestingly, robots that show speech entrainment and self-disclosure (e.g., shared personal information about the robot's poor speech and hearing abilities) were also more beneficial for learning target words than robots that did not show such behaviors (Kory-Westlund & Breazeal, 2019b). Although different levels of a robot's verbal behavior produced similar learning gains in the second language learning of children (Kennedy et al., 2016), the correctness of verbal demonstrations of objects was also vital to children's learning, and children were more likely to ask the accurate robot for information (Brink & Wellman, 2020). Accordingly, social behaviors aimed at increasing a robot's social fidelity are able to enhance language learning outcomes.

The above studies have revealed the important role of robot social behaviors in teaching language, but they have rarely discriminated between the types of social behavior. Robots can use different modalities to show social behaviors, including at least verbal and gestural modalities (Sinatra et al., 2021), but it is unknown whether they influence language learning outcomes similarly and whether they have additive influences on the outcomes of language learning. In human teaching, demonstrations of social behaviors to learners can increase learning outcomes, for example, by using verbal communications (Csibra & Gergely, 2009; Soudek & Soudek, 1985; Sutiyatno, 2018). According to the media equation hypothesis, social robots or automatic machines can be treated as social actors instead of simple tools; thus, people can extend the responses they exhibit toward real humans to social robots and are sensitive to the social cues delivered by robots (Reeves & Nass, 1996; Rothstein et al., 2020). Therefore, we hypothesize that robot social behaviors affect language learning.

Seemingly, the more social behaviors robots display, the greater the learning of the content taught. However, previous studies have shown that gestural behaviors sometimes do not benefit or even disrupt learning in human teaching because language-related teaching is usually implemented using the verbal modality. Furthermore, since currently available robots are more limited in their motor degrees of freedom than humans and perform gestures that look different from those of humans (de Wit et al., 2021; Vogt et al., 2017), their gestures often shift learners' attention to unrelated information (Kamlasi & Sahan, 2022; Yuka, 2009). For instance, Schleihauf et al. (2021) found that children paid attention to a robot's actions that were causally unrelated to their learning goal and copied these unrelated actions, displaying over-imitation. Children who interacted with a robot using social adaptive behaviors, including iconic gestures during teaching, did not learn a significant amount (Kennedy et al., 2015b); such a robot's effects on supporting second language tutoring were mixed when this robot performs deictic and/or iconic gestures (de Wit et al., 2021). Hence, a robot's gestural behavior during teaching, at least, would not benefit learning outcomes as much as verbal behavior and may not elicit any learning gain, as verbal behavior shares the same communicative channel as teaching language (Sinatra et al., 2021).

1. 1. The present study

The present study investigates whether the type of robot social behavior influences language learning outcomes and whether different types of social behavior have additive influences on the outcomes of language learning. To answer these questions, we used a social robot to teach Chinese preschool children (aged 6–7 years) new English words with nonsocial or minimally social utterances; importantly, during teaching, the robot displayed various types of social behavior: verbal behavior (e.g., asking children to look at this robot), gestural behavior (e.g., pointing to the word's referent), and a combination of verbal and gestural behaviors. Our reason for testing preschoolers[1] aged 6 to 7 years is that this age is usually at the beginning of second language learning in China, and the appropriate applications of robots to assist children at this age require evidence of the influential patterns of robot social behaviors when teaching language. The learning outcomes were measured by pre- and posttests of the children's vocabulary of target words, and these measurements for the various types of social behavior were compared with those at baseline simply by saying target words to the children to assess how social behavior type influences learning outcomes in a second language. Based on the above discussions on the media equation hypothesis and the different characteristics of social verbal and gestural behaviors, we specifically expected that the robot showing social verbal behavior would increase preschoolers' learning outcomes compared to the baseline, with only teaching words. However, concerning gestural behavior, there is inconsistent evidence in the relevant literature on its benefits during teaching; thus, we remained open to other possibilities.

2. Methods

2.1. Participants

Eighty-seven children aged 6–7 years were recruited from a preschool (kindergarten) located in the downtown area of a mid-sized city in China. They were in their last year before entering primary school and were from families with middle socioeconomic status. All children belonged to the Han ethnicity, the majority ethnic group in China. Four children were excluded because two of them were afraid of robots and two of them knew more than approximately 80% of the to-be-learned words before the robot teaching, resulting in eighty-three valid participants. The included participants were randomly assigned to one of four conditions as described below: 21 in the control condition (10 girls and 11 boys; $M_{age} = 6.53$, $SD = 0.50$), 21 in the verbal behavior condition (10 girls and 11 boys; $M_{age} = 6.57$, $SD = 0.51$), 20 in the gestural behavior condition (10 girls and 10 boys; $M_{age} = 6.57$, $SD = 0.51$), and 21 in the verbal and gestural behavior condition (10 girls and 11 boys; $M_{age} = 6.50$, $SD = 0.47$). Based on a meta-analysis about the influence of robot behaviors on cognitive outcomes (Belpaeme et al., 2018), the effect size can be medium ($f = 0.40$). To obtain such an effect size, this sample size was determined by using G*Power 3.1 with an alpha level of 0.05 and a power of 0.80 for a between-subjects design to compare the learning gains across four groups of participants. The suggested sample size was at least 19 individuals in each condition. Parents who agreed to their children's participation signed an informed consent form, and children verbally agreed to participate. The experimental protocol titled "social robots and relationships" was approved by the institutional review board at the Department of Psychology at Ningbo University.

2.2. Robots

The robot used was Alpha Mini (i.e., named Wukong in Chinese) with a size of 245 mm × 149 mm × 112 mm (UBTECH, Shenzhen, China; see Figure 1). This robot is highly portable and is capable of bringing fun interactions and many expressions and functionalities, including voice interaction, face recognition, and object recognition. It has fourteen servo motors and can display a variety of flexible movements, such as dancing and getting up after a fall. It can show emotion and expressions with two adorable LCD eyes. This robot's speech and movements can be controlled by scripts and accordingly are suitable for various social behaviors with verbal and gestural signals.

2.3. Materials and design

The pictures used for English word learning were taken from the book "My Body" and "Get Ready" for kids, which

Figure 1. The robot used in the present experiment.

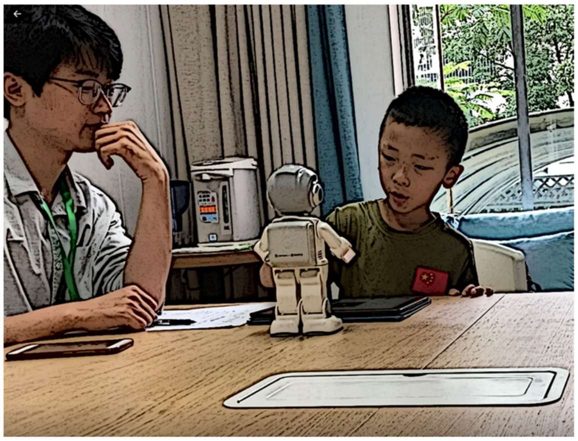

Figure 2. The setup of the robot, the child and the experimenter.

helped children learn words related to body parts and activities and were published by Harcourt School Publishers. According to the recommendations of well-experienced kindergarten teachers and prior investigations into preschoolers' English vocabularies, we selected twelve pictures (see the Supplementary figure) and accompanied English words, including "ear," "tie," "eye," "nose," "mouth," "hand," "arm," "foot," "eat," "drink," "brush," and "stretch." The two words 'ear' and 'tie' were always used as practice to help the children familiarize themselves with the procedure, and the remaining ten words were used as the target words taught by the robot. The words were spoken by the robot with the pitch of a human boy, and the pictures were displayed by a tablet that was placed in front of the robot (see Figure 2). The voice was automatically controlled by scripts, and the presentation for the picture matching the meaning of a word was remote controlled by an experimenter.

The robot was set to teach English words to children using speech, and the children were asked to learn and acquire as many words as possible. When teaching, the robot was designed to show various types of social behaviors to the children, including only speech when teaching the meaning of words (i.e., control condition without or with less social behavior), verbal behavior, gesture behavior, and both verbal and gestural behavior, based on previous studies (Brink & Wellman, 2020; Chen et al., 2020). Only the words with bold font were spoken in English, and the other words and sentences were spoken in Chinese in the descriptions below for each condition.

In the control condition, the robot gave simple greetings to the children before the teaching began and then taught ten words successively. Specifically, the robot said, "Hello, I am Wukong; it is nice to see you! Our world has a nation in which all children speak English, and let us learn this language together! Please look at the picture shown on the tablet in front of me. The child in this picture is looking at a butterfly with her eyes, and if we want to say eyes in English, it should be spelled "**eye,**" "**eye.**" The Chinese "yanjing" (眼睛) is "**eye**" in English, and the English "**eye**" is "yanjing" in Chinese. Let us see the next picture… [The remaining nine words were taught with similar speech styles] OK! We have completed our learning, and I am happy you have learned these new words!"

In the verbal behavior condition, what the robot said was almost the same as in the control condition, except that more words with social demonstrations were used to simulate verbal communication. Specifically, the robot said, "… if we want to say eyes in English, it should be spelled "**eye.**" The **Chinese** "yanjing" (眼睛) "**eye**" in English, and the English "**eye**" is "yanjing" in Chinese. Please look at me. My eyes are big. To express this in English, we should say that my "**eye**" is big. We should protect eyes, as they are the windows to the soul… "

In the gestural behavior condition, what the robot said was the same as in the control condition, except that when saying "please look at the picture shown on the tablet in front of me" it exhibited some gestural behaviors by bowing its head to look at the picture and then lifting its head to look at the child, and when saying "Let us see the next picture," it behaved by pointing to the picture with its right hand and then returning to the default gesture.

In the verbal and gesture behavior condition, what the robot said was almost the same as in the control condition, except that it showed social utterances as in the verbal behavior condition and simultaneously performed the same gestures as in the gestural behavior condition.

2.4. Procedure

The following four phases were included: the practice phase, the pretest phase, the learning phase, and the posttest phase. All of these were conducted in a separate room and video-recorded. After completing all tests, the children were asked to report what they remembered about the robot's behaviors.

In the practice phase, the children were asked to talk about the twelve displayed pictures in Chinese successively to ensure that they knew about the objects in the pictures. Meanwhile, the experimenter asked the children to point to the body part and report what the kid in the picture was doing in Chinese. If the child answered incorrectly, the experimenter corrected the child and then asked again until the correct answer was provided. Next, the experimenter removed two pictures (i.e., showing "ear" and "tie") and taught the students which picture each English word referred to, with the same utterances as in the control condition but in the experimenter's voice. Finally, the experimenter asked the child about the Chinese meaning of "ear" and "tie." The practice finished when the child was able to correctly report the Chinese meaning. This phase lasted approximately 3 minutes.

In the pretest phase, the experimenter tested how many words among the ten to-learned words the children knew in two tasks (i.e., vocabulary size before robot teaching): word translation and picture mapping. If the child knew more than eight of the words in either task, this child was excluded from further tests; hence, two children were ultimately excluded. Specifically, in the word translation task, the experimenter said "Hello, [the child's name]! Do you know the Chinese meaning of [English word]? Please say it out loud!," for each word to be learned. In the picture mapping task, ten pictures were presented on the tablet screen, and the experimenter said "Hello, [the child's name]! Do you know which picture [English word] maps onto? Please point it out!," for each word to be learned. Regardless of correctness, the experimenter did not provide any feedback. This phase lasted approximately 2 minutes.

In the learning phase, the child sat in front of the robot face-to-face, approximately 50 cm away from the robot. There was a tablet between the child and the robot to present the picture. The experimenter first introduced the robot, saying "This is Wukong. He will teach you some English words, and you should try your best to learn these words. After Wukong completes his teaching, you will be tested on how many words you have learned. You will receive a gift if you learn words carefully. Let's start." Then, the robot taught the words according to the assigned social behavior condition. The experimenter sat on the left of the child and did not do anything during the robot teaching unless the child touched the tablet incorrectly or played with the robot (Figure 2). This phase lasted approximately 5 minutes.

In the posttest phase, the children were first tested to determine how many words they had learned according to the same procedure as in the pretest phase (i.e., vocabulary size after robot teaching). This phase lasted approximately 2 minutes. To test whether the children perceived the robot's mind differently, we referred to a previous measurement of perception of mind with two dimensions: agency and experience (Brink & Wellman, 2020; Severson & Lemm, 2016). For agency, the children were asked whether the robot could think and make decisions with three items. They were also asked whether the robot could experience emotions and perceive feelings with three items. The questions were formatted on a 4-point scale coded as 0 (no), 1 (yes—a little bit), 2 (yes—a medium amount), and 3 (yes—a lot) using the stimuli of thumbs-down (no), thumbs-up (yes), and the scale with bars shown in the Supplementary figure.

2.5. Analysis

The vocabulary size was calculated for each task in the pretest and posttest phases by two undergraduates who did not know the aim of this study and sat together to watch the recorded videos to identify whether the child reported the Chinese meaning and pointed to the correct picture. To better measure learning performance according to robot teaching, we calculated the learning gain for each condition by subtracting the vocabulary size in the posttest from that in the pretest. We averaged the mean item ratings to attain a mental ability rating (Cronbach $\alpha = 0.69$) for each condition, and the mean rating for each dimension was also computed (Cronbach $\alpha_{agency} = 0.59$; Cronbach $\alpha_{experience} = 0.60$), given that agency and experience measured different perceived mental abilities. All data, analysis codes, and supplementary videos showing how the various types of social behavior were presented were made publicly available via the Open Science Framework and can be accessed at https://osf.io/upqw9/?view_only=19aec8ad2923440684c1efdff6d462a8. The data were analyzed and visualized using Python version 3.9.7 (Van Rossum & Drake, 2009) with the package Pingouin version 0.5.0 (Vallat, 2018) and the package Plotnine version 0.8.0 (https://github.com/has2k1/plotnine).

3. Results

First, we examined whether the children had the same vocabulary size before learning (i.e., in pretest) across conditions. Then, we assessed the learning gain after robot teaching. Finally, we examined their perceived mental abilities toward the robot.

3.1. Vocabulary size before robot teaching

Both tasks showed the same pattern of vocabulary size before the robot teaching (see Table 1). A one-way analysis of variance (ANOVA) was conducted on this measurement and revealed that the vocabulary size before the robot teaching did not significantly differ across types of social behavior in either task (word translation task: $F(3, 79) = 0.54$, $p = 0.657$, $\eta_p^2 = 0.02$; picture mapping task: $F(3, 79) = 0.54$, $p = 0.657$, $\eta_p^2 = 0.02$). Hence, the vocabulary size of the words to be learned before robot teaching was matched across groups of children.

3.2. Learning gain after robot teaching

Both tasks showed almost the same pattern of learning gain after robot teaching (see Table 1 and Figure 3). Concerning the word translation task, the learning gain for each type of social behavior was more than zero (control: 95% CI of difference = [0.13, 0.82], $t(20) = 2.91$, $p = 0.012$, $d = 0.64$; verbal behavior: 95% CI of difference = [1.19, 2.62], $t(20) = 5.53$, $p < 0.001$, $d = 1.21$; gestural behavior: 95% CI of difference = [0.20, 1.40], $t(19) = 2.79$, $p = 0.012$, $d = 0.62$; verbal and gestural behavior: 95% CI of difference = [0.99, 2.06], $t(20) = 5.98$, $p < 0.001$, $d = 1.31$; P values were corrected with the False Discovery Rate method, FDR), suggesting that robot teaching can help Chinese children learn English words. The ANOVA showed that the learning gain after robot teaching was significantly different across the four types of social behaviors in the word translation task ($F(3, 79) = 5.91$, $p = 0.001$, $\eta_p^2 = 0.18$). The post hoc multiple comparisons (FDR correction) found that the learning gains in both the verbal behavior condition (95% CI of difference = [0.66, 2.20], $t(40) = 3.75$, $p = 0.003$, $d = 1.16$) and the verbal and gestural behavior condition (95% CI of difference = [0.44, 1.66], $t(40) = 3.46$, $p = 0.004$, $d = 1.07$) were larger than those in the control condition, and the learning gain in the verbal behavior condition was larger than that in the gestural behavior condition (95% CI of difference = [0.20, 2.01], $t(39) = 2.47$, $p = 0.037$, $d = 0.77$); no other effects were found to be significant (ts < 1.89, ps > 0.100, ds < 0.60).

Table 1. Descriptive results of vocabulary size in different test phases for each condition (M ± SE).

	Control	Verbal behavior	Gestural behavior	Verbal and gesture behavior
Word translation task				
Before robot teaching	0.81 ± 0.49	0.48 ± 0.24	0.70 ± 0.33	1.19 ± 0.51
After robot teaching	1.29 ± 0.61	2.38 ± 0.51	1.50 ± 0.52	2.71 ± 0.63
Learning gain	0.48 ± 0.16	1.90 ± 0.34	0.80 ± 0.29	1.52 ± 0.25
Picture mapping task				
Before robot teaching	0.81 ± 0.49	0.48 ± 0.24	0.70 ± 0.33	1.19 ± 0.51
After robot teaching	1.33 ± 0.61	2.43 ± 0.51	1.55 ± 0.51	2.71 ± 0.63
Learning gain	0.52 ± 0.18	1.95 ± 0.35	0.85 ± 0.28	1.52 ± 0.25

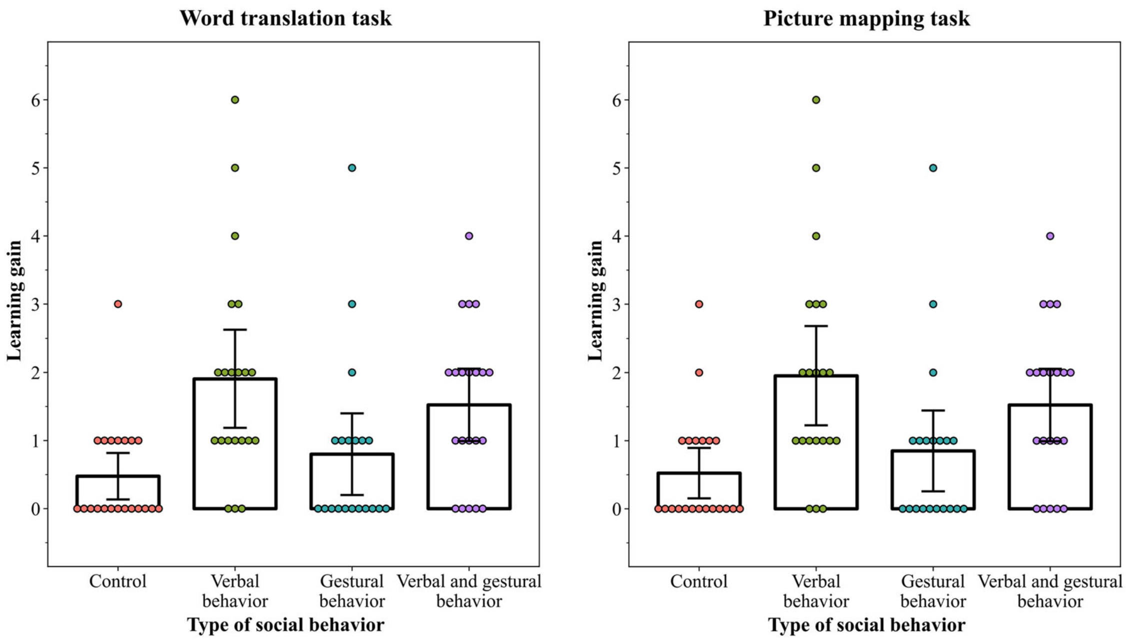

Figure 3. Learning gain after robot teaching as a function of social behavior type in each task. The dots indicate the learning gain of each participant, and the black vertical lines indicate the confidence intervals (CI) of the means.

Regarding the picture mapping task, we found the same results as in the word translation task. The learning gain for each type of social behavior was greater than zero (control: 95% CI of difference = [0.15, 0.89], $t(20) = 2.95$, $p = 0.008$, $d = 0.64$; verbal behavior: 95% CI of difference = [1.23, 2.68], $t(20) = 5.61$, $p < 0.001$, $d = 1.22$; gestural behavior: 95% CI of difference = [0.26, 1.44], $t(19) = 3.00$, $p = 0.008$, $d = 0.67$; verbal and gestural behavior: 95% CI of difference = [0.99, 2.06], $t(20) = 5.98$, $p < 0.001$, $d = 1.31$), again showing that robot teaching influences children's English learning. The ANOVA confirmed that the type of social behavior influenced the learning gain ($F(3, 79) = 5.64$, $p = 0.001$, $\eta_p^2 = 0.18$). The post hoc multiple comparisons found that the learning gains in both the verbal behavior condition (95% CI of difference = [0.64, 2.22], $t(40) = 3.65$, $p = 0.004$, $d = 1.13$) and the verbal and gestural behavior condition (95% CI of difference = [0.37, 1.63], $t(40) = 3.22$, $p = 0.008$, $d = 0.99$) were larger than those in the control condition, and the learning gain in the verbal behavior condition was larger than that in the gestural behavior condition (95% CI of difference = [0.19, 2.01], $t(39) = 2.46$, $p = 0.038$, $d = 0.76$); no other effects were found to be significant ($ts < 1.77$, $ps > 0.125$, $ds < 0.56$).

3.3. Perceived mental ability

ANOVA was conducted on the mental ability rating and found that the perceived mental ability was not significantly different across types of social behavior ($F(3, 79) = 0.69$, $p = 0.563$, $\eta_p^2 = 0.03$). For the rating of each dimension (see Table 2), no significant effect was found (agency: $F(3, 79) = 1.19$, $p = 0.319$, $\eta_p^2 = 0.04$; experience: $F(3, 79) = 0.64$, $p = 0.588$, $\eta_p^2 = 0.02$), but agency was perceived as higher than experience ($F(1, 79) = 103.58$, $p < 0.001$, $\eta_p^2 = .58$), which was not modulated by the type of social behavior ($F(3, 79) = 1.36$, $p = 0.261$, $\eta_p^2 = 0.05$). Importantly, the rated agency in each condition was significantly different from 0 ($ts > 8.94$, $ps < 0.001$, $ds > 1.95$), as was the rated experience ($ts > 3.60$, $ps < 0.003$, $ds > 0.80$).

4. Discussion

In this study, we explored how various types of robot social behavior influence preschoolers' (aged 6–7 years) learning outcomes for English words. During robot teaching, the type of social behavior was operationalized as the verbal modality, the gestural modality, and the combination of these two. It was found that any type of robot social behavior benefits

Table 2. Descriptive results of perceived mental ability for each condition (M ± SE).

	Control	Verbal behavior	Gestural behavior	Verbal and gesture behavior
Agency	2.30 ± 0.15	1.87 ± 0.17	1.95 ± 0.17	1.94 ± 0.22
Experience	1.03 ± 0.19	1.11 ± 0.19	0.73 ± 0.20	1.00 ± 0.22
Mental ability	1.67 ± 0.15	1.49 ± 0.14	1.34 ± 0.15	1.47 ± 0.20

word learning. Moreover, when robot social behavior included verbal demonstrations, the preschoolers learned more vocabulary than when the robot displayed other social behaviors lacking verbal interactions (i.e., control and gestural behavior); however, gestural behavior showed no greater learning outcomes than the control condition. These results imply that the efficacy of robot tutoring on preschoolers' second language learning depends upon the type of social behavior exhibited by the robot. Social verbal behavior is the most effective in increasing learning gains, while gestural behavior does not show a beneficial effect, at least in our study settings.

Concerning the lack of difference in learning gains between the gestural behavior and the baseline, it may be argued that the preschoolers only paid attention to the verbal signals but did not notice the presented gestural behavior. However, post hoc interviews of the participants did not support this possibility, as all participants reported that they saw the robot pointing to the pictures and turning its head. Importantly, although not statistically significant, the learning gain of communicating with both social verbal and gestural behaviors was slightly lower than that of the robot displaying only social verbal behavior. As a result, both lines of evidence demonstrated that the preschoolers who participated were aware of social gestural behavior. Since the robot's gestures clearly revealed its mechanical characteristics, is it possible that preschoolers treated the robot showing gestures more mechanically than they treated other behaviors? Nevertheless, the measurements of perceived mental abilities ruled out this alternative. Particularly, we did not find a significant difference in perceived agency or experience between verbal behavior and gestural behavior. In summary, at least in our study's settings, the robot's gestures seemed unlikely to increase learning gains.

Consistent with previous studies (Alemi & Haeri, 2020; Chen et al., 2020; Kanero et al., 2018; Lee et al., 2011; van den Berghe et al., 2019), in this study, we have demonstrated that robot teaching can benefit children, even with minimal social behavior in the control. We have also expanded on previous findings by directly manipulating various types of social behavior. Interestingly, we did not find that robot social gestural behavior increased the learning gain, even when combined with verbal behavior. This finding is in accordance with the results of previous studies that found that nonverbal social behavior did not improve learning outcomes and sometimes diverted students' attention away from the taught contents (Demir-Lira et al., 2020; Kennedy et al., 2015b; Yadollahi et al., 2018). In fact, many studies have shown that social behavior that improves learning outcomes mainly sets verbal signals to copy social fidelity (Brink & Wellman, 2020; Davison et al., 2021; Kory-Westlund & Breazeal, 2019a; 2019b; Rintjema et al., 2018). The difference in word learning between the verbal and gestural behavioral conditions may be due to the compatibility of the modality between delivering learning contents and presenting social behavior. Since the robot used verbal utterances to teach the preschoolers English words, this modality was compatible with showing social behavior. In this case, the learners easily integrated these two kinds of information, showing improved learning gains, as with human teachers. While the modality of showing gestural behavior was different from that of teaching words and given that our brain has evolved to integrate two types of behavioral signals for real human demonstrations, the imperfect robot's gestural demonstrations would make it difficult for preschoolers to integrate them; accordingly, the gestural cues did not work to assist teaching. This speculation was supported by findings that people have difficulty integrating two modalities of signals generated by automatic machines (Nocentini et al., 2019; Onyeulo & Gandhi, 2020). Further study is needed to address this speculation.

Our findings are consistent with the media equation hypothesis (Reeves & Nass, 1996; Rothstein et al., 2020), showing that children are as sensitive to social signals during robot teaching as when learning from a human (Csibra & Gergely, 2009). The different effects of verbal and gestural behaviors could also be explained in the framework of the media equation hypothesis. It is possible, considering that the verbal signals generated by the study's robot were more humanlike than the gestural signals under certain technical limitations, the robot's social verbal behavior evoked similar responses to humans and thereby exerted more influence on the learning gains of English words than its social gestural behavior. Moreover, the findings have important implications for designing social robots used in education. At least from a short-term efficacy perspective, one-to-one tutoring robots could be used in preschoolers' word learning, and any form of social behavior is able to improve learning gains. To achieve better learning outcomes, social verbal behavior should be critically considered to enhance robot tutoring functions in education, such as by using personalized words (e.g., "I," "we") and slang (e.g., We should protect eyes, as they are the windows to the soul …). Importantly, caution is required when applying social gestural behavior to a robot in a tutoring context.

4.1. Limitations and future directions

In this study, we measured learning outcomes immediately after training children in the focal words. It has been suggested that children need rest time (e.g., sleep) to consolidate newly acquired knowledge (Axelsson et al., 2016). The effect of robot social behavior on learning outcomes may be modulated by the time gap between the posttest phase and the learning phase. Future studies could thus examine the short-term effects after a short delay and even explore whether these effects last for a long time.

Furthermore, in the pretest and posttest phases, the images that were also used when training the words with the robot were used. The acquisition of new words not only shows that each word is successfully linked with the referent presented in a learning session, but also reflects that each acquired word is mapped onto the referred object in each new scene. Future studies should therefore consider the design when using a different set of images for the same words during posttest, which might identify the ability of word generalization and the deep connection between words and referred objects. This conclusion could be examined in other learning contexts as well, for instance, when learning grammar and sentences.

Although the participating children were aware of the gestural behaviors during learning, the robot (i.e., Alpha Mini) used in this study is quite small, and its arms are quite short, especially in comparison with a human finger that can point closely to the referent of a word. Thus, perhaps, its gestures were limited in guiding the attention of a learner to the educational content. Future studies should thus use a larger social robot, such as NAO, whose gesturing arms and hands are larger and more salient than Alpha Mini's, to examine the generalizability of our findings.

Moreover, social behaviors are common in more natural activities, such as storytelling. It is possible that the way that the learning activity was designed in this study may have influenced the value of certain robot gestures, such as pointing. Social behaviors could be instigated by other actions, for example, saying the name of a child and shaking the child's hand. It is also important to test how different levels of social behavior within the same modality modulate learning outcomes.

To conclude, in this study, we have demonstrated that the type of robot social behavior influences learning gains when a robot teaches preschoolers second language words, social verbal behavior is the most effective at increasing learning gains, and gestural behavior during robot tutoring should be used with caution, at least in settings similar to those of our study.

Note

1. In China, children, usually aged 7 years, start a formal and compulsory school program as primary school students. Before enrolling in primary school, 4- to 6-year-olds enter kindergarten in China, which is a preschool educational approach based on playing, singing, and practical activities, such as drawing. In our study, we tested children who were in their last year of kindergarten. Hence, we still refer to these children as preschoolers.

Ethics statement

The research was conducted in accordance with the relevant APA and authors' national ethical guidelines.

Supplementary materials

All data, analysis codes, and supplementary videos showing how the various types of social behavior were presented were made publicly available via the Open Science Framework and can be accessed at https://osf.io/upqw9/?view_only=19aec8ad2923440684c1efdff6d462a

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Fundamental Research Funds for the Provincial Universities of Zhejiang [Grant no. SJWZ2020001].

References

Alemi, M., & Haeri, N. (2020). Robot-assisted instruction of L2 pragmatics: Effects on young EFL learners' speech act performance. *Language Learning & Technology*, *24*(2), 86–103. http://hdl.handle.net/10125/44727

Axelsson, E. L., Williams, S. E., & Horst, J. S. (2016). The effect of sleep on children's word retention and generalization. *Frontiers in Psychology*, *7*, 1192. https://doi.org/10.3389/fpsyg.2016.01192

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, *3*(21), eaat5954. https://doi.org/10.1126/scirobotics.aat5954

Breazeal, C., Harris, P. L., DeSteno, D., Kory Westlund, J. M., Dickens, L., & Jeong, S. (2016). Young children treat robots as informants. *Topics in Cognitive Science*, *8*(2), 481–491.

Brink, K. A., & Wellman, H. M. (2020). Robot teachers for children? Young children trust robots depending on their perceived accuracy and agency. *Developmental Psychology*, *56*(7), 1268–1277. https://doi.org/10.1037/dev0000884

Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. *Annual Review of Psychology*, *68*(1), 627–652. https://doi.org/10.1146/annurev-psych-010416-043958

Chen, H., Park, H. W., & Breazeal, C. (2020). Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement. *Computers & Education*, *150*(5), 103836–103858. https://doi.org/10.1016/j.compedu.2020.103836

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Davison, D. P., Wijnen, F. M., Charisi, V., van der Meij, J., Reidsma, D., & Evers, V. (2021). Words of encouragement: How praise delivered by a social robot changes children's mindset for learning. *Journal on Multimodal User Interfaces*, *15*(1), 61–76. https://doi.org/10.1007/s12193-020-00353-9

de Wit, J., Willemsen, B., de Haas, M., van den Berghe, R., Leseman, P., Oudgenoeg-Paz, O., Verhagen, J., Vogt, P., & Krahmer, E. (2021). Designing and evaluating iconic gestures for child-robot second language learning. *Interacting with Computers*, *33*(6), 596–626. https://doi.org/10.1093/iwc/iwac013

Demir-Lira, Ö. E., Kanero, J., Oranç, C., Koşkulu, S., Franko, I., Göksun, T., & Küntay, A. C. (2020). L2 Vocabulary teaching by Social Robots: The role of gestures and on-screen cues as scaffolds. *Frontiers in Education*, *5*. https://doi.org/10.3389/feduc.2020.599636

Kamlasi, I., & Sahan, A. (2022). An analysis on politeness of non-verbal communication in an English as a foreign language classroom. *VELES Voices of English Language Education Society*, *6*(1), 258–270. https://doi.org/10.29408/veles.v6i1.5190

Kanero, J., Geçkin, V., Oranç, C., Mamus, E., Küntay, A. C., & Göksun, T. (2018). Social robots for early language learning:

Current evidence and future directions. *Child Development Perspectives*, *12*(3), 146–151. https://doi.org/10.1111/cdep.12277
Kennedy, J., Baxter, P., & Belpaeme, T. (2015a). Comparing robot embodiments in a guided discovery learning interaction with children. *International Journal of Social Robotics*, *7*(2), 293–308. https://doi.org/10.1007/s12369-014-0277-4
Kennedy, J., Baxter, P., & Belpaeme, T. (2015b). The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning. *ACM/IEEE International Conference on Human-Robot Interaction*, 67–74.
Kennedy, J., Baxter, P., Senft, E., & Belpaeme, T. (2016). Social robot tutoring for child second language learning. *ACM/IEEE International Conference on Human-Robot Interaction,* 231–238.
Kory-Westlund, J. M., & Breazeal, C. (2019a). A long-term study of young children's rapport, social emulation, and language learning with a peer-like robot playmate in preschool. *Frontiers in Robotics and AI*, *6*(81), 81. https://doi.org/10.3389/frobt.2019.00081
Kory-Westlund, J. M., & Breazeal, C. (2019b). Exploring the effects of a social robot's speech entrainment and backstory on young children's emotion, rapport, relationship, and learning. *Frontiers in Robotics and AI*, *6*(54), 54. https://doi.org/10.3389/frobt.2019.00054
Lee, S., Noh, H., Lee, J., Lee, K., Lee, G. G., Sagong, S., & Kim, M. (2011). On the effectiveness of robot-assisted language learning. *ReCALL*, *23*(1), 25–58. https://doi.org/10.1017/S0958344010000273
MacDorman, K. F., & Chattopadhyay, D. (2016). Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition*, *146*, 190–205. https://doi.org/10.1016/j.cognition.2015.09.019
Mori, M. (1970). The uncanny valley. *Energy*, *7*(2), 33–35. https://cir.nii.ac.jp/crid/1573950400963190528
Movellan, J., Eckhardt, M., Virnes, M., & Rodriguez, A. (2009). Sociable robot improves toddler vocabulary skills. *ACM/IEEE International Conference on Human Robot Interaction*, 307–308. https://doi.org/10.1145/1514095.1514189
Nocentini, O., Fiorini, L., Acerbi, G., Sorrentino, A., Mancioppi, G., & Cavallo, F. (2019). A survey of behavioral models for social robots. *Robotics*, *8*(3), 54. https://doi.org/10.3390/robotics8030054
Onyeulo, E. B., & Gandhi, V. (2020). What makes a social robot good at interacting with humans? *Information*, *11*(1), 43. https://doi.org/10.3390/info11010043
Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places.* Cambridge University Press.
Rintjema, E., Van Den Berghe, R., Kessels, A., De Wit, J., & Vogt, P. (2018). A robot teaching young children a second language: The effect of multiple interactions on engagement and performance. *ACM/IEEE International Conference on Human-Robot Interaction*, 219–220.
Roesler, E., Manzey, D., & Onnasch, L. (2021). A meta-analysis on the effectiveness of anthropomorphism in human-robot interaction. *Science Robotics*, *6*(58), eabj5425. https://doi.org/10.1126/scirobotics.abj5425
Rothstein, N., Kounios, J., Ayaz, H., & Visser, E. J. D. (2020). Assessment of human-likeness and anthropomorphism of robots: A literature review. *International Conference on Applied Human Factors and Ergonomics*, 190–196.
Schleihauf, H., Hoehl, S., Tsvetkova, N., König, A., Mombaur, K., & Pauen, S. (2021). Preschoolers' motivation to over-imitate humans and Robots. *Child Development*, *92*(1), 222–238. https://doi.org/10.1111/cdev.13403
Severson, R. L., & Lemm, K. M. (2016). Kids see human too: Adapting an individual differences measure of anthropomorphism for a child sample. *Journal of Cognition and Development*, *17*(1), 122–141. https://doi.org/10.1080/15248372.2014.989445
Sinatra, A. M., Pollard, K. A., Files, B. T., Oiknine, A. H., Ericson, M., & Khooshabeh, P. (2021). Social fidelity in virtual agents: Impacts on presence and learning. *Computers in Human Behavior*, *114*, 106562. https://doi.org/10.1016/j.chb.2020.106562
Soudek, M., & Soudek, L. I. (1985). Non-verbal channels in language learning. *ELT Journal*, *39*(2), 109–114. https://doi.org/10.1093/elt/39.2.109
Sutiyatno, S. (2018). The effect of teacher's verbal communication and non-verbal communication on students' English achievement. *Journal of Language Teaching and Research*, *9*(2), 430–437. https://doi.org/10.17507/jltr.0902.28
Tanaka, F., & Matsuzoe, S. (2012). Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *Journal of Human-Robot Interaction*, *1*(1), 78–95. https://doi.org/10.5898/JHRI.1.1.Tanaka
Toh, L. P. E., Causo, A., Tzuo, P. W., Chen, I. M., & Yeo, S. H. (2016). A review on the use of robots in education and young children. *Journal of Educational Technology & Society*, *19*(2), 148–163. http://www.jstor.org/stable/jeductechsoci.19.2.148
Vallat, R. (2018). Pingouin: Statistics in python. *Journal of Open Source Software*, *3*(31), 1026. https://doi.org/10.21105/joss.01026
van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., Van der Ven, S., & Leseman, P. (2019). Social robots for language learning: A review. *Review of Educational Research*, *89*(2), 259–295. https://doi.org/10.3102/0034654318821286
Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual.* CreateSpace.
Vogt, P., De Haas, M., De Jong, C., Baxter, P., & Krahmer, E. (2017). Child-robot interactions for second language tutoring to preschool children. *Frontiers in Human Neuroscience*, *11*, 73. https://doi.org/10.3389/fnhum.2017.00073
Yadollahi, E., Johal, W., Paiva, A., & Dillenbourg, P. (2018). When deictic gestures in a robot can harm child-robot collaboration. *ACM Conference on Interaction Design and Children*, 195–206. https://doi.org/10.1145/3202185.3202743
Yin, J., Wang, S., Guo, W., & Shao, M. (2021). More than appearance: The uncanny valley effect changes with a robot's mental capacity. *Current Psychology*. https://doi.org/10.1007/s12144-021-02298-y
Yuka, A. (2009). Positive politeness strategies in oral communication I textbooks: Focusing on terms of address. *The Economic Journal of Takasaki City University of Economics*, *52*(1), 59–70. https://iss.ndl.go.jp/books/R000000004-I10432758-00

About the authors

Jun Yin is a professor in Department of Psychology, Ningbo University, China, and received his PhD in Applied Psychology from the Zhejiang University, China in 2013. His research interests include human-robot interaction, user-centered design, group behaviors, and social norms.

Wenjiao Guo is a Master degree candidate in Department of Psychology, Ningbo University, China, and received her Bachelor in the same university in 2021. Her research interests include human-robot interaction and user-centered design.

Wei Zheng is a Master degree candidate in Department of Psychology, Ningbo University, China, and received her Bachelor in Marketing Management from Ningbo University of Technology, China in 2021. Her research interests include human-robot interaction and user-centered design.

Mingyu Ren is a Master degree candidate in Department of Psychology, Ningbo University, China, and received her Bachelor in Marketing from North University of China, in 2019. Her research interests include human-robot interaction and user-centered design.

Simin Wang is a Master degree candidate in Department of Psychology, Ningbo University, China, and received her Bachelor in Ningbo University, China in 2022. Her research interests include human-robot interaction and user-centered design.

Yingdong Jiang received his Master in Applied Psychology from Ningbo University, China in 2022. His research interests include human-robot interaction and user-centered design.