Journal of Experimental Child Psychology 203 (2021) 105040

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Can a robot teach me that? Children's ability to imitate robots

Kristyn Sommer [a,b,*], Virginia Slaughter [a], Janet Wiles [b], Kathryn Owen [c], Andrea A. Chiba [d], Deborah Forster [e], Mohsen Malmir [f], Mark Nielsen [a,g]

[a] *Early Cognitive Development Centre, School of Psychology, University of Queensland, St. Lucia, QLD 4067, Australia*
[b] *ARC Centre of Excellence for the Dynamics of Language, School of Information Technology and Electrical Engineering, University of Queensland, St. Lucia, QLD 4067, Australia*
[c] *Early Care and Education, University of California, San Diego, La Jolla, CA 92093, USA*
[d] *Department of Cognitive Science and Program in Neuroscience, University of California, San Diego, La Jolla, CA 92093, USA*
[e] *Contextual Robotics Institute and Design Lab, University of California, San Diego, La Jolla, CA 92093, USA*
[f] *Soroco, Boston, MA 02110, USA*
[g] *Faculty of Humanities, University of Johannesburg, Auckland Park 2006, South Africa*

A R T I C L E I N F O

Article history:
Received 20 September 2019
Revised 3 September 2020
Available online 7 December 2020

Keywords:
Child–robot interaction
Imitation
Robot deficit
observational learning
social learning
human-robot interaction

A B S T R A C T

Commensurate with constant technological advances, social robots are increasingly anticipated to enter homes and classrooms; however, little is known about the efficacy of social robots as teaching tools. To investigate children's learning from robots, 1- to 3-year-olds observed either a human or a robot demonstrate two goal-directed object manipulation tasks and were then given the opportunity to act on the objects. Children exhibited less imitation from robotic models that varied with task complexity and age, a phenomenon we term the "robot deficit." In addition, the more children engaged with the robot prior to administration of the imitation task, the more likely they were to replicate the robot's actions. These findings document how children are able to learn from robots but that ongoing design of robotic platforms needs to be oriented to developing more socially engaging means of interacting.

* Corresponding author at: Early Cognitive Development Centre, School of Psychology, University of Queensland, St. Lucia, QLD 4067, Australia.

E-mail address: kristyn.hensy@uqconnect.edu.au (K. Sommer).

https://doi.org/10.1016/j.jecp.2020.105040
0022-0965/

Introduction

As research endeavors attempt to catch up with technological advances, empirical attention is being oriented to one of the most prominent forms of new educational technology: *social robots* (Belpaeme, Kennedy, Ramachandran, Scassellati, & Tanaka, 2018). A social robot is defined as any robot that engages and communicates with humans by following behavioral norms of human interaction (Bartneck & Forlizzi, 2004). Delineating such norms, for learning in young children specifically, can be guided by the tenets of *natural pedagogy* (Csibra & Gergely, 2009), a system that supports the fast and efficient learning of cultural knowledge, incorporating social and communicative cues such as eye gaze, contingent responsivity, and child-directed speech. These cues are thought to prompt children to adopt a "pedagogical stance," preparing them to receive and encode information for later retrieval.

Some social robots are capable of employing the social and communicative cues that natural pedagogy demands, and there is evidence that children can learn a range of skills and behaviors from robotic platforms that do so (Belpaeme et al., 2018; Kanero et al., 2018; Mubin, Stevens, Shahid, Mahmud, & Dong, 2013; Toh, Causo, Tzuo, Chen, & Yeo, 2016). One of the most fundamental ways children learn from others is by imitating (Legare & Nielsen, 2015), the high-fidelity replication of an observed sequence of actions. Will children imitate robots?

In a rare study, Itakura et al. (2008) presented 24- to 35-month-old children with a video depicting a robot who either successfully completed an action (e.g., placing beads in a cup) or attempted but failed to complete the action (e.g., placing the beads on the edge of the cup, where they proceeded to fall off) (modeled from the seminal work of Meltzoff, 1995). Before the demonstration, the robot either gazed at a human experimenter, also depicted on the video (social condition), or did not (nonsocial condition). The children imitated the robot's successful action regardless of condition, suggesting that gaze behaviors were not needed for successful learning. However, children only inferred the robot's intentions, by fulfilling its intended but unsuccessful action (e.g., placing the beads in the cup when the robot failed to do so), when the robot first gazed at the human. This finding contrasts somewhat with that of Meltzoff (1995), who identified that children did not infer the intended (but failed) actions of mechanical pincers. Itakura et al. (2008) concluded that children can imitate the actions of robots even without the demonstration of social behaviors such as eye-gaze and that children can attribute intentions to robots so long as they employ social and communicative cues.

The findings from Itakura et al. (2008) suggest that children can perceive robots as social and communicative agents. However, Itakura et al. employed only single-step actions, with each step having a very clear affordance. Natural pedagogy enables very young children to acquire opaque cultural knowledge where observational learning alone might not be sufficient. It is possible that children simply observed the movement of the objects without engaging in imitation or relying on social cues. Furthermore, Itakura et al. only compared children's imitation of the robot relative with a nonsocial version of the same robot and with a no-demonstration baseline. Hence, this study does not inform us *how well* children learn through imitation in comparison with human teachers. Importantly, children are so driven to copy other humans that they will reproduce another's obviously causally redundant actions (see Hoehl et al., 2019). Do they treat robot models in the same way?

In attempting to answer this question, Sommer et al. (2020) had 4- to 6-year-old children observe either a human model or a humanoid robot demonstrate a series of functional and redundant actions. The children replicated the functional actions of both the human and robot with the same fidelity; however, they were significantly less likely to imitate the redundant actions of the robot than those of the human. This study suggested that young children can imitate a robot but are less inclined to do so when they perceive the actions as irrelevant to completing the task. However, this study was silent as to the proclivities of younger children, and given that a capacity for imitation can be found as early as 6 months of age (Barr, Dowden, & Hayne, 1996) and that children under 3 years learn less from screens than from real-life interaction (Barr, 2010; Kuhl, 2007, 2011; Nielsen, Simcock, & Jenkins, 2008), arguably due to a lack of social engagement, charting social learning during early development is critical.

The current study thus aimed to investigate whether 1- to 3-year-old children imitate a social robot and a human adult equivalently. To extend the findings of Itakura et al. (2008), we employed multistep actions with some opacity and without obvious object affordances for some or all actions. Furthermore, we conducted face-to-face demonstrations from both a human and a robot in an early childhood education center with a humanoid robot platform that possesses many social capacities and communicative cues such as eye gaze, contingent responsivity, and child-directed speech. We reasoned that if children only require the social and communicative cues denoted by natural pedagogy to learn from robots, they should imitate the robot at rates that are equivalent to those of imitating the human. However, if children require something more that is inherently human than the use of social and communicative cues to enable imitation from robots, they should imitate the robot with lower fidelity.

Method

Participants

The final sample included 64 typically developing children aged 12–36 months (30 boys and 34 girls; M_{age} = 25.42 months, *SD* = 6.81, range = 12.23–36.36). Children were recruited in large metropolitan cities either from an early childhood education center in the United States (n = 49) or from a university laboratory in Australia (n = 15). Because the robot was only available in the United States, all children in the robot condition were tested at that location (n = 33). Due to insufficient numbers of children being available at the U.S. site, further testing was conducted for the human condition in Australia (this was due to convenience with regard to recruiting and testing) in order to balance the two conditions, resulting in a total of 16 children in the United States and the 15 children in Australia being tested in the human condition. To account for the potentially diverse samples, an analysis of location was included in the final models. An additional 3 children participated in the experiment but were excluded due to fussiness (n = 1) or failure by the robot to complete the task (n = 2).

Design

This study was a mixed design. Each child was assigned to one of two conditions (between participants), with presentation order of the two imitation tasks counterbalanced across participants (between participants). The two conditions included a live human and a live robot, and the two tasks involved a rattle and a light switch (within participants).

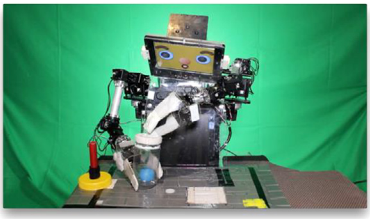
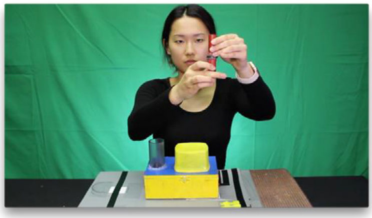
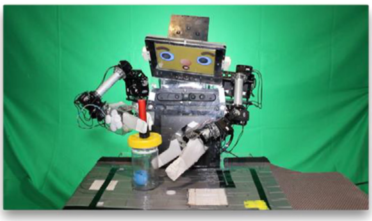
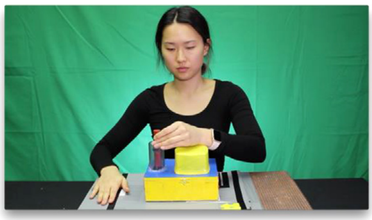
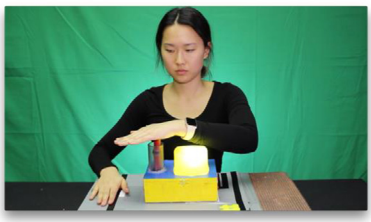
Materials

Apparatus: imitation tasks

Rattle. The rattle task consisted of three pieces: a base, a shaker, and a handle. The base included a small clear plastic jar with a lid. The lid of the jar had a hole cut into the middle and was partially covered by a piece of foam. The outside of the lid was covered with Velcro. A small blue ping-pong ball was used to turn the jar into a rattle. A large yellow plastic jar lid was fixed to the dowel to act as the handle. The underside of the yellow lid was also covered in Velcro. This apparatus has been successfully used in imitation tasks previously (e.g., Simcock & Dooley, 2007) and required three actions, performed in the correct sequence, to achieve the goal state (see Table 1).

Light switch. The light switch task consisted of three pieces: two pieces of a single tool and a wooden box. The tool consisted of two short red dowels with a round magnet affixed to one end of each dowel that when fixed together made a single long dowel. The wooden box was a multicolored MDF (medium-density fiberboard) box with several fixtures. A button was fixed to one side of the box, encased by a piece of long plastic tubing. Opposite the button, a light bulb was affixed to the box and covered by a yellow case. The light switch task required children to switch a light on. The task required three actions to achieve the goal state (see Table 1).

Table 1
Sequence of actions required for each task and scores assigned for imitation.

Score	Task	
	Rattle (presented by the robot)	Light switch (presented by the human)
Start state Score = 0		
Action 1 Score = 1		
Action 2 Score = 2		
Action 3 Score = 3		

Note. For video footage of RUBI-6 performing the demonstration of each task, see online supplementary material.

Apparatus: Robot

The robot used in this study was a custom-built research platform called RUBI-6 (Fig. 1). The RUBI-6 robot is the latest in a line of humanoid social robots built for close proximity social interaction with toddlers (Malmir, Forster, Youngstrom, Morrison, & Movellan, 2013; Malmir, Sikka, Forster, Movellan, & Cottrell, 2015; Movellan, Eckhardt, Virnes, & Rodriguez, 2009; Movellan, Tanaka, Fortenberry, & Aisaka, 2005; Tanaka, Movellan, Fortenberry, & Aisaka, 2006). RUBI-6′s face was an Apple iPad Mini (7.9 inches) with an animated cartoon face (Malmir, 2017). RUBI-6 spoke, played games, and sang nursery rhymes from a Windows tablet embedded in the torso. RUBI-6′s torso sat atop a wide drive-train. However, in this experiment, RUBI-6 did not locomote at any stage. The mechanical components of RUBI-6 were composed of two arms with 6 degrees of freedom and a head with 2 degrees of freedom (pitch and yaw). RUBI-6 possessed a unique capacity for fluid object manipulation. RUBI-6′s physical motion somewhat closely resembles human movement, which is a current challenge for

Fig. 1. RUBI-6 robotic research platform.

many social robots, leading RUBI-6 to be an excellent candidate for the demonstration of imitable actions.

The RUBI-6 robot was capable of close contact social interaction with very young children. RUBI-6 spoke and sang to children and interacted with them physically. Children could play games with RUBI-6 on its torso-based touchscreen tablet, wherein RUBI-6 would respond to children's touch, congratulating children on their performance or pronouncing words to objects they had selected. Occasionally the games presented on the robot's torso screen were nursery rhymes. Children could also engage with the robot through dance to common nursery rhymes such as "Wheels on the Bus" and "Five Little Monkeys." Children could also hand RUBI-6 a toy, which the robot grasped in its pincers and thanked children for sharing, rotating and gazing at the toy and exclaiming in delight before releasing the toy. This interaction gave children sufficient experience with RUBI-6 as both socially contingent and goal directed. The culmination of behaviors that RUBI-6 was able to express made the robot highly social, capable of constructing a social context for children in order for learning to take place.

Procedure

Pretest play session

Before the experiment commenced, children spent at least 10 min with the experimenter, human model, and robot in their classroom in small groups. Children were allowed to engage with the robot for as short or long a time as they felt comfortable, with children being requested to spend at least 10 min around the robot and human model. However, some children engaged with the robot for only about 5 min before leaving while presenting negative affect toward the robot, whereas others returned on multiple occasions for up to 2 h and needed to be removed from the robot.

For each child's first play session, children were brought into the robot's space in groups of 3 or 4. Children were encouraged to approach the robot and were shown how to play with the robot by the experimenter. Once children began approaching the robot, the experimenter acted as a minimal facilitator, engaging with and prompting children only when play had slowed.

During this period, children had time to learn how the robot interacted, could physically engage with the robot, and could observe the robot's high level of contingent responsivity. Children could observe other children or engage themselves with the robot through one of two methods: either the give and take of toys or the digital games on the torso of the robot. The robot's games were presented at random; however, children needed to trigger each new game. Therefore, across different children and sessions, children experienced a different order and assortment of games, with all children receiving sufficient experience with the robot as a socially contingent and communicative agent.

Imitation test

For each task, the protocol followed the same pattern. Children were seated on the lap of an experimenter, and then the agent (robot or human) began the demonstration. The demonstrating agent would gaze at children, state "Watch what I can do with these things," and redirect attention toward the apparatus before demonstrating the target actions.

At the conclusion of a single demonstration, children were relocated to a small table away from the agent. From this position, the demonstrating agent was not in children's direct line of sight. On the table was a duplicate set of objects from the demonstration. The experimenter asked children, "What can you do with these things?" Children were then given 60 s to interact with the objects. The phrase "What can you do with these things?" was selected in order to avoid insinuating to children that they should copy the actions of the model. Each time children acted on the objects, the experimenter said, "That's great! What else can you do with these things?" If children did not touch the objects for a period of 30 s, children were prompted by the experimenter, who said, "Can you show me what you can do with these things?" At the conclusion of the 60 s, children and the experimenter relocated to the opposite side of the classroom and engaged in free play while the agent set up the second set of toys. The procedure was repeated with the second set of toys, and then the experiment was concluded.

Coding and reliability

Imitation coding

All coding was conducted from video footage of children's imitation session. Imitation for the two tasks was scored as the total number of correct actions performed in the correct sequence (see Table 1). For each task, this resulted in a score from 0 to 3. A score of 0 implied that the child performed no target actions in the correct sequence. This was the case even if the child produced some of the task components but did so out of sequence (e.g., for the rattle task, if the child put the lid on the rattle and shook it but did not insert the ball). We chose to code only actions imitated in the correct sequence, rather than any actions performed irrespective of sequence, following previous research employing the rattle apparatus (Barr & Hayne, 1999; Nielsen et al., 2008). We were interested in investigating the highest-fidelity form of imitation, whereby children faithfully copy both specific actions and the exact sequence of actions, to capture any subtle nuances in imitation that may be present between human and robot models.

Alternate imitation coding

We acknowledge that there are multiple ways of coding imitative behavior. Thus, in addition to our focal interest of sequential imitation of target actions, we also coded target action performance regardless of sequence. In this case, target actions were coded if they were observed at any point during the 60-s time frame irrespective of the sequence in which the actions were performed.

Interrater reliability

A second coder independently observed and coded behavior for a random subset (20%) of the sample. According to two-way mixed, consistency, absolute agreement intraclass correlation (ICC;

McGraw & Wong, 1996), interrater reliability was in the excellent range for both the rattle task (ICC = 1.00) and the light switch task (ICC = 1.00) (Cicchetti, 1994).

Results

Imitation analysis

Children's imitative performance was analyzed in generalized linear mixed models on a multinomial distribution using the GLIMMIX procedure in SAS 9.3 (Stroup, 2016), with results summarized in Table 2. The following full factorial model was analyzed: Condition (human or robot) × Age (mean-centered continuous variable ranging from 12.23 to 36.36 months) × Task (rattle or light switch). All models also included a random intercept to account for individual participant differences. Fixed effects of sex (male or female), task order (rattle first or light switch first), and testing location (United States or Australia) were also added as control variables to both models in order to partial out variance associated with them.

The full factorial main effects model (i.e., Model 1) revealed a significant effect of condition, indicating that children in the human condition imitated with higher fidelity than children in the robot condition (b = 1.49, t = 2.41, p = .019) (Fig. 2). A significant effect of age was also identified, indicating that as children increased by 1 month in age, imitative fidelity increased by 0.19 (b = 0.19, t = 4.13, p < .001). Finally, a significant effect of task was found, indicating that children imitated with higher fidelity in the rattle task than in the light switch task (b = 1.29, t = 3.06, p = .003).

Model 2 examined whether imitation fidelity varied as a function of Condition × Age, Condition × Task, or Task × Age. This was accomplished by including three two-way interaction terms involving these variables in the main-effects model. As with Model 1, Model 2 contained significant main effects of condition, age, and task (see Table 2). Model 2 contained a significant two-way interaction between condition and age (F = 4.09, p = .048), indicating that children's imitation varied between conditions as a function of age (Fig. 3). The interaction was followed up by calculating age slopes for both conditions in order to examine how children's imitation changed over age for each condition. The model suggests that as children increased in age by 1 month, imitation significantly increased by 0.32 in the robot condition (b = 0.32, t = 3.97, p < .001, 95% confidence interval (CI) [0.16, 0.49]) and also increased significantly by 0.13 in the human condition (b = 0.13, t = 2.04, p = .046, 95% CI [0.00, 0.26]).

Table 2
Fixed effects for children's imitative performance.

	Model 1			Model 2		
Variable	*b*	*SE*	*p*	*b*	*SE*	*p*
Fixed effects						
Intercept	−2.63	0.97	**.008**	−4.14	1.23	**.001**
Gender (control)	−0.64	0.50	.205	−0.65	0.56	.256
Location (control)	0.03	0.69	.965	0.03	0.76	.968
Task order (control)	0.51	0.49	.305	0.55	0.55	.327
Condition	1.49	0.62	**.019**	2.05	0.73	**.007**
Age (months)	0.19	0.05	**<.001**	0.23	0.06	**<.001**
Task	1.29	0.42	**.003**	1.69	0.50	**.001**
Condition × Age	–	–	–	*F* = 4.09		**.048**
Condition × Task	–	–	–	*F* = 7.96		**.007**
Age × Task	–	–	–	*F* = 2.85		.097
Random effects						
Intercept	1.41	1.09	.098	2.11	1.38	.064
Akaike information criterion	298.48			290.97		

Note. The generalized linear mixed model is modeling the likely that a child will score higher values on the imitation task. Bolded values indicate statistical significance.

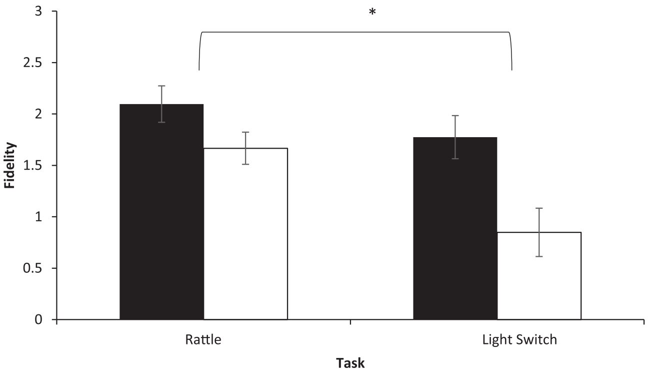


Fig. 2. Mean imitation fidelity on each task in the human condition (black bars) and robot condition (white bars).

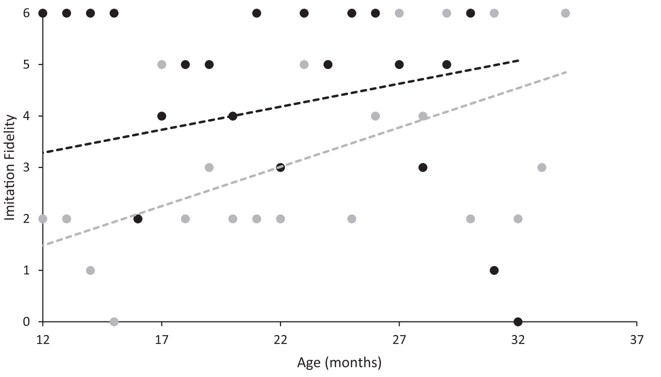


Fig. 3. Combined imitation fidelity across both imitation tasks in the human condition (black) and robot condition (gray).

Model 2 also contained a significant two-way interaction between condition and task (F = 7.96, p = .007), indicating that children's imitation varied between tasks as a function of condition. The interaction was followed up by calculating the difference in imitation scores for each condition on each task. The model suggests that children imitated the human more than the robot on the light switch task (b = 3.38, t = 3.42, p = .001). However, children's imitation did not vary between the human and robot on the rattle task (p = .336).

Alternate imitation analysis

It is possible that children imitate humans and robots equivalently when the additional cognitive load of remembering the sequence of target actions is removed. Thus, to evaluate children's imitation of robots under this broader scope of imitation, a secondary type of imitation was coded, whereby children's performance on the target actions alone was considered irrespective of the sequence in which they were exhibited (see Appendix B in online supplementary material for alternate analysis). This analysis identified a significant effect of condition (b = 1.39, t = 2.18, p = .033), such that children imitated the target actions of the human more than those of the robot; however, there was no significant effect of task (p = .347) or significant interaction between condition and task (p = .078), such that children's imitation did not vary across task or as a factor of the agent across the two tasks.

Post hoc analyses

As highlighted above, children appear to imitate robots less than human models, albeit only for specific tasks, and children do improve with age. However, we do not yet know *why* children may imitate robots with lower fidelity than humans. We reasoned that if the robot prompts children into a pedagogical stance through social and communicative cues, children would imitate the robot and human equivalently. Because children did not imitate equivalently, this suggests that social and communicative cues employed by the robot might not have been as efficient as a human at triggering the pedagogical stance. However, this is not the only explanation for the impoverished imitation of robots.

Children have been documented to imitate mechanical pincers with lower fidelity than human hands (Slaughter & Corbett, 2007; Slaughter, Nielsen, & Enchelmaier, 2008). The robot, RUBI-6, possessed mechanical pincers. One theory of why children imitate pincers with lower fidelity than hands is that children are not capable of mapping their own biology onto the pincers in order to encode and reproduce the target actions (Paulus, Hunnius, Vissers, & Bekkering, 2011; Slaughter & Corbett, 2007; Slaughter et al., 2004). The biological mismatch may also have been an issue for the robot in the current study. If children struggle to map their own biology onto mechanical pincers, the deficit on imitation may have emerged regardless of whether the robot in the current study employed social and communicative cues.

Another explanation aligns with findings that children imitate ingroup members more than outgroup members (e.g., Howard, Henderson, Carrazza, & Woodward, 2015; Wilks, Kirby, & Nielsen, 2018), high-status models more than lower-status models (McGuigan, 2013), adults more than peers (Flynn, 2008; but see Wood et al., 2016), and peers more than puppets (McGuigan & Robertson, 2015). In these cases, it is theorized that children's imitation is affected by heightened social motivation to imitate a more desirable group or agent, suggesting that beyond social and communicative cues the status of the model also plays an integral role in what and from whom children learn. Because we do not know how children perceive robots—whether as outgroups, peers, or something else—we do not know whether impoverished imitation in this study is due to a lack of social motivation or a biological mismatch between pincers and a child's hands. Some insight comes from Sommer et al. (2020), who found that 4- to 6-year-olds can imitate the functional actions of humans and robots equivalently but elect to imitate redundant actions of robots with significantly lower fidelity than humans, suggesting that the biological mismatch explanation lacks veracity. However, it is unknown whether the governing principles that apply to older children's imitation of a robot's redundant actions also apply here to younger children's imitation of a robot's functional actions as a learning mechanism. To extricate the above two alternative hypotheses, we analyzed children's play sessions with the robot.

We rated the videotaped data for children's social engagement with the robot during play as an index of their level of interaction with it. Motivation and engagement are popular measures in research on child–robot interaction (Hong, Huang, Hsu, & Shen, 2016; Kim & Kim, 2011), although such research has primarily focused on children above 5 years of age, typically relying on verbal responses. Children in the current experiment were either preverbal or capable of only short sentences to express themselves. As such, we were unable to employ typical self-report measures of engagement. Therefore, we looked for nonverbal measures of social engagement behavior. Most appropriate for our purpose were measures designed for children with autism spectrum disorder (ASD), who are less capable than typically developing (TD) children of expressing themselves verbally (Kim, Paul, Shic, & Scassellati, 2012; Rudovic, Lee, Mascarell-Maricic, Schuller, & Picard, 2017). We used three dimensions of social engagement: behavioral engagement, valence, and arousal.

In the current experiment, children's play with the robot was unstructured and up to 5 children engaged with the robot during the play sessions. As such, the footage was not a candidate for some typical computerized methods of engagement analysis. Nonetheless, from the footage, through manual rating we were able to reliably extract five dimensions of behavioral engagement for each child: eye contact, touch, communication, location, and negative response. Eye contact and communication were selected as measures of engagement for this age group specifically because it has been documented that children's eye contact functions to construct joint attention, a necessary interaction in natural pedagogy that facilitates learning (Csibra & Gergely, 2009; Tomasello, Carpenter, Call, Behne, & Moll, 2005). In addition, children with social deficits such as ASD show gaze aversion and significantly less eye contact behaviors that can significantly reduce rates of imitation (Hutt & Ounsted, 1966; Volkmar & Mayes, 1990; Williams, Whiten, & Singh, 2004). Communication—both verbal and nonverbal—is often incorporated into combined engagement measures prevalent in child development literature (Aguiar & McWilliam, 2013; De Kruif & McWilliam, 1999).

The less traditional measures of touch, location, and negative response were also included due to observations made during the experiment or due to particular capacities of the robot. Touch was employed specifically in relation to our robot, whose main features are a give-and-take capacity and a touchscreen, both of which elicit or afford touch. Location of children within the testing area was included because children were able to navigate the testing area freely, specifically assessing whether children remained in the testing space or departed. Finally, negative response was included following observations made by the experimenter during children's interaction with the robot that appeared to affect the rest of children's interaction with the robot for that session. It should be noted that the negative response of children during the initial sessions analyzed here did not extend to later sessions. Children's negative affect toward the robot usually was resolved by the next opportunity to engage with the robot during play times.

The first measure, behavioral engagement, was measured on a 6-point scale, where 5 = *high engagement,* 4 = *medium engagement,* 3 = *low engagement,* 2 = *indifferent,* 1 = *noncompliance,* and 0 = *negative response* (Rudovic et al., 2017). All five dimensions were used at each level of engagement (see Table A1 in supplementary material for exemplars of each level for each dimension). In addition to behavioral engagement, two other measures of social engagement were employed: valence and arousal. Valence was measured on a 5-point scale from 2 (*high positive valence*) to − 2 (*high negative valence*) with a midpoint of 0 (*neutral*) (see Table A2 in supplementary material for exemplars). Arousal was also measured on a 5-point scale from 2 (*very high arousal*) to − 2 (*very low arousal*) with a midpoint of 0 (*neutral*) (see Table A3 in supplementary material for exemplars).

We analyzed the first 10 min of each child's play from the robot condition for each scale in 10-s increments. We took an average of the full 10-min session for each child for behavioral engagement, valence, and arousal and analyzed whether any of the three measures predicted imitation from the robot. See Table 3 for descriptive statistics on behavioral engagement, valence, and arousal scores.

Interrater reliability

A second rater independently observed and rated behavior for a random subset (20%) of the sample. According to two-way mixed, consistency, absolute agreement ICC (McGraw & Wong, 1996), interrater reliability was in the excellent range for behavioral engagement (ICC = .981), valence (ICC = .975), and arousal (ICC = .773) (Cicchetti, 1994).

Social engagement analysis

The social engagement model was analyzed using the GLIMMIX procedure in SAS 9.3 (Stroup, 2016) on the imitative performance of children in the robot condition only, with results summarized in Table 4. Behavioral engagement, valence, and arousal all were treated as continuous measures. Because the observational analyses of the social engagement sessions were a post hoc analysis, we included the variables from Model 2 (task, age, and Task × Age) in the social engagement models with the addition of the three new variables: behavioral engagement (mean-centered continuous variable ranging from −1.42 to 0.95), valence (mean-centered continuous variable ranging from −0.36 to 0.38), and arousal (mean-centered continuous variable ranging from −0.55 to 0.44) in Model 3. Model 3 also included a random intercept to account for individual participant differences. In addition to all

Table 3
Descriptive statistics of engagement, valence, and arousal scores.

	M	*SE*	Min	Max
Behavioral engagement (*n* = 25)	3.81	1.47	2.40	4.77
Valence (*n* = 25)	0.09	0.04	−0.27	0.47
Arousal (*n* = 25)	0.13	0.04	−0.42	0.58

Table 4
Fixed effects for children's imitation of the robot as predicted by their social engagement behaviors in the pretest play session.

	Model 3			Model 4		
Variable	*b*	*SE*	*p*	*b*	*SE*	*p*
Fixed effects						
Intercept	−3.45	1.19	**.009**	−6.52	2.41	**.016**
Gender (control)	−1.70	0.96	.090	−1.87	1.18	.130
Task order (control)	−0.18	0.94	.846	−0.77	1.51	.617
Age (months)	0.23	0.10	**.037**	0.40	0.19	.051
Task	2.50	0.93	**.013**	5.53	2.09	**.017**
Behavioral engagement	1.90	0.90	**.047**	3.81	1.56	**.025**
Valence	1.97	3.40	.568	3.28	5.74	.575
Arousal	−4.97	3.38	.155	−10.63	5.53	.071
Age × Task	–	–	–	*F* = 0.84		.372
Behavioral Engagement × Task	–	–	–	*F* = 2.57		.127
Valence × Task	–	–	–	*F* = 0.38		.547
Arousal × Task	–	–	–	*F* = 4.07		.059
Behavioral Engagement × Age	–	–	–	*F* = 0.01		.907
Valence × Age	–	–	–	*F* = 0.01		.913
Arousal × Age	–	–	–	*F* = 0.63		.437
Random effects						
Intercept	0.69	1.36	.306	1.29	2.52	.305
Akaike information criterion	102.05			103.51		

Note. The generalized linear mixed model is modeling the likely that a child will score higher values on the imitation task. Bolded values indicate statistical significance.

variables included in Model 3, Model 4 included the original Age × Task interaction term from Model 2 as well as six new interaction terms: behavioral engagement, valence, and arousal on task and behavioral engagement, valence, and arousal on age.

Model 3 revealed a significant effect of age and task as observed in Model 1 (Table 2). Model 3 revealed a significant effect of behavioral engagement (b = 1.90, t = 2.10, p = .047), indicating that children with higher engagement behaviors toward the robot had higher imitation scores (Table 4). Model 4 also revealed a significant effect of task and behavioral engagement (b = 3.81, t = 2.44, p = .025) but no significant effect of age, contrary to Models 1 and 2 (Fig. 4). There were no significant interaction terms.

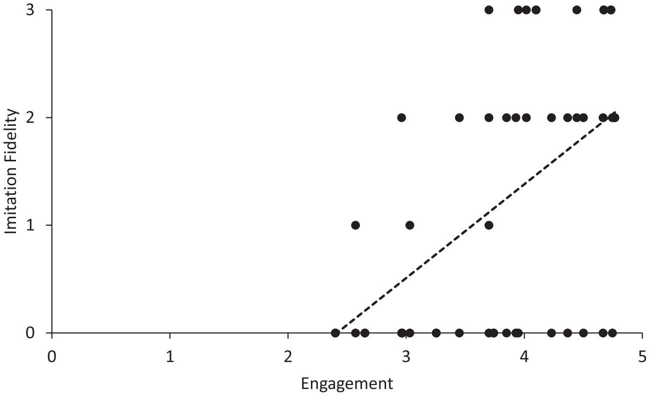


Fig. 4. Imitation fidelity against engagement for each child in the robot condition.

Discussion

Children tested here imitated a social robot with lower fidelity than a human, suggesting that children during the first few years of life do not learn from robots as well as from human models, not even robots equipped with cues to encourage natural pedagogy. However, this finding comes with several caveats. Children did learn from the human and robot equivalently on one of the two tasks employed in this study. This suggests that, at least for young children, social learning can be affected by task characteristics. In addition, older children imitated the robot more than younger children, an effect that was larger in magnitude than that observed in imitation of the human. This suggests that children's imitation of robots compared with humans may "catch up" with development, potentially leading to equivalency at some point. It is possible that older children may imitate robots with greater fidelity than younger children because they are more adept at learning and can generalize their ability to learn to nonhuman sources. Nevertheless, this implies that children first need other humans to learn how to learn, thereby resulting in poorer learning from robots in younger children.

As already alluded to, children imitated the rattle task at equivalent rates whether the model was a human or robot; however, they imitated the light switch task with lower fidelity for the robot than for the human. This difference in imitation between the two tasks was unexpected. Although the tasks were designed to be similarly complex, requiring three discrete actions to achieve the end-state goal, unintended nuances in their design may have led to the differential rates of imitation found. The rattle task comprised a transparent base, whereby all actions enacted on the apparatus were visible to children. Therefore, children could observe the ball pass through the lid and into the base, before the handle was affixed to the lid. This may have inadvertently facilitated learning the affordances of each component of the object, diluting the need for social input. In contrast, the light switch task required children to infer that the robot intended to switch the light on through its preceding actions. Once the two dowels were assembled into a single tool, children observed the dowel enter the semi-opaque tube enclosing the hidden button. Children then observed the robot press the button that triggered the light to illuminate on the opposite side of the box. Although children appear able to infer a robot's intentions (Itakura et al., 2008), children in this study may have found the light switch task to be more difficult because they needed to both observe the robot's actions and infer the robot's intention to switch the light on. That is, the mechanism of switching on the light was opaque because the button was concealed and the link between the button press (which is quite a small, potentially imperceptible, biological motion when made by a robot) and the light appearing on was not as transparent as, for example, a robot dropping a loop directly in front of a hook, as in Itakura et al. (2008).

An alternative explanation for the variation in tasks may relate to the choice in prompting phrase for children to begin engaging with the apparatus. Children were asked "What can you do with these things?" Although this was done to avoid insinuating to children that they ought to copy the model, children may instead have interpreted this phrase as "What *else* can you do with these things?" Thus, the source of variation between the light switch and rattle may have been due to the number of other actions possible with the combination of objects available. The rattle objects may have had a more limited number of actions available than the light switch objects. Given that children's actions were measured only for the first 60 s of interaction with the apparatus, children may have been more likely to cycle through all available actions and back to the original imitation task within the time frame in the rattle task than in the light switch task. To rectify this issue in future work, an alteration to the prompting phrase to a phrase such as "It's your turn now" would be valuable.

Another cause for variation in imitation between the two tasks may be differences in familiar scripts and action schemas. The rattle task arguably involved more actions that are familiar to children (e.g., shaking and rattling objects) than the light switch task. The light task, in addition to having an opaque causal structure, also involved many potentially novel actions with objects. Thus, whereas the rattle task may have primed familiar scripts and action schemas, the light task might not have done so, leading to lower fidelity imitation. Further investigation is required into what social and communicative cues a robot may need to present in order to facilitate the imitation of opaque action sequences by young children and whether the priming of familiar scripts and action schemas may interact with the opacity of actions.

After observing a difference in imitation between the human and robot, at least for the light switch task, we attempted to identify the "unknown factors" reducing imitation from the robot. Children have been found to imitate mechanical pincers with lower fidelity than human hands, a finding attributed to a biological mismatch between children and robots (Slaughter & Corbett, 2007; Slaughter et al., 2004). Furthermore, imitation is also documented to be affected by the extent to which observed actions can be mapped onto infants' own motor repertoire (Paulus et al., 2011); thus, it is plausible that mapping robotic actions is a more difficult task than mapping human actions. However, children's imitation can also be impoverished by group membership or status of the agent (Howard et al., 2015; McGuigan, 2013; Wilks et al., 2018) due to lower social motivation to adhere to the teachings of these agents. This phenomenon has been observed to some degree in older children's overimitation of robots, whereby children will imitate the redundant actions of a robot, but at reduced rates relative to a human model, while imitating the functional actions of a human and robot equivalently (Sommer et al., 2020). Therefore, we investigated whether children's level of social engagement during play with the robot, as a proxy for social motivation, would predict rates of imitation. We found that it did; higher behavioral engagement during play was associated with higher imitation.

In light of this, we conjecture that if children require more than just social and communicative cues to enable high fidelity imitation, social motivation may be the modulating factor. In previous research, children who attended more to a demonstrator's social cues were found to be more likely to imitate with higher fidelity than children who attended less (Óturai, Kolling, & Knopf, 2013). However, this interpretation of the data may be problematic because our analysis highlights only correlational effects between behavioral engagement and imitation. There is currently little research investigating children's individual differences in their propensity to view robots as social agents. Some children may be more inclined to see a robot as social and even as a high-status figure from which they can learn. Conversely, other children may find it difficult to perceive a robot as social or may perceive it as an outgroup member or low-status figure. Because children's imitation can be modulated by these factors (Howard et al., 2015; McGuigan, 2013; Nielsen & Blank, 2011; Wilks et al., 2018), future research should investigate children's perceptions of robots and whether or not these perceptions predict how children learn.

Furthermore, the amount of time children engaged with the robot was unconstrained and challenging to measure accurately in the current study. The robot existed in children's environment for some weeks, and children could come and go as they pleased. Although the length of time children spent with the robot was primarily dictated by the children, it is possible that length of exposure to the robot may have also affected children's learning from the robot. Future research should endeavor to measure whether familiarity with a robot also affects children's learning.

This study fills a gap in our knowledge of toddlers' learning from social robots and humans. We identified that young children can successfully imitate robots equivalently to humans in some tasks but not in others and that there is a developmental trajectory in which children's performance improves with age when learning from robots. This study also highlights the importance of social engagement and is crucial in designing and assessing robots as teachers, whereby robotic designers should aim to implement social bonding behaviors that are effective for all children in order to yield successful learning on a variety of tasks. These findings have implications for educators wishing to use robots in a teaching capacity. Here we present some initial steps toward identifying what, how, and why children learn from robots. As these new technological agents become increasingly prevalent in the lives of children and are granted status as facilitators of learning, the need for research such as this will only become more urgent and more compelling.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2020.105040.

References

Aguiar, C., & McWilliam, R. A. (2013). Consistency of toddler engagement across two settings. *Early Childhood Research Quarterly, 28*, 102–110.

Barr, R. (2010). Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice. *Developmental Review, 30*, 128–154.
Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6- to 24-month-old infants. *Infant Behavior and Development, 19*, 159–170.
Barr, R., & Hayne, H. (1999). Developmental changes in imitation from television during infancy. *Child Development, 70*, 1067–1081.
Bartneck, C., & Forlizzi, J. (2004). A design-centred framework for social human–robot interaction. In *ROMAN 2004: 13th IEEE International Workshop on Robot and Human Interactive Communication* (pp. 591–594). https://doi.org/10.1109/ROMAN.2004.1374827.
Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics, 3*(21), eaat5954.
Cicchetti (1994). Guidlines, criteria, and rules of thumb for evalauting normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284–290.
Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*, 148–153.
De Kruif, R. E. L., & McWilliam, R. A. (1999). Multivariate relationships among developmental age, global engagement, and observed child engagement. *Early Childhood Research Quarterly, 14*, 515–536.
Flynn, E. (2008). Investigating children as cultural magnets: Do young children transmit redundant information along diffusion chains? *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363*, 3541–3551.
Hoehl, S., Keupp, S., Schleihauf, H., McGuigan, N., Buttelmann, D., & Whiten, A. (2019). "Over-imitation": A review and appraisal of a decade of research. *Developmental Review, 51*, 90–108.
Hong, Z. W., Huang, Y. M., Hsu, M., & Shen, W. W. (2016). Authoring robot-assisted instructional materials for improving learning performance and motivation in EFL classrooms. *Journal of Educational Technology and Society, 19*, 337–349.
Howard, L. H., Henderson, A. M. E., Carrazza, C., & Woodward, A. L. (2015). Infants' and young children's imitation of linguistic in-group and out-group informants. *Child Development, 86*, 259–275.
Hutt, C., & Ounsted, C. (1966). The biological significance of gaze aversion with particular reference to the syndrome of infantile autism. *Behavioral Science, 11*, 346–356.
Itakura, S., Ishida, H., Kanda, T., Shimada, Y., Ishiguro, H., & Lee, K. (2008). How to build an intentional android: Infants' imitation of a robot's goal-directed actions. *Infancy, 13*, 519–532.
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*, 146–151.
Kim, E. S., Paul, R., Shic, F., & Scassellati, B. (2012). Bridging the research gap: Making HRI useful to individuals with autism. *Journal of Human-Robot Interaction, 1*, 26–54.
Kim, J., & Kim, J.-K. (2011). The effectiveness of robot pronunciation training for second language acquisition by children: Segmental and suprasegmental feature analysis approaches. *International Journal of Robots, Education and Art, 1*, 1–17.
Kuhl, P. K. (2007). Is speech learning "gated" by the social brain?. *Developmental Science, 10*, 110–120.
Kuhl, P. K. (2011). Social mechanisms in early language acquisition: Understanding integrated brain systems supporting language. In J. Decety & J. T. Cacioppo (Eds.), *The Oxford handbook of social neuroscience* (pp. 649–667). New York: Oxford University Press.
Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in Cognitive Sciences, 19*, 688–699.
Malmir, M. (2017). *Perceptual algorithms for social robots in early childhood education* Unpublished dissertation. San Diego: University of California.
Malmir, M., Forster, D., Youngstrom, K., Morrison, L., & Movellan, J. R. (2013). Home alone: Social robots for digital ethnography of toddler behavior. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 762–768). https://doi.org/10.1109/ICCVW.2013.104.
Malmir, M., Sikka, K., Forster, D., Movellan, J., & Cottrell, G. (2015). In *Deep Q-learning for active recognition of GERMS: Baseline performance on a standardized dataset for active learning*. https://doi.org/10.5244/c.29.161 (pp. 161.1–161.11).
McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation. *Psychological Methods, 1*, 30–46.
McGuigan, N. (2013). The influence of model status on the tendency of young children to over-imitate. *Journal of Experimental Child Psychology, 116*, 962–969.
McGuigan, N., & Robertson, S. (2015). The influence of peers on the tendency of 3- and 4-year-old children to over-imitate. *Journal of Experimental Child Psychology, 136*, 42–54.
Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*, 838–850.
Movellan, J. R., Eckhardt, M., Virnes, M., & Rodriguez, A. (2009). Sociable robot improves toddler vocabulary skills. In Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction—HRI '09 (p. 307). doi:10.1145/1514095.1514189.
Movellan, J. R., Tanaka, F., Fortenberry, B., & Aisaka, K. (2005). The RUBI/QRIO project: Origins, principles, and first steps. In Proceedings of the 4th IEEE International Conference on Development and Learning, 2005 (pp. 80–86). doi:10.1109/DEVLRN.2005.1490948.
Mubin, O., Stevens, C. J., Shahid, S., Mahmud, A. Al, & Dong, J. (2013). A review of the applicability of robots in education. *Technology for Education and Learning*. https://doi.org/10.2316/Journal.209.2013.1.209-0015.
Nielsen, M., & Blank, C. (2011). Imitation in young children: When who gets copied is more important than what gets copied. *Developmental Psychology, 47*, 1050–1053.
Nielsen, M., Simcock, G., & Jenkins, L. (2008). The effect of social engagement on 24-month-olds' imitation from live and televised models. *Developmental Science, 11*, 722–731.
Óturai, G., Kolling, T., & Knopf, M. (2013). Relations between 18-month-olds' gaze pattern and target action performance: A deferred imitation study with eye tracking. *Infant Behavior and Development, 36*, 736–748.
Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011). Imitation in infancy: Rational or motor resonance? *Child Development, 82*, 1047–1057.

Rudovic, O., Lee, J., Mascarell-Maricic, L., Schuller, B. W., & Picard, R. W. (2017). Measuring engagement in robot-assisted autism therapy: A cross-cultural study. *Frontiers in Robotics and A, I*, 4. https://doi.org/10.3389/frobt.2017.00036.
Simcock, G., & Dooley, M. (2007). Generalization of learning from picture books to novel test conditions by 18- and 24-month-old children. *Developmental Psychology, 43*, 1568–1578.
Slaughter, V., & Corbett, D. (2007). Differential copying of human and nonhuman models at 12 and 18 months of age. *European Journal of Developmental Psychology, 4*, 31–45.
Slaughter, V., Heron, M., Tilse, E., Jenkins, L., Muller, U., & Liebermann, D. (2004). Origins and early development of human body knowledge. Monographs of the Society for Research in Child Development, 69(2, Serial No. 276).
Slaughter, V., Nielsen, M., & Enchelmaier, P. (2008). Interacting socially with human hands at 24 months of age. *Infancy, 13*, 185–195.
Sommer, K., Davidson, R., Armitage, K. L., Slaughter, V., Wiles, J., & Nielsen, M. (2020). Preschool children overimitate robots, but do so less than they overimitate humans. *Journal of Experimental Child Psychology, 191* 104702.
Stroup, W. W. (2016). Generalized linear mixed models: Modern concepts, methods and applications. (Chapman & Hall/CRC Texts in Statistical Science). Boca Raton, FL: CRC Press.
Tanaka, F., Movellan, J. R., Fortenberry, B., & Aisaka, K. (2006). Daily HRI evaluation at a classroom environment: Reports from dance interaction experiments. In *Proceedings of the 1st Annual Conference on Human-Robot Interaction (HRI)* (pp. 3–9). https://doi.org/10.1145/1121241.1121245.
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. *Educational Technology & Society, 19*, 148–163.
Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*. 675–691; discussion 691–735.
Volkmar, F. R., & Mayes, L. C. (1990). Gaze behavior in autism. *Development and Psychopathology, 2*, 61–69.
Wilks, M., Kirby, J., & Nielsen, M. (2018). Children imitate antisocial in-group members. *Developmental Science, 21* e12675.
Williams, J. H. G., Whiten, A., & Singh, T. (2004). A systematic review of action imitation in autistic spectrum disorder. *Journal of Autism and Developmental Disorders, 34*, 285–299.
Wood, L. A., Harrison, R. A., Lucas, A. J., McGuigan, N., Burdett, E. R. R., & Whiten, A. (2016). "Model age-based" and "copy when uncertain" biases in children's social learning of a novel task. *Journal of Experimental Child Psychology, 150*, 272–284.