



Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement

Huili Chen^{*}, Hae Won Park, Cynthia Breazeal

Personal Robots Group, MIT Media Lab, Cambridge, MA, USA

ARTICLE INFO

Keywords:

Intelligent tutoring systems
Interactive learning environments
Evaluation of CAL systems
Cooperative/collaborative learning

ABSTRACT

Pedagogical agents are typically designed to take on a single role: either as a tutor who guides and instructs the student, or as a tutee that learns from the student to reinforce what he/she knows. While both agent-role paradigms have been shown to promote student learning, we hypothesize that there will be heightened benefit with respect to students' learning and emotional engagement if the agent engages children in a more peer-like way — adaptively switching between tutor/tutee roles. In this work, we present a novel active role-switching (ARS) policy trained using reinforcement learning, in which the agent is rewarded for adapting its tutor or tutee behavior to the child's knowledge mastery level. To investigate how the three different child-agent interaction paradigms (tutee, tutor, and peer agents) impact children's learning and affective engagement, we designed a randomized controlled between-subject experiment. Fifty-nine children aged 5–7 years old from a local public school participated in a collaborative word-learning activity with one of the three agent-role paradigms. Our analysis revealed that children's vocabulary acquisition benefited from the robot tutor's instruction and knowledge demonstration, whereas children exhibited slightly greater affect on their faces when the robot behaves as a tutee of the child. This synergistic effect between tutor and tutee roles suggests why our adaptive peer-like agent brought the most benefit to children's vocabulary learning and affective engagement, as compared to an agent that interacts only as a tutor or tutee for the child. This work sheds light on how fixed role (tutor/tutee) and adaptive role (peer) agents support children's cognitive and emotional needs as they play and learn. It also contributes to an important new dimension of designing educational agents — actively adapting roles based on the student's engagement and learning needs.

1. Introduction

Early childhood is a critical period of development that sets the foundation for children's future academic success and aspiration. Unfortunately, only about 30% of eligible 4-year-old children are enrolled in state pre-K programs every year ([National Institute of Early Education Research, 2018](#)). Many young children do not have access to quality preschool programs or equivalent home schooling, and consequently do not achieve kindergarten readiness prior to entering the formal education system ([Nores & Barnett, 2014](#); [U.S. Department of Education, 2015](#)). Statistics show many children who start off below readiness level have hard time catching up ([Garcia & Weiss, 2017](#)). Access to extracurricular support (e.g., after-school or summer programs) could help reduce this gap, but resources are limited and can be very costly ([Grossman, Lind, Hayes, McMaken, & Gersick, 2009](#)).

^{*} Correspondence to: Media Lab, Massachusetts Institute of Technology, E14-474, 75 Amherst St, Cambridge, MA 02139, USA.

E-mail addresses: hchen25@media.mit.edu (H. Chen), haewon@media.mit.edu (H.W. Park), cynthiab@media.mit.edu (C. Breazeal).

In at-risk communities, it is very challenging for kindergartens to offer a curriculum that is cognitively and academically leveled to every student in their classrooms. Children enter school with a wide range of cognitive and pre-literacy starting points, as each child has a unique distribution of the various cognitive, visual, social and linguistic skills needed to be a successful reader (Dehaene, 2009; Wolf & Gottwald, 2016). Hence, there is a real need and compelling opportunity to develop adaptive educational technologies to supplement the learning experiences that diverse learners receive at school and augment early childhood education — especially given their promise to deliver personalized education and be cost-effective at scale.

A wide variety of technological interventions, such as intelligent tutoring agents, game apps and computer simulations, have been designed to support students across a range of ages and in a variety of academic domains (Belpaeme, Kennedy, Ramachandran, Scassellati, & Tanaka, 2018; Breazeal, Morris, Gottwald, Galyean, & Wolf, 2016; D'mello & Graesser, 2013; Lindgren, Tscholl, Wang, & Johnson, 2016; Vaala, Ly, & Levine, 2015). Inspired by Blooms 2-sigma effect (Bloom, 1984), intelligent tutoring systems (ITS) is a well-established field where the computer instructs, provides feedback, and guides student learning in 1:1 interaction (Graesser, Hu, & Sottolare, 2018). ITS has investigated diverse cognitive elements to support student's needs encompassing assessing mastery (Baker et al., 2010), modeling student's cognitive states (Corbett, Kauffman, Maclaren, Wagner, & Jones, 2010), adapting content to individual student needs (Manickam, Lan, & Baraniuk, 2017), as well as looking at student affect and motivation (D'mello & Graesser, 2013). In this learning paradigm, the computer or intelligent agent interacts with students as their tutor or teacher. Alternatively, teachable agents have also been designed to emulate a younger, less capable playmate. This approach is grounded in a widely studied and practiced concept in education called "learning-by-teaching" (Roscoe & Chi, 2007) that enables children to improve and consolidate their learning by teaching another (Biswas, Leelawong, Schwartz, Vye, & at Vanderbilt, 2005; Brophy, Biswas, Katzlberger, Bransford, & Schwartz, 1999; Chin, Dohmen, & Schwartz, 2013; Park & Howard, 2014; Tanaka & Matsuzoe, 2012).

To date, pedagogical agents are mostly designed to serve a single role: either as tutor or as tutee. However, we argue that a flexible interaction paradigm, in which an educational agent can adaptively switch between roles at appropriate times, holds great potential to leverage the benefits of both as human peers often provide to one another. Given the promise of AI agents for early childhood education, and the different interaction paradigms that are possible, it is important to understand how different designs influence young children's learning and emotional experience. Resulting insights will help inform the design of effective and emotionally engaging educational interventions that support the diverse cognitive, social, emotional and physical learning needs in early childhood.

The main research questions in this paper are (1) how do young children learn or engage when interacting with different types of pedagogical agents, and (2) how can we leverage the benefits of each pedagogical agent paradigm to provide a synergistic impact on young children's learning and engagement. More specifically, we are particularly interested in comparing different child-agent interaction paradigms where the agent takes on a different role (i.e., tutor, tutee, or reciprocal peer) and how each impacts children's learning and affective experience.

Our work makes the following contributions. First, we introduce a new agent-role paradigm where the educational agent acts as a reciprocal peer. Tutor agents, designed to explicitly teach students via instruction, demonstration and feedback, is a well-established paradigm. More recently, pedagogical agents that act as a tutee to engage children in a learning-by-teaching paradigm have been proposed. To our knowledge, no prior work studied how a pedagogical agent switched its role between tutor and tutee when interacting with children. In this work, we designed a novel adaptive role switching (ARS) model whereby the robot can flexibly change its role in a reciprocal interaction to teach and learn from each child. We developed the ARS model using reinforcement learning to maximize children's exposure to both tutor and tutee roles at appropriate times. We first pre-trained the model using a pilot study dataset where the robot randomly switched between roles as it played the vocabulary game with children. We used this baseline model to seed the ARS policy training for the model's faster convergence and adaptation to each child in our main study. The robot, in effect, learned when to switch its role to stay in sync with each child's real-time learning performance.

Second, our work is the first experiment to our knowledge that directly compares the impact of a tutor, tutee, or reciprocal peer robot on young children's learning and affective engagement. We designed a between-subjects experiment with 59 children aged 5–7 years old divided into three counterbalanced groups, and assigned to one of our three experimental conditions. An expressive and appealing social robot, Tega, was used as the learning companion in our work (Fig. 1(a)). In the *tutee* condition, the robot behaved as a curious learner who lacked vocabulary knowledge and needed the child's help. In the *tutor* condition, the robot behaved as an expert that never made mistakes and always gave the child feedback and guidance. In the *peer* condition, the robot used the ARS policy to determine which role to exhibit at each turn of game play. Participants played a vocabulary learning game with a Tega robot for two 30-min sessions, learning new target vocabulary in each. Each group was compared and evaluated with respect to children's word learning performance and facial expressions as an indicator of affective engagement. We found that children's vocabulary learning and affective behavior were the most enhanced with the reciprocal, adaptive peer robot.

Our analysis found that children's vocabulary acquisition benefited from the robot's instruction and knowledge demonstration, whereas children's facial affect were slightly more expressive when the robot behaved as a tutee of the child. This synergistic effect between tutor and tutee roles suggests why our adaptive peer-like agent brought the most benefit. Hence, our third contribution is a novel user experience paradigm for peer-like educational agents that can successfully engage children in reciprocal and adaptive tutor-tutee roles. In light of these findings, we provide design guidance for effective and engaging peer-like learning companions that promote the growth and welfare of young children.

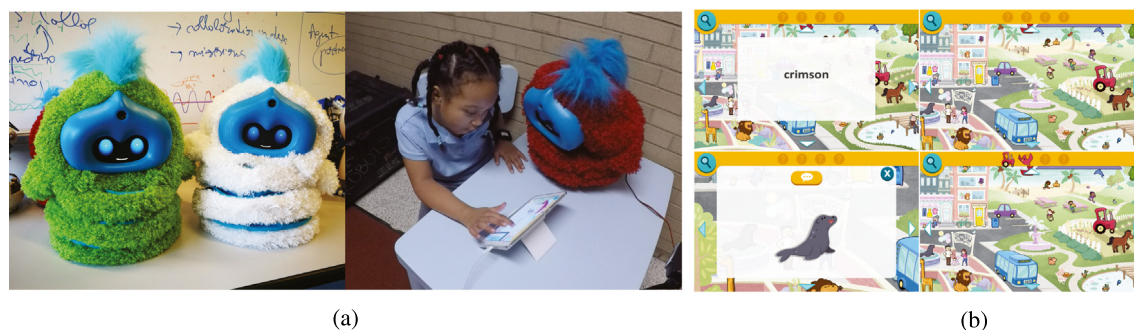


Fig. 1. The social robot, Tega, engages young children in educational activities on a touchscreen tablet as their learning companion. (a): Tega and a child are playing the *Word Quest* game on a tablet together. (b): The *Word Quest* game is a collaborative game in which a child and a robot take turns identifying objects called out by a quest mission.

2. Background

2.1. Pedagogical agents for young children

Oral language development prior to entering kindergarten can significantly impact children's acquisition of early literacy skills, and thereby influence their later educational success (Hart & Risley, 1995; Pérez, Tabors, & López, 2007). Young children are wired to learn a language from friendly others through social interaction and play (Kuhl, 2007, 2011). Recent studies show that exposure to a sufficient quantity of spoken language without social interaction may not be the most effective approach to promote children's language or vocabulary development (Romeo et al., 2018). It turns out that the dynamic conversational back-and-forth between the child and others also matters in children's language development, as this conversational turn-taking is associated with the activation of children's left inferior frontal cortex (Broca's area), which plays an important role in neural language processing, according to Romeo et al. (2018). The importance of multi-modal social interaction poses interesting design challenges for creating intelligent technologies that can promote oral language and literacy skills for young children.

Prior work has shown that students across different grades can learn from chatbots and virtual avatars (D'mello & Graesser, 2013; D'Mello, Olney, Williams, & Hays, 2012). However, in order to engage the young learners and support them long-term, an automatic assessment and intervention should be implemented for personalized intervention (Woolf, 2008). Social robots hold great promise to promote young children's learning, as young children are particularly receptive to learning from expressive social robots (Hyun, Kim, Jang, & Park, 2008; Kanero et al., 2018). Social robots have been developed to help young children in a variety of educational contexts including STEM, second language, vocabulary, and literacy skills (Brown & Howard, 2015; Fridin, 2014; Gordon & Breazeal, 2015; Kennedy, Baxter, Senft, & Belpaeme, 2016; Park, Gelsomini, Joo Lee, & Breazeal, 2017). Compared with virtual agents, the attentive and expressive co-present behaviors of social robots are more likely to elicit rich social behaviors from children that benefit and enhance their engagement and learning (Kennedy, Baxter, & Belpaeme, 2015; Leyzberg, Spaulding, Toneva, & Scassellati, 2012). For instance, young children readily engage with social robots as peer-like companions that emulate children's behavior via language and affective mirroring to build rapport and advance their oral language skills (Gordon et al., 2016; Kory-Westlund, 2019). Given the greater potential of social robots to serve as emotionally appealing learning companions for young children in a social context, we used the Tega robot in our study.

2.2. Pedagogical robot as tutor

A significant body of work has explored the use of social robots or computer agents as tutors where they engage children as more knowledgeable partners to foster the acquisition of new knowledge and skills (Belpaeme et al., 2018; Chang, Lee, Chao, Wang, & Chen, 2010; Kennedy et al., 2016). According to Vygotsky's Zone of Proximal Development theory, a more capable partner can guide and scaffold interaction to lead the edge of children's intellectual growth (Vygotsky, 1978, p86). Thus, a tutor agent has great potential to foster children's learning from an educational standpoint.

Tutor agents can be presented either as adults, or being of a similar age to the student (e.g., peer tutoring) (Kanda, Hirano, Eaton, & Ishiguro, 2004; Ryokai, Boulanger, & Cassell, 2003), as well as sharing other group affinity attributes with the student (e.g., speech, gesture, body) (Zaga, Lohse, Truong, & Evers, 2015). Social robot tutors do not try to appear human, but rather exhibit anthropomorphic qualities that appeal to children, and are often presented as knowledgeable playmates. For the language and literacy domain, social robot tutors have been effective in helping children learn new vocabulary through playing games (Movellan, Eckhardt, Virnes, & Rodriguez, 2009), as well as have been designed to facilitate their second language (Chang et al., 2010) and engage in 1:1 personalized dialogic storytelling (Park, Grover, Spaulding, Gomez, & Breazeal, 2019).

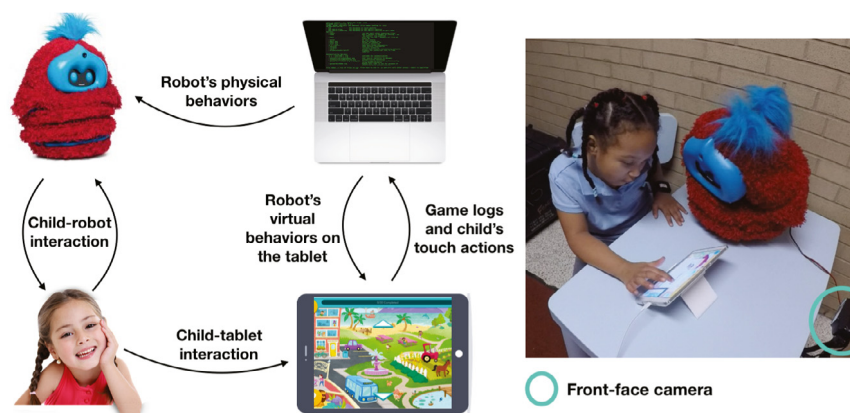


Fig. 2. The integrated system for a robotic learning peer consists of (1) a computer hub that communicates between a robot and a tablet, (2) the *WordQuest* vocabulary game that runs on a touchscreen tablet, (3) *Tega* social robot, and (4) a camera that captures children's facial expression as an indicator of their affective engagement.

2.3. Pedagogical robot as tutee

In other educational contexts, virtual agents and social robots have been framed as a supportive, but less knowledgeable playmate whom children can teach (Biswas et al., 2005; Chin et al., 2013; Hood, Lemaignan, & Dillenbourg, 2015; Park & Howard, 2014; Tanaka & Matsuzoe, 2012). In this approach, albeit less common than the tutor paradigm (Belpaeme et al., 2018), it has been successful in helping to consolidate children's learning and improve their learning retention (Tanaka & Matsuzoe, 2012).

Interacting with an agent presented as a younger and less-capable tutee of the child has been shown to elicit greater enjoyment and a higher tolerance of the robot's technical limitations and errors (Kanda et al., 2004). From a computational perspective, it also enables the robot to infer an estimate of the child's ability from their demonstrations. This student model can then be used to personalize learning content that is matched to the child's knowledge level (Park & Howard, 2015). However, learning with a robot tutee without involving a human teacher to provide instruction may hinder children's learning of new things. For example, in Tanaka and Matsuzoe (2012), a classroom teacher was needed to first teach a set of novel words before children could then teach those new words to a robot. Furthermore, the teacher needed to repeat the vocabulary lesson multiple times before some children could comprehend the words well enough to teach the robot. Thus, it seems a robot tutee is effective in emotionally engaging children, can be used to infer what children understand from their demonstrations, or can reinforce and improve the retention of what children already know. However, it is less effective in helping children learn new things.

2.4. Pedagogical robot as peer

Children's relationships with their human peers provide opportunities for learning through observing peers, teaching other peers, being in conflict with peers, and cooperating with peers (Bandura & Walters, 1963; De Lisi & Golbeck, 1999; Rubin, Bukowski, & Parker, 2007; Tudge & Rogoff, 1989). Peer learning is a bi-directional reciprocal learning activity in which students acquire knowledge and skill through actively helping and supporting each other (Topping, 2005). Learning with peers has great potential in bringing unique motivational and cognitive benefits for participating children (Damon, 1984). Guided reciprocal peer-questioning has been shown to lead students to ask more critical thinking questions, give more explanations and achieve greater learning (King, 1990). Additionally, peer tutoring, one form of peer learning, has been shown to benefit both a peer tutor and a peer tutee, improving both the tutor's and tutee's self-esteem and social adjustment (Allen, 1976). Thereby, when children actively take both roles in peer learning, it is very likely that they will benefit the most from the learning interaction.

Reciprocal peer interaction with computer agents or social robots is a far less explored paradigm, but has shown positive impact on student learning (Howard, Jordan, Di Eugenio, & Katz, 2017; Ros et al., 2016). When a student learns creative dance with a robot, the robot's role switch between acting as the lead and follower promotes the student's intrinsic motivation to learn (Ros et al., 2016). However, the robot's role-switching mechanism in Ros et al. (2016) was hardwired to be a leader in the first half and follower in the second half of the session, so it was not adaptive to children's learning progress and needs. Another study designed a mixed-initiative peer-like dialog agent to aid college students in learning Computer Science concepts (Howard et al., 2017). The agent behaved as either a less knowledgeable or a more knowledgeable peer by automatically tracking and shifting task initiative (i.e., who is making a contribution to achieve a goal) according to the student's knowledge and problem solving initiative history. To implement this reciprocal peer agent model, classifiers were trained to recognize task initiative, and then a fixed rule-based system was used to determine when to switch initiative during a collaborative problem solving context. This system presented in Howard et al. (2017) is probably the closest prior work to ours. Our work is, however, differentiated in two important ways. First, Howard's system did not adapt its policy to individual students whereas ours learns a personalized policy. Personalizing a peer robot's behavior

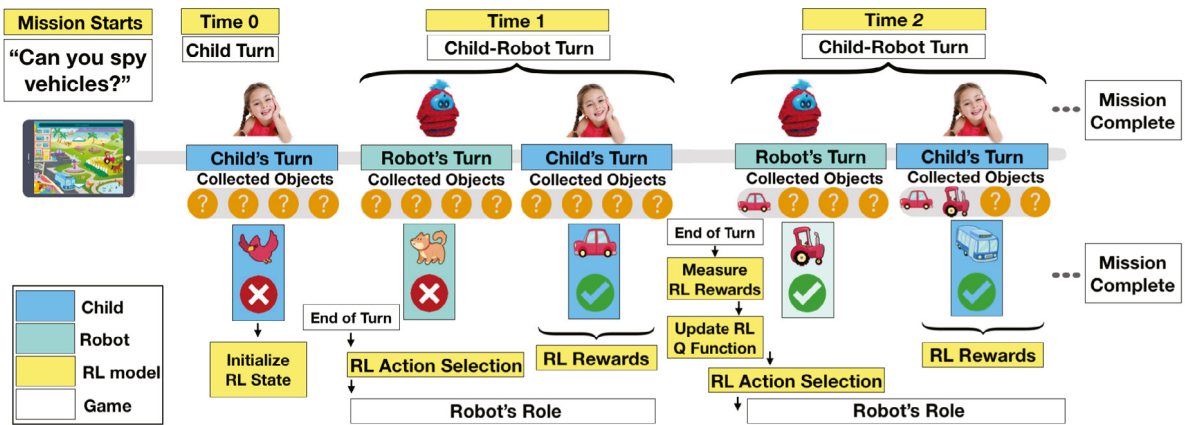


Fig. 3. A child and a robot take turns finding game objects that match the target category word in a given game mission. In the reinforcement learning (RL) model, one robot-child turn corresponds to one timestamp. At the beginning of each robot-child turn, the robot receives a role assignment (i.e., RL action) from the model and performs role-associated actions. At the end of the robot-child turn, the total reward of the RL action is calculated, and the Q-function is updated. When four objects are collected either by players, the mission is completed.

is an important dimension to explore in reciprocal child-agent interaction, as it has been shown to increase children's acceptance of the robot and have a positive influence on learning (Baxter, Ashurst, Read, Kennedy, & Belpaeme, 2017). Second, our system is designed to support young children's learning where collaborative play and emotional engagement are important, whereas Howard's system relies on dialog for much older students.

In light of this prior work and pedagogical theory, we hypothesize that reciprocal interactions with a robot peer, which allow a child to be both a tutor and a tutee of the robot, will more effectively cultivate his/her learning and social-emotional engagement than with a robot that is only a tutor or tutee. Furthermore, the role-switching policy should be adaptive and personalized to the needs of each child. In the following sections we present our system design, interaction design, and evaluation study to investigate this hypothesis.

3. Interaction design and materials

3.1. Child-robot educational game play

Children learn from different types of experiences, including playing games with others. Digital games have become increasingly prevalent and influential for language education, such as vocabulary learning (Chen, Tseng, & Hsiao, 2016; Hassinger-Das et al., 2016; Smith et al., 2013; Zou, Huang, & Xie, 2019). Given the great promises in game-aided learning, we designed a game-based vocabulary learning scenario, in which a child plays a collaborative vocabulary game on a tablet with a robot playmate. An overview of our integrated system is shown in Fig. 2. It consists of (1) a computer hub, (2) the *WordQuest* vocabulary game that runs on a touchscreen tablet, (3) *Tega* social robot, and (4) a front facing camera that captures children's facial expression during the interaction.

Each system component publishes its states and subscribes to data topics published by other components. The computer hub manages this communication using an open-source protocol called Robot Operating System (Quigley et al., 2009). It governs the game logic, i.e., when and how the robot should act based on the user input and game states, and records time synchronized interaction data. In the peer condition, the algorithm that guides the robot's adaptive role actions is also housed in the computer hub.

The *WordQuest* game app shown in Fig. 1(b) is similar to the classic *I Spy* game where a child and a robot take turns identifying virtual objects in a scene specified by a quest mission. A player (either child or robot) can pan around the scene, zoom in and out, click the objects, and read the object word. The current version of the game has 50 animated clickable objects that are age appropriate for young children. The game specifies quest missions for the child and robot to complete. Each quest mission is a challenge word that is likely to be unknown by young children (e.g., "can you find objects that are in crimson?"). The goal is for a child to learn the meaning of the challenge word (e.g., "crimson") by taking turns with a robot and finding the corresponding objects in the game scene (e.g., red objects in the scene). When a total of four correct objects are collected by the child and the robot, the mission is completed. The game has 11 missions – 5 from the *indoor scene* and 6 from the *outdoor scene*. The challenge words are: azure, gigantic, minuscule, garment, lavender, vehicle, delighted, crimson, soar, aquatic, and recreational activity. Fig. 3 depicts the game play scenario.

The social robotic platform we chose for the study is *Tega*, an appealing, expressive and fluffy robot designed and deployed as a learning companion for young children. The robot about 11 inches tall with a squash-and-stretch body and plush exterior (Fig. 1). An Android smartphone is used to graphically display the robot's animated face, control the motors, and function as the

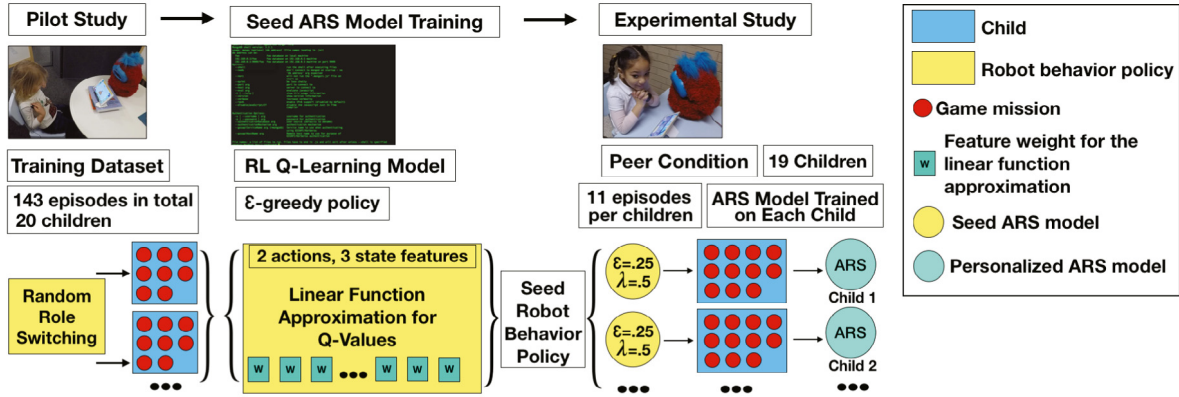


Fig. 4. During the pilot study, 20 children played with a robot that randomly switched its role between *tutor* and *tutee*, and the data of 143 game missions were collected. During the training phase, a reinforcement learning (RL) model was trained on the collected dataset using a linear function approximation for its Q-values, and obtained a pre-trained model as a seed adaptive role-switching (ARS) model. During the experimental study, the *peer* robot started with this pre-trained ARS model but further personalized it to each of the 19 children in the peer condition.

communication hub between the physical robot and the computer hub. In addition, *Tega* has a child-like voice, can move with emotive body gestures, and can adjust its speaking rate. It is also able to recognize simple verbal responses to questions it asks by recording children's speech via built-in microphones and using Google's Automatic Speech Recognition service to decode a child's utterance and perform simple natural language processing to extract his/her spoken intent. The robot has been used successfully in various educational settings in studies lasting several weeks for different learning tasks such as storytelling, vocabulary learning, puzzle solving, etc. (Gordon et al., 2016; Park et al., 2019).

The sensor modules collect children's interaction data from touch and vision modalities. We record all touch actions on the tablet (e.g., dragging, tapping) to track the game state along with other task-related data (e.g., interaction duration). An external USB camera is located behind the tablet and oriented toward the child's face to record the child's facial expression during interaction (Fig. 2).

3.2. Three child-agent interaction paradigms

The robot can adopt one of three roles, performing different sets of behaviors in each (Table 1). For a given robot-child turn, the robot exhibits three behaviors during its turn at three game event triggers: (1) when it is searching for an object, (2) when it selects an object, and (3) after it receives its result from the game on the tablet. Similarly, it demonstrates its behaviors during the child's turn at two game event triggers: (1) when the child is searching for an object, and (2) when the child receives their result from the game. Each robot behavior (e.g., "hint providing") has a set of 3–5 specific verbal expressions each accompanied by the robot's emotive body movements and facial expressions. When one behavior is executed, one of its associated verbal expressions (e.g., "Azure is a color") will be randomly selected to diversify the robot's speeches and actions.

- In the **tutor** paradigm, the robot knows the meanings of all words and behaves as a more skillful partner who demonstrates knowledge and gives informative feedback to the child without ever making a mistake throughout the entire game play. Sometimes the tutor robot may ask the child whether they need any help.
- In the **tutee** paradigm, the robot is situated as a novice who lacks the knowledge of all vocabulary words. The tutee robot occasionally asks the child for help, asks for an explanation for why it got something wrong, and shows curiosity and a positive attitude toward learning. Hence, the design encourages reflection and consolidation. The robot has a 0.4 probability of selecting a correct object, but neither explains the word's meaning nor provides explanation to the child. The intuition behind this design is that the tutee robot is only able to find correct objects by guessing. Thus, the child can only learn the word's meaning through trial-and-error with the robot.
- In the **peer** paradigm, the robot is situated as a reciprocal and adaptive partner that adapts its interaction style (tutor or tutee) to match the child's knowledge within each turn exchange. The sets of behaviors for tutor/tutee roles are the same as for the tutor-only or tutee-only conditions. For example, the robot can dynamically take the tutor role, proactively demonstrating where a correct object is and explaining its meaning when a child is really struggling with the game mission. Alternatively, the robot can switch to a tutee role – expressing intellectual curiosity, asking the child for help or misidentifying an object – when the child needs more practice to consolidate their learning. The role-adaptation mechanism is implemented using a reinforcement learning model, and its technical details are presented in the next section.

Table 1

Tutor and *tutee* robots only perform behaviors associated with the *tutor* and *tutee* roles throughout entire learning interaction, respectively. A *peer* robot is assigned with a role (tutor/tutee) by the adaptive role-switching (ARS) model at the beginning of each robot-child turn, and then performs the behaviors associated with the assigned role during the turn.

Role	Turn	Robot behavior	Definition	Verbal expression example
<i>Tutor</i>	Robot	Keyword Definition	Explain the meaning of the mission word	"Vehicle is something you can drive, steer or ride in."
		Object Searching	Express confidence when searching for an object	"I know which object is correct."
		Game Object Selection	Always select a correct object	NA
		Vocabulary Explanation	Explain why the object the robot chooses is correct	"Train is a vehicle, because we can ride in it."
		Response to Robot's Correct Answer	Confidently confirms the result	"Aha. I found it."
	Child	Offering Help	Offer to help the child find a correct object	"Do you need my help?"
		Keyword Definition	Explain the meaning of the mission word to the child	"Color azure means blue"
		Providing Hints	Share hints on the meaning of the word	"Azure is a color"
		Response to Child's Correct Answer	Express positivity and excitement	"We got one more done."
		Response to Child's Incorrect Answer	Express encouragement and support	"You will get this next time. I believe in you.", "Practice makes perfect, right?"
<i>Tutee</i>	Robot	Asking for Help	Ask the child to help find an object	"Can you help me find a correct object?", "Do you have any suggestions for me?"
		Object Searching	Express curiosity and enthusiasm to learn	"I love learning new words that I don't know.", "I am very curious whether this is correct."
		Object Selection	Often select a wrong object (probability is 0.6)	NA
		Asking for Explanation	Ask the child why the object the robot chooses is wrong	"Can you tell me why I am wrong?"
		Response to Robot's Correct Answer	Express excitement	"Yeah, I got it right!"
	Child	Response to Robot's Incorrect Answer	Express positivity and hope	"I will learn as the game goes. Don't you think so?"
		Learn from Child	Express enthusiasm to learn from the child	"Why did you choose this one? I want to learn from you."
		Curiosity toward Child Actions	Express curiosity in what the child is going to find	"I am curious of what you will find!", "I am excited to see what you spy"
		Response to Child's Correct Answer	Express excitement and hopes to learn from the child	"You got it right. Congratulations.", "I would like to learn. Do you want to teach me?"
		Response to Child's Incorrect Answer	Express encouragement and support	"We will learn this together.", "We all make mistakes but we improve from it too."

4. Adaptive role switching model

We use reinforcement learning (RL) to construct an active role-switching (ARS) policy for a reciprocal peer robot. The pipeline for developing, training, and testing the ARS model is outlined in Fig. 4. The resulting policy tells the robot which role it should take for each student given their specific state to promote vocabulary learning.

4.0.1. Reinforcement learning

Reinforcement learning is traditionally defined as part of a Markov Decision Processes (MDP) (Sutton & Barto, 1998). An MDP is a tuple $\langle S, A, P, R \rangle$ such that S is a state space, A is an action space, and P and R are the distribution of probabilities and rewards respectively. Value-function RL methods can be in general categorized into two types: model-free methods (e.g., Q-Learning) and model-based methods (e.g., R-MAX) (Brafman & Tennenholtz, 2001; Sutton & Barto, 1998). Since it is nearly impossible to construct a model accurately reflecting how children would learn in our context before any interaction happens, we used a model-free method (i.e., Q-learning) to learn the ARS policy for the adaptive peer robot.

4.0.2. Active role-switching policy formulation

An overview of how we formulate the ARS policy is shown in Fig. 3. We define one timestamp in the RL model as one robot–child turn pair. At the beginning of a robot–child turn, the robot receives a role assignment from the ARS policy based on the current RL state, and the robot performs a set of behaviors associated with that role throughout the turn. During the child’s turn, the child’s performance and reactions to the robot’s behaviors are used as the RL action’s rewards. At the end of the robot–child turn, the RL actions’ total reward is calculated, and the Q-function is updated. Each game mission starts with the child’s turn to initialize the RL model’s first state, and ends when four correct objects are collected irrespective of who collected the objects. The RL model’s action space A consists of two actions: *tutor* role and *tutee* role. The state representation S consists of three state features: (1) the current number of robot–child turns, $s_{turn} \in \{0, \dots, n\}$, (2) the current number of *tutor*-role occurrences, $s_{tutor} \in \{0, \dots, m\}$, $m \leq n$, $m \leq 4$, and (3) the current number of correct objects found by the child $s_{correct} \in \{0, \dots, q\}$, $q \leq 4$. These three state features are chosen because they indicate different aspects of children’s learning: learning progress (s_{turn}), the amount of coaching received (s_{tutor}), and learning performance ($s_{correct}$). Since the game is completed when four objects are collected, the game state also satisfies $s_{correct} + s_{tutor} \leq 4$.

4.0.3. Value function formulation

Given the large state space and sparsity of the real-time child–robot interaction samples, we use a linear value function approximation (LVFA) instead of tabular solutions. The LVFA method has the advantage of simple update rules, allowing easy interpretation of each state feature’s contribution to the RL model’s action selection. It also enables efficient generalization from a limited subset of the state space to a larger subset of the state space (Sutton & Barto, 1998). Therefore, the LVFA solution is chosen as the ARS model’s value function.

To implement the LVFA function, the stochastic gradient descent (SGD) learning method is used. The value function $v_{\pi}(S_t)$ is represented as $\hat{v}_{\pi}(s) = \mathbf{w}^T \boldsymbol{\phi}(s) = \sum_{i=1}^d w_i \phi_i(s)$, where s is the 3-dimensional RL state ($s_{turn}, s_{tutor}, s_{correct}$) at time t , $\mathbf{w}$ is a weight vector in the linear SGD approximation, d is $|\mathbf{w}|$, and $\boldsymbol{\phi}(s)$ is a feature vector representing state s that forms a linear basis for the set of approximate functions. To construct $\boldsymbol{\phi}(s)$, we take a classic approach – a second-order multivariate polynomial regression – to capture a curvilinear relationship in our model’s value function. Specifically, the regression formula is $\phi_i(s) = \prod_{j=1}^3 s_j^{c_{j,i}}$ where j is the RL state space dimension, and $c_{j,i}$ is an integer in the set $\{0, 1, 2\}$ (Sutton & Barto, 1998, p172). To avoid the over-fitting issue, we only keep the regression terms with the sum of its exponents lower than or equal to 3. Hence, the regression $\phi_i(s)$ has the following terms: $\{s_{turn}, s_{tutor}, s_{correct}, s_{turn}s_{tutor}, s_{turn}s_{correct}, s_{tutor}s_{correct}, s_{turn}s_{tutor}s_{correct}, s_{turn}^2s_{tutor}, s_{turn}^2s_{correct}, s_{tutor}^2s_{correct}, s_{tutor}^2s_{turn}, s_{correct}^2s_{turn}, s_{correct}^2s_{tutor}\}$.

Lastly, we implement an ϵ -greedy algorithm, where $\epsilon = 0.25$ with a discount factor $\lambda = 0.5$ after manually testing different values of ϵ and λ on the pilot dataset.

4.0.4. Reward function formulation

We define the act of finding a game object as an attempt $atp_t \in \{1, 0\}$ with 1 representing correct and 0 representing incorrect attempts, and the presence of the child offering help to the robot is denoted as $help_{child} \in \{1, 0\}$. The number of child’s consecutive incorrect attempts is $s_{consec_incorrect}$. The total reward r_{total} at t is $r_{T1} + r_{T2}$ when the robot takes a tutor role, and is r_{L1} when the robot takes a tutee role.

$$r_{T1} = \begin{cases} 10 * (s_{correct} - 0.5 * s_{turn}), & \text{if } atp_t = 1 \text{ and } s_{correct} \geq 0.5 * s_{turn}. \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

$$r_{T2} = \begin{cases} s_{consec_incorrect}, & \text{if } atp_t = 1. \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

$$r_{L1} = \begin{cases} 4 - s_{correct} + help_{child}, & \text{if } atp_t = 1. \\ 4 - s_{turn} + help_{child}, & \text{otherwise.} \end{cases} \quad (3)$$

The model’s reward function is designed to synchronize the robot’s tutor/tutee role with a child’s knowledge level in order to provide scaffolding actions at appropriate times while allowing him/her to experience the benefits of both roles. The reward functions, r_{T1} and r_{L1} , reward the robot’s actions that match the child’s knowledge level. For r_{T1} , when the child’s correct attempt accuracy is above 50% ($s_{correct} > 0.5 * s_{turn}$), the tutor role is increasingly rewarded to match the child’s mastery level.

Similarly, r_{L1} is larger when the child is more likely to be a novice. To assess how novice the child is, the robot uses (1) the child’s correctness in the past turns ($s_{correct}$) and (2) game mission progress (s_{turn}). The child is probably a novice and r_{L1} will be larger when the child cannot find correct objects in the beginning of a mission ($4 - s_{turn}$ if $atp_t = 0$), or s/he had a low accuracy rate but just found a correct object by chance ($4 - s_{correct}$ if $atp_t = 1$). In addition, a binary factor $help_{child}$ is added to r_{L1} , because when the child helps the robot find an object during the robot’s turn, the robot has more opportunities to observe how well the child understands a word. The constant factor in r_{L1} is set to 4, as it matches the total number of objects allowed for retrieval in a game mission. The scale factor in r_{T1} is set to 10 to keep the value ranges for both r_{T1} and r_{L1} consistent ($r_{T1}, r_{L1} \in [0, 5]$). In addition to rewarding the actions that elicit the synchronized knowledge level, we also add r_{T2} in order to ensure the robot will scaffold the child’s learning as a tutor only when they are really struggling to learn. Specifically, the robot’s tutor role is increasingly rewarded when the child struggles harder – this is manifested in the number of their consecutive incorrect attempts ($s_{consec_incorrect}$).

Table 2

Participant children in our between-subjects experiment were randomly divided into three conditions (*tutor*, *tutee* and *peer*), counterbalanced by age, gender, prior knowledge of the target vocabulary, and English proficiency (native or ELL).

Condition	Num. children	Mean age (SD)	Female (%)	ELL (%)	Mean Pre-Test Score (SD)
<i>tutee</i>	19	6.00(0.74)	57.89%	47.37%	2.68(1.37)
<i>tutor</i>	21	5.85(0.65)	61.9%	52.38%	2.44(1.43)
<i>peer</i>	19	5.95(0.60)	55.9%	47.37%	2.58(1.43)

4.0.5. Pre-training a seed active role-switching policy

Prior to the main study, a pilot study was conducted to collect an initial set of samples from children to train a seed ARS policy. The model was trained on 143 game missions collected from 20 children who interacted with a random role-switching robot. Using this pilot dataset, a seed model was trained off-line on 404 episodes of robot–child turns (recall each episode is one robot–child turn pair within a game mission). Overall, the robot took balanced tutor and tutee roles (tutor: 48%; tutee: 52%). In this pilot, the ϵ -greedy algorithm employed $\epsilon = 1$, so it always chose exploration over exploitation. It is evident that the policy did not converge after 404 episodes, but a gradually decreasing trend of the value error $U_t - \hat{v}(S_t, \mathbf{w}_t)$ was observed where $U_t = R + \gamma \hat{v}(S_{t+1}, \mathbf{w}_t)$. This seed policy served to accelerate the training of the ARS policy to each child during the main study.

5. Main study design

How do different child–agent interaction paradigms influence children’s learning and affective experience? Can an adaptive peer robot successfully leverage the advantages of both tutor and tutee roles to better promote children’s learning and affective experience? To investigate these questions, we designed a between-subjects experiment where participant children were randomly divided into three conditions, counterbalanced by age, gender, prior knowledge of the target vocabulary, and English proficiency (native or ELL). The three conditions correspond to the robot engaging the child either as a *tutor*, a *tutee*, or an adaptive and reciprocal *peer* (as described in 3.2).

5.1. Hypotheses

We expected that all three interaction paradigms would affect children’s vocabulary learning and affective experience, albeit in different ways. We hypothesized that children would learn vocabulary more effectively when interacting with a tutor robot than a tutee robot, while a tutee robot would be more emotionally engaging for children (as revealed through facial expressions) than a tutor robot. We also hypothesized that an adaptive robot peer, who reciprocally embodies the advantages of tutor and tutee roles, will be the most effective in fostering children’s learning and emotional engagement. The list of hypotheses in regards to children’s learning and affective experience is as follows:

- **H1:** Children interacting with the *tutor* robot will learn more target vocabulary words than children interacting with the *tutee* robot.
- **H2:** Children interacting with the *tutee* robot will be more expressive in their facial affect display and more engaged than children interacting with the *tutor* robot.
- **H3:** Children will learn the most target words when interacting with the *peer* robot across the three conditions.
- **H4:** Children interacting with the *peer* robot will be the most expressive in their facial affect display and show the highest enjoyment across the three conditions.

5.2. Participants

Sixty-four children between the ages of 5–7 were recruited for the study from a greater Boston public school. Five out of 64 children who were excluded from the data analysis since they did not provide informed consent, two had difficulty understanding how to play the vocabulary learning game *WordQuest* during the practice round and withdrew from the study, and three did not complete the entire protocol due to early school leave. Thus, a total of 59 children completed the study and their data were used for the quantitative analysis in this paper (Table 2). Thirty participants were in kindergarten (ages 5–6) and 29 participants were in 1st grade (ages 6–7). We found no statistically significant difference in children’s average age, grade level, pre-test score, and English proficiency across the three study groups.

5.3. Protocol and procedure

The study protocol consisted of a pre-test, two sessions with the robot followed by immediate post-tests, and a three-week delayed post-test (Fig. 5). The study was conducted in the Spring semester of 2018, between February and March. All children completed the pre-test within one week prior to the first robot session. Similarly, each robot session for all participants was conducted within three weeks.

In the first session, the experimenter guided the child through a practice round for five minutes, teaching the basics of the game mechanism. When teaching children how to play the game, the experimenter also discussed with them about how the robot could

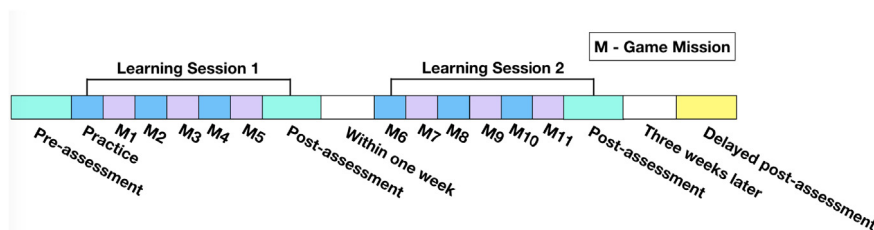


Fig. 5. The experiment consisted of two learning sessions, a pre-assessment, a post-assessment and a three-week delayed post-assessment. Children received five and six missions in session 1 and session 2, respectively.

Table 3

The results of children's vocabulary assessments by experimental condition were measured in terms of their mean and standard deviation scores. Children in the *peer* condition learned the most across all vocabulary acquisition measures. The distribution of top-performing children by experimental conditions shows that the top-performing children in the *peer* condition consistently outnumbered the top-performing children from other two conditions with respect to the immediate Δ score (IMMIT-CHANGE) and delayed Δ score (DLAY-CHANGE).

Condition	PRE-TEST	IMMIT-TEST	DLAY-TEST	IMMIT-CHANGE			DLAY-CHANGE		
	M (SD)	M (SD)	M (SD)	M (SD)	Top 9	Top 20	M (SD)	Top 11	Top 18
<i>Tutee</i>	2.68 (1.38)	4.63 (1.74)	4.74 (1.65)	1.95 (2.27)	1 (11%)	2 (10%)	2.05 (1.36)	1 (9%)	2 (11%)
<i>Tutor</i>	2.43 (1.43)	6.00 (2.07)	5.10 (1.95)	3.57 (2.33)	3 (33%)	7 (35%)	2.67 (1.94)	3 (27%)	6 (33%)
<i>Peer</i>	2.58 (1.43)	7.37 (2.41)	6.05 (2.14)	4.79 (2.51)	5 (56%)	11 (55%)	3.47 (2.14)	7 (64%)	10 (56%)

tap and move objects on the tablet without any hands and how it recognizes things on the screen – i.e., how the robot ‘talks’ to the tablet via WiFi and Data. Then, children played the *WordQuest* game with the robot for 20–30 min to learn a new set of words. In session 1, the child and the robot had to finish five missions, each mission introducing a new challenge word, to complete a game. Session 2 was comprised of 6 missions, with a different set of challenge words, to complete a game. In each mission, the child and the robot took turns to find four objects associated with the challenge word. If they were unable to complete a specific mission within six minutes, the game automatically terminated the current mission and loaded the next mission. This ensured that the child and robot would be able to try all required missions and see all the challenge words in given session. It also prevented them from being stuck in a mission for too long. For details of the game mechanics and interaction design, see Section 3.

6. Data analysis and results for vocabulary acquisition

6.1. Vocabulary learning analysis

We analyzed children's vocabulary acquisition performance per condition immediately after each session and three weeks after the learning interaction. Then, we examined the effect of vocabulary difficulty on children's learning acquisition (dividing the 11 challenge words into *Easy* and *Advanced* based on how many children got them right on the pre-test), and the interaction effect on children's learning between vocabulary difficulty and experiment conditions. We also measured the effect of children's vocabulary level prior to learning (*Top 50% on the pre-test* versus *Bottom 50% on the pre-test*) on learning outcomes to see how effective these three roles were on promoting vocabulary learning of children starting at different levels of proficiency.

The vocabulary test was comprised of the 11 challenge words that appeared in the *WordQuest* game. It followed the format of the Peabody Picture Vocabulary Test (PPVT) (Dunn & Dunn, 2007), where the child was shown pages with four pictures on each, and had to point to the picture that illustrated the meaning of the stimulus word spoken by the examiner. To reduce false positive answers, the examiner also encouraged the child to inform the examiner if the child did not know the meaning of the stimulus word. When the child selected a correct picture, the examiner asked the child the meaning of the stimulus word to avoid random guesses. The pre-test result was used to form a baseline of participants' knowledge before the robot intervention. The immediate post-tests were administered at the end of each learning session, with words that appeared in the given session. The three-week delayed post-test measured children's long-term retention of all 11 challenge words.

6.2. Vocabulary learning results

We analyzed children's learning outcomes on the target vocabulary using their pre-test scores (PRE-TEST), immediate and delayed post-test scores (IMMIT-TEST, DLAY-TEST), and Δ scores between the pre-test and the two post-tests (IMMIT-CHANGE, DLAY-CHANGE). For each analysis, we first did a Shapiro–Wilk test for normality and Levene Test for equality of variances, and found both tests statistically insignificant. Then, the one-way ANOVA analysis and trend analysis, which is a contrast-coded generalized linear model used to predict learning outcomes (Davis, 2010), was performed on each analysis category. See Table 3.

The ANOVA results show a significant effect of the robot's role in both the immediate post-test score and Δ score with children in the *peer* group obtaining the highest scores (IMMIT-TEST: $F(2, 56) = 8.14$, $p = 8e-4$; IMMIT-CHANGE: $F(2, 56) = 6.02$, $p = 0.002$).

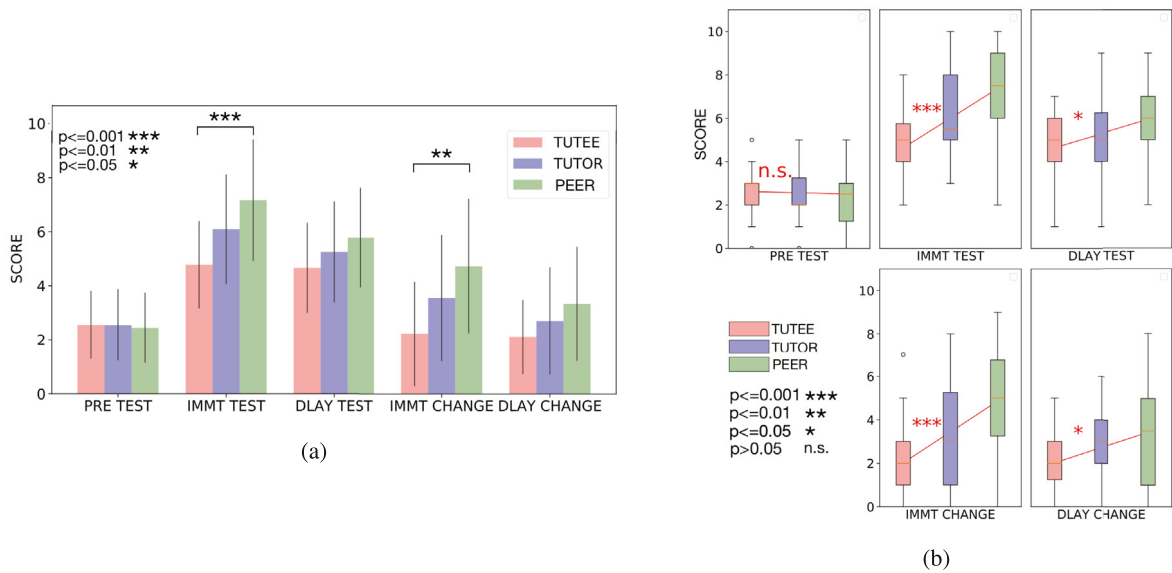


Fig. 6. Analyses on children's target vocabulary learning consist of (a) ANOVA analyses and (b) trend analyses. The ANOVA analyses showed a statistically significant difference with respect to the immediate post-test and Δ scores, and children in the *peer* condition achieved greater scores than children in the *tutee* condition according to the statistically significant difference between them found in the post-hoc analyses. The trend analyses showed a statistically significant trend of increase: *tutee* < *tutor* < *peer* for all four post-test vocabulary measurements.

Furthermore, their post-hoc analyses show that the differences in immediate post-test and Δ scores between *peer* and *tutee* conditions were statistically significant with children in the *peer* condition achieving higher scores. In contrast, the differences between *peer* and *tutor* conditions, and between *tutor* and *tutee* were not found statistically significant. See Fig. 6a.

The trend analysis showed a statistically significant trend of increase: *tutee* < *tutor* < *peer* for all four post-test vocabulary measurements (IMMIT-TEST: $F(2, 56) = 16.57$, $p = 2e-4$, DLAY-TEST: $F(2, 56) = 4.27$, $p = 0.043$, IMMIT-CHANGE: $F(2, 56) = 13.85$, $p = 4.6e-4$, DLAY-CHANGE: $F(2, 56) = 5.45$, $p = 0.023$). See Fig. 6b. This result suggests that children in the *tutor* condition consistently learned more vocabulary than children in the *tutee* condition. Also, interacting with the *peer* robot helped children learn the most words. This result supports H1 and H3.

6.3. Distribution of top-performing children by condition

We ranked children based on their Δ scores between the pre-test and the two post-tests, respectively (Table 3). Since multiple students were ranked as top 10 with respect with the immediate Δ score, top 9 students were selected as the threshold. For the same reason, top 11 and top 18 were selected for the delayed Δ score. We found that 5 of top-performing 9 students and 11 of top-performing 20 students are from the *peer* condition in terms of the immediate Δ score. With respect to the delayed Δ score, 7 of top 11 students and 10 of top 18 students are from the *peer* condition. The results showed that the top-performing children in the *peer* condition consistently outnumbered the top performing children from the other two conditions, and strengthens the validity of H3.

6.4. Effect of children's demographics on learning

We did two-way mixed ANOVA analyses to measure how children's age, grade level, language proficiency might affect their vocabulary acquisition and retention, and found no statistically significant results.

6.5. Effect of vocabulary difficulty on children's learning performance

To assess the impact of robot's role on helping children learn vocabulary words with varying difficulty, we split the words into two sets (*Easy* and *Advanced*) based on the number of students who correctly identified the challenge word prior to its mission. The *Easy* set ended up including *gigantic*, *vehicle*, *soar*, *recreational*, *lavender*. The *Advanced* set included *azure*, *minuscule*, *garment*, *delighted*, *crimson*, and *aquatic*.

For each set, we performed a 3×2 two-way mixed ANOVA with Experimental Condition (*tutee*, *tutor*, *peer*) as a between-subjects variable and Vocabulary Difficulty (*Easy*, *Advanced*) as a within-subject variable (Fig. 7(a)). The summary statistics of children's learning performance on the *Advanced* and *Easy* vocabulary sets are reported in Table 4. The results on the ANOVA analysis and its pairwise post-hoc analysis using Bonferroni adjustments were reported in Tables 5a and 6, respectively.

Table 4

The 11 target vocabulary words were split into *Easy* and *Advanced* sets, and the performance of children's vocabulary acquisition was measured with respect to how well they learned each vocabulary set. Similarly, participants were split into *Top* and *Bottom* sets based on their pre-test score. The performance of children's vocabulary acquisition per condition was measured with respect to how well each participant set learned 11 target words.

Category	Condition	N	Immediate Δ score			Delayed Δ score		
			Mean	SD	Median	Mean	SD	Median
Vocabulary difficulty								
Easy	Tutee	19	0.68	1.42	1.00	0.63	1.13	1.00
	Tutor	21	1.57	1.47	1.00	1.38	1.33	1.00
	Peer	19	1.47	1.46	1.00	1.11	1.37	1.00
	Total	59	1.25	1.50	1.00	1.05	1.32	1.00
Advanced	Tutee	19	1.21	1.06	1.00	1.37	0.93	1.00
	Tutor	21	1.86	1.55	1.00	1.14	1.21	1.00
	Peer	19	3.26	1.52	3.00	2.32	1.45	2.00
	Total	59	2.10	1.63	2.00	1.59	1.32	1.00
Student performance								
Bottom	Tutee	8	3.25	2.05	3.00	2.63	1.60	3.00
	Tutor	12	4.50	1.98	4.00	3.08	1.93	3.00
	Peer	9	5.56	2.83	5.00	4.00	2.45	4.00
Top	Tutee	11	1.00	2.00	2.00	1.64	1.12	2.00
	Tutor	9	2.33	2.29	1.00	2.11	2.03	2.00
	Peer	10	4.10	2.08	4.00	3.00	1.94	3.00

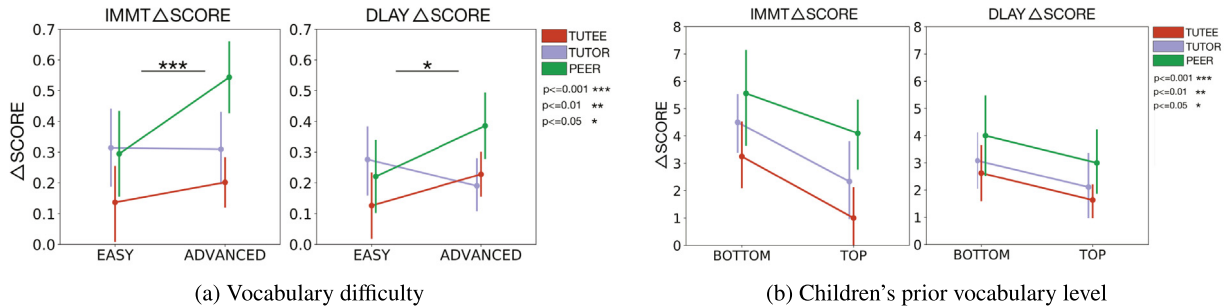


Fig. 7. (a): The 2×3 mixed ANOVA analyses with vocabulary difficulty (VD) and experiment condition (EC) on Δ vocabulary scores show a statistically significant interaction effect of EC with VD on both the immediate and delayed Δ scores, and statistically significant main effects of both EC and VD on the immediate Δ score. The results show that children learned and retained easy words equally well per condition, but children in the *peer* condition scored higher than children in the other two conditions when learning challenging words. (b): The 2×3 ANOVA analyses with children's prior vocabulary level (PVL) and experiment condition (EC) show statistically significant main effects of both EC and VL on the immediate Δ score, but no interaction effect between EC and PVL was observed.

We found a statistically significant interaction effect of Experimental Condition with Vocabulary Difficulty on both the immediate and delayed Δ scores. More specifically, children were able to learn easy words equally well per condition, but they learned advanced words differently per condition. The pairwise post-hoc analysis shows statistically significant differences in the immediate Δ score between *peer* and *tutor/tutee* conditions, where children in the *peer* condition learned more advanced words than children in the other two conditions. They also retained more advanced words in the *peer* condition than children in the *tutor* condition, a difference in the delayed Δ score that was statistically significant according to the post-hoc analysis. The results show that learning advanced words were harder for children in the *tutor* and *tutee* conditions than for those students in the *peer* condition.

We also found a significant main effect of Vocabulary Difficulty on the Immediate Δ Score. The post-hoc analysis shows that children interacting with the *peer* robot learned more advanced words than easy words, a difference that was statistically significant. The difference can be partially explained by the ceiling effect when children were learning the easy words in the *peer* condition.

Overall, these results show that children interacting with a *peer* robot learned and retained words best compared to the other conditions. This validates **H3**, i.e., the *peer* robot will promote children's vocabulary learning more than the other conditions, especially for the advanced vocabulary words.

6.6. Effect of children's prior vocabulary level on their learning performance

To assess whether children starting at different vocabulary levels learned the target words equally well across the three conditions, we split all 59 participants into two groups based on their pre-assessment vocabulary scores. In the *Bottom 50%* group children scored below the median in the pre-test, and in the *Top 50%* group children scored above the median (Table 4). For

Table 5

(a) The 2×3 mixed ANOVA analyses with vocabulary difficulty (VD) and experiment condition (EC) on Δ vocabulary scores show a significant interaction effect of EC with VD on both the immediate and delayed Δ scores, and significant main effects of both EC and VD on the immediate Δ score. The results showed that children learned and retained easy words equally well per condition, but advanced words differently per condition. (b) The 2×3 ANOVA analyses with prior vocabulary level (PVL) and experiment condition (EC) on Δ vocabulary scores show statistically significant main effects of both EC and PVL on the immediate effect Δ score. No statistically significant interaction effect between EC and PVL was observed.

(a)								
Source	Immediate Δ score				Delayed Δ score			
	df	MS	F	p	df	MS	F	p
EC	2	0.598	6.159	.004**	2	0.152	2.427	.098
VD	1	0.292	5.985	.018*	1	0.090	1.813	.184
EC * VD	2	0.169	3.470	.038*	2	0.172	3.443	.039*
Within groups	56	.049			56	0.050		

(b)										
Source	Immediate Δ Score					Delayed Δ Score				
	df	SS	MS	F	p	df	SS	MS	F	p
EC	2	77.290	38.650	7.951	<.001***	2	19.310	9.655	2.748	.073
PVL	1	55.810	55.811	11.482	<.001***	1	14.120	14.123	4.019	.050
EC * PVL	2	1.811	0.910	0.187	.830	2	0.000	0.001	0.000	.999
Residuals	53	257.620	4.86			53	186.230	3.514		

* = $p < 0.05$.

** = $p < 0.01$.

*** = $p < 0.001$.

Table 6

The post-hoc pairwise comparison analyses between experimental condition (EC) and vocabulary difficulty (VD) on Δ vocabulary scores show that children in the *peer* conditions learned statistically significantly more advanced words than both children in the *tutor* and *tutee* conditions (Immediate Δ Score), and retained significantly more advanced words than children in the *tutor* condition (Delayed Δ Score). In addition, children in the *peer* condition learned more advanced words than easy words, a statistically significant difference that can be partially explained by the ceiling effect when children were learning the easy words in the *peer* condition.

Source		Immediate Δ score					Delayed Δ score				
		Estimate	SE	df	t-value	p-value	estimate	SE	df	t-value	p-value
VD	EC comparison										
	Tutee-Tutor	-0.177	0.086	101	-2.076	.121	-0.150	0.075	111	-1.995	.146
	Peer-Tutee	0.158	0.088	101	1.802	.223	0.095	0.077	111	1.231	.663
	Peer-Tutor	-0.020	0.086	101	-0.229	1.000	-0.055	0.075	111	-0.734	1.000
	Tutee-Tutor	-0.108	0.086	101	-1.261	0.631	0.038	0.075	111	0.500	1.000
	Peer-Tutee	0.342	0.088	101	3.905	<.001***	0.158	0.077	111	2.051	.128
EC	Peer-Tutor	0.234	0.086	101	2.741	.022*	0.196	0.075	111	2.602	.032*
	VD comparison										
	Tutee	-0.065	0.072	56	-0.906	.369	-0.102	0.073	56	-1.404	.166
	Tutor	0.005	0.068	56	0.070	.945	0.086	0.069	56	1.244	.219
Peer	Easy-Advanced	-0.249	0.072	56	-3.478	.001**	-0.165	0.073	56	-2.276	.027*

* = $p < 0.05$.

** = $p < 0.01$.

*** = $p < 0.001$.

each set of students, we performed a 3×2 two-way ANOVA with Experimental Condition (*tutee*, *tutor*, *peer*) as a between-subjects variable and children's Prior Vocabulary Level (*Bottom*, *Top*) as a between-subjects variable on the immediate and delayed Δ scores, respectively (Fig. 7(b)).

The results of the ANOVA analysis are displayed in Table 5b. We found a significant main effect of Experimental Condition and a significant main effect of Vocabulary Level on the immediate Δ score. For the delayed Δ score, neither the Experimental Condition nor the Vocabulary Level had any statistically significant main effect. The results show that children from the *Bottom* group learned statistically significantly more vocabulary words than children in the *Top* group in the immediate post-test. A ceiling effect may exist in the *Top* group to explain the difference between the *Top* and *Bottom* groups. Last, we did not observe a statistically significant interaction of Experimental Condition with Vocabulary Level on the immediate and delayed Δ scores.

These results indicate that children starting at different vocabulary levels are all able to learn words by interacting with the robot. In addition, children starting with low vocabulary knowledge are able to learn challenge words with a robot at least as effectively as children with high prior knowledge.

Table 7

The ANOVA and Kruskal–Wallis ANOVA analyses on children's affective display by condition showed that children that interacted with a *peer* robot exhibited stronger and more varied facial expressions, particularly positive valence, during the interaction.

Affect Metric	Tutee (M $\pm$ SD)	Tutor (M $\pm$ SD)	Peer (M $\pm$ SD)	F	p-value	post-hoc results (p-value)
valence	-0.198 $\pm$ 3.280	0.376 $\pm$ 4.688	4.651 $\pm$ 6.809	10.168	0.006	[tutee-peer:0.018],[tutor-peer:0.011]
joy	1.659 $\pm$ 1.698	2.411 $\pm$ 3.576	5.382 $\pm$ 6.101	9.946	0.007	[tutee-peer:0.019],[tutor-peer: 0.012]
smile	2.645 $\pm$ 2.545	3.435 $\pm$ 4.202	7.354 $\pm$ 6.962	9.910	0.007	[tutee-peer:0.015],[tutor-peer:,0.015]
cheekRaise	0.467 $\pm$ 0.426	0.979 $\pm$ 1.799	1.519 $\pm$ 1.509	9.404	0.009	[tutee-peer:0.023],[tutor-peer: 0.023]
jawDrop	29.757 $\pm$ 20.121	18.345 $\pm$ 15.610	31.495 $\pm$ 16.262	7.638	0.022	[tutor-peer: 0.030]
mouthOpen	19.716 $\pm$ 17.362	9.474 $\pm$ 8.885	19.651 $\pm$ 14.330	7.148	0.028	[tutor-peer: 0.029]
eyeClosure	18.953 $\pm$ 13.774	10.311 $\pm$ 7.819	10.023 $\pm$ 11.484	6.659	0.036	[tutee-peer: 0.038]
browFurrow	0.576 $\pm$ 0.885	0.347 $\pm$ 0.326	0.934 $\pm$ 0.904	6.653	0.036	NA
lidTighten	0.643 $\pm$ 0.895	0.251 $\pm$ 0.432	0.409 $\pm$ 0.984	6.629	0.036	[tutee-tutor: 0.035]

7. Data analysis and results for emotional engagement

7.1. Children's facial affect measurement

Human facial expression is one of the most powerful channels to sense and detect affective states due to the rich expressiveness of human face. Many resources that support automatic data collection and video analysis (Calvo & D'Mello, 2011; De la Torre & Cohn, 2011; Ekman, Friesen, & Ellsworth, 2013) also help researchers using facial affect data. Recent advances in pattern recognition and multimodal sensing continue to improve technology's ability to automatically detect and analyze subtle human affect at a higher frequency over a longer period of time. The affect sensing technology has been widely used in research and commercial settings, such as analyzing children's engagement during learning and studying gender differences in human facial expression (Kory Westlund et al., 2017; McDuff, Kodra, Kaliouby, & LaFrance, 2017). We selected Affdex SDK¹ to evaluate our participants' facial affective display during learning after carefully reviewing prior works in the field. Affdex SDK is trained on the world's largest dataset of facial expressions to accurately code the facial expressions of diverse population (McDuff et al., 2016). We extracted 30 facial affect features for 59 children from videos (30 fps) recorded by the front-view camera. Affdex outputs 30 affect-related metrics including 7 emotions (e.g., joy), 21 facial features (e.g., brow furrow), in addition to arousal and valence. Each metric ranges between [0,100] that indicates the intensity of the feature except for Valence which has a scale of [-100,100] to represent intrinsic positive and negative emotions. To filter out artifacts or dropped frames in the raw metric vectors, we applied a median filter over a sliding window of 15 frames per second and extracted the mean values of each affect metric.

7.2. Children's affective display per condition

The mean values of affect features were significantly skewed except those for *attention* and *engagement*, according to Shapiro–Wilk test of normality and Levene's test for the equality of variance. Given this, we performed an ANOVA to measure the mean values of *attention* and *engagement* by condition. We used a Kruskal–Wallis ANOVA (KW-ANOVA) to measure the other 28 features per condition given that it does not assume normality of population distributions. We then performed a Dunn's test with Bonferroni correction for the post-hoc analysis following the rejection of the KW-ANOVA test.

The ANOVA and KW-ANOVA analyses show that 9 of the 30 affect metrics were significantly different by experimental condition (see Table 7). The nine affect metrics are *valence*, *joy*, *smile*, *cheek raise*, *jaw drop*, *mouth open*, *eye closure*, *brow furrow*, and *lid tighten*. Neither *attention* nor *engagement* yielded significant ANOVA results.

Our post-hoc analyses show that children in the *peer* condition scored statistically significantly higher than children in either the *tutor* or *tutee* conditions with respect to *valence*, *joy* and *smile*. Furthermore, children in the *peer* condition had statistically significantly higher scores on *jaw drop* and *mouth open* than children in the *tutor* condition. Children in the *peer* condition also had statistically significantly lower scores on *eye closure* than children in the *tutee* condition. Last, children interacting with a *tutee* robot exhibited a greater score on *lid tighten* than children interacting with a *tutor* robot.

These results show that children interacting with the *peer* robot exhibited stronger and more varied facial expressions, particularly positive valence, during the interaction. This finding confirms H4: a *peer-adaptive* robot will elicit children's greatest affective display and enjoyment among three conditions. Only the *lid tighten* feature has a statistically significant difference between the *tutee* and *tutor* conditions (being stronger in the *tutee* condition). Children in the *tutee* condition had higher trending scores on 5 of 9 affect metrics than the *tutor* condition, though the post-hoc analyses did not reveal significance. The results suggest that children who interacted with the *tutee* robot demonstrated slightly greater affective displays than children with the *tutor* robot – providing weak support to H2: i.e., children interacting with a *tutee* robot will be more affectively expressive and engaged than children with a *tutor* robot.

¹ Affectiva, Affdex SDK <https://www.affectiva.com/product/emotion-sdk/>.

8. Discussion

8.1. Children's vocabulary learning

Overall all, we observe that the *peer* robot promoted children's greatest learning among the three robots. Notably, our analysis of children's prior vocabulary knowledge showed that learning with the *peer* robot promoted children's learning, irrespective of their prior competence.

We found that children in the *tutor* condition learned more words than children in the *tutee* condition, and this difference was statistically significant. In addition, the *tutor* condition had more top-performing children than the *tutee* condition. These results validate **H1**. Children who interacted with a *tutee* robot probably performed worse because they were unable to grasp the meanings of new words simply through their own trial-and-error and observing the robot's trial-and-error, without receiving explicit guidance or informative feedback from the robot. Hence, our results support that a pedagogical agent's knowledge demonstration is important for learning something new, and is also consistent with prior work using social robots in a learning-by-teaching paradigm (Tanaka & Matsuzoe, 2012).

Compared with the *tutor* and *tutee* conditions, children in the *peer* condition learned the most vocabulary, learned the most advanced vocabulary, retained the advanced vocabulary better, and had the most top-performing children. These results validate **H3**. Children's improved performance on learning and retaining advanced words suggests that the reciprocal and adaptive *peer* robot is particularly effective at facilitating learning when it becomes challenging. Children who interacted with the *tutor* robot performed worse than the *peer* robot, probably because children in the *tutor* condition were given correct answers immediately and could not benefit enough from opportunities to explore, reflect, or try to actively infer the word's meaning (e.g., the benefits of mutual peer-exploration), and then consolidate what they learned (e.g., the benefits of learning-by-teaching). Such limitations may have become especially evident when game missions had challenging words that required more effort to acquire and were also easier to forget without sufficient practice. Unlike the *tutor* robot, the *peer* robot sometimes played as a novice who also occasionally made mistakes, asked questions, requested explanations, and made bids for the child's help. This reciprocal interaction design enabled children to explore the meaning of challenging words together, reinforce word meaning immediately once learned, and leverage the benefits of the *learning-through-teaching* paradigm to foster children's consolidation, reflection, and deeper learning.

In addition, we acknowledge that children in our study might have a diverse range of cognitive development and linguistic skills. Cognitive development, for example, between kindergarteners and first graders differs hugely. Such great cognitive and linguistic diversity in our participants may also explain why children learned the most target words in the *peer* condition. The *peer* robot was able to personalize its interaction style to each child based on her/his individual learning progress, whereas the two fixed-role robots used the same interaction strategies to every child. This finding provides greater motivation for developing adaptive educational technology for diverse young learners.

Given these reciprocal benefits, our findings support that a peer-like pedagogical agent that can adaptively engage children as either an expert or novice at appropriate times offers compelling interaction benefits that foster children's learning. To our knowledge, this is the first work to explore this paradigm in early childhood education.

8.2. Children's affective display

Our results show that the *peer* robot brought children the greatest enjoyment during vocabulary learning. Children in the *peer* condition were also more expressive than children in the *tutor* and *tutee* conditions. Specifically, children interacting with the reciprocal peer robot exhibited statistically significantly stronger facial displays in six affect features, three of which had positive valence. This finding confirms **H4**, that children in the *peer* condition would exhibit the greatest affective diversity and positive affect. Furthermore, because the reciprocal peer robot also exhibited *tutee* behaviors (e.g., making mistakes, asking for help, showing curiosity and positivity in learning), this may have served to support children's active learning and boost children's self-esteem when learning became challenging. Such differences in children's affect could have contributed to children's improved learning performance in the *peer* condition, particularly when learning advanced words.

In addition, we found the differences in children's *mouth open* and *jaw drop* between *peer* and *tutor* conditions statistically significant (both being stronger in the *peer* condition). This is probably because children talked to the robot more often when interacting with the *peer* robot. Greater verbalization and conversational turn-taking with the robot may also have helped to foster deeper learning. As suggested in prior work, active learning via a think-aloud strategy can promote young children's persistent learning gains and strengthen their engagement when performing cognitively demanding tasks (Ramachandran, Huang, Gartland, & Scassellati, 2018). For this study, we only collected videos of children's facial expressions, and thus did not run analysis on children's vocalization and speech due to the lack of audio data.

Interestingly, we observed that children who played with the novice robot exhibited only slightly stronger affect than children with the expert robot. Only one of the 30 affect features, *lid tighten* showed a statistically significant post-hoc result between the two conditions, being stronger in the *tutee* condition. This may be because children who played with the novice robot did not receive any help, causing children to feel more confused and frustrated, in contrast to children who received guidance and knowledge demonstration from playing with an expert robot. This finding only provides weak support to **H2**, i.e., children interacting with a *tutee* robot will be more expressive and engaged than children with a *tutor* robot. It is quite possible that children in the *tutor* and *tutee* conditions felt bored and disengaged from time to time, albeit for different reasons. Playing with an expert robot who never makes mistakes could potentially get a bit repetitive and predictable for children. In contrast, playing with a novice robot who always guesses and cannot help when the child struggles could potentially lead children to become disengaged due to the frustration and confusion from time to time.

8.3. Building rapport with a reciprocal peer

In light of these observations, we argue that the behavior of an adaptive, reciprocal peer is more engaging, interesting, and fun for children because the robot is not totally predictable and it encourages social reciprocity between child and robot through mutual support. For instance, it becomes interesting for children when a robot makes mistakes as a capable playmate – it draws the child in – especially when the robot asks the child for help. The child also feels more compelled to help the robot, because the robot can help the child when they struggle from time to time. This builds a sense of camaraderie during game play, and the rapport also helps children stick with it when learning becomes challenging.

Building social support and rapport may be particularly important for young children from a developmental standpoint as they are wired to learn from playful interaction with friendly others. For instance, building social rapport has other benefits such as peer-modeling and emulation that have shown to promote children's growth mindset (Park, Rosenberg-Kima, Rosenberg, Gordon, & Breazeal, 2017), curiosity (Gordon, Breazeal, & Engel, 2015), and creativity (Ali, Moroso, & Breazeal, 2019). Prior work has also shown that building a positive relationship with a social robot over longer term, repeated encounters is also correlated with increased learning outcomes (Kory-Westlund, 2019). Hence, beyond presenting a pedagogical agent as a peer through appearance, backstory, linguistic behavior, and other stylistic attributes – we argue that the adaptive, reciprocal qualities of how a pedagogical agent engages with a child to create a sense of camaraderie and social rapport has a multitude of potential benefits through mechanisms of social learning and the psychology of social engagement to benefit not only skill learning and emotional engagement, but also to foster broader developmental benefits as well.

9. Limitations and future work

There are a number of ways this work could be extended and deepened. First, the current study design exposes children to target vocabulary in only a single session. The overall amount of exposure is relatively short – only two 30-min interaction sessions with 11 target vocabulary words. We may see a greater difference in terms of children's learning performance across the three conditions over a longer term study and opportunities for repeated practice.

Longer term encounters would also open up the possibility to explore the effects of personalization across all three conditions. We used reinforcement learning to train an adaptive policy for when to switch roles between tutor and tutee. The interactions with each child were, however, too brief to achieve deeper personalization. This is why we use term *adaptive* rather than *personalized*. In the long-term deployment, we had expect all roles to benefit from becoming more personalized to specific children's needs and behaviors.

Second, our current study implemented the role-switching policy using a Q-learning model. A variety of other models can be potentially used to implement the personalization policy, such as Gaussian Process models. Therefore, we plan to implement multiple role-switching models and compare their performance and impact on children's learning and engagement in a future long-term study.

In future work, it would be interesting to examine how children's affective displays differ when they work on game missions containing easily comprehensible words versus challenging words per experimental condition. It is plausible that there may be an interaction effect between robot's role and word difficulty on children's affective display. More specifically, children in the *tutee* condition may exhibit greater affective displays than those in the *tutor* condition only when the game mission contains easy words instead of challenging words. In addition, our current work analyzed children's affective display over the entire learning sessions, but future work can also investigate how different within-interaction events (e.g., when the robot makes a mistake, when the robot shows encouragement) impact children's affective displays. More contextual information can be integrated into the interpretation of children's facial expressions.

It would also be interesting to examine children's social rapport with the robot playmate in all three conditions, and investigate how this impacts vocabulary learning, emotional engagement, as well as other developmental benefits such as promoting curiosity. We had expect to see personalization boost social rapport and these possible benefits based on prior work. In addition, while prior work has tended to focus on 1:1 interaction, it would be interesting to explore small group interaction and the effects on rapport and learning outcomes.

Finally, our work has focused on supporting the learning and engagement needs of young children given the critical importance of early childhood education as well as the high economic and academic impact of early intervention. However, we believe the interaction affordances and support of a reciprocal, adaptive peer will benefit older students, too. Hence, it would be interesting to study the effects of older students across conditions and for a range of subject matter and learning attitudes. We also argue it is important to develop best practices and ethical guidelines for the use of intelligent pedagogical agents for young children to help ensure they contribute positively to children's developmental needs.

10. Conclusion

This work is the first to implement and evaluate a reinforcement learning-based reciprocal child-agent peer-learning paradigm for children, and is the first to compare the impact of different roles of pedagogical agents *tutor*, *tutee*, *peer* on children's learning and affect together. We developed a novel bidirectional child-agent peer-learning paradigm, inspired by children's peer-to-peer interaction and built using a reinforcement learning model. The robot was rewarded for synchronizing its behaviors to a child's knowledge level, and for its scaffolding actions when the child is struggling. This enabled the child and robot to be exposed to both tutor and tutee roles dynamically, creating a socially rich interaction experience that builds a sense of camaraderie. In fact,

in the videos we annotated, we do see children positively responding to the robot's verbal and nonverbal encouragement cues and offering the same to the robot, evidencing their emotional and relational engagement with the *peer* robot. We explored the impact of different educational agent's roles (*tutor*, *tutee* and *peer*) on children's learning and emotional engagement through collaborative play in the context of an educational game. We found that children who interacted with a reciprocal, adaptive *peer* agent showed the greatest vocabulary learning, varied face-based affect, and positive valence among the three types of pedagogical agents.

In sum, our technical contribution and real-world evaluation study in a public school adds to a growing body of work exploring how different student-agent interaction paradigms impact young children's behavior, emotional engagement, and learning. Given the importance of effective interventions during early childhood and the importance of creating an emotionally engaging experience for young children, we hope this work contributes to the realization of intelligent and emotionally engaging pedagogical agents for this important, yet relatively under-served learner population with AI-enabled solutions.

CRedit authorship contribution statement

Huili Chen: Conceptualization, Methodology, Software, Validation, Data curation, Writing - original draft, Visualization, Writing - review & editing, Investigation. **Hae Won Park:** Conceptualization, Writing - original draft, Project administration, Writing - review & editing. **Cynthia Breazeal:** Funding acquisition, Supervision, Writing - review & editing.

Acknowledgment

This work was supported by NRI IIS 1734443, NSF, USA IIP-1717362 and Intel Graduate Fellowship. We sincerely thank Shang-Yun Wu for developing game assets and helping run the study. We also thank teachers and staff from the Boston public schools for supporting our study.

References

- Ali, S., Moroso, T., & Breazeal, C. (2019). Can children learn creativity from a social robot? In *The 12th conference on creativity & cognition*.
- Allen, V. L. E. (1976). *Children as teachers: Theory and research on tutoring*. New York: Academic Press.
- Baker, R. S. J. d., Corbett, A. T., Gowda, S. M., Wagner, A. Z., McLaren, B. A., Kauffman, L. R., et al. (2010). Contextual slip and prediction of student performance after use of an intelligent tutor. In P. De Bra, A. Kobsa, D. Chin (Eds.), *User modeling, adaptation, and personalization* (pp. 52–63). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Bandura, A., & Walters, R. (1963). *Social learning and personality development*. Holt, Rinehart and Winston, URL <https://books.google.com/books?id=WhHAAAAMAAJ>.
- Baxter, P., Ashurst, E., Read, R., Kennedy, J., & Belpaeme, T. (2017). Robot education peers in a situated primary school study: Personalisation promotes child learning. *PLoS One*, 12(5), e0178126.
- Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, 3(21), URL <http://robotics.sciencemag.org/content/3/21/eaat5954>.
- Biswas, G., Leelawong, K., Schwartz, D., Vye, N., & at Vanderbilt, T. T. A. G. (2005). Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence*, 19(3–4), 363–392.
- Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4–16, URL <http://www.jstor.org/stable/1175554>.
- Brafman, R. I., & Tennenholtz, M. (2001). R-MAX - A general polynomial time algorithm for near-optimal reinforcement learning. In *IJCAI international joint conference on artificial intelligence*.
- Breazeal, C., Morris, R., Gottwald, S., Galyean, T., & Wolf, M. (2016). Mobile devices for early literacy intervention and research with global reach. In *Proceedings of the third (2016) ACM conference on learning @ scale* (pp. 11–20). New York, NY, USA: ACM.
- Brophy, S., Biswas, G., Katzlberger, T., Bransford, J., & Schwartz, D. (1999). Teachable agents: Combining insights from learning theory and computer science. In S. P. Lajoie, & M. Vivet (Eds.), *Artificial intelligence in education* (pp. 21–28). IOS Press.
- Brown, L., & Howard, A. M. (2015). Engaging children in math education using a socially interactive humanoid robot. In *IEEE-RAS international conference on humanoid robots*.
- Calvo, R. A., & D'Mello, S. K. (2011). *New perspectives on affect and learning technologies* (Vol. 3). Springer Science & Business Media.
- Chang, C. W., Lee, J. H., Chao, P. Y., Wang, C. Y., & Chen, G. D. (2010). Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. *Educational Technology and Society*.
- Chen, M., Tseng, W.-T., & Hsiao, T.-Y. (2016). The effectiveness of digital game-based vocabulary learning: A framework-based view of meta-analysis: The effectiveness of DGBL. *British Journal of Educational Technology*, 49.
- Chin, D., Dohmen, I., & Schwartz, D. (2013). Young children can learn scientific reasoning with teachable agents. *IEEE Transactions on Learning Technologies*, 1, 6, 248–257.
- Corbett, A., Kauffman, L., McLaren, B., Wagner, A., & Jones, E. (2010). A cognitive tutor for genetics problem solving: Learning gains and student modeling. *Journal of Educational Computing Research*, 42(2), 219–239.
- Damon, W. (1984). Peer education: The untapped potential. *Journal of Applied Developmental Psychology*, 5, 331–343.
- Davis, M. J. (2010). Contrast coding in multiple regression analysis: Strengths, weaknesses, and utility of popular coding structures. *Journal of Data Science*.
- De la Torre, F., & Cohn, J. F. (2011). Facial expression analysis. In *Visual analysis of humans* (pp. 377–409). Springer.
- De Lisi, R., & Golbeck, S. L. (1999). Implications of Piagetian theory for peer learning. In *Cognitive perspectives on peer learning*.
- Dehaene, S. (2009). *Reading in the brain: The new science of how we read*. New York, USA: Penguin Books.
- D'mello, S., & Graesser, A. (2013). Autotutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems*, 2(4), 23:1–23:39, <http://doi.acm.org/10.1145/2395123.2395128>.
- D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies*, 70(5), 377–398, URL <http://www.sciencedirect.com/science/article/pii/S1071581912000250>.
- Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test*. Pearson Assessments.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (2013). *Emotion in the human face: Guidelines for research and an integration of findings* (Vol. 11). Elsevier.
- Fridin, M. (2014). Storytelling by a kindergarten social assistive robot: A tool for constructive learning in preschool education. *Computers and Education*.

- Garcia, E., & Weiss, E. (2017). *Education inequalities at the school starting gate: Gaps, trends, and strategies to address them: Tech. rep.*, Economic Policy Institute.
- Gordon, G., & Breazeal, C. (2015). Bayesian active learning-based robot tutor for children's word-reading skills. In *AAAI*.
- Gordon, G., Breazeal, C., & Engel, S. (2015). Can children catch curiosity from a social robot? In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction* (pp. 91–98). New York, NY, USA: ACM.
- Gordon, G., Spaulding, S., Westlund, J. K., Lee, J. J., Plummer, L., Martinez, M., et al. (2016). Affective personalization of a social robot tutor for children's second language skills. In *Proceedings of the 30th conference on artificial intelligence*.
- Graesser, A. C., Hu, X., & Sottolare, R. (2018). Intelligent tutoring systems. In *International handbook of the learning sciences*. Routledge, Ch. 24.
- Grossman, J. B., Lind, C., Hayes, C., McMaken, J., & Gersick, A. (2009). *The cost of quality of out-of-school time programs: Tech. rep.*, Public/Private Ventures.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Paul H Brookes Publishing.
- Hassinger-Das, B., Ridge, K., Parker, A., Golinkoff, R., Hirsh-Pasek, K., & Dickinson, D. (2016). Building vocabulary knowledge in preschoolers through shared book reading and gameplay. *Mind, Brain, and Education*, 10.
- Hood, D., Lemaignan, S., & Dillenbourg, P. (2015). When children teach a robot to write. In *Proceedings of the tenth annual ACM/IEEE International conference on human-robot interaction*.
- Howard, C., Jordan, P., Di Eugenio, B., & Katz, S. (2017). Shifting the load: a peer dialogue agent that encourages its human collaborator to contribute more to problem solving. *International Journal of Artificial Intelligence in Education*.
- Hyun, E. J., Kim, S. Y., Jang, S., & Park, S. (2008). Comparative study of effects of language instruction program using intelligence robot and multimedia on linguistic ability of young children. In *Proceedings of the 17th IEEE international symposium on robot and human interactive communication*.
- Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A field trial. *Human-Computer Interaction*.
- 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*.
- Kennedy, J., Baxter, P., & Belpaeme, T. (2015). Comparing robot embodiments in a guided discovery learning interaction with children. *International Journal of Social Robotics*, 7(2), 293–308.
- Kennedy, J., Baxter, P., Senft, E., & Belpaeme, T. (2016). Social robot tutoring for child second language learning. In *ACM/IEEE international conference on human-robot interaction*.
- King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal*, 27(4), 664–687.
- Kory-Westlund, J. M. (2019). *Relational AI: Creating long-term interpersonal interaction, rapport, and relationships with social robots* (Ph.D. dissertation), Massachusetts Institute of Technology.
- Kory Westlund, J. M., Jeong, S., Park, H. W., Ronfard, S., Adhikari, A., Harris, P. L., et al. (2017). Flat vs. expressive storytelling: Young children's learning and retention of a social robots narrative. *Frontiers in Human Neuroscience*, 11, 295, URL <https://www.frontiersin.org/article/10.3389/fnhum.2017.00295>.
- Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, 10(1), 110–120.
- Kuhl, P. K. (2011). Social mechanisms in early language acquisition: Understanding integrated brain systems supporting language.
- Leyzberg, D., Spaulding, S., Toneva, M., & Scassellati, B. (2012). The physical presence of a robot tutor increases cognitive learning gains. In *CogSci*.
- Lindgren, R., Tscholl, M., Wang, S., & Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers & Education*, 95, 174–187.
- Manickam, I., Lan, A. S., & Baraniuk, R. G. (2017). Contextual multi-armed bandit algorithms for personalized learning action selection. In *2017 IEEE international conference on acoustics, speech and signal processing* (pp. 6344–6348).
- McDuff, D., Kodra, E., Kaliouby, R. e., & LaFrance, M. (2017). A large-scale analysis of sex differences in facial expressions. *PLoS One*, 12(4), 1–11, <https://doi.org/10.1371/journal.pone.0173942>.
- McDuff, D., Mahmoud, A., Mavadati, M., Amr, M., Turcot, J., & Kaliouby, R. e. (2016). AFFDEX SDK: A cross-platform real-time multi-face expression recognition toolkit. In *Proceedings of the 2016 CHI conference extended abstracts*. ACM.
- Movellan, J. R., Eckhardt, M., Virnes, M., & Rodriguez, A. (2009). Sociable robot improves toddler vocabulary skills. In *2009 4th ACM/IEEE international conference on human-robot interaction* (pp. 307–308).
- National Institute of Early Education Research (2018). *State of preschool yearbooks 2018: Tech. rep.*
- Nores, M., & Barnett, S. (2014). *Access to high quality early care and education: Readiness and opportunity gaps in America (CEELO policy report)*. Center on Enhancing Early Learning Outcomes.
- Páez, M. M., Tabors, P. O., & López, L. M. (2007). Dual language and literacy development of spanish-speaking preschool children. *Journal of Applied Developmental Psychology*, 28(2), 85–102.
- Park, H. W., Gelsomini, M., Joo Lee, J., & Breazeal, C. (2017). Telling stories to robots: The effect of backchanneling on a child's storytelling. In *International conference on human-robot interaction*.
- Park, H. W., Grover, I., Spaulding, S., Gomez, L., & Breazeal, C. (2019). A model-free affective reinforcement learning approach to personalization of an autonomous social robot companion for early literacy education. In *Proceedings of the 33th conference on artificial intelligence*.
- Park, H. W., & Howard, A. M. (2014). Robot learners: Interactive Instance-based learning and its application to therapeutic tasks. In *AI for human-robot interaction, 2014 AAAI fall symposium*.
- Park, H. W., & Howard, A. M. (2015). Retrieving experience: Interactive instance-based learning methods for building robot companions. In *2015 IEEE international conference on robotics and automation* (pp. 6140–6145).
- Park, H. W., Rosenberg-Kima, R., Rosenberg, M., Gordon, G., & Breazeal, C. (2017). Growing growth mindset with a social robot peer. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction* (pp. 137–145). New York, NY, USA: ACM.
- Quigley, M., Conley, K., Kerkey, B. P., Faust, J., Foote, T., Leibs, J., et al. (2009). ROS: an open-source Robot Operating System. In *ICRA workshop on open source software*.
- Ramachandran, A., Huang, C.-M., Gartland, E., & Scassellati, B. (2018). Thinking aloud with a tutoring robot to enhance learning. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction* (pp. 59–68). New York, NY, USA: ACM, <http://doi.acm.org/10.1145/3171221.3171250>.
- Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., et al. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, 29(5), 700–710, PMID: 29442613. <http://dx.doi.org/10.1177/0956797617742725>.
- Ros, R., Oleari, E., Pozzi, C., Sacchitelli, F., Baranzini, D., Bagherzadhalimi, A., et al. (2016). A motivational approach to support healthy habits in long-term child-robot interaction. *International Journal of Social Robotics*.
- Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-Building and knowledge-telling in peer tutors explanations and questions. *Review of Educational Research*, 77(4), 534–574.
- Rubin, K. H., Bukowski, W. M., & Parker, J. G. (2007). Peer interactions, relationships, and groups. In *Handbook of child psychology*.
- Ryokai, K., Boulanger, C., & Cassell, J. (2003). Virtual peers as partners in storytelling and literacy learning. *Journal of Computer Assisted Learning*, 19, 195–208.
- Smith, G. G., Li, M., Drobisz, J., Park, H.-R., Kim, D., & Smith, S. D. (2013). Play games or study? Computer games in ebooks to learn English vocabulary. *Computers & Education*, 69, 274–286.
- Sutton, R. S., & Barto, A. G. (1998). Introduction to reinforcement learning. *Learning*.
- 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*.

- Topping, K. J. (2005). Trends in peer learning. *Educational Psychology*, 25(6), 631–645, <https://doi.org/10.1080/01443410500345172>.
- Tudge, J., & Rogoff, B. (1989). Peer influences on cognitive development: Piagetian and Vygotskian perspectives. In *Crosscurrents in contemporary psychology, Interaction in human development* (pp. 17–40). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- U. S. Department of Education (2015). *A matter of equity: Preschool in America: Tech. rep.*, U.S. Department of Education.
- Vaala, S., Ly, A., & Levine, M. H. (2015). *Getting a read on the app stores: A market scan and analysis of children's literacy apps*. New York, NY: The Joan Ganz Cooney Center at Sesame Workshop.
- Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Wolf, M., & Gottwald, S. (2016). *Tales of literacy in the 21st century. Literary agenda series*. Oxford, UK: Oxford University Press.
- Woolf, B. (2008). *Building intelligent interactive tutors, student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann.
- Zaga, C., Lohse, M., Truong, K. P., & Evers, V. (2015). The effect of a robot's social character on children's task engagement: peer versus tutor. In A. Tapus, E. André, J.-C. Martin, F. Ferland, & M. Ammi (Eds.), *Social robotics* (pp. 704–713). Cham: Springer International Publishing.
- Zou, D., Huang, Y., & Xie, H. (2019). Digital game-based vocabulary learning: where are we and where are we going? *Computer Assisted Language Learning*, 1–27.