“We Make a Great Team!”: Adults with Low Prior Domain Knowledge Learn more from a Peer Robot than a Tutor Robot

Nicole Salomons, Kaitlynn Taylor Pineda, Adérónké Adéjàre, Brian Scassellati
Yale University
{nicole.salomons, kaitlynn.pineda, aderonke.adejare, brian.scassellati}@yale.edu

***Abstract*—In peer tutoring, the learner is taught by a colleague rather than by a traditional tutor. This strategy has been shown to be effective in human tutoring, where students have higher learning gains when taught by a peer instead of a traditional tutor. Similar results have been shown in child-robot interactions studies, where a peer robot was more effective than a tutor robot at teaching children. In this work, we compare skill increase and perception of a peer robot to a tutor robot when teaching adults. We designed a system in which a robot provides personalized help to adults in electronic circuit construction. We compare the number of learned skills and preferences of a peer robot to a tutor robot. Participants in both conditions improved their circuit skills after interacting with the robot. There were no significant differences in number of skills learned between conditions. However, participants with low prior domain knowledge learned significantly more with a peer robot than a tutor robot. Furthermore, the peer robot was perceived as friendlier, more social, smarter, and more respectful than the tutor robot, regardless of initial skill level.**

***Index Terms*—human-robot interaction; peer learning; robot tutoring**

I. Introduction

In peer-to-peer tutoring, children or adults teach each other rather than being taught by a teacher [27]. There are benefits of peer instruction for both the student who is teaching and the student who is learning. Throughout this paper, we will call a student instructing another student as the *peer-teacher* or *peer*, the student who is being instructed as the *learner*, and a traditional teacher as *tutor*. When a peer-teacher prepares content and teaches a colleague, they demonstrate higher learning gains than when they only learn the content for themselves [6]. Likewise, the learner who is taught by a peer frequently learned more than the one who is taught by a teacher, especially if they had higher prior domain knowledge [25]. Furthermore, peer instruction lowers failing rates [29], creates an increased sense of community [39], and increases student self-esteem [18].

Most work in robot tutoring has focused on having the robot take on the role of a traditional tutor [8]. While a few studies investigated the robot’s role as a peer, these focused on child-robot interactions [13], [28], [42]. Similar to peer tutoring among adults, a peer robot tutor also had several advantages. When interacting with a peer robot compared to a tutor robot, children learned more [7], became more engaged [42], and developed a stronger growth mindset [28].

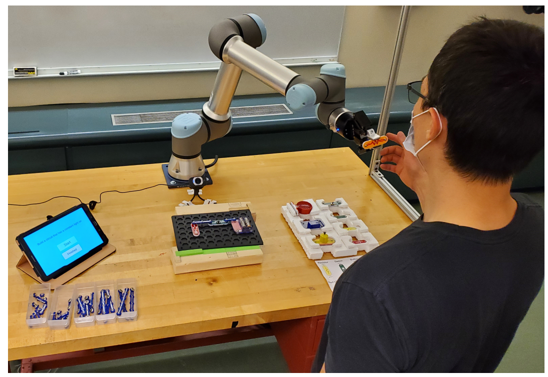

Fig. 1. Participants built electronic circuits with either a peer robot or a tutor robot. The robot would provide personalized help based on the user’s skills. In the figure, we see the robot suggesting the user add a resistor to the board.

While children appear to benefit from peer-based child-robot interactions, it is unclear whether the same results will hold for adults. Adults may have higher expectations for a peer that a robot could struggle to meet. Adults may also have more practical experience working with peers and as such might not easily accept a robot peer. Lastly, working with adults will also require working in a more challenging educational domain, with harder and more complex problems to be taught, which may not easily transfer to a robot. To study how peer robots are viewed by adults, we designed an in-between participant study where participants interacted with either a peer robot or a tutor robot, during an electronic circuit building task.

Participants built ten electronic circuits with an autonomous robot, as seen in Figure 1. The system modeled the user’s skills throughout each task using Continuous Bayesian Knowledge Tracing [37]. The robot then provided personalized help to the user depending on their skill state and the skills needed for each task. The robot provided nearly identical advice in both conditions. However, in the peer condition, the advice was delivered using pronouns that indicated that the robot was an equal and invested colleague (“we/us”), and in the other indicated that the robot was a more knowledgeable authority figure (“you”). Participants completed a pre-test and a post-test to detect skills learned. They also completed questionnaires about their perceptions of the robot.

While participants in both conditions demonstrated significantly more skills in the post-test than in the pre-test, there were no significant increases of skills between the conditions. However, when analyzing only participants with low initial pre-test scores, they learned significantly more in the peer condition than in the tutor condition. Furthermore, participants viewed the peer robot as more intelligent, more social, and friendlier than the tutor robot, independent of prior skill knowledge. Additionally, participants who interacted with the peer robot felt more respected by it than the tutor robot.

II. Background

In this section, we introduce background on peer-to-peer tutoring and show some advantages of this learning strategy. We then review literature on how the peer-to-peer tutoring strategy has been extended to human-robot tutoring.

A. Peer-to-Peer Tutoring

Peer-to-peer learning is defined as "an educational practice in which students interact with other students to attain educational goals" [27]. It is typically used as a supplement to the classroom learning process between a teacher and students. Peer-to-peer learning is a favorable educational practice because it prepares students for learning from others in workplaces and communities [27].

There are numerous benefits to peer-to-peer learning. Peer-teachers who studied in preparation to teach something and then taught the information generally scored higher on a retention test than students who prepared only for themselves [4], [6], [19]. In reciprocal peer tutoring (RPT), students are paired together to review content and to practice skills [18]. This strategy resulted in greater improvements in cognitive gains, lower levels of subjective distress, and higher course satisfaction [18]. In addition to increasing student achievement, peer-to-peer learning has many social benefits, including positive race relations in desegregated schools, mutual concern among students, and student self-esteem [40]. Interventions were most effective with younger, urban, low-income, and minority students [35].

Peer-to-peer learning is also effective for adult learning. Lasry et al. showed that peer-taught university students had higher learning gains than traditionally taught students [25]. Additionally, they show a significant increase in learning gains for students with high background knowledge but not for students with low background knowledge. Another potential benefit for peer tutoring is that adults learn better in an informal environment and need to be respected when learning new things [15]. Awan [5] commends the use of peer-to-peer learning in radiology residencies because it promotes active and relevant learning. This practice also prepares future physicians for explaining medical topics to their patients [9].

B. Robotic Peer-to-Peer Tutoring

Social robots have been found to be effective tutors via individualized tutoring interactions [30], [31], [38]. Tutoring robots can take on several different roles, including a learner, a peer, or a teacher [1], [12], [13], [42]. However, approximately 86% of the studies conducted with robots to facilitate human learning consist of the robot taking on the role of tutor [8].

Although the use of robots as peers represents a minority of the literature, a peer-teacher social robot has been shown to positively benefit a child's language learning [23]. A peer robot can also enhance a child's own creative thinking [2]. Zaga et al. showed how children demonstrated increased engagement when playing with a peer robot to complete a Tangram puzzle compared to when playing with a tutor robot [42]. In a long-term study, a peer-teacher humanoid social robot with the ability to personalize its interactions with children in a classroom increased the children's learning of novel topics [7]. Park et al. [28] determined that children who played with a social robotic peer that exhibited a growth mindset (a belief that success arises from effort and perseverance) developed a stronger growth mindset of their own. Chen et al. [12] noted that the children who interacted with their adaptive peer-teacher robot not only had more expressive faces than the children who interacted with their tutor robot, but they also learned more and retained advanced vocabulary.

While robots can be used as a peer-teacher during child learning scenarios, they can also play the role of a tutee or naive peer. In this case, a person takes the role of a peer-teacher and educates their robotic peer, resulting in the enhancement of the person's learning through the reinforcement of concepts. Japanese children at an English language school improved their spontaneous learning of new English vocabulary words after teaching them to a robot [41]. The forms of teaching naturally implemented by the children involved direct teaching, gesturing, and verbal teaching. Robots have been used in the role of the tutee where the children taught the robot handwriting [11], [21], [22]. The children demonstrated to an autonomous robotic agent how to write certain letters or words, helping develop their own writing ability.

Although these studies focused on interactions with children, a peer robot can nevertheless provide unique learning benefits separate from those of a tutor robot.

III. Methodology

Higher learning gains and positive traits have been seen when interacting with a peer-teacher both in human tutoring [6], [25] and in robot-child tutoring [13], [28], [42]. Therefore, we predict that participants interacting with a peer robot will learn more new skills than participants interacting with a tutor robot. Furthermore, an adult peer-teacher was especially beneficial when the adult learner had higher prior knowledge in the domain [25]. Therefore, we also predict that adults with higher knowledge will benefit more from a peer robot.

Research has shown that adults learn better in an informal environment and highlight the importance of feeling respected [15]. We hypothesize that adults interacting with the peer robot will feel treated more as an equal and therefore feel more respected than those interacting with a tutor robot.

Prior work has shown that students who interacted with a peer robot were more engaged [42] and were able to create rapport with it [23]. Therefore, we hypothesize that

Section Part	Tutor Robot	Peer Robot
Introduction to Robot	Experimenter: "*Hello, This is Urie, the robot. Urie will be teaching you a bit about eletronic circuits today.*"	Experimenter: "*Hello, This is Urie, the robot. You and Urie will be collaborating in building some eletronic circuits today.*"
Reinforcement Question	"*When should you add a speaker to the circuit?*"	"*When should we add a speaker to the circuit?*"
Wrong Piece	"*Can you explain to me what a button does? Do you think you need it for the current task?*"	"*Can you explain to me what a button does? Do you think we need it for the current task?*"
Piece Recommendation	"*Here, try to add the resistor to the circuit.*"	"*Here, let's try to add the resistor to the circuit.*"
Help Utterance	"*To power the music circuit you need to make sure that its positive port is connected to the positive port of the battery.*"	"*To power the music circuit we need to make sure that its positive port is connected to the positive port of the battery.*"
Task Finished Correctly	"*Awesome, you are a good student.*"	"*Awesome, we make a good team.*"

Fig. 2. We present some of the different utterances between conditions. The robot was introduced differently to the participant depending on the condition. The remaining utterances were very similar and often only differed in the pronoun used. Some examples of help actions included asking questions to reinforce a correctly applied skill, pointing out a wrong piece on the board, recommending a piece, and giving a description of an incorrectly applied skill.

participants interacting with a peer robot will be more engaged and view the robot more positively than participants with the tutor robot. Lastly, people often have very high expectations out of robots [3], [24], [33]. When the participant is told the robot is a tutor, those expectations might be higher, as the robot is presented as an expert at the task. Therefore, we hypothesize that a peer robot might be seen as more intelligent than a tutor robot as there will be lower expectations put on it.

We have five main hypotheses for this study:

Hypothesis 1a: *Adults in both conditions will show significant improvement in electronic circuit skills from pre-test to post-test.*

Hypothesis 1b: *Adults will learn more from a peer robot than a tutor robot.*

Hypothesis 1c: *Adults with high initial knowledge will especially benefit from a peer robot, compared to adults with high initial knowledge interacting with the tutor robot.*

Hypothesis 2: *Adults will view a peer robot more positively than a tutor robot.*

Hypothesis 3: *Adults will be more engaged with a peer robot than a tutor robot.*

Hypothesis 4: *Adults will feel more respected by a peer robot than a tutor robot.*

Hypothesis 5: *Adults will see a peer robot as more intelligent than a tutor robot.*

To test our hypotheses, we split participants into two conditions: one where they interact with the robot as a tutor, and one where they interact with the robot as a peer. Participants built electronic circuits using a modular circuit-building toy called Snap Circuits [17]. We chose circuit design because it is a task that is challenging for most adults; there are varying levels of initial knowledge, with some participants having high initial knowledge and others low initial knowledge on circuits; and a robot can model the task using a sensing system.

A. Conditions

Participants were split randomly into two conditions:

- **Tutor Robot** - The robot acted as a traditional tutor and provided instructions to the participant. The robot was introduced as a teacher to the participant, and its utterances towards the participant used second person singular pronouns like "you".
- **Peer Robot** - The robot acted like a peer who is working together with the participant on the circuit. The robot was introduced as a collaborator, and its utterances used first-person plural pronouns ("we/us").

The difference between conditions was minimal, especially considering the robot had few anthropomorphic features. In both conditions, the robot's utterances were very similar, mostly changing pronouns from "we" in the peer condition, and "you" in the tutor condition. The robot always presented correct help suggestions independent of condition. Some examples of utterances can be seen in Figure 2.

B. Robot System

Participants interacted with the robot on a large table. Figure 3 shows an illustration of the experimental setup. Participants were given each task via a tablet, and on the tablet, they could indicate that they had finished the current task and start the next task. The tablet provided no help with the task. Participants used wires and electronic circuit pieces to build their circuits on a board in the middle of the table. An overhead Kinect Azure camera detected what pieces were on the board and how they were connected. A green hand strip at the bottom of the board was used to detect when the participants' hands were on top of the board, and therefore the camera's observations would be inaccurate. A second camera faced the participant to record the interaction.

A UR5e robot from Universal Robots was used in this study. It is a lightweight industrial robotic arm with 6-DOF. It could pick up the snap circuit pieces with its gripper and hand them to the participant. The robot was able to communicate to the participant via a text-to-speech voice. Additionally, the robot displayed idling behavior with random movements every few

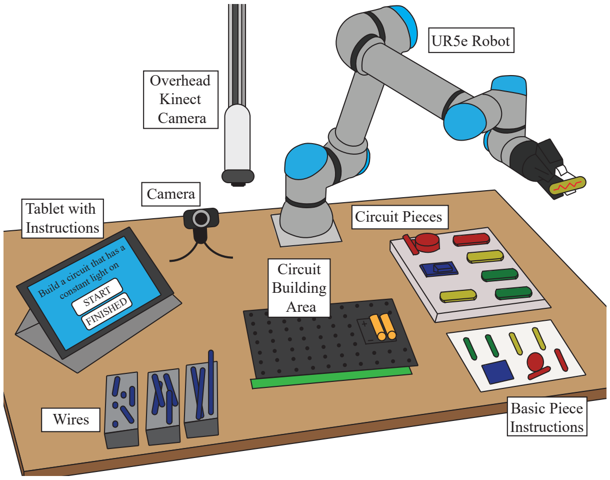

Fig. 3. The experimental setup. Participants were given tasks via a tablet application. In the middle of the table, they built circuits using wires and circuit pieces. They were provided basic instructions with the piece names. An overhead camera focused on the circuit and modeled which skills were correctly applied. A UR5e robot provided them with help every 30 seconds based on what was needed for the current task. An additional camera collected video and audio data of the participant.

seconds, occasionally looking at the circuit board, pieces, or the participant, by pointing the gripper at it. The robot acted completely autonomously throughout the study.

C. Snap Circuits Tasks and Skills

We created 32 different electronic circuit tasks of varying difficulty, of which participants completed ten. There were more tasks than the number the participant completed, so the robot could adjust to each person's skill level. Section III-F describes how tasks were chosen for each participant. For each task, the participant was given an empty circuit board with only a battery on it, many wires of different sizes, and seven pieces: an LED, a switch, a button, a motor, a resistor, a music circuit, and a speaker. Each piece could be snapped together on the board to form circuits. An example of a completed circuit can be seen in Figure 4.

The participant was instructed what task to complete next via a tablet. They were given three minutes for each task unless they correctly completed it before the time expired. Some examples of tasks are: "Build a circuit that plays music when a switch is turned on" and "Build a circuit that spins a motor when a switch is turned on or a button is pressed".

Each task required the participant to demonstrate different skills. Some examples of these skills are: adding a speaker when it is needed, creating a closed circuit, knowing the directionality of an LED, powering the music circuit, and creating AND and OR gates. A total of 17 skills were tested. Section III-E explains how we model participants' skills throughout the tasks.

D. Experimental Procedure

Each session took approximately 60 minutes. The procedure of the experiment consisted of the following:

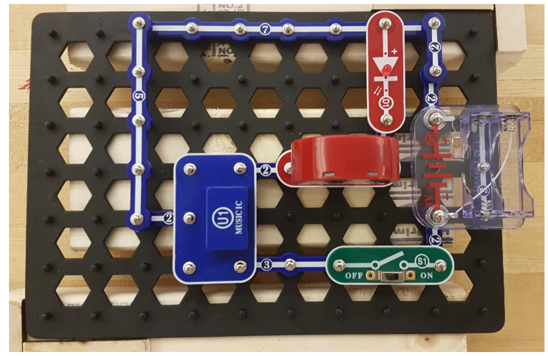

Fig. 4. An example of a completed circuit. This circuit plays music and blinks a light in the rhythm of the music when the switch is turned on.

1) The participant completed the consent form and a demographic questionnaire.
2) The participant completed six pre-test electronic circuit tasks and questions. These are detailed in Section III-H.
3) The experimenter introduced the robot. Depending on the condition, the robot was introduced as either a peer/collaborator or as a teacher.
4) The participant built ten electronic circuits alongside the robot. The robot provided personalized help actions.
5) The participant answered post-study questionnaires.
6) The participant did six post-test electronic circuit tasks.
7) At the end of the interaction, participants were debriefed and paid $10 for their time.

E. Skill Estimation

A computer vision system with an overhead camera observed the user as they placed pieces on the board. It tracked which pieces were on the board, and which pieces were connected to each other. User skills were modeled using an extension of Bayesian Knowledge Tracing called Continuous Bayesian Knowledge Tracing (C-BKT) [37]. C-BKT was used as it allows skill modeling of complex tasks where the observations are noisy, and skills vary in the amount of time needed to demonstrate them. The system individually modeled each of the 17 skills by creating an estimate of whether the user had mastered each skill. We represented this estimate as a vector b, where each skill was initialized to 0.5, representing complete uncertainty of the user's skill state. Each second, b was updated using observations from the computer vision system detailing which skills were applied correctly and which ones were not.

F. Task Selection

Prior work shows that selecting tasks with appropriate difficulty leads to higher learning gains [16], [36]. Therefore tasks were chosen for each participant according to their demonstrated capabilities. To rate the difficulty of each task, each of the 17 skills was given a difficulty rating from a scale of 1.0 to 5.0, with 5.0 being the most difficult. These were determined by consulting an electronic engineering major. The ratings were stored in a difficulty vector d. For example, the skill for whether a participant knew when to use an LED was given a difficulty rating of 1, while the skill for whether the participant knew how to create an OR gate was given a 4.5. The current belief estimate b was used to select the next task.

In order to determine which task to give next to a participant, all remaining tasks are assigned a difficulty rating R based on the skills Sk that a task t incorporated. The rating was calculated based on the difficulty of each skill and the participant's current belief value b. Participants with higher belief values would likely find the task easier. Therefore, we used $1 - b(i)$ to measure how difficult the task would be for the participant. As we are summing over the difficulty of each skill for a task, the more skills a task tests, the more difficult it will likely be. The difficulty rating R for a specific task is calculated as follows:

$$R_t = \sum_{i \in Sk} (1 - b(i)) * d(i) \tag{1}$$

There is also a fixed ideal rating value V that was set equal to five after initial trial and error. The V is intended to help ensure that an appropriate task is selected next for the respective participant so that the task is not too easy nor too overwhelming [26]. The task whose r value is closest to V is selected as the next task and removed from the possible remaining tasks for the next iteration.

$$NextTask = \min_{t \in T}(|R_t - V|) \tag{2}$$

In the case where several tasks are equally close to V, one of these potential tasks is selected at random. The process is repeated until the interaction with the participant ends.

G. Help Action Selection

Personalizing help in tutoring systems leads to higher learning gains [14], [32]. Therefore the robot provides assistance to the participant according to the skills they had demonstrated during the current task. The robot provided a help action every 30 seconds. There were six different types of help actions, of which the system selected one at random. The different kinds of help actions were:

- **Reinforcement Question** - The robot asked a question about a skill the participant had demonstrated.
- **Reinforcement Utterance** - The robot confirms that a skill the participant had demonstrated is correct.
- **Wrong Piece Point** - The robot pointed to a piece on the board and said it was not needed.
- **Wrong Piece Utterance** - The robot said that one of the pieces on the board was not needed.
- **Help Movement** - The robot gave help to the participant by explaining something about a skill they had not demonstrated. While explaining it, the robot pointed to something on the board or handed the participant a piece.
- **Help Utterance** - The robot gave help to the participant by explaining a skill the participant had not demonstrated. The robot did not move in this case.

The robot did not select a reinforcement help action if the participant did not have any correct skills displayed during the task. Likewise, it would also remove wrong piece help actions from the randomization options if all the current pieces on the board were needed. Additionally, when the participant pressed the finished button on the tablet, but the task was incorrect or incomplete, the robot randomly selected either a *help movement* or a *help utterance* for one of the skills that were demonstrated incorrectly. Examples of different types of robot help actions can be seen in Figure 2.

H. Metrics

We had three different types of metrics: test metrics, behavioral metrics, and survey metrics.

1) Test Metrics:

- **Pre-test and Post-test** - The pre-test and post-test were composed of six very similar questions. The first two questions on both tests were the same. They asked participants to build from scratch a circuit that shines a constant light and a circuit that plays music, respectively. Participants were given five minutes to do both tasks. The third and fourth tasks on both tests required participants to add pieces to the board to complete the circuits. These tasks were identical between pre-test and post-test, other than the circuit boards being rotated 180 degrees to the participant in the post-test. For the fifth and sixth tasks, we presented pictures of pre-built circuits and asked participants to write down what the circuits did. These were similar between pre-test and post-test, but the pieces were arranged differently on the board. Participants were given five minutes to complete tasks three through six.
  We classify participants into either having high prior circuit knowledge or low prior circuit knowledge based on their pre-test. If participants got less than half of the skills correctly on the pre-test, they were considered to have low prior circuit knowledge. Otherwise, they were considered to have high prior circuit knowledge.

2) Behavioral Metrics:

- **Speaking** - We measured how much participants talked to the robot. Participants were classified as engaging in conversation with the robot if they said at least ten sentences to it, and as not engaging in conversation if they talked to it less.

3) Survey Metrics:

- **Demographics** - Before the interaction, we administered a demographics questionnaire that asked participants questions about their gender, age, occupation or major (if student), and country of origin. We also asked how often they used a computer, their familiarity with robots, and their level of expertise on electrical circuits.
- **Post Interaction Questionnaire** - We administered the RoSAS questionnaire about their feelings towards robots [10]. The RoSAS measured participants' perceptions of the robot's warmth, competence, and discomfort. We also asked the participant to rate the following on a 1-7 Likert

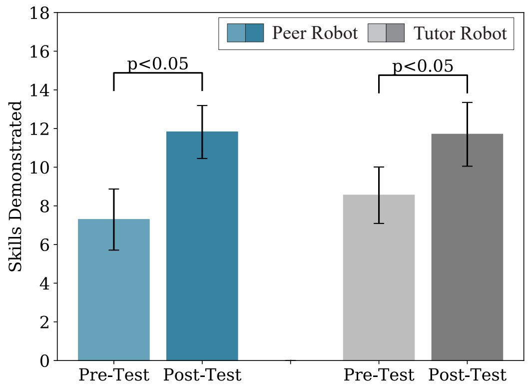

Fig. 5. (a) Participants significantly improved their circuit knowledge skills from pre-test to post-test in both conditions. (b) There were no significant differences in number of skills learned between conditions.

Scale with 1 being "Not Applicable" and 7 being "Most Applicable": The robot acted like my colleague; The robot treated me like an equal; I felt like I was being judged by the robot; I felt like the robot respected my capabilities; The robot was friendly; I felt engaged while interacting with the robot; I felt like the robot was boring; I felt like the robot was smart; I felt like the robot was good at electronic circuits; The robot was better than me at electronic circuits. Finally, we had an open-ended question for participants: Is there anything you wished the robot would have done differently?

I. Participants

There were 37 participants who completed the experiment. Interactions with the robot lasted on average 37 min and 37s (SD = 5 min and 6s). The university's Institutional Review Board approved the study, and participants signed a consent form agreeing to participate. There were nine male and eight female participants in the peer condition, and their average age was 28.00 (SD = 12.90). There were nine male, ten female, and one gender-fluid participants in the tutor condition, and their average age was 25.00 (SD = 10.60). Participants in the peer condition rated themselves as an average of 2.47 (SD = 1.23) on a 1-5 scale on their prior electronic circuit knowledge, whereas participants in the tutor condition rated themselves an average of 2.40 (SD = 1.60). There were no significant differences in gender, age, or prior circuit expertise between conditions.

IV. Results

A. Manipulation Check

First, we check whether the peer robot and the tutor robot were perceived differently. We asked each participant on a scale of 1-7 whether they perceived the robot as a peer and whether they felt like they were treated as an equal. Participants in the peer robot perceived it significantly more as a peer ($M = 4.59, SD = 2.18$), than participants who interacted with the tutor robot ($M = 2.65, SD = 1.39$), $t(37) = 3.27, p = 0.002$. And participants in the peer condition ($M = 5.00, SD = 1.83$) perceived the robot as treating them as an equal significantly more than participants in the tutor condition ($M = 3.58, SD = 2.09$), $t(35) = 4.50, p = .041$. Therefore, we believe our manipulation check was successful.

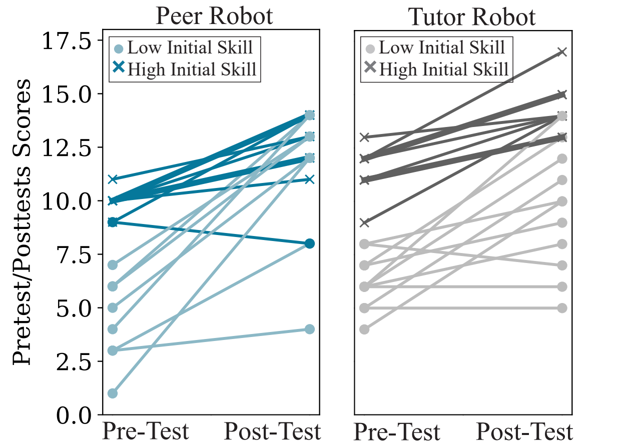

Fig. 6. The pre-test and post-test scores for the peer and tutor conditions. There were no significant differences in skills gained between conditions. However, participants with low skill knowledge improved their skills significantly more with the peer robot than the tutor robot.

B. Test Results

We compare the skill increase in each condition from pre-test to post-test on the 17 skills. A skill is attributed as known when the participant has correctly applied it at least half of the time. On average, participants in the peer robot condition scored $7.29 (SD = 3.16)$ on the pre-test and $11.82 (SD = 2.74)$ on the post-test. Participants in the tutor condition scored on average $8.55 (SD = 2.93)$ on the pre-test and $11.70 (SD = 3.31)$ on the post-test. An ANOVA comparing moment (pre-test and post-test) and condition found significant differences $F(3, 74) = 10.01, p < 0.001$. A Tukey HSD test revealed that both the peer condition ($p = 0.001$) and the tutor condition ($p = 0.009$) significantly improved from pre-test to post-test, as seen in Figure 5. There were no significant differences between conditions for the pre-test ($p = 0.587$) or the post-test ($p = 0.900$).

Participants in the peer condition learned on average 4.53 ($SD = 3.22$) new skills, whereas participants in the tutor condition learned on average 3.15 ($SD = 2.37$) new skills. These differences were not significant $t(37) = 1.46, p = 0.154$. Next, we compare participants with prior low electronic circuit knowledge and participants with high prior electronic circuit knowledge. We compared the number of learned skills between condition and prior knowledge using an ANOVA and found significant differences $F(3, 37) = 5.47, p = 0.004$. A Tukey HSD test revealed that participants with low prior knowledge learned significantly more in the peer condition than the tutor condition (p=0.023), but no significant differences were found for high prior knowledge participants (p=0.900). These results can be seen in Figure 6.

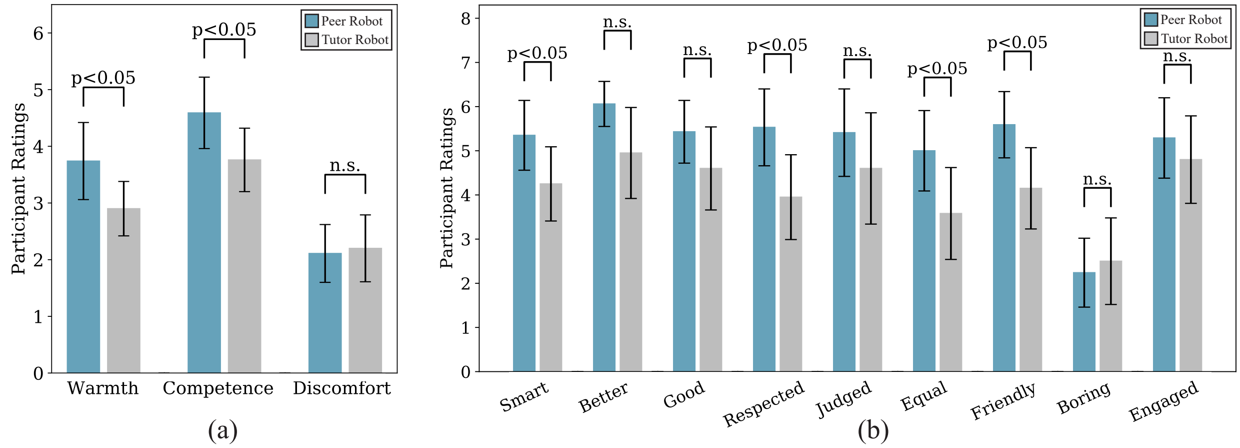

Fig. 7. Questionnaire results. (a) The peer robot was perceived as significantly more social and intelligent than the tutor robot. (b) The peer robot was perceived as significantly smarter, more respectful, friendly than the tutor robot, in addition to participants feeling more like they were treated as an equal.

C. Behavioral Results

In the peer condition, there were six participants who engaged in conversation with the robot, and nine participants who did not. There were two participants whose audio data was corrupted. In the tutor condition there were eight participants who engaged in conversation with the robot and twelve who did not. Using a Chi-Squared test, these results were not statistically significantly different from each other $X^2(1, N = 35) = 0, p = 1.000$.

D. Questionnaire Results

On the RoSAS questionnaire, participants rated the robot as more warm in the peer condition ($M = 3.74, SD = 1.37$) compared to the tutor condition ($M = 2.90, SD = 0.96$). Participants also rated the robot as more competent in the peer condition ($M = 4.59, SD = 1.27$) than the tutor condition ($M = 3.76, SD = 1.12$). Lastly, participants rated the robot similarly in regards to discomfort between the peer ($M = 2.11, SD = 1.03$) and tutor conditions ($M = 2.20, SD = 1.19$). Their ratings were significantly different for warmth $t(37) = 2.18, p = 0.036$, and for competence $t(37) = 2.09, p = 0.044$, but not for discomfort $t(37) = -0.25, p = 0.804$. The RoSAS questionnaire results are seen in Figure 7(a).

On the post-experiment questionnaire, participants in the peer condition ($M = 5.35, SD = 1.58$) rated the robot significantly smarter than the tutor condition ($M = 4.25, SD = 1.68$), $t(37) = 2.04, p = 0.049$. Participants rated the robot as being better than them at electronic circuits in the peer condition ($M = 6.06, SD = 1.03$) than the tutor conditions ($M = 4.95, SD = 2.06$), but these differences were not quite significant $t(37) = 2.01, p = 0.052$. Lastly, participants rated the robot a $5.42 (SD = 1.42)$ on being good at circuits in the peer condition, and a $4.60 (SD = 1.88)$ in the tutor condition. These differences were not significant $t(37) = 1.46, p = 0.152$.

Participants felt more respected by the peer robot ($M = 5.53, SD = 1.74$) than the tutor robot ($M = 3.95, SD = 1.93$), and this difference was significant $t(37) = 2.59, p = 0.014$. Participants in the peer condition rated the robot a 3.06 (SD = 1.98) for feeling judged and a 3.60 (SD=2.52) in the tutor condition, $t(37) = -0.72, p = 0.479$. Lastly, participants in the peer condition ($M = 5.00, SD = 1.83$) perceived the robot as treating them as an equal significantly more than participants in the tutor condition ($M = 3.58, SD = 2.09$), $t(35) = 4.50, p = .041$.

Participants viewed the robot as more friendly in the peer condition ($M = 5.59, SD = 1.50$) than the tutor condition ($M = 4.15, SD = 1.84$), $t(37) = 2.57, p = 0.015$. Participants did not think the robot was more boring in one condition than the other (peer: $M = 2.24, SD = 1.56$; tutor: $M = 2.50, SD = 1.96$), $t(37) = -0.45, p = 0.657$, nor did they feel more engaged in one condition than another (peer: $M = 5.29$, $SD = 1.83$; tutor: $M = 4.80$, $SD = 1.99$; $t(37) = 0.77$, $p = 0.447$). These results can be seen in Figure 7(b).

V. Discussion

The work is the first to show multiple benefits in peer tutoring while only manipulating one very small aspect (pronouns used and how the experimenter presented the robot). Most research in peer robot tutoring was conducted with children and was either not focused on peer vs. tutor [2], [24] or had many differences between conditions [13], [42]. Additionally, we believe this is the first HRI work comparing peers and tutors that shows significant differences in learning using a pre-test and post-test rather than other measures.

A. Hypotheses

Participants in both the peer condition and the tutor condition significantly improved their electronic circuit skills from pre-test to post-test. This shows that the robot in both conditions successfully taught the adults. Therefore Hypothesis 1a is true: *Adults in both conditions showed significant improvement in electronic circuit skills from pre-test to post-test*. Participants did not learn more skills in the peer condition compared to the tutor condition. Therefore we cannot confirm Hypothesis 1b; *Adults did **not** learn more from a peer robot than a tutor robot*. Participants with high skill knowledge did not have significantly different skill increase between conditions. Therefore in regards to Hypothesis 1c, *Adults with high initial knowledge did **not** especially benefit from a peer robot, compared to adults with high initial knowledge interacting with the tutor robot*. On the contrary, participants with low circuit knowledge learned significantly more with the peer robot. These results differ from those seen in adult peer-to-peer tutoring. Therefore, the robot taking on the role of a peer should be especially considered in scenarios where the person likely has low prior knowledge in the domain.

Participants did rate the peer robot more positively in several dimensions. On the post-experiment questionnaire, participants rated the peer robot as significantly more social and as significantly friendlier than the tutor robot. Therefore we believe that Hypothesis 2 is true: *Adults viewed a peer robot more positively than a tutor robot.* It is important that people view the robot positively when interacting with it, as they will likely be more engaged and learn from the robot in the long term.

In most Human-Robot Interactions studies, engagement is assessed using gaze patterns [34]. However, due to the current COVID-19 pandemic, participants wore masks during the interactions, which made computer vision systems that tracked participant faces unreliable. Therefore we measured engagement using the amount participants talked to the robot and their self-assessed engagement on the questionnaire. Participants did not significantly view the robot as being more boring in one condition than another. Neither did they report being more engaged. There were also no significant differences between engaging in conversation with the robot between conditions. Therefore, we do not support Hypothesis 3; *Adults were **not** more engaged with a peer robot than a tutor robot*.

Participants reported feeling significantly more respected by the peer robot compared to the tutor robot. Additionally, participants felt that they were treated significantly more as an equal when interacting with the peer robot. Therefore, we confirm Hypothesis 4: *Adults felt more respected from a peer robot than a tutor robot*. This is important, as feeling respected is an essential factor in learning success [15].

Participants in the peer condition viewed the peer robot as significantly smarter than participants in the tutor condition. Additionally, participants rated the peer robot significantly more competent than the tutor robot. Therefore we confirm Hypothesis 5: *Adults saw a peer robot as more intelligent than a tutor robot.*

B. Expectations of A Tutoring Robot

One possible reason the peer robot was rated as more intelligent than the tutor robot was because participants had lower expectations of a peer than a tutor. A tutor is presented as an expert, whereas there is more uncertainty involved in the capabilities of a peer-teacher. An open-ended question asking participants whether they wished the robot had done anything differently confirmed that many wished the robot had additional capabilities. Participants wished the robot had given examples of completed circuits, had the ability to answer participants' questions, and had given step-by-step instructions for the more complicated circuits.

Domains in adult tutoring are often more complex than those seen in children's tutoring, with many requiring computer vision systems to model the interactions. Therefore giving personalized advice is not as straightforward as giving help during child-robot interactions. Presenting the robot as a peer could lower expectations. In consequence, people might be more willing to receive advice from it than they would from a tutor robot whose expectations are not met.

C. In-group/Out-group effects

In the peer condition, the robot presents itself as being in-group with the participant when using the pronouns "we/us". Whereas in the tutor condition, the robot presents itself as an authority figure by placing itself in the out-group when using the "you" pronoun. People evaluate robots more positively when they are in-group than when they are out-group [20]. This is one potential confound in our work, where part of the results could be due to in-group/out-group membership.

D. Limitations and Future Work

Our work has several limitations, which could be potential areas for future research. First, it is unclear why a peer robot was especially beneficial for low prior knowledge participants. Therefore, a future study that focused on this would greatly benefit the community. Second, due to the pandemic, we could not look in-depth into engagement and facial expressions, and future work should analyze these once masks are unnecessary. Lastly, testing different tasks and settings would bring further insight into the benefits of varying robot roles.

VI. Conclusion

This paper explored different roles a robot can take when teaching people about electronic circuits. The robot would either take on the role of a peer or the role of a tutor. Participants with low prior circuit knowledge learned significantly more with the peer robot than with the tutor robot. This shows the benefits of peer robots, especially in domains where the user is likely lower-skilled. Additionally, participants who interacted with the peer robot viewed it as more friendly, more social, more intelligent, and felt more respected than participants who interacted with the tutor robot, independent of prior knowledge. These are all essential qualities for a robot to have, such that participants would be willing to have long-term interactions with it. Therefore, we recommend more exploration into robots who interact with the user as peers versus as teachers, especially when teaching adults.

Acknowledgment

This work was funded by the National Science Foundation (NSF) under grants No. 1955653, 1928448, 2106690, and 1813651.

References

[1] Timothy Adamson, Debasmita Ghose, Shannon C Yasuda, Lucas Jehu Silva Shepard, Michal A Lewkowicz, Joyce Duan, and Brian Scassellati. Why we should build robots that both teach and learn. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pages 187–196, 2021.

[2] Safinah Ali, Tyler Moroso, and Cynthia Breazeal. Can children learn creativity from a social robot? In *Proceedings of the 2019 on Creativity and Cognition*, pages 359–368. ACM, 2019.

[3] Patrícia Alves-Oliveira, Tiago Ribeiro, Sofia Petisca, Eugenio Di Tullio, Francisco S Melo, and Ana Paiva. An empathic robotic tutor for school classrooms: Considering expectation and satisfaction of children as end-users. In *International Conference on Social Robotics*, pages 21–30. Springer, 2015.

[4] Linda F Annis. The processes and effects of peer tutoring. *Human Learning: Journal of Practical Research & Applications*, 1983.

[5] Omer A. Awan. Peer to peer learning: Its importance and benefits. *Academic Radiology*, 28(5):747–748, 2021.

[6] John A Bargh and Yaacov Schul. On the cognitive benefits of teaching. *Journal of Educational Psychology*, 72(5):593, 1980.

[7] Paul Baxter, Emily Ashurst, Robin Read, James Kennedy, and Tony Belpaeme. Robot education peers in a situated primary school study: Personalisation promotes child learning. *PLOS ONE*, 12(5):1–23, 2017. Publisher: Public Library of Science.

[8] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. Social robots for education: A review. *Science robotics*, 3(21), 2018.

[9] Carolien Bulte, Aaron Betts, Kathryn Garner, and Steven Durning. Student teaching: views of student near-peer teachers and learners. *Medical teacher*, 29(6):583–590, 2007.

[10] Colleen M Carpinella, Alisa B Wyman, Michael A Perez, and Steven J Stroessner. The robotic social attributes scale (rosas) development and validation. In *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*, pages 254–262, 2017.

[11] Shruti Chandra, Raul Paradeda, Hang Yin, Pierre Dillenbourg, Rui Prada, and Ana Paiva. Do children perceive whether a robotic peer is learning or not? In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 41–49. ACM, 2018.

[12] Huili Chen, Hae Won Park, and Cynthia Breazeal. Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement. *Computers & Education*, 150:103836, 2020.

[13] Huili Chen, Hae Won Park, Xiajie Zhang, and Cynthia Breazeal. Impact of interaction context on the student affect-learning relationship in child-robot interaction. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pages 389–397, 2020.

[14] Benjamin Clement, Didier Roy, Pierre-Yves Oudeyer, and Manuel Lopes. Multi-armed bandits for intelligent tutoring systems. *arXiv preprint arXiv:1310.3174*, 2013.

[15] Jannette Collins. Education techniques for lifelong learning: principles of adult learning. *Radiographics*, 24(5):1483–1489, 2004.

[16] Yossi Ben David, Avi Segal, and Ya'akov Kobi Gal. Sequencing educational content in classrooms using bayesian knowledge tracing. In *Proceedings of the sixth international conference on Learning Analytics & Knowledge*, pages 354–363. ACM, 2016.

[17] Elenco. Snap circuits. https://www.elenco.com/brand/snap-circuits/, 2021. Accessed: 2019-9-10.

[18] John W Fantuzzo, Ronald E Riggio, Sharon Connelly, and Linda A Dimeff. Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: A component analysis. *Journal of educational psychology*, 81(2):173, 1989.

[19] Logan Fiorella and Richard E Mayer. The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, 38(4):281–288, 2013.

[20] Markus Häring, Dieta Kuchenbrandt, and Elisabeth André. Would you like to play with me? how robots' group membership and task features influence human–robot interaction. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 9–16. IEEE, 2014.

[21] Deanna Hood, Séverin Lemaignan, and Pierre Dillenbourg. When children teach a robot to write: An autonomous teachable humanoid which uses simulated handwriting. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, HRI '15, pages 83–90. Association for Computing Machinery, 2015.

[22] Alexis Jacq, Séverin Lemaignan, Fernando Garcia, Pierre Dillenbourg, and Ana Paiva. Building successful long child-robot interactions in a learning context. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 239–246, 2016. ISSN: 2167-2148.

[23] Jacqueline M Kory-Westlund and Cynthia Breazeal. A long-term study of young children's rapport, social emulation, and language learning with a peer-like robot playmate in preschool. *Frontiers in Robotics and AI*, 6:81, 2019.

[24] Minae Kwon, Malte F Jung, and Ross A Knepper. Human expectations of social robots. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 463–464. IEEE, 2016.

[25] Nathaniel Lasry, Eric Mazur, and Jessica Watkins. Peer instruction: From harvard to the two-year college. *American journal of Physics*, 76(11):1066–1069, 2008.

[26] Janet Metcalfe and Nate Kornell. A region of proximal learning model of study time allocation. *Journal of Memory and Language*, 52(4):463–477, 2005.

[27] Angela M O'Donnell and Alison King. *Cognitive perspectives on peer learning*. Routledge, 2014.

[28] Hae Won Park, Rinat Rosenberg-Kima, Maor Rosenberg, Goren Gordon, and Cynthia Breazeal. Growing growth mindset with a social robot peer. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pages 137–145, 2017.

[29] Leo Porter, Cynthia Bailey Lee, and Beth Simon. Halving fail rates using peer instruction: a study of four computer science courses. In *Proceeding of the 44th ACM technical symposium on Computer science education*, pages 177–182, 2013.

[30] Aditi Ramachandran, Chien-Ming Huang, Edward Gartland, and Brian Scassellati. Thinking aloud with a tutoring robot to enhance learning. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 59–68. ACM.

[31] Aditi Ramachandran, Sarah Strohkorb Sebo, and Brian Scassellati. Personalized robot tutoring using the assistive tutor POMDP (AT-POMDP). *Proceedings of the AAAI Conference on Artificial Intelligence*, 33:8050–8057, 2019.

[32] Aditi Ramachandran, Sarah Strohkorb Sebo, and Brian Scassellati. Personalized robot tutoring using the assistive tutor pomdp (at-pomdp). 2019.

[33] Céline Ray, Francesco Mondada, and Roland Siegwart. What do people expect from robots? In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3816–3821. IEEE, 2008.

[34] Charles Rich, Brett Ponsler, Aaron Holroyd, and Candace L Sidner. Recognizing engagement in human-robot interaction. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 375–382. IEEE, 2010.

[35] Cynthia A Rohrbeck, Marika D Ginsburg-Block, John W Fantuzzo, and Traci R Miller. Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of educational Psychology*, 95(2):240, 2003.

[36] Nicole Salomons, Emir Akdere, and Brian Scassellati. Bkt-pomdp: Fast action selection for user skill modelling over tasks with multiple skills.

[37] Nicole Salomons and Brian Scassellati. Continuous bayesian knowledge tracing - modelling user skills over time. *In Submission*, 2022.

[38] Thorsten Schodde, Kirsten Bergmann, and Stefan Kopp. Adaptive robot language tutoring based on bayesian knowledge tracing and predictive decision-making. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '17, pages 128–136. Association for Computing Machinery, 2017.

[39] Beth Simon, Sarah Esper, Leo Porter, and Quintin Cutts. Student experience in a student-centered peer instruction classroom. In *Proceedings of the ninth annual international ACM conference on International computing education research*, pages 129–136, 2013.

[40] Robert E Slavin. Cooperative learning. *Review of educational research*, 50(2):315–342, 1980.

[41] Fumihide Tanaka and Shizuko Matsuzoe. Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *JHRI*, pages 78–95, 2012.

[42] Cristina Zaga, Manja Lohse, Khiet P. Truong, and Vanessa Evers. The effect of a robot's social character on children's task engagement: Peer versus tutor. In Adriana Tapus, Elisabeth André, Jean-Claude Martin, François Ferland, and Mehdi Ammi, editors, *Social Robotics*, Lecture Notes in Computer Science, pages 704–713. Springer International Publishing, 2015.