Supporting Interest in Science Learning with a Social Robot

Joseph E Michaelis and Bilge Mutlu
University of Wisconsin–Madison, Madison, WI, USA
{jemichaelis,bmutlu}@wisc.edu

ABSTRACT

Education research offers strong evidence that *social supports*, learning interventions situated in meaningful social interaction, during learning can aid in developing interest and promote understanding for the content. However, children are often asked to complete homework tasks in isolation. To address this discrepancy, we build on prior work in social robotics to demonstrate the effectiveness of a *socially adept* robot, as compared to a *socially neutral* robot to generate situational interest and improve learning while reading a science textbook. We conducted a randomized controlled experiment ($N = 63$) of one reading interaction with either the socially adept or socially neutral robot. Our results show that children who read with a socially adept robot found the robot to be friendlier and more attractive, reported a higher level of closeness and mutual-liking for the robot, had higher situational interest, and made more scientifically accurate statements on a concept-map activity. We discuss the practical and theoretical implications of these findings.

Figure 1: Child reading an augmented book with the learning companion robot, "Minnie."

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in HCI*; • **Social and professional topics** → *K-12 education*; • **Applied computing** → *Collaborative learning*.

KEYWORDS

Social robots; education; interest development; science learning

ACM Reference Format:
Joseph E Michaelis and Bilge Mutlu. 2019. Supporting Interest in Science Learning with a Social Robot. In *Interaction Design and Children (IDC '19), June 12–15, 2019, Boise, ID, USA.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3311927.3323154

1 INTRODUCTION

Education researchers and learning scientists have emphasized that impactful student learning requires *deep understanding*, knowledge that goes beyond simple acquisition of facts and procedures [10, 70]. Interest in academic content, such as science, helps promote increased effort, self-regulation, and perseverance through challenge that aid in deep understanding [65]. Advances in learning sciences and interest development research emphasize the importance of personalized learning and *social supports*, learning activities situated in meaningful social interaction, to develop interest and improve learning, often expanding the level of difficulty of work that the learner can be successful in [36, 76]. While there is clear demand for increasing student interest and deep understanding in science [16], personalized learning and social supports that can benefit this type of learning take considerable resources and planning to include in a learning environment and are often lacking in learning opportunities that take place in the home (i.e., homework). To address the lack of social support for children during science learning at home, we have developed a social robot to act as an in-home learning companion (See Figure 1). The robot is designed to augment traditional science textbook reading by incorporating *socially situated interest scaffolds* that are personalized learning supports that specifically emphasize the cognitive and affective needs of children, including positive social interaction, for learning and developing interest.

Students who have interest in what they engage in experience deeper cognitive processing, persist longer through challenges, and focus longer on the learning activity [29, 31]. Developing student interest during learning does not automatically occur. Rather, interest requires support and opportunity to develop. Interest can be positively influenced by learning activities that employ *interest scaffolding* [36], instruction with supports tailored to meet the specific interest development needs of individual students, such as giving high interest students opportunities to pursue their own ideas while learning [66]. These interest scaffolds can also be *socially situated* to include meaningful social interactions, such as sharing insights during learning, that are powerful supports for interest development. [2, 59, 61]. Designing learning environments

IDC '19, June 12–15, 2019, Boise, ID, USA

ACM ISBN 978-1-4503-6690-8/19/06…$15.00
https://doi.org/10.1145/3311927.3323154

with effective socially situated interest scaffolding has the potential to be a transformative mechanism to enhance student achievement.

Social robots may be an especially effective educational technology to provide socially situated interest scaffolding while learning in academic content because of their ability to rapidly adapt their interactions to tailor learning supports and their ability to behave in socially meaningful and adept ways, such as using non-verbal cues and personal stories, during those interactions [5]. This effect may be particularly powerful for learning at home, where homework is often done in isolation, and could benefit from social interaction and interest support. Given recent advancements in educational robotics [5, 6], it is crucial that their potential for having a positive impact on interest and learning in academic content is examined through rigorous study. We believe that socially situated interest scaffolds, delivered by a learning companion robot programmed to provide meaningful and adept social interactions, will make an especially effective learning technology to improve educational outcomes such as learning and interest development. As we begin to explore this possibility, our initial goal is to test how impactful socially adept robot behaviors, as part of the socially situated interest scaffolds, are to learning and interest development.

2 RELATED WORK

In this work, we draw on theories of interest development, social learning and human-robot interaction that are reviewed here.

2.1 Interest Development and Social Learning

There is evidence that interest is closely related to learning and deeper understanding of content [31, 65]. As such, it is important to foster long-term interest in academic subjects to facilitate learning in those domains over time. Researchers studying interest development describe *situational interest* as a momentary psychological state during engagement, evident from observed increases in focus and attention [31, 55], that are classified as being either *catch* (or triggered) or *hold* (or maintained). Catch situational interest, often elicited by novel or exciting environmental features, leads to increased engagement and positive feelings in the moment. Hold situational interest—through increasing a learner's knowledge of the content and supporting them to see what they are learning as valuable and rewarding—will lead to sustained engagement and re-engagement over time [34, 65]. Repeated activation of situational interest, particularly instances of hold situational interest, is thought to lead to the development of *individual interest*, a stable pre-disposition to re-engage with content [30, 31, 65]. While increased situational interest does not always lead to improved learning in the moment [21], situational interest is the mechanism for developing the long-term individual interest in the content that is related to improved learning outcomes over time [31].

Interest development supports are more deeply impactful when they are *socially situated*, where the supports are presented in the context of meaningful social interaction [2, 7]. Socially situated learning promotes positive affect, an important early vehicle for triggering situational interest, and more importantly, supports developing value and knowledge for the content that facilitates hold situational interest [65]. Value for the content is supported through social interaction during learning, because these interactions give the learner opportunities to build relationships, share values, and feel a sense of social involvement during academic learning [2, 18, 34]. Value for the content is promoted when others share excitement for the work, provide access to additional information and guidance with ideas about how and what to pursue next [7]. These social connections aid an individual in internalizing values for the content through finding shared purpose, focus, and values [18]. During learning, children benefit from social interactions with others who are supportive and approachable, friendly, humorous and enthusiastic, because these interactions promote feelings of belonging and relatedness to the content and to others engaging in learning that content [49]. For example, in a longitudinal study of young readers, Nolen [59] found a social reading environment, where children engaged in partner readings and shared book recommendations, and teachers emphasized the "social value of reading," promoted developing individual interest in reading over time. Social supports that aid learners in finding content to be valuable and rewarding have also been used to promote interest in science using positive social influence. Jackson et al. [35] found that, for women, positive social recognition, such as encouraging and understanding the person's interest in science, is related to increased career interest in science fields. Even simple encouragement can serve as an effective social support; Hulleman et al. [34] found that the more parents talked to their children about the importance of STEM, the more STEM classes students took in high school.

Content knowledge is also supported through social interactions, because learning is transformed into a collaborative activity that allows for shared knowledge construction and learning through modeling that promote deeper understanding of the content [4, 58, 75]. Learning for deeper understanding entails going beyond superficial knowledge to make rich interconnections in the content with other prior knowledge—where the learner understands the 'big ideas' or systemic connections—and generating new ideas that relate to the context of the content [12, 63]. Social interactions influence cognition by integrating perspectives through shared language, and distributing cognitive processes among group members that helps generate new ideas and insights that result from the interaction [19, 32]. This type of social interaction is effective for improving reading comprehension when children are able to discuss what they are reading with others. Social others can perform part of the cognitive activity required for comprehension by summarizing and rephrasing what they are reading or pointing out key features of the text. Social interactions also promote deeper understanding from reading through discussions that prompt new ideas about the content or connect that content to other prior knowledge, that may improve long-term knowledge retention [37].

Social influences can increase positive affect, value, and knowledge in academic domains that improve learning and interest. However, students do not always have the opportunity to experience positive social interactions while learning, particularly at home.

2.2 Human-Robot Interaction

Human-computer interaction (HCI) and Human-robot interaction (HRI) research has found that humans respond to computers and technology in social ways and that robots can be designed in ways to elicit a strong social response [20, 24]. In general, working with a non-human agent enhances learning more than working without one [71]. These effects are greater with the use of conversational and polite speech and human-like actions by the agent [38]. These agents are most often designed as digital companions, but a robotic learning companion is best suited to provide social support, because their physical presence enhances their capacity for greater social connection-making than other technologies [3, 5, 40]. While robotic platforms come with many inherent challenges, including cost and complexity, the potential benefit of greater social connection-making by promoting interest and learning motivates us to overcome these challenges.

Even very simple robots that incorporate few if any social cues, here-after referred to as a *socially neutral robot* as done by Saerbeck et al. [69], can be perceived in a social way. However, *socially adept robots* that utilize existing human-human social cues and behaviors for *social connection-making* are capable of establishing and maintaining strong social connections that last over time [5]. These cues include eye gaze, such as eye contact [56] and well-timed gaze-aversion patterns [1], addressing the person by name during interaction or acknowledging what happened in prior interaction (e.g., recalling who won or lost a previous game) [46], making individualized recommendations for that person [48], and speaking in an expressive [44] or informal tone [8]. Leite et al. [45], based on a synthesis of social robotics design research, suggest that social robots should be capable of identifying users, recalling previous interactions, personalizing interactions based on the user and prior interactions, and incrementally disclosing personal characteristics and demonstrating novel behavior. These design considerations help social robots overcome some of the difficulties found in HRI such as a potential drop in engagement after a novelty effect has worn off and programmed interactions becoming repetitious over time [23, 46]. Implementing these recommendations into the design of a social companion robot will encourage broader acceptance of the technology [14, 27] and aid in social connection-making between the robot and a learner that benefit learning [5, 69].

One aspect of learning that a social robot may be especially useful for is the support of in-home science learning. In-home learning has long been an established part of school-based education, and homework completion and accuracy contribute to academic performance [15], particularly in pre-college math and science classes [22]. While reading from a textbook is a common homework assignment for middle school science classrooms, many middle school students find textbook reading boring and difficult [28]. Including socially situated interest scaffolds addresses this problem by improving a student's situational interest in textbook reading, that creates a path towards increased learning and individual interest in science. To support interest during science textbook reading, Guthrie and Klauda [28] suggest supporting student self-efficacy with supplemental materials, demonstrating the value of reading a textbook for understanding the content, and using social interaction around reading. There is also evidence that including value and relevance connections while learning from a science textbook have a positive effect on perceived value and topic interest in the content [78]. This prior work suggests that middle school students' situational interest and learning would benefit from reading with a social robot that is designed for socially situated interest scaffolding.

The last decade has seen an increase in research into how social robots can support motivation and promote learning through interactions with children in and out of schools [5]. These studies have typically looked at robots designed to improve basic skills such as vocabulary and language learning [42, 48, 69, 74], handwriting [33], and test taking [11]; or to make learning more enjoyable [39] and positively influence a child's curiosity [26]. A promising study by Saerbeck et al. [69] found that a robot programmed with socially supportive behaviors, when compared to a socially neutral robot, improved learning and motivation in a language learning task. Similarly, Kory-Westlund et al. [44] found that a robot using expressive speech was more effective than flat speech, when the robot read stories and asked children questions about the stories, for promoting vocabulary learning as well as concentration and engagement. In math learning activities, studies show that children prefer working with a social robot over a workbook [47] and that dynamic vocalizations from a robot improve rapport and social presence [50], but neither study showed any learning gains. One example of a study in a science learning environment found that children that routinely asked science questions of a social robot in their classroom had increased levels of curiosity for science [72]. While studies have shown that socially supportive behavior by a robot is beneficial to learning and motivation in some areas, little work has been done to examine the use of a social robot to facilitate learning and motivation related to science content or to augment reading non-fiction textbook materials.

The role of the robot as a *listener* during reading activity, where the child is responsible for the reading and the robot augments the reading with verbal responses, has also been under-explored. In prior work [51, 53], we designed a learning companion robot to augment in-home reading activities with children. This work found that children reading with the robot at home made a social connection with the robot and described liking the activity because the robot acted as a social companion with whom they could share the reading. The social connection made with the robot appeared to positively influence the child's interest and was found to deepen over a period of two weeks. These results suggest that augmenting existing learning activities can be a feasible means of integrating a social robot with existing curriculum to support classroom educational goals at home. While this work is promising, there are several limitations that limit the robot's ability to provide socially meaningful interactions. First, the robot used a classic inexpressive text-to-speech engine that children found detracted from the reading experience. Second, these studies included a small sample size and did not compare experiences between two similar, but distinct, robots. Third, the studies only considered the student experience and their situational interest while reading with the robot and did not examine *learning* as an outcome of the interaction. Thus, the study is limited in it's capacity for more generalizable claims and does not investigate whether it is the socially adept activity of the robot or simply the effect of having any robot that drives social connection-making, the development of interest, or learning. Finally, the studies have been limited to reading activities structured around casual reading of popular fiction and non-fiction books. It may be easier for the robot to establish a social connection and promote interest and learning while sharing this casual type of reading experience than in the context of reading an academic textbook.

We expand on findings from prior work to further our understanding of learning and interest development with a social robot during a shared reading experience of academic content. We examine the ability of a socially adept robot to promote social connection-making, situational interest, and learning in the context of science textbook reading as compared to a socially neutral robot. We conducted this examination as a randomized controlled experiment as a first step in assessing the feasibility of a socially adept robot to positively affect interest and learning that can later be tested in more authentic settings. In this study, we posited and pre-registered[1] the following hypothesis:

Hypothesis. *Interacting with a socially adept robot will result in greater social connection-making with the robot, situational interest, and learning for children reading a science textbook when compared to a socially neutral robot.*

3 METHOD

To explore the effects of socially adept behavior from a robot learning companion on social connection-making, situational interest, and learning in the context of science textbook reading, we conducted a laboratory study in which we asked participants to read a science textbook with a learning companion robot. The robot, which we called "Minnie," was programmed to demonstrate interaction behavior that was either socially adept or socially neutral. The socially adept robot incorporated design elements from our previous work [51] as socially situated interest scaffolds, and the neutral robot reduced the use of the elements where applicable. In the following paragraphs, we describe the design of the robot §3.1.1, the experimental conditions §3.1.2, the design of the robot's comments §3.1.3, and the selection of reading materials §3.1.4. We then explain our procedure §3.2 and our evaluation measures §3.3.

3.1 Robot Design

Minnie (See Figure 1) is a modified version of the freely available 3D-printable Maki robot design from Hello Robo.[2] The robot has a 13.5-inch-tall body including a static torso with servo-controlled moveable head and eyes. We modified the Maki design to improve the interaction methods for the robot by adding a small camera with a fish-eye lens and a seven-inch touchscreen to the robot. To do so, we modified the 3D-printable files to set a mounting space for the camera near the top of the torso, extruded a mounting plate from the front torso of the robot and added a custom case for the screen to connect to the mounting plate. The camera was used to allow the robot to process facial recognition software and read scannable ID tags. The touchscreen was used to display the graphical user interface (GUI) with an image across the top three-quarters of the screen and five interactive color-coded buttons along the bottom quarter. Each colored button also included the text indicating its function. The child could indicate "yes" or "continue" with the green button, "no" or "stop" with the red button, "repeat" with the blue button, "pause" with the yellow button, and ask for "help" with the purple button ("help" was disabled for this study).

3.1.1 General Interaction Design of the Robot. In both conditions, the interaction with Minnie begins with the child reading out-loud

[1] Pre-registration available at: http://aspredicted.org/blind.php?x=ub7de5.
[2] https://www.hello-robo.com/

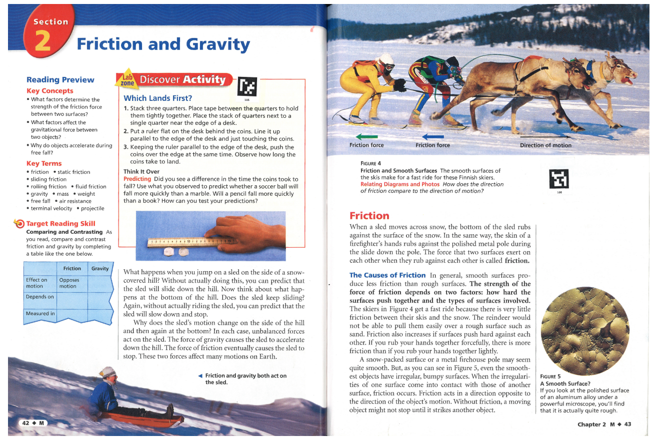

Figure 2: Pages from the middle school physics text. The left page includes an AprilTag at the top right of the text and a hands-on "discover activity" to be completed during the interaction. The right page includes an AprilTag that prompts variable comments related to the image above it.

from a custom-made introductory book that is used to demonstrate how to interact with the robot. During this reading, children learn how to scan special AprilTags,[3] embedded as visual identifiers (IDs), in the books and how to use the color-coded buttons to respond to the robot. During the introduction book, children are also asked to complete a short topic sorting task. Using 10 reading topic interest cards labeled with AprilTags (e.g., animals, science, art), children choose all of the types of books or topics they like to read about and scan the ApriTags on those cards. These topics are saved as topic preferences in each child's user profile, and these preferences are later used to select tailored responses during reading.

After the introductory book, children are then asked to read, out-loud, 13 pages of text from a middle school physics textbook [60] that has AprilTags embedded on each page (See Figure 2). The robot tells them that reading this section is their goal for the day and to begin when they are ready. As the children progress through the book, they scan AprilTags on each page (See Figure 3), and the robot responds with pre-programmed comments that correspond to the material on that page. These comments include discussing the main text that the child is reading, directing the child's attention to illustrations on the page, or asking them to complete small activities in the textbook (See Figure 2). There are also two comments during the reading where the robot will tailor the comment to incorporate one of the child's topic preferences from their user profile. At the end of the reading activity, the robot indicates the child completed their reading goal and shuts down.

3.1.2 Comparing Robot Conditions During Interaction. During the interaction, the learning companion robot is programmed to behave according to one of two conditions. In the socially adept condition (treatment condition), the robot behaves according to social design principles identified in prior work, including the following:

[3] https://april.eecs.umich.edu/software/apriltag

Figure 3: Child scanning an AprilTag by pointing the textbook towards the robot's camera (red circle on robot torso).

- *Expressive speech*: Using a natural and expressive speech synthesis tool, the Salli voice from AWS Polly,[4] we prerecorded each comment for the socially adept robot and tested each for accurate pitch, tone, and expression [44].
- *Nonverbal cues*: The robot moved its head and eyes to track the position of the child and to make eye contact using facial recognition software; averted its gaze at semi-regular intervals during interaction to avoid overly long eye contact; and averted its gaze prior to speaking in instances when the comment should be received as thoughtful [1, 56].
- *Personal comments*: During reading, the robot makes comments that incorporates personal connections and reactions to the reading based on a fictional back-story, demonstrates enthusiasm for the reading, refers to itself as a partner in the reading, and calls the child by their name [46, 48].

In the socially neutral robot condition (control condition), the robot followed the same interaction protocol as the the robot in the treatment condition, but we removed the socially adept behaviors. All speech in the control condition was synthesized using SVOX Pico[5] text-to-speech to provide clear speech that is unexpressive in tone and pitch. The robot does not track the child's face during their interaction, nor does the robot use gaze aversion in conjunction with any spoken comments. Finally, the robot's comments were revised to remove any personal connections or emotions from the robot, such as referring to its fictional back-story or showing enthusiasm for reading. The robot also does not refer to itself as a partner in the activity, such as stating "you should keep reading," instead of "we should keep reading," nor does it refer to the child by their name. Each comment for the socially neutral robot maintained the general content of the comment as written in the socially adept robot condition (e.g., both conditions include comments that direct the child to look at a picture or re-phrasings of a key portion of the text on the page) and is similar in length of speech as in the treatment condition. While we removed as many socially adept behaviors as possible to create a contrast between conditions, intermittent blinking of the robot's eyelids and small semi-randomized head motions during the interaction ensured that the children were aware that the robot was functioning properly in both conditions. When either an AprilTag is scanned or a button is pressed, the robot blinks and makes a "bleep" noise to indicate receiving the input.

[4]https://aws.amazon.com/polly/
[5]https://packages.debian.org/source/jessie/svox

3.1.3 Robot Comments. Each robot comment was designed to appear as if the robot were following along with the child's reading, and each corresponding AprilTag is placed in proximity to the location of the text or graphic to which the comment refers. For example, for the page on the left side of Figure 2, scanning the AprilTag on that page prompts the child to complete the "discover activity" immediately to the left of the tag. For each comment, a control and treatment condition version are written where we controlled for length of the comment, basic content of the comment, and any cognitive supports included in the content.

To illustrate how the robot provided comments during the reading activity, we provide an example of the robot comments from one page of the textbook. On page 43 of the textbook (Figure 2, right page) there is a figure that includes an image of Finnish Reindeer Skiers in action to demonstrate the effect of friction on a smooth surface. For this page, we placed an AprilTag just below this figure and when the tag is scanned, the robot selects one of ten comments pre-programmed to relate the image in the figure to a topic preference for the child. In this case, if the child indicated that they prefer to read about sports during their topic sort, the robot's comment makes a connection to the sport depicted in the image. In the control condition, this comment read, "The skiing in this picture is a unique sport called Reindeer racing. The force from the reindeer must be greater than the force of friction from the skiers," and in the treatment condition, this comment read, "Racing with reindeer! What an incredible sport. I bet the force from the reindeer must be greater than the force of friction from the skiers." Both comments tailor the content to the reader's topic preferences, but only the treatment condition uses language that is personal and enthusiastic.

3.1.4 Science Textbook Reading. The science textbook was the book titled, "Motion, Forces, and Energy," in the Prentice Hall Science Explorer series [60]. This book was chosen because it is a commonly used textbook for U.S. middle school (7th and 8th grade) science classes and is written at a reading level appropriate for 10–12 year-old children, but it was unlikely that our sample would be overly familiar with the content, as none had begun their 7th grade year. We asked children to read section 2.1, Nature of Forces, that covered an introduction to forces including a definition of force, balanced and unbalanced forces in two-dimensions, and friction and gravity. There were two activities in this section that the robot asked the children to complete. Activity materials were provided for the child by the researcher, and after each activity the robot added clarifying comments designed to support comprehension of the activity outcomes. For example, on page 42 (See Figure 2, left page) the robot prompts the child to complete the "discover activity" on the page. The activity involves the child placing two different sized stacks of quarters on the edge of a table and pushing them off simultaneously to observe which falls faster. Since Newtonian Physics would determine that both stacks would fall at the same rate, for clarity, the robot is programmed to comment after the activity that both stacks of quarters fell at the same time.

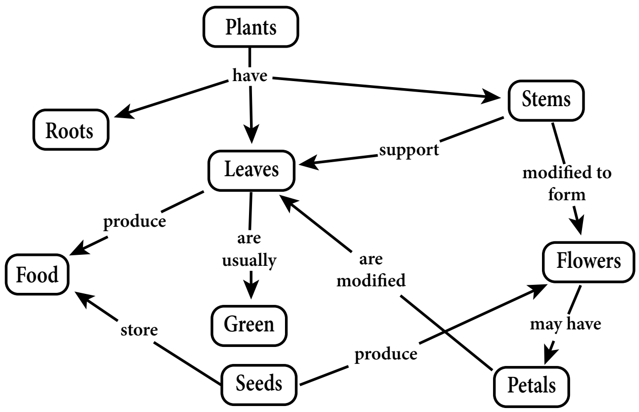

Figure 4: Sample content map depicting connections between biology concepts.

3.2 Procedure

Participants (N = 63) were children from the greater Madison, Wisconsin area (ages 10–12; M = 10.9; 36 male, 27 female) who had not begun their 7th grade year. Children were randomly assigned to either the treatment or control condition. One child in the control condition could not complete the study due to technical problems with the robot, and four children (two in each condition) did not complete all study activities in the time allotted for the visit. Data from these five participants were removed from analysis, leaving 29 children in each condition. Children completed study activities in a campus lab office after parent consent. Parents of the children were compensated $25 USD for the study, and study protocols were reviewed and approved by an institutional review board.

Each participant began the study by completing a pre-test to assess prior knowledge of the physics concepts included in the reading, followed by measures of individual interest in science and reading and a reading ability assessment. Children then followed the robot interaction protocol that included reading the introduction book and 13 pages of the textbook chapter. After completing the reading session with the robot, children were asked to complete a concept-map activity, a survey of their social evaluation of the robot, a post-test to assess learning from the reading activity, a survey of situational interest, and provide demographic information. Finally, children were interviewed about their experience with the robot. In this analysis, we include results from quantitative measures from the study. All qualitative data are currently undergoing further analysis and will be reported in detail in future publications. All measures are described in detail in the next section.

3.3 Measures

3.3.1 Learning Measures. We administered two types of assessments to estimate content knowledge for each child. The first was a content knowledge test similar to classroom quizzes commonly administered to assess a child's understanding of academic content. The content knowledge tests in this study consisted of 12 questions (9 multiple-choice and 3 open-ended) designed for the study to be balanced in complexity from basic comprehension to synthesis [9] and align with the science content in the textbook. Two similar versions of the test were created to be used as either pre- or post-tests. The test order was randomized to control for test difficulty. Each question was scored on a scale of 0-1. Several multiple choice items had more than one correct answer choice. Partial credit was given for selecting each correct answer choice, and no penalties were given for incorrect selections. Open-ended questions were scored by a researcher using a scoring rubric. Inter-rater reliability was tested for 20% of the data and found to be high for all three items (κ = 1.0, 0.85, 0.88). Pre- and post-tests were scored by summing points across all items with a range of possible scores from 0 to 12.

Since this type of content knowledge test often reflects a shallow understanding of content knowledge that can be easily repeated from reading materials, we also chose to incorporate a concept-map activity (See Figure 4) to assess deeper understanding. Making a concept-map consists of drawing nodes that represent main ideas and drawing lines between nodes including text to describe how the ideas in the nodes are related. The concept-map assessment represents an alternative method for assessing student learning and comprehension to that of a quiz score, because a concept-map requires a depth of understanding of how ideas are interrelated with little guidance from the assessment itself [68]. Thus, using a concept-map is intended to give insight into deeper content understanding.

Each child was given instructions and an example (Figure 4) on how to complete a concept-map prior to beginning their own concept-map of the ideas contained in the textbook reading. Children where then instructed to, "draw a concept-map about what you just read on the paper below. Start out with what you believe to be the main ideas of what you read and connect other ideas from there." They were given five minutes to complete their concept-map and were not allowed to refer back to the text. Each concept-map was analyzed using a rubric based on a standardized rubric used in prior work and scored for accuracy [64, 73]. Concept-maps scores were calculated by summing the number of scientifically accurate connections made. Inter-rater reliability was calculated for 20% of the data and found to be high (κ = 0.83).

3.3.2 Situational Interest. To measure situational interest for the reading session, we included a survey of situational interest with two factors, catch and hold situational interest, that is based on a previously validated and reliable measure [43]. We modified the language of each item to relate to experiences with reading activities, and in some cases made the reading level of the items more age appropriate. These two factors each included 6 Likert-style items on a scale of 1 (*strongly disagree*) to 7 (*strongly agree*). An example of a situational interest item includes: "I think the topics I learned about in the reading activity matter to me."

3.3.3 Social Connection-Making. After each session, children completed a survey to measure their perception of the robot along several socially meaningful traits to assess the level of social connection-making the child experienced with the robot. The measures of children's perception of the robot included a survey of 14 items developed from of a measure of social and intelligence evaluations of digital agents that includes five factors: sociable (5 items), mutual liking (2 items), attractiveness (2 items), human-likeness (1 item), and intelligence (4 items) [57, 62]. This survey included 14 adjective pair items where children were asked to rate their feelings about the robot on a scale of 1 (first adjective, e.g., *unhappy*) to 7 (second adjective, e.g., *happy*). For example, one item for the sociable factor asked the child "How bored or excited was your partner?," and the child was asked to choose between bored and excited on the adjective pair scale. Examples of adjective pairs for the other factors on the survey include very little to very much (human-like), not-cute to cute (attractive), not-smart to smart (intelligence). We also asked how much the child liked the robot, and how much the robot liked the child with adjective pairs of very little to very much. We also included one item to assess the child's feeling of closeness with the robot based on a previously validated measure [25] that found the single item to be equally appropriate for measuring closeness as several other larger measurement tools. This measure asks children to choose one set of increasingly overlapping circles that "best shows your relationship with the robot."

Each survey was scored as a mean score of all items on the survey to reduce the scale to a range (1–7) more easily interpreted than a simple raw score. For each survey, we calculated Crohnbach's alpha, with $\alpha > 0.70$ as an indicator of reliability [17], and conducted a factor analysis to verify the factor structure of each measure was consistent with models from prior work.

3.3.4 Randomization Checks. We also included several measures to estimate whether the randomized groups were relatively equivalent. These measures included the pre-test of content knowledge, reading ability, and self-report Likert-style surveys of individual interest for science (science interest) and individual interest for reading (reading interest). We also asked children to provide demographics including age and gender.

For the reading ability assessment, children read a one-page article and answered six questions (four multiple-choice, two open-ended) on the article. The assessment was based on the 5th-grade Florida Comprehensive Assessment Test for reading comprehension. Children were given as much time as they needed to read the article and were then given five minutes to answer the questions while referencing the article. Each question was scored on a scale of 0–1 with partial credit for partially correct answers with a maximum possible score of 6.

The science interest and reading interest scales, developed and tested in prior work [52], were based on a previously validated reliable measure of individual interest [54]. The reading interest scale was modified by incorporating language from the Motivation for Reading Questionnaire [77]. Each of these scales included 10 Likert-style items, and children were asked to rate their agreement with each statement on a scale of 1 (*strongly disagree*) to 7 (*strongly agree*). Examples of science and reading interest items include: "if my science homework is interesting, I will keep working at it even if it is difficult," and "I like to read hard, challenging books."

3.3.5 Statistical Analysis. As a check of random assignment to condition groups, all group characteristics were compared for mean differences, and a comparison of gender ratios for each group was also considered. To test our hypothesis, we then conducted an independent t-test to compare mean differences between conditions for each of the 10 outcome measures. Since these tests all represented tests of the primary hypothesis, we use a significance test of $\alpha = 0.05$ for each test, without including family-wise error correction as it is unnecessary in this type of testing [41]. We do also report the results of significance testing using a Holm-Bonferonni correction for family-wise error rate in the case that our readers would like to apply this more stringent criteria in considering our results. We also consider the effect size for each result using Cohen's d with conventional effect sizes of small ($d > 0.20$), medium ($d > 0.40$), and large ($d > 0.80$) [13].

Table 1: Comparison of group characteristics between control and treatment groups.

	Control	**Treatment**
Measure	*Mean (SD)*	*Mean (SD)*
Age	10.79 (0.62)	11.03 (0.82)
Male/Female	17/12	16/13
Reading Interest	5.40 (0.87)	5.63 (0.75)
Science Interest	4.39 (1.23)	4.61 (1.00)
Reading Ability	4.21 (1.36)	4.54 (1.17)
Pre-test	3.73 (1.56)	3.45 (1.76)

4 RESULTS

All group characteristics were found to be similar between the control and treatment groups (See Table 1), satisfying our randomization check. We also found all survey measures to have high internal reliability and a factor structure similar to prior work. Individual interest in reading ($\alpha = 0.75$; $M = 5.50$; $SD = 0.80$) was higher than individual interest in science ($\alpha = 0.87$; $M = 4.54$; $SD = 1.09$), and both distributions were centered around the mean and similar for both groups ($t(58) = 0.75$, $p = 0.46$). For pre-test assessments, low scores ($M = 3.64$; $SD = 1.70$) indicated that prior knowledge for the physics content was low but similar between groups ($t(58) = 0.62$, $p = 0.53$). Measures of reading ability ($M = 4.36$; $SD = 1.30$) indicated a relatively high reading level that skewed to the left and was even between groups ($t(58) = 1.01$, $p = 0.32$). Both groups also had similar ratios of males to females. Males ($M = 4.07$, $SD = 1.76$) had higher prior knowledge than females ($M = 3.07$, $SD = 1.47$; $t(60) = 2.38$, $p = 0.02$) did, but no other differences were found between genders. We also found differences in reading ability and prior knowledge between ages groups. The 12-year old children had higher prior knowledge ($M = 4.72$, $SD = 2.03$) than 10-year-olds ($M = 2.93$, $SD = 1.16$, $p = 0.007$), and 12-year-olds ($M = 5.16$, $SD = 0.46$, $p = 0.0007$) and 11-year-olds ($M = 4.49$, $SD = 1.19$, $p = 0.03$) had higher reading ability than 10-year-olds ($M = 3.59$, $SD = 1.37$). No other main effects were found based on age.

To examine group differences in social connection-making with the robot we compared measures of children's perceptions of the robot for each group (See Table 2 and Figure 5). We found a large effect size for differences in closeness, with the treatment group ($M = 3.17$, $SD = 0.89$) significantly higher than the control ($M = 2.41$, $SD = 0.91$, $t(56) = 3.22$, $p = 0.002$, $d = 0.84$). For perception of the robot's sociability, we found a large effect size where those in the treatment group ($M = 5.63$, $SD = 0.98$) had higher mean scores than those in the control group ($M = 4.51$, $SD = 1.11$, $t(56) = 4.09$, $p = 0.0001$, $d = 1.07$). For the measure of mutual-liking, we found a large effect size where the treatment group $M = 5.53$, $SD = 1.27$) had significantly higher mean scores than the control group ($M = 4.40$,

Table 2: Perception of robot measures, including Cohen's d measure of effect size. * denotes $p < 0.05$, and † denotes Bonferroni corrected α. For all tests, $df = 56$.

Measure	*Group*	*M (SD)*	*t*	*d*	*p*
Closeness	Control	2.41 (0.91)	3.22	0.84	0.002∗ †
	Treatment	3.17 (0.89)			
Sociable	Control	4.51 (1.11)	4.09	1.07	0.0001 ∗ †
	Treatment	5.63 (0.98)			
Mutual–Liking	Control	4.40 (1.42)	3.22	0.84	0.002∗ †
	Treatment	5.53 (1.27)			
Attractive	Control	3.36 (1.46)	2.19	0.57	0.03 ∗
	Treatment	4.19 (1.42)			
Human	Control	3.45 (1.30)	1.07	0.2	0.29
	Treatment	3.83 (1.39)			
Intelligent	Control	5.83 (1.02)	0.52	0.14	0.60
	Treatment	5.71 (0.72)			

$SD = 1.42$, $t(56) = 3.22$, $p = 0.002$, $d = 0.84$). We found a medium effect for children describing the robot as attractive with treatment scores ($M = 4.19$, $SD = 1.42$) significantly higher than the control ($M = 3.36$, $SD = 1.46$, $t(56) = 2.19$, $p = 0.03$, $d = 0.57$). Finally, no significant differences were found for describing the robot as human-like ($t(56) = 1.07$, $p = 0.29$, $d = 0.20$) or intelligent ($t(56) = 0.52$, $p = 0.60$, $d = 0.14$).

We then compared situational interest and learning for each group (See Table 3 and Figure 6). For hold situational interest, we found a medium effect and significant difference in mean scores, where the treatment group ($M = 5.03$, $SD = 1.27$) was higher than the control group ($M = 4.19$, $SD = 1.49$, $t(56) = 2.3$, $p = 0.03$, $d = 0.60$). For catch situational interest, there was a medium effect size and marginal difference between groups, where the treatment group ($M = 5.00$, $SD = 1.42$) was higher than the control group ($M = 4.21$, $SD = 1.64$, $t(56) = 1.95$, $p = 0.06$, $d = 0.51$). For Post-Test measures, there was no significant difference between control ($M = 5.29$, $SD = 2.16$) and treatment groups ($M = 5.40$, $SD = 2.32$, $t(56) = 0.19$, $p = 0.85$, $d = 0.05$). There was a significant increase in test scores with a large effect size from pre-test to post-test for both control ($t(56) = 3.16$, $p = 0.0025$, $d = 0.83$) and treatment groups ($t(56) = 3.5$, $p = 0.0008$, $d = 0.93$). For scientifically accurate connections made on the concept-map activity, we found a medium effect size where those in the treatment group ($M = 3.48$, $SD = 2.96$) had significantly higher mean scores than those in the control group ($M = 1.79$, $SD = 1.95$, $t(56) = 2.57$, $p = 0.01$, $d = 0.67$).

5 DISCUSSION

In this paper we present quantitative measures to test the hypothesis that: Interacting with a socially adept robot will result in greater social connection-making with the robot, situational interest, and learning for children reading a science textbook when compared to a socially neutral robot. In this short-term study, we found statistical support favoring the socially adept robot for measures of social connection-making, catch and hold situational interest, and learning measured by concept-maps, but no treatment differences on the written post-test.

Table 3: Differences in learning and interest measures, including Cohen's d measure of effect size. * denotes $p < 0.05$, and † denotes Bonferroni corrected α. For all tests, $df = 56$.

Measure	*Group*	*M (SD)*	*t*	*d*	*p*
Concept-map	Control	1.79 (1.95)	2.57	0.67	0.01 ∗
	Treatment	3.48 (2.96)			
Post-Test	Control	5.29 (2.16)	0.19	0.05	0.85
	Treatment	5.40 (2.32)			
Hold Interest	Control	4.19 (1.49)	2.3	0.60	0.03∗
	Treatment	5.03 (1.27)			
Catch Interest	Control	4.21 (1.64)	1.95	0.51	0.06
	Treatment	5.00 (1.42)			

5.1 Social Connections, Interest, and Learning

Our quantitative measures demonstrate that reading with the socially adept robot supports greater social connection-making that benefits situational interest and deep understanding than the socially neutral robot. These findings are consistent with prior work that suggests socially adept behavior for a robot enhances social connection-making [46], socially supportive behaviors enhance learning and motivation [69], and that socially situated interest supports in educational activities benefit interest and learning [36].

We found evidence that children found the socially adept robot to be more attractive and sociable and that they felt greater levels of closeness and mutual-liking for the socially adept robot. However, we did not find differences in how the children perceived each robot's intelligence or human-likeness. These results are consistent with the design of our study, and with previous work [45, 48, 57]. We designed the main differences between the robots to be social differences that were driven by verbal and non-verbal social cues. These socially adept behaviors included head and eye gaze behaviors designed to increase the perception that the robot was making eye contact and averting its gaze in socially appropriate ways, and verbal patterns that included more personal dialogue and more natural and expressive speech patterns. At least in short-term interactions, these differences seemed to increase the perception that the robot was sociable and attractive, and may have contributed to feeling closeness and mutual-liking with the robot. Interestingly, we did not find differences in how the children perceived the robot's intelligence or it's human-likeness. While this finding is in contrast to previous work suggesting non-verbal cues such as proper eye gaze movements increases the perception that the robot is thoughtful [1], it may be that the identical physical design of the robot and input methods contributed more strongly to these perceptions than did any of the social differences between the robots. It also may be the case that children were less aware of the robot's physical movements because they were attending the reading from the textbook. However, these explanations and inferences require further study, and studies of long-term interactions would be beneficial.

As expected, the differences in social aptitude of the robots led to differences in situational interest, where scores for catch and hold situational interest were higher for children working with the socially adept robot. While the differences in catch situational interest were only marginally significant, we did find a medium

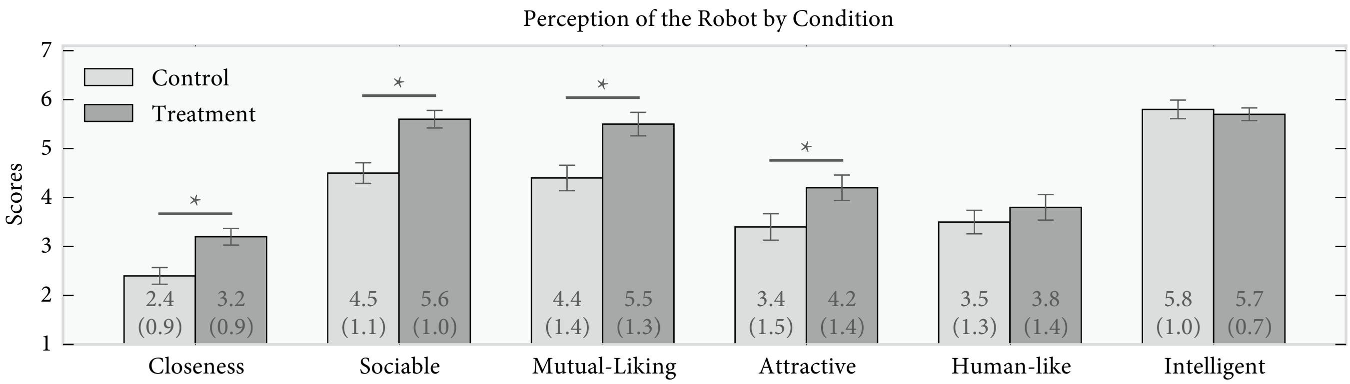

Figure 5: Bar graphs comparing perceptions of the robot by condition with means and (standard deviation) appearing in each bar. A solid line with * indicates $p < 0.05$, and a dotted line indicates $p < 0.10$. Whiskers represent standard error.

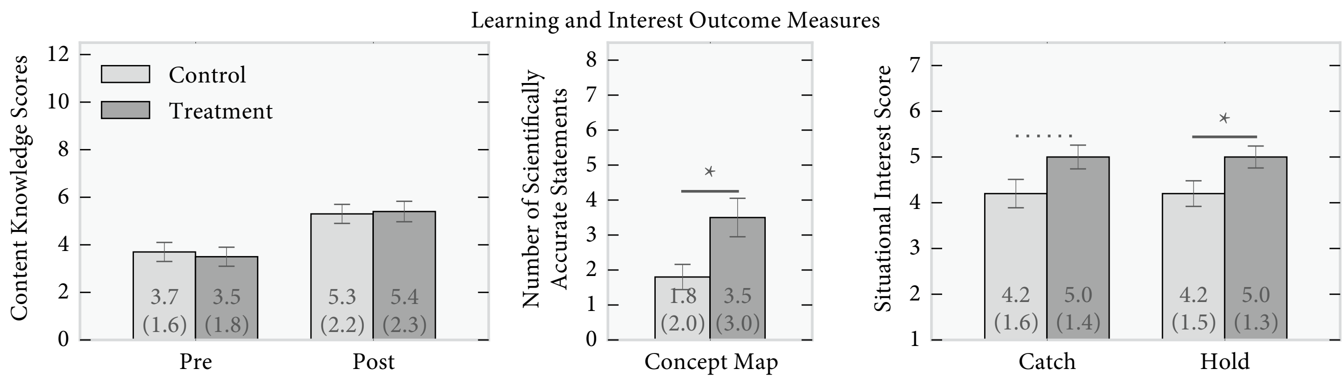

Figure 6: Bar graphs comparing learning and interest outcome measures by condition. Means and (standard deviation) appear in each bar. A solid line with * indicates $p < 0.05$, and a dotted line indicates $p < 0.10$. Whiskers represent standard error.

effect size, which indicates a fairly strong impact. We believe that liking the robot, feeling closer to the robot, and seeing the robot as sociable and attractive all contributed to the feeling of positive social interaction that improve situational interest [7, 49]. This finding is important, because we are able to demonstrate that social behaviors from the robot contribute to situational interest beyond simple effects from novelty or the learning supports of working with a robot in general. We also find the promotion of hold situational interest is particularly important, because this type of situational interest is thought to be related to developing long-term individual interest. This outcome supports the feasibility of this type of social robot to promote individual interest development over a long-term interaction consistent with our prior work that shows engagement with a socially adept robot can be maintained over the course of two weeks[53]. What remains to be seen is how well a socially adept robot can maintain those interactions and continue to support interest and learning in science during long-term interactions.

Finally, we also found that working with the socially adept robot impacted learning outcomes, but the positive impact on learning was only found from comparison of the concept-map scores rather than the more traditional quiz type of learning measures of the post-test. Both groups had similar mean scores on their post-test quizzes that were significant increases when compared to their pre-test scores. These improvements shows augmenting reading a science textbook with an interactive robot may be an effective means for content learning in this context, but further testing is needed to compare learning gains to a child reading on their own. Our finding of similar post-test scores between groups may be because some of the mechanisms for improved learning through social supports are related to reducing cognitive load during learning via distributing cognitive activity and pointing out relevant information during the learning activity [32]. In our study, robots in both conditions provided similar cognitive supports, because we controlled for the amount of verbal interaction and the parts of the textbook reading that were referred to by the robot.

Our finding of higher concept-maps scores for learners in the socially adept condition may suggest that heightened social connections and situational interest were more beneficial in an unstructured measure of learning like the concept-map assessment than they are in a traditional post-test quiz. The post-test assessment was designed to assess content knowledge across a variety of levels of difficulty, but the test structure gave students support for utilizing concepts as these concepts were referred to in other portions of the test. In contrast, the concept-map was an entirely unstructured activity, in which the children were asked to start with a blank paper and recall concepts and organize them according to their connections to each other. This may explain the differences in our measures of learning outcomes, and some contradictory results from prior work on the effect of situational interest on learning [49]. However, this relationship requires further exploration.

5.2 Summary, Limitations, and Future Work

Increasing individual interest and promoting deeper understanding in science are both sorely needed in an age of increasing demand for scientific and technical expertise in the workforce [16]. Our findings provide early evidence that learning companion robots designed to maximize socially adept behaviors and utilize socially situated interest scaffolds help address this problem by improving social connection-making, situational interest and learning in science. Of particular importance, is the relationship between socially adept robot behaviors and increased hold situational interest that contributes to developing individual interest in science. These benefits of working with a social robot, and their ability to perform personalized learning tasks and easily integrate into educational activities may eventually justify the additional cost associated with robotic technology needed to provide social learning supports [5].

For long-term development of individual interest in science, repeated experiences with high levels of *hold situational interest* are crucial [67], and a learning companion robot appears more capable of inducing this type of situational interest during a science homework reading activity when it utilizes socially adept behavior. Integrating socially adept robot interactions as part of a science curriculum over a series of in-home readings appears to be a feasible way of influencing student interest in science while not disrupting the existing curriculum or homework activity. This recommendation has practical significance for curriculum designers and educators, because it presents a model for designing social robots to be integrated into existing science curriculum, especially in the role of supporting children while reading a textbook for homework.

We also make a significant theoretical contribution to the fields of HCI and HRI by demonstrating the impact of socially adept robot behaviors on social connection-making, interest, and learning. We have shown that social aptitude appears to impact children's perception of the robot as sociable and attractive as well as their feelings of closeness and mutual-liking of the robot, but that this aptitude does not necessarily have an impact on how human-like or intelligent the robot appears. This finding is in contrast to other findings in HRI that have found that social behaviors, such as eye gaze aversion while speaking can lead to feeling the robot is more thoughtful [1], and suggests more research in this area is needed to fully flesh out this complicated interaction. We also contribute to theory in interest development and HCI by demonstrating that a robotic agent is effective in playing the role of a social other in interest development, and that the social aptitude of the robot positively impacts both hold and catch situational interest in an activity in similar ways to human-human interactions. This is an important finding, because it supports the continued exploration of social robots as learning companions to benefit both learning and interest development in academic content.

There are several limitations of the study that need to be addressed to further explore social robots as learning companions. First, this study represents successful outcomes for social connection-making, situational interest, and learning, but the interaction with the robot was confined to a controlled laboratory setting during one short reading session. This short-term design allows us to assess the feasibility of the robot to positively impact learning, but limits our ability to project these effects into long-term interactions. A better understanding of learning with the robot in real environments will require long-term *in situ* testing, which will allow for examination of the stability of the social connections and learning, and the development of individual interest. Other studies might include a delayed post-test as a measure of knowledge retention that prior work suggests benefits from social learning [37]. A second limitation is that the sample size of our study is somewhat low for conducting a randomized controlled experiment and leaves some of our analysis underpowered. We believe the medium to large effect sizes found are indicators of the differences between socially adept versus socially neutral robot behaviors, but future work should increase the sample to increase generalizability. We also suggest future work increase the sample size to include additional comparison groups with children reading by themselves or with another technology, such as a tablet computer, so that the learning gains found could be compared against other methods of learning. We would also suggest studies with a larger sample examine the differences in outcomes based on existing interest in science or in reading ability that our study was too underpowered to explore. There are also limitations in interpreting the results of our learning measures. It is not entirely clear why working with the socially adept robot would promote higher concept-map scores but not post-test scores. In this paper, we suggest this may be due to the differences in structure of these assessments and that the concept-map activity may represent a deeper type of understanding, but this conclusion requires further validation. This complex result may also be related to the order that these assessments were administered. Each child completed the concept-map before the post-test, and this order may have induced a priming effect on the post-test scores. Again, further study that balances the order of the learning measures are needed to better interpret our results. Finally, the physical and programming design of the robot will need to be improved to better realize it's potential as a learning companion, but these limitations are common in this field [5]. These improvements include better methods of speech recognition for children that would allow for more natural verbal interactions between the child and robot, and methods of expanding a robot's capacity for evaluating and utilizing information about the environment and the learner's social and emotional states. We intend to address many of these limitations in our future work.

In summary, based on our work, we believe that a learning companion robot is well suited to successfully augment traditional science learning activities, such as reading informational texts as homework, and that equipping these robots with socially adept behaviors provides additional benefits for social connection-making, situational interest, and learning.

6 SELECTION AND PARTICIPATION OF CHILDREN

To identify participants for this study, we recruited families with children, aged 10–12, from the local community through email recruitment using university mass email services, and through posting recruitment flyers on and near campus. Children's parents gave written informed consent and all children assented to participate prior to the start of the study. Parents and children were informed that researchers may use and share data, including video, images, and transcribed speech, in publications or presentations for academic purposes with all identifiable information in transcribed speech removed. All children had the opportunity to interact with our learning companion robot. The protocol was approved by the University's Internal Review Board.

ACKNOWLEDGMENTS

We thank Zhi Tan for the development of an initial version of our robot platform, Murad Jaber, Songwong Tasneeyapant, Tianyi Shan, Tavishi Gupta, and Thana Pongcharoenyong for their contributions to further development, and Nicole Jagemann and Matt Kaufman for their research assistance. This work was supported by the National Science Foundation under award 1149970 and the University of Wisconsin–Madison Education Graduate Research Scholars Fellowship.

REFERENCES

[1] Sean Andrist, Xiang Zhi Tan, Michael Gleicher, and Bilge Mutlu. 2014. Conversational Gaze Aversion for Humanlike Robots. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-robot Interaction (HRI '14)*. ACM, New York, NY, USA, 25–32. https://doi.org/10.1145/2559636.2559666

[2] Flavio S. Azevedo. 2015. Sustaining interest-based participation in science. In *Interest in Mathematics and Science Learning*, K. Ann Renninger, Martina Nieswandt, and Suzanne Hidi (Eds.). American Educational Research Association, Washington D.C., 281–297.

[3] Wilma A. Bainbridge, Justin W. Hart, Elizabeth S. Kim, and Brian Scassellati. 2011. The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents. *Int J of Soc Robotics* 3, 1 (Jan. 2011), 41–52. https://doi.org/10.1007/s12369-010-0082-7

[4] Albert Bandura. 1977. *Social learning theory.* Prentice Hall, Englewood Cliffs, NJ.

[5] Tony Belpaeme, James Kennedy, Adiri Ramachandran, Brian Scassellati, and Fumihide Tanaka. 2018. Social robots for education: A review. *Science Robotics* 2, 3 (2018), eaat5954.

[6] Fabiane B. V. Benitti. 2012. Exploring the educational potential of robotics in schools: A systematic review. *Computers & Education* 58, 3 (April 2012), 978–988. https://doi.org/10.1016/j.compedu.2011.10.006

[7] David A. Bergin. 2016. Social Influences on Interest. *Educational Psychologist* 51, 1 (Jan. 2016), 7–22. https://doi.org/10.1080/00461520.2015.1133306

[8] Maria Blancas, Vasiliki Vouloutsi, Klaudia Grechuta, and Paul F. M. J. Verschure. 2015. Effects of the Robot's Role on Human-Robot Interaction in an Educational Scenario. In *Biomimetic and Biohybrid Systems (Lecture Notes in Computer Science)*, Stuart P. Wilson, Paul F.M.J. Verschure, Anna Mura, and Tony J. Prescott (Eds.). Springer International Publishing, 391–402.

[9] Benjamin S. Bloom and I. De V. Heyns. 1956. Development and Applications of Tests of Educational Achievement. *Review of Educational Research* 26, 1 (Feb. 1956), 72–88. https://doi.org/10.3102/00346543026001072

[10] John D. Bransford, Brigid Barron, R. D. Pea, Andrew Meltzoff, Pat Kuhl, Philip Bell, Reed Stevens, Daniel Schwartz, Nancy Vye, Byron Reeves, Jeremy Roschelle, and Nora Sabelli. 2005. Foundations and opportunities for an interdisciplinary science of learning. In *The Cambridge Handbook of the Learning Sciences* (1 ed.), R. Keith Sawyer (Ed.). Cambridge University Press, 39–77.

[11] LaVonda Brown, Ryan Kerwin, and Ayanna M. Howard. 2013. Applying Behavioral Strategies for Student Engagement Using a Robotic Educational Agent. In *IEEE International Conference on Systems, Man, and Cybernetics*. Manchester, UK. https://doi.org/10.1109/SMC.2013.744

[12] Douglas Clark and Marcia C. Linn. 2003. Designing for Knowledge Integration: The Impact of Instructional Time. *Journal of the Learning Sciences* 12, 4 (2003), 451–493.

[13] Jacob Cohen. 1988. *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates, Hillsdale, N.J. OCLC: 17877467.

[14] Alexandre Coninx, Paul Baxter, Elettra Oleari, Sara Bellini, Bert Bierman, Olivier Blanson Henkemans, Lola Cañamero, Piero Cosi, Valentin Enescu, Raquel Ros Espinoza, Antoine Hiolle, Rémi Humbert, Bernd Kiefer, Ivana Kruijff-Korbayová, Rose-marijn Looije, Marco Mosconi, Mark Neerincx, Giulio Paci, Georgios Patsis, Clara Pozzi, Francesca Sacchitelli, Hichem Sahli, Alberto Sanna, Giacomo Sommavilla, Fabio Tesser, Yiannis Demiris, and Tony Belpaeme. 2016. Towards Long-term Social Child-robot Interaction: Using Multi-activity Switching to Engage Young Users. *J. Hum.-Robot Interact.* 5, 1 (March 2016), 32–67. https://doi.org/10.5898/JHRI.5.1.Coninx

[15] Harris Cooper, Saiying Steenbergen-Hu, and Amy L. Dent. 2012. Homework. In *APA educational psychology handbook, Vol 3: Application to learning and teaching*, Karen R. Harris, Steve Graham, Tim Urdan, Adriana G. Bus, Sonya Major, and H. Lee Swanson (Eds.). Vol. 3. APA, 475–495.

[16] National Research Council. 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* The National Academies Press, Washington, DC.

[17] Linda Crocker and James Algina. 2006. *Introduction to Classical and Modern Test Theory.* Cengage Learning, Mason, OH.

[18] Edward L. Deci and Richard M. Ryan. 1990. A Motivational Approach to Self: Intergration in Personality. Lincoln, NE, 237–288.

[19] Pierre Dillenbourg. 1996. Distributing cognition over humans and machines. In *International Perspectives on the Design of Technology-supported Learning Environments*. Routledge, New York, NY, 165–182.

[20] Brian R. Duffy. 2003. Anthropomorphism and the social robot. *Robotics and Autonomous Systems* 42, 3 (March 2003), 177–190. https://doi.org/10.1016/S0921-8890(02)00374-3

[21] Amanda M. Durik and Judith M. Harackiewicz. 2007. Different strokes for different folks: How individual interest moderates the effects of situational factors on task interest. *Journal of Educational Psychology* 99, 3 (2007), 597–610. https://doi.org/10.1037/0022-0663.99.3.597

[22] Huiyong Fan, Jianzhong Xu, Zhihui Cai, Jinbo He, and Xitao Fan. 2017. Homework and students' achievement in math and science: A 30-year meta-analysis, 1986-2015. *Educational Research Review* 20 (Feb. 2017), 35–54. https://doi.org/10.1016/j.edurev.2016.11.003

[23] Ylva Fernaeus, Maria Håkansson, Mattias Jacobsson, and Sara Ljungblad. 2010. How Do You Play with a Robotic Toy Animal?: A Long-term Study of Pleo. In *Proceedings of the 9th International Conference on Interaction Design and Children (IDC '10)*. ACM, New York, NY, USA, 39–48. https://doi.org/10.1145/1810543.1810549

[24] Julia Fink. 2012. Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction. In *Social Robotics (Lecture Notes in Computer Science)*, Shuzhi Sam Ge, Oussama Khatib, John-John Cabibihan, Reid Simmons, and Mary-Anne Williams (Eds.). Springer Berlin Heidelberg, 199–208.

[25] Simon Gächter, Chris Starmer, and Fabio Tufano. 2015. Measuring the Closeness of Relationships: A Comprehensive Evaluation of the 'Inclusion of the Other in the Self' Scale. *PLOS ONE* 10, 6 (2015), e0129478.

[26] Goren Gordon, Cynthia Breazeal, and Susan Engel. 2015. Can Children Catch Curiosity from a Social Robot?. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '15)*. ACM, New York, NY, USA, 91–98. https://doi.org/10.1145/2696454.2696469

[27] Maartje M. A. de Graaf, Somaya Ben Allouch, and Jan A. G. M. van Dijk. 2016. Long-term evaluation of a social robot in real homes. *Interaction Studies* 17, 3 (Jan. 2016), 462–491. https://doi.org/10.1075/is.17.3.08deg

[28] John T. Guthrie and Susan L. Klauda. 2012. Making textbook reading meaningful. *Educational Leadership* 69, 6 (2012), 64–68.

[29] Judith M. Harackiewicz, Jessie L. Smith, and Stacy J. Priniski. 2016. Interest Matters: The Importance of Promoting Interest in Education. *Policy Insights from the Behavioral and Brain Sciences* 3, 2 (2016), 220–227. https://doi.org/10.1177/2372732216655542

[30] Suzanne Hidi and Judith M. Harackiewicz. 2000. Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research* 70, 2 (2000), 151–179. https://doi.org/10.3102/00346543070002151

[31] S. Hidi and K. A. Renninger. 2006. The four-phase model of interest development. *Educational Psychologist* 41, 2 (2006), 111–127. https://doi.org/10.1207/s15326985ep4102_4

[32] James Hollan, Edwin Hutchins, and David Kirsh. 2000. Distributed Cognition: Toward a New Foundation for Human-computer Interaction Research. *ACM Trans. Comput.-Hum. Interact.* 7, 2 (June 2000), 174–196. https://doi.org/10.1145/353485.353487

[33] Deanna Hood, Severin Lemaignan, and Pierre Dillenbourg. 2015. When children teach a robot to write: An autonomous teachable humanoid which uses simulated handwriting. In *HRI '15*. Portland, OR, 83–90.

[34] Chris S. Hulleman, Dustin B. Thoman, Anna-Lena Dicke, and Judith M. Harackiewicz. 2017. The Promotion and Development of Interest: The Importance of Perceived Values. In *The Science of Interest*, Paul A. O'Keefe and Judith M. Harackiewicz (Eds.). Springer International Publishing, Cham, 189–208. https:

//doi.org/10.1007/978-3-319-55509-6_10
[35] Matthew C. Jackson, Christina C. Leal, Jeanette Zambrano, and Dustin B. Thoman. 2018. Talking about science interests: the importance of social recognition when students talk about their interests in STEM. *Soc Psychol Educ* (Oct. 2018). https://doi.org/10.1007/s11218-018-9469-3
[36] Sanna Järvelä and K. Ann Renninger. 2014. Designing for learning: Interest, motivation, and engagement. In *The Cambridge Handbook of The Learning Sciences* (2 ed.), R. Keith Sawyer (Ed.). Cambridge University Press, New York, NY, 358–376.
[37] David W. Johnson and Roger T. Johnson. 2009. An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning. *Educational Researcher* 38, 5 (June 2009), 365–379.
[38] W. Lewis Johnson and James C. Lester. 2016. Face-to-Face Interaction with Pedagogical Agents, Twenty Years Later. *Int J Artif Intell Educ* 26, 1 (March 2016), 25–36. https://doi.org/10.1007/s40593-015-0065-9
[39] Takayuki Kanda, Michihiro Shimada, and Satoshi Koizumi. 2012. Children learning with a social robot. In *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 351–358. https://doi.org/10.1145/2157689.2157809
[40] James Kennedy, Paul Baxter, and Tony Belpaeme. 2015. Comparing Robot Embodiments in a Guided Discovery Learning Interaction with Children. *Int J of Soc Robotics* 7, 2 (April 2015), 293–308. https://doi.org/10.1007/s12369-014-0277-4
[41] Geoffrey Keppel and Thomas D. Wickens. 2004. *Design and Analysis: A Researcher's Handbook, 4th Edition*. Pearson, Needham, MA.
[42] Yanghee Kim. 2013. Digital Peers to Help Children's Text Comprehension and Perceptions. *Journal of Educational Technology & Society* 16, 4 (2013), 59–70. http://www.jstor.org/stable/jeductechsoci.16.4.59
[43] Maximillian Knogler, Judith M. Harackiewicz, Andreas Gegenfurtner, and Doris Lewalter. 2015. How situational is situational interest? Investigating the longitudinal structure of situational interest. *Contemporary Educational Psychology* 43 (2015), 39–50. https://doi.org/10.1016/j.cedpsych.2015.08.004
[44] Jacqueline M. Kory-Westlund, Sooyeon Jeong, Hae W. Park, Samuel Ronfard, Aradhana Adhikari, Paul L. Harris, David DeSteno, and Cynthia L. Breazeal. 2017. Flat vs. Expressive Storytelling: Young Children's Learning and Retention of a Social Robot's Narrative. *Front. Hum. Neurosci.* 11 (2017). https://doi.org/10.3389/fnhum.2017.00295
[45] Iolanda Leite, Carlos Martinho, and Ana Paiva. 2013. Social Robots for Long-Term Interaction: A Survey. *Int J of Soc Robotics* 5, 2 (April 2013), 291–308. https://doi.org/10.1007/s12369-013-0178-y
[46] Iolanda Leite, Carlos Martinho, Andre Pereira, and Ana Paiva. 2009. As Time goes by: Long-term evaluation of social presence in robotic companions. In *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*. 669–674. https://doi.org/10.1109/ROMAN.2009.5326256
[47] Karina R. Liles, De'Aira G. Bryant, and Jenay M. Beer. 2017. How Can Social Robots Motivate Students to Practice Math?. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17)*. ACM, New York, NY, USA. https://doi.org/10.1145/3029798.3034814
[48] Gi H. Lim, Seung W. Hong, Inhee Lee, Il H. Suh, and Michael Beetz. 2013. Robot recommender system using affection-based episode ontology for personalization. 155–160. https://doi.org/10.1109/ROMAN.2013.6628437
[49] Lisa Linnenbrink-Garcia, Erika A. Patall, and Emily E. Messersmith. 2013. Antecedents and consequences of situational interest. *British Journal of Educational Psychology* 83, 4 (2013), 591–614. https://doi.org/10.1111/j.2044-8279.2012.02080.x
[50] Nichola Lubold, Erin Walker, and Heather Pon-Barry. 2016. Effects of voice-adaptation and social dialogue on perceptions of a robotic learning companion. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 255–262. https://doi.org/10.1109/HRI.2016.7451760
[51] Joseph E. Michaelis and Bilge Mutlu. 2017. Someone to Read with: Design of and Experiences with an In-Home Learning Companion Robot for Reading. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 301–312. https://doi.org/10.1145/3025453.3025499
[52] Joseph E. Michaelis and Bilge Mutlu. 2017. Someone to Read with: Design of and Experiences with an In-Home Learning Companion Robot for Reading. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 301–312.
[53] Joseph E. Michaelis and Bilge Mutlu. 2018. Reading socially: Transforming the in-home reading experience with a learning-companion robot. *Science Robotics* 3, 21 (2018), eaat5999. https://doi.org/10.1126/scirobotics.aat5999
[54] Joseph. E. Michaelis and Mitchell J. Nathan. 2015. The Four-Phase Interest Development in Engineering Survey. Seattle, WA.
[55] Mathew Mitchell. 1993. Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology* 85, 3 (1993), 424–436. https://doi.org/10.1037/0022-0663.85.3.424
[56] Bilge Mutlu, Takayuki Kanda, Jodi Forlizzi, Jessica Hodgins, and Hiroshi Ishiguro. 2012. Conversational Gaze Mechanisms for Humanlike Robots. *ACM Trans. Interact. Intell. Syst.* 1, 2 (Jan. 2012), 12:1–12:33. https://doi.org/10.1145/2070719.2070725
[57] Bilge Mutlu, Steven Osman, Jodi Forlizzi, Jessica Hodgins, and Sara B. Kiesler. 2006. Task Structure and User Attributes as Elements of Human-Robot Interaction Design. https://www.semanticscholar.org/paper/Task-Structure-and-User-Attributes-as-Elements-of-Mutlu-Osman/0f191b59ac8e5ae3b09db2aafd57159ebc1f864c
[58] Piia Näykki and Sanna Järvelä. 2008. How Pictorial Knowledge Representations Mediate Collaborative Knowledge Construction In Groups. *Journal of Research on Technology in Education* 40, 3 (2008), 359–387.
[59] Susan Bobbitt Nolen. 2007. Young Children's Motivation to Read and Write: Development in Social Contexts. *Cognition and Instruction* 25, 2 (2007), 219–270.
[60] Michael J. Padilla, Martha Cyr, Ioannis Miaoulis, and T. Griffith Jones. 2005. *Motion, Forces, and Energy*. Pearson Prentice Hall, Needham, MA.
[61] David H. Palmer. 2009. Student Interest Generated During an Inquiry Skills Lesson. *Journal of Research in Science Teaching* 46, 2 (2009), 147–165. https://doi.org/10.1002/tea.20263
[62] Salvatore Parise, Sara B. Kiesler, Lee S. Sproull, and Keith Waters. 1999. Cooperating with life-like interface agents. *Computers in Human Behavior* 15, 2 (March 1999), 123–142. https://doi.org/10.1016/S0747-5632(98)00035-1
[63] William R. Penuel and Lawrence P. Gallagher. 2009. Preparing Teachers to Design Instruction for Deep Understanding in Middle School Earth Science. *Journal of the Learning Sciences* 18, 4 (2009), 461–508.
[64] Rich Radcliffe, David Caverly, Cynthia Peterson, and Matt Emmons. 2004. Improving textbook reading in a middle school science classroom. *Reading Improvement* 41, 3 (2004), 145–156. http://link.galegroup.com/apps/doc/A122922868/AONE?sid=googlescholar
[65] K. Ann Renninger and Suzanne Hidi. 2016. *The power of interest for motivation and engagement*. Routledge, New York, NY.
[66] K. Ann Renninger, Yanyan Ren, and Heidi M. Kern. 2018. Motivation, Engagement, and Interest. In *International Handbook of the Learning Sciences*, Frank Fischer, Cindy E. Hmelo-Silver, Susan R. Goldman, and Peter Reimann (Eds.). Routledge, New York, NY, 116–126.
[67] K. Ann Renninger and Stephanie Su. 2012. Interest and its development. In *Oxford Handbook of Human Motivation*, Richard M. Ryan (Ed.). Oxford University Press, 167–187.
[68] William R. Robinson. 1999. A View from the Science Education Research Literature: Concept Map Assessment of Classroom Learning. *J. Chem. Educ.* 76, 9 (Sept. 1999), 1179. https://doi.org/10.1021/ed076p1179
[69] Martin Saerbeck, Tom Schut, Christoph Bartneck, and Maddy D. Janse. 2010. Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. ACM Press, Atlanta, Georgia, USA, 1613. https://doi.org/10.1145/1753326.1753567
[70] R. Keith Sawyer. 2014. The new science of learning. In *The Cambridge Handbook of the Learning Sciences* (2 ed.), R. Keith Sawyer (Ed.). Cambridge University Press, New York, NY, 1–18.
[71] Noah L. Schroeder, Olusola O. Adesope, and Rachel Barouch Gilbert. 2013. *How Effective Are Pedagogical Agents for Learning? A Meta-Analytic Review.*
[72] Masahiro Shiomi, Takayuki Kanda, Iris Howley, Kotaro Hayashi, and Norihiro Hagita. 2015. Can a Social Robot Stimulate Science Curiosity in Classrooms? *Int J of Soc Robotics* 7, 5 (Nov. 2015), 641–652. https://doi.org/10.1007/s12369-015-0303-1
[73] Trish Stoddart, Robert Abrams, Erika Gasper, and Dana Canaday. 2000. Concept maps as assessment in science inquiry learning - a report of methodology. *International Journal of Science Education* 22, 12 (Dec. 2000), 1221–1246. https://doi.org/10.1080/095006900750036235
[74] Fumihide Tanaka and Shizuko Matsuzoe. 2012. Children Teach a Care-Receiving Robot to Promote Their Learning: Field Experiments in a Classroom for Vocabulary Learning. *Journal of Human-Robot Interaction* 1, 1 (2012), 78–95.
[75] Adam Toon. 2014. Friends at last? Distributed cognition and the cognitive/social divide. *Philosophical Psychology* 27, 1 (2014), 112–125. https://doi.org/10.1080/09515089.2013.828371
[76] Lev Vygotsky. 1978. *Mind in society: The development of higher psychological processes*. Harvard University Press, Cambridge, MA.
[77] Allan Wigfield and John T. Guthrie. 1997. Relations of Children's Motivation for Reading to the Amount and Breadth of Their Reading. *Journal of Educational Psychology* 89, 3 (1997), 420–32.
[78] Jeanette Zambrano, Gary Lee, Christina C. Leal, and Dustin B. Thoman. 2018. Helping others is interesting! Developing and sustaining science interest with prosocial utility value in textbooks.. In *AERA 2018*. New York.