Can Children Catch Curiosity from a Social Robot?

Goren Gordon
Personal Robots Group, MIT
Media Lab
20 Ames Street E15-468
Cambridge, MA 02139
goren@gorengordon.com

Cynthia Breazeal
Personal Robots Group, MIT
Media Lab
20 Ames Street E15-468
Cambridge, MA 02139
cynthiab@media.mit.edu

Susan Engel
Department of Psychology
Bronfman Science Center
Williams College
Williamstown MA 01267
Susan.engel@williams.edu

ABSTRACT

Curiosity is key to learning, yet school children show wide variability in their eagerness to acquire information. Recent research suggests that other people have a strong influence on children's exploratory behavior. Would a curious robot elicit children's exploration and the desire to find out new things? In order to answer this question we designed a novel experimental paradigm in which a child plays an education tablet app with an autonomous social robot, which is portrayed as a younger peer. We manipulated the robot's behavior to be either curiosity-driven or not and measured the child's curiosity after the interaction. We show that some of the child's curiosity measures are significantly higher after interacting with a curious robot, compared to a non-curious one, while others do not. These results suggest that interacting with an autonomous social curious robot can selectively guide and promote children's curiosity.

Categories and Subject Descriptors

K.3.1 [**Computers and Education**]: Computer Uses in Education; I.2.9 [**Artificial Intelligence**]: Robotics

Keywords

children education, autonomous robot behavior, dragonbot

1. INTRODUCTION

Curiosity is the basic drive to ask questions and to better understand events. Even in infancy and early childhood, curiosity enables young learners to acquire evidence and develop models of how the world works [19]. As children get older and enter school, intrinsic curiosity is still the main drive for efficient learning even with great teachers [14, 6]. The question then arises: what influences the basic curiosity drive in young children? One relatively unmapped influence is social interaction, i.e. how interactions with other individuals, be they more curious or less, influence the internal motivation to learn [7].

HRI'15, March 2–5, 2015, Portland, Oregon, USA.

It has been shown that social cues are of paramount importance in language learning in children [17], as well as changing their mindset [4] and consequently their academic achievements [3]. Recursively, a decrease in intrinsic motivation changes the mindset into a fixed one [9]. Furthermore, during late childhood, adolescence and adulthood, the curiosity drive declines [6, 7]. Though previous research has identified some of the ways in which adults encourage or discourage children's expressions of curiosity [12], very little research has examined the effects of peers on a child's expressions of curiosity. We thus wish to explore whether curiosity in young children can be manipulated and increased by robotic peer interaction.

Social robots have been used in recent years as educational companions to children, teaching them new vocabulary, math concepts and social skills. It has been shown that children can treat robots as informants [18], but how will a child react to a robot's curiosity driven behavior? Can a child "catch" curiosity from a curious social robot? In other words, can interaction with a curious robot promote children's curiosity?

To address this question, we performed an exploratory study in which we manipulated the behavior of an autonomous social robot interacting with children in an educational setting. The interaction involved a novel Story-maker app co-played on a tablet by the child and robot, wherein the child manipulates characters on the tablet and the robot tells an appropriate story. The story is also written on the tablet, thus promoting reading skills. Furthermore, the robot is portrayed as a younger peer that tries to learn to read, prompting the child to teach it new words.

During the interaction, the robot is either curious or not, where in our experimental paradigm a curious robot behaves with enthusiasm about learning and exploration, challenges the child and suggests novel moves on the tablet app. The non-curious robot plays with the child, asks her to show it words, yet does not express any overt or explicit desire to learn new things. We quantified children's curiosity after the interaction via three different measures: free exploration, question generation and uncertainty seeking tasks. We show that free exploration and uncertainty seeking are significantly higher after interacting with a curious robot compared to a non-curious one, whereas question generation is unaffected by the manipulation. These results confirm that at-least some aspects of children's curiosity can be increased by interacting with an autonomous social robot.

2. RELATED WORK

Children's curiosity has been studied using different quantifiable measures, which we adapted to our study. **Free exploration:** One measure relates to actively seeking information by opening novel boxes, quantifying different aspects of the behavior, e.g. approach time to the box, number of different boxes opened [11]. We have developed a novel digital version of this free exploration measure, to be used in tablet-related interactions. **Fish task:** We used a recently developed uncertainty seeking tablet app, called "The Fish Task" [13], which addresses children's desire to choose options containing unknown, probabilistic results more than certain and deterministic ones. **Question generation:** We used an established and more qualitatively-derived measure, namely, the question generation task in which the child is prompted to ask as many questions about a topic, without providing answers [10]. The latter condition is imperative, since it was shown that answering the question generates conversation irrespective of the intrinsic motivation to know [5]. We used all three measures to address the question of whether there are different aspects of curiosity [15] and whether they can be manipulated by a social robot.

Children's change in curiosity has been investigated in previous studies. In Ref. [12], the presence of an adult and the context of her behavior was shown to influence a child's free exploration behavior. A series of studies [19] have shown that infants explore more if their prior beliefs are violated, i.e. if they see evidence that contradicts their expectations. Furthermore, it was shown, through manipulation of child-toy-experimenter interaction that if evidence fails to distinguish among competing beliefs, infants explore more to disambiguate their beliefs [19]. In another study, the effects of personal curiosity traits and the school environment on academic achievements have been shown to be complex, namely, high curiosity children in challenging schools had the highest performance, whereas high curiosity children in non-challenging school had the lowest [14]. We ask whether a robotic peer, as opposed to a parent, experimenter or school environment, can change a child's curiosity.

Social robots have been used previously to teach children new material. In [18], RUBI-4, a humanoid robot with a tablet embedded in its midsection, played simple vocabulary games with preschool children. In [21] the experimenter asked either the preschool child or the robot to act out novel verbs.They found that teaching the robot helped children remember the verbs, as well as inspiring further teaching-verbs play. In contrast to these studies, we wish to not only make the children learn new things, but promote their drive to learn and explore, i.e. increase their inherent curiosity.

3. EXPERIMENTAL SETUP

The experimental setup is composed of the robot, tablet, cameras and microphones, Fig. 1(left). The tablet had three apps that were used during the experiment, namely, the main Story-maker app and two curiosity-assessment apps: Free exploration app and the Fish task app [13]. All taps and interactions with the tablet were recorded.

3.1 Robotic platform

For the social robotic platform we used Dragonbot [20], a squash-and-stretch Android smartphone based robot. The facial expression, sound generation and part of the logic is generated on the smartphone, which is mounted on the face of the robot. The robot appears to be a soft, furry, fanciful creature that is designed to engage children. Dragonbot is a very expressive platform and has a large repertoire of possible facial expressions and actions (indicated in *italics* below). We installed a commercial child-like voice for the text-to-speech software on the smartphone, to facilitate a more generic and engaging interaction. The robot was autonomous and was not controlled by a remote operator. It reacted to the child's interaction with the tablet.

Figure 1: Experimental setup (left) and screenshot from the Story-maker app (right)

3.2 Story-maker app

The main tablet app was a novel Story-maker app we developed for this study, which enables the child to co-create a story with the robot, Fig. 1(right). The game contains several characters that the child can move. After each movement, a sentence is automatically generated using a novel auto-generation mechanism, which (i) randomly selects an adjective for the character; (ii) detects the closest other character for the story interaction; (iii) follows an xml script of the plot of the story; and (iv) uses an open-source natural language generation library [8] to construct a full sentence.

The xml plot files are constructed in a generic fashion, such that (i) each character has a list of possible adjectives (e.g. red, big), motions (e.g. fly, jump) and speech (e.g. roar, squeak); (ii) the plot line is constructed of a sequence of movements, speech, feelings and resolutions; and (iii) the story conversation is constructed such that any sequence of character selection generates a coherent story line. The result of each movement is thus a full sentence that describes the progression of the story plot. After several such sentences, the scene changes and new characters are introduced, while some of the old ones are taken away. There are three scenes to the story, which ends with a final resolution sentence.

This is a sample text from a generated story: *Butterfly whispers to bird: What happened to dragon? The purple dragon flies to the bird. Dragon roars to bird: I lost my ball. Dragon says to the wild butterfly: Can you help me find my ball? Butterfly whispers: I think I saw the ball near the pine tree.*

This app was designed to promote the child's feeling of control over the interaction, in the sense that the story told correlated to the child's actions. Furthermore, the pace of the story was dictated by the child. Several features were integrated into this app to maintain engagement and interest, important aspects that affect curiosity [1]. The first was the selection of the story's protagonist. The child could select out of two options, namely, a dragon or a bird. The story plot followed that selection throughout. The second feature was the introduction of multiple scenes, each with its unique characters, wherein only the protagonist and its sidekick moved from one scene to the next. This introduced variability and novelty in the story. The third was an insertion of a "silly' sentence into the story, e.g. "the dragon burps". This kept the child more engaged in the story. The fourth feature was the variability of language complexity throughout the story. More complex words were deliberately introduced, e.g. "spacious", "anxious", to keep even older children engaged. The fifth feature was the introduction of an antagonist, i.e. a character who is detrimental to the main plot.

3.3 Subjects

71 subjects participated in the study. 7 of the subjects did not speak or cooperate during the task and 1 had a technical difficulty and were thus excluded. Out of the remaining 63, only 48 (21 female, 27 male) completed the initial assessment and the task and were included in the analysis. Subjects were randomized across conditions and ages, but analysis and exclusion occurred after the study was finished, resulting in slightly different number of subjects across conditions. The average age was 6 (1.23 standard-deviation), with children ages 3.4-8.4 years old, Fig. 2(inset).

The subjects were recruited from a compiled mailing list of local social media family groups. 13 of the subjects previously interacted with the same robotic platform in an earlier study in our lab. These were evenly distributed across conditions.

4. PROTOCOL

Initial assessment: reading skills. During the introduction to the study, the child is asked to spell her name, and is informed that she is going to play word games with the experimenter and then with the robot. We informed the participants that if they are bored or do not wish to continue, they can stop the interaction whenever they wanted. The first "word game" is the TOWRE word assessment test [22], in which the experimenter asks the child to read words from lists, as fast as she could, for 45 seconds. The raw TOWRE score is defined as the total number of correctly read words during these 45 seconds. We administered both sight and phonetic word lists, where the total raw score is the sum of the two tests' raw score.

Robot introduction. The child then sits next to a small table upon which there is a tablet and the robot, Fig. 1(left). The robot is "sleeping", i.e. its eyes are closed, and it is introduced as "Parle, a young robot that just learned how to speak and wants to learn to read". This was said since the text-to-speech module used was sometimes unclear to the children, and we wished them to feel comfortable to ask when they did not understand it. The robot awakens, *yawns* (an overt motion and sound), and introduces itself: "I am Parle, we are going to play word games together." It then makes a *shy* facial expression.

Reading pre-test. The next phase of interaction is a pre-test, during which the robot asks the child to teach it some words. It verbally asks the child to show it a word, e.g. "dragon", whereupon the word, and four distractors appear on the tablet. The child then needs to tap on the correct word. The words asked are selected from the entire vocabulary, according to a novel expected information-gain algorithm that attempts to maximize the knowledge the robot has on the child's reading skill. The algorithm is based on Bayesian updates of a vocabulary of words, based on the child's answers of the prompted questions and is reported in detail in another publication under review. The four distractors are also selected from the same vocabulary: two words which are most similar; one word that the child should know, according to the assessment algorithm; and one word that the child should not know how to read. This is repeated ten times, to get a thorough assessment of the child's reading knowledge.

Child-robot co-play. The main phase of the interaction is based on the Story-maker app described above. The robot asks the child "do you want to play with me and create our own story?" The child needs to tap on a "yes" or a "no". If the child taps "no", the robot makes a *shy* face and asks again "do you want to play with me?" (The robot does not take no for an answer). When the child taps "yes" the robot *laughs*, says "that's great" and prompts the child to select the protagonist of the story. After it is selected, the robot instructs the child how to play: "you move the characters around and I will tell the story. The game will help us if we have trouble reading." The tablet then speaks, in a different voice, "move one of the characters. A sentence will be written on top." This creates a clear separation between robot and tablet, in the sense that the robot plays with the child and the tablet informs them about the written words and sentences.

During the child-tablet-robot interaction, when the child moves a character, the autonomous robot first speaks the generated sentence, and then the sentence appears on the tablet above the scene. In 50% of the sentences, the robot asks the child to show it a word, e.g. "I don't know how to read the word dragon. Can you show it to me?". This resulted in an average of 11 words per interaction. In the first two questions, the robot also says "look in the sentence above the colorful picture" in order to direct the child where to look. The child is then required to tap on the correct word. Each tapping on a word on the tablet results in the tablet speaking that word. In this sense, the tablet is an informant, whereas the child and robot are both the students. If the child is correct, the robot says "yes" in an excited voice, thanks the child and the story continues. If the child is wrong, the robot expresses *frustration* and asks the word again. If the child is wrong again, the tablet shows the correct word in an emphasized manner and speaks it. If the child moves a character instead of tapping the word, the robot makes a *thinking* expression and says "ok, let's continue". Additionally, at the beginning of a new scene, the robot says "move a character to hear the next sentence" so as to direct the child what to do in this new scene. The game continues until the end of the story, when the robot says: "The End. That was a great story."

Reading post-test. In the last stage of the interaction, the post-test, the robot again asks the child to teach it some words, similar to the pre-test phase. During this phase, the words that were asked during the story phase are asked again, starting with the incorrectly identified words, then the correctly identified and finally random words. A total of ten words are asked during this phase.

Robot behavior. In order to increase believability and engagement with the autonomous robot, we inserted randomness to the expressions and sentences the robot asked, so as to avoid boring repetition. During the pre- and post-test phases, the robot asked: "Can you show me the word X?", "That is a new word, X. Can you tap on it?", "I don't know the word X. Can you show it to me." This increased diversity and randomness in the robot's behavior is suggested to be essential for the children's engagement, a major factor in educational interactions [1].

During the story phase, the robot made some silly comments, to increase plausibility of it being a younger peer. During the game, the robot followed 10% of the child's movements on the tablet with "that motion is funny. Maybe the story should go like this." or "I'm not sure what to say now. Let me try something." In 10% of the scene change, the robot says "I might surprise you with some silliness". Additionally, during the silly sentence the Story-maker app generated, the robot *laughed* and said "this is silly", so as to be part of the game and interaction.

When the child pressed the correct word and the tablet also spoke it, the robot said "it helps me that when you touch a word the game speaks it" or "did you hear the game speak the word?" This was designed to encourage the child to press on more words in order to hear them.

Interaction end. The robot interaction ended with the robot yawning and saying: "I am tired now. I think I will go to sleep. It was great playing with you. I hope we can do it again sometimes. See you. Bye bye" The robot's eyes closed and it did not move anymore. This part was a clear delineation between the robot interaction and the final assessment, which we required the child to perform alone. All the children accepted this termination, e.g. none of them asked "why is it sleeping?", or "can we wake it up?".

Post-interaction assessments. After the robot went to sleep, we administered the Free exploration task, during which the child played the task and we questioned the parent about the interaction: "Do you think your child learned something new from this interaction?" When the Free exploration task was over, we questioned the child about the interaction and the robot. At the end of the questionnaire, we administered the question generation task, followed by the Fish task.

5. CONDITIONS

There were three conditions to the study, namely, curious tablet ($n = 13$, 7 male, 6 female), curious robot ($n = 16$, 11 male, 5 female) and non-curious robot ($n = 19$, 9 male, 10 female), where only subjects that completed each task were considered. The non-curious robot behaved as a compelling playmate, as described in the previous section, without any overt expressions of curiosity. The curious tablet and the curious robot were identical in all but two aspects: (i) in the curious tablet condition the robot was covered by a box, i.e. the child did not see the robot at all and; (ii) the experimenter introduction was "you are going to play with a virtual agent, Parle."

The curious robot had several behaviors that expressed curiosity. In its own introduction, the curious robot said "I want to learn to read. I hope you can teach me some words." The non-curious robot, on the other hand, said "lets start." Another curious behavior was the overt expression of enthusiasm of learning. It said: "I love to learn" or "I want to know more" before asking a new question; "it is always great to learn something new" or "that is a great word to know" after it was shown a correct word; after the robot addressed the tablet's spoken words, the curious robot said "this way we can both learn how to read" or "I love to learn this way"; "I love getting it wrong sometimes. This is how you learn new things" after the tablet showed the correct word; after the child moved a character instead of tapping on a word, the robot said "I am also eager to see what happens".

Another expression of curiosity is the wonder and imagination of future events. 25% of the time, prior to the child's movement the curious robot said either "I wonder what would happen if dragon goes to talk to butterfly", where "dragon" and "butterfly" are characters in the Story-maker app and the sentence changed in respect to the characters in each scene; or "I love trying new things. Can you move another character?" Furthermore, at the beginning of each new scene, it exclaimed enthusiastically and said "a new scene. I wonder what you would do now."

Lastly, the curious robot selected specific words to ask the child during the interaction, based on the assessment algorithm. It asked about the word that had the closest probability to 50% that the child knows how to read it, in spirit to Vygotsky's zone of proximal development. If the child knew perfectly how to read, e.g. an older child, the robot asked about the longest word in the sentence. This behavior guaranteed to challenge the child during the interaction [14], another characteristic of a curious peer. The non-curious robot, on the other hand, asked about a random word in the sentence, thus sometimes asking too hard a word and most of the time too easy.

The difference between the curious and non-curious conditions during the entire 15 minute interaction amounted to a total of roughly 10-15 expressions.

6. CURIOSITY MEASURES

6.1 Free exploration

The first measure was the Free exploration task, which used the same graphics as the main Story-maker app. In this app there were four characters from the story app and three new ones. The child could move any character and that generated a spoken sentence (by the tablet) as well as a written one. Each movement generated a new sentence, even of the same character, such that each new active interaction with the app generated a novel sentence the child was exposed to. The child could also tap on a word and hear it. The game lasted for two minutes and then ended.

This measure was used right after the robot interaction ended with the sleeping robot. The experimenter said to the child: "I am going to ask your parent some questions. In the meantime you can play this game. Do whatever you want with it." This measure and its framing had several reasons: (i) The adults, i.e. parent and experimenter, were removed from the scene so as to allow the child to express her own inner curiosity [12]. (ii) There were no limitations or suggestions on the things the child could do, thus enabling free exploration. (iii) The child had control over the amount of information she was exposed to.

Children started the game at different times, sometimes not understanding they should play the game while their parent has gone, while others waited for more confirmation on when they should start. Hence, each child that played the Free exploration game, played for a different amount of time. We thus considered only the first 60 seconds after the first interaction with the app, such that the measure indicates the true interaction with the game.

Several measures were considered, such as whether the subject first interacted with the new characters, total number of characters moved, etc. These measures were too discretized and constant across all subjects and were thus not used. A more theory-based measure was selected, namely, the active seeking of new information, here experienced via hearing new sentences. The normalized curiosity measure we used was the portion of the 60 seconds that the tablet spoke, i.e. that the child was exposed to new information. Thus a normalized measure of 0.5 means that 30 seconds out of the first 60 seconds of interaction the child moved characters that prompted the tablet to speak new sentences, each approximately spoken for 2 seconds. A higher measure means the child moved many characters that resulted in many different sentences, whereas a lower measure means the child did not move many characters and hence was not exposed to novel information. 8 children quit playing before 60 seconds have passed and were thus excluded from the analysis.

6.2 Question generation

The second measure was question generation [10, 6]. The experimenter said to the subject: "I am going to make a movie on the robot and game so that people who can't come here can learn of them. I want you to ask me as many questions as you can about the robot and I will answer them in the movie. What do you want to know about the robot?" The framing of the question was done for the following reasons: (i) The questions should not be answered during this assessment. (ii) The "movie" was the rationale of why to ask and why the experimenter did not answer. (iii) The entire framing was sometimes too complex for young children, hence we added the last sentence to clarify that we want them to ask questions.

The question generation task came after an extensive questionnaire administered to the participants. Hence, some of the children simply repeated these questions when prompted to ask questions about the robot or agent. We only considered *novel* questions, i.e. questions that were different than the ones the experimenter asked, as "questions generated" by the participants, since we were interested in their self-generated inquiries. Thus, the measure was taken to be binary: zero for no questions asked and one for any number of novel questions asked. 3 children did not talk or cooperate during this task and were thus excluded from the analysis.

6.3 Uncertainty-seeking

The third and last measure was the Fish task app [13]. This app is portrayed as a game in which the children are in a submarine with two windows. They can open one window and see a fish through it. The two windows differ in the uncertainty of which fish will be outside. Thus one window is presented with one fish next to it, indicating that with certainty that fish is outside the window. Another window can be presented with, e.g. 5 fish next to it, indicating that one of those five fish is outside the window, but the child cannot know until she opens it. Thus, the child needs to select which window to open, i.e. which amount of uncertainty she seeks. The app is cleverly designed to explore many differences in uncertainty in a repetitive yet step-wise manner. There are 18 turns, i.e. selections to be made, in the game, wherein the largest amount of uncertainty is 7 fish.

We used the normalized measure of the total amount of uncertainty selected, i.e. number of fish next to the window opened, divided by the maximum amount of uncertainty possible, i.e. $18 \times 7 = 126$. Thus, a child who seeks only uncertainty will always select the window with the maximal number of fish next to it and will get the score of 1. A child who seeks certainty will always select the window with one fish next to it and receive the score of 0.1417.

15 children did not complete the Fish task game and were thus excluded from the analysis. Due to its length, this task was administered at the end of the session, that lasted around 30 minutes. This may account for the high drop rate of this task.

7. HYPOTHESIS

The main hypothesis of this study is that interacting with a curious social robot is contagious, i.e. that curiosity of children, quantified by the curiosity measures described above, will be significantly higher for children interacting with a curious robot than with a non-curious robot.

However, we hypothesize a more subtle result, based on our understanding of the different types of curiosity [15]. We measured three measures, each one corresponding to a different aspect of curiosity, namely, free exploration, question generation and uncertainty seeking. However, the curious robot in our study, while expressing several curiosity-driven behaviors, did not express all types measured. More concretely, the robot exhibited free exploration by expressing love to learn new things and suggesting moving new characters. It also expressed uncertainty-seeking behavior, by wondering about new situations that could happen. However, at no point did the robot ask the child any question, novel or otherwise. Hence, our second and stronger hypothesis is that only the free exploration and uncertainty seeking measures will be increased by the interaction with the curious robot compared to the non-curious one, whereas the question generation measure will not be affected. In this study, we did not address the issue of whether the question generation measure can also be manipulated by a social robot behavior.

Regarding the learning gains, since we did not construct the manipulation to differentiate learning new words, i.e. the same Story-maker app is used in exactly the same manner across conditions, we hypothesize that the children will learn new words, but not differently with respect to the curious and non-curious conditions. Nevertheless, we hypothesize that the robot conditions will have larger effects than the curious tablet condition, due to the physical embodiment of the robot [2].

8. RESULTS

We first analyzed the measures themselves, across conditions. We tested whether the Free exploration measure had a normal distribution across conditions, using the Shapiro-Wilk normality test. The hypothesis that the data comes from a normal distribution was confirmed ($p = 0.41$, Shapiro-Wilk). Hence, we analyzed the Free exploration measure using analysis-of-variance (ANOVA) tests. For the question generation measure, we performed Fisher's exact test. We further tested whether the Fish task measure had a normal distribution across conditions, using the Shapiro-Wilk normality test. The hypothesis that the data comes from a normal distribution was not confirmed ($p = 0.01$, Shapiro-Wilk). Hence, we analyzed the Fish task measure using the a-parametric Kruskal-Wallis test.

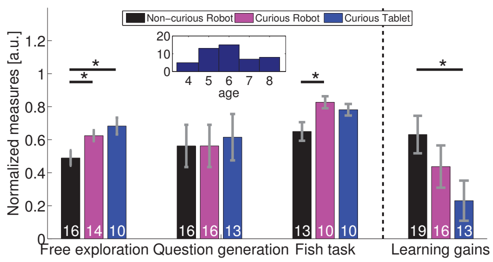

Figure 2: Average normalized curiosity (left) and learning-gains (right) measures across conditions (error bars denote SEM). White number indicates number of subjects. * $p < 0.05$. Inset: age histogram.

We then compared the curiosity measures between all conditions, Fig. 2. There was no significant difference in any curiosity measure between the curious tablet and the curious robot condition ($p = 0.47, 1.0, 0.45$ for the Free exploration, question generation and Fish task, respectively). The curious robot and the non-curious robot conditions, however, had mixed results. The curious robot resulted in significantly higher Free exploration and Fish task measures, compared to the non-curious robot ($F(1, 32) = 5.4, p = 0.027$ and $\chi^2(1, 21) = 3.9, p = 0.047$, respectively). However, there was no significant difference in the question generation task ($p = 1$, Fisher's exact test).

These results validate our main hypothesis: (i) Curiosity can be higher after interaction with a curious robot, compared to an interaction with a non-curious one. (ii) The curious robot impacted children's curiosity only on those behaviors that the robot models for the child, i.e., Free exploration and Fish task measures increased, whereas the question generation measure did not.

When comparing the attitudes children had towards the robot across conditions, as measured by our post-assessment questionnaire, we found that children found both robots to be equally engaging, whereas the tablet was less so. For example, when asked "if you were to play the Story-maker game again, would you prefer to play by yourself, with your mom/dad or with Parle?", 69%, 56% and 23% of the children preferred to play with Parle in the curious robot, non-curious robot and curious tablet conditions, respectively. This indicates that the non-curious robot was as a compelling playmate as the curious one, suggesting that the curiosity-driven behavior was the cause of curiosity measures, and not the engagement or affects of the children towards the robot. While there were no significant differences between robots and tablet, the tablet had lower likeability scores.

We then analyzed the learning gains, i.e. whether the children learned new words. For this, we considered only words that the child misidentified during the main phase, i.e. when the Story-maker app showed the child the correct word. For these misidentified words we analyzed whether in the post-test they were correctly identified; if so we labeled them "learned words". Over all conditions, children learned on average 1.2 (± 0.8) words. Since the post-test was a multiple-choice one, we tested whether the learned words were identified above chance level. We performed a t-test on the null hypothesis that for each child the learned word was identified with chance level, i.e. 20%. The alternative hypothesis, that it is above chance level was confirmed (average correct probability 62%, $p < 0.001$). These results indicate that the interaction, even though short, was sufficient to teach the children new words.

We further wanted to test whether the learning gains were condition-dependent, Fig. 2. First, we binarized the data such that a child was labeled "learned" if she learned at least one word, and "not-learned" if not. Then we compared the percentage of children that learned new words in each condition: non-curious robot 63%, curious robot 44% and curious tablet 23%. We then performed a Fisher's exact test on each condition pair, and found that only non-curious robot and curious tablet results in a significant difference ($p = 0.036$).

These results partially validate our learning-gain hypothesis: (i) Children learned how to identify new words during the interaction. (ii) Physical presence of the non-curious robot resulted in significantly higher learning gains than the curious tablet. However, the curious robot did not result in significantly higher learning gains, suggesting a complex interplay between learning and curiosity gains.

9. DISCUSSION

9.1 Curious tablet vs. curious robot

The fact that the curious tablet and the curious robot conditions were found to be virtually identical is somewhat puzzling. Previous studies have shown differences in both attitude and learning gains in a similar comparison [2, 16]. However, in our study the two conditions were much more similar: the "virtual agent" had the same emotional sounds of excitement and frustration; the sound was emitted in both conditions from the speakers, for increased volume; there was no virtual character on the tablet. We asked the children at the end of the interaction in the tablet condition to point to where they thought the virtual agent is. Only 4 pointed to the box, whereas 4 pointed to the tablet and 3 pointed to the speakers and some even to the curtain behind the box. We can conclude that the children did not treat the virtual agent as a virtual character in the tablet, but rather as a dis-embodied voice. Moreover, the interaction was focused on the tablet, and not on the robot, e.g. it was not designed as a face-to-face interaction, but rather as co-play on the tablet. Furthermore, the lack of difference is mainly in respect to the curiosity measures, whereas learning gains behave somewhat differently. We thus attribute the lack of difference in curiosity measures between these conditions to the fact that the tablet condition was perceived as playing a tablet game with a hidden real robot, e.g. like discussing a document with someone over the phone. Nevertheless, in order to fully address these issues, a full manipulation of the tablet condition is in order, e.g. adding a non-curious tablet condition, a tablet condition with no affective expressions, designing a face-to-face co-play interaction. These are beyond the scope of the current study.

9.2 Curiosity measures generalizability

Each of the measures we used has pros and cons relating to their usage in other studies. Free exploration is an important aspect of curiosity, as it is an independent activity for seeking information. The app developed and used can assess several components, e.g. interaction with new characters, patterns of exploration, etc. The quantitative measure used here reflects the amount of actively exposed information, which is at the core of free exploration. Furthermore, it taps into a "low cognitive level" of curiosity, in the sense that actions are finger manipulation and perception is passive listening. Hence, while it is easily applicable in a wide range of interactions, its individual nature does not account for the social aspect of curiosity and it assesses only low cognitive processing. The Fish Task has identical considerations.

On the other hand, question generation is a very social aspect of curiosity as well as requires a higher cognitive processing. The subject has to mentally think of novel questions, without any information from the experimenter. Furthermore, the interaction between shyness and curiosity comes into play and may confound the assessment. Nevertheless, the open-endedness of the measure can yield enlightening results and is easily administered in any study.

9.3 Different types of curiosity

Studies from the 70's and 80's have shown that curiosity is not a unitary characteristic, but rather a composite one that aggregates different types of aspects [15, 23]. Consistent with these views, we have measured three distinct quantitative aspects of curiosity, namely, free exploration, question generation and uncertainty seeking. We have shown that the curious robot can effect those measures that its behavior explicitly models for the child. While we did not show that the question generation can be similarly manipulated by a social curious robot, we interpret these results as a relation between the type of behavior the robot exhibited and the curiosity aspect affected. Thus, for example, the fact that the curious robot says "I love trying new things. Could you move another character?" is a direct manipulation on the Free exploration task wherein the child can move different characters and learn new things about them from the spoken sentences. Furthermore, most of its curiosity-driven utterances refer to seeking the unknown, e.g. "a new scene. I wonder what you would do now", thus manipulating the uncertainty seeking measure. However, the curious robot never asks a question related to the child, the story, their interaction or anything else [5, 10]. It does not express its knowledge thirst via the direct channel of verbal interaction.

One confound to this interpretation is the measure medium. The interaction with the robot was via the tablet throughout the entire session, i.e. the robot did not respond to verbal or non-verbal communication from the child. Furthermore, the Free exploration and Fish task measures were presented in a similar medium, i.e. a tablet app, whereas the question generation one was via the verbal interaction with the experimenter. Thus, the difference between manipulation effects may be due to the measure medium.

One could consider a different type of manipulation wherein the robot deliberately asks the child, or the tablet in the presence of a child, knowledge-seeking questions. We hypothesize that this manipulation, controlled-for by a social robot that either does not ask questions or asks irrelevant, repetitive or boring questions, will increase the question generation measure. Performing a cross-manipulation paradigm can raise the medium confound and is intended for future work.

9.4 Learning gains

We have shown that the non-curious robot has the highest learning gains, significantly higher than the curious-tablet. This suggests that learning gains are higher with a physical robot [2, 16]. Moreover, learning outcomes for such a short encounter are suggestive of longer term gains, but a longitudinal study is required in order to assess learning outcomes of curious vs. non-curious and virtual vs. physical interventions.

The trend in learning gains, i.e. non-curious tablet highest, followed by curious robot and curious tablet lowest, is reversed for the increase in the Free exploration measure. This raises the possibility that the two aspects interact. In other words, it may be the case that during a single, short interaction one cannot achieve both learning gains and increased curiosity, as measured by the Free exploration task. A longitudinal study, with repeated encounters may illuminate the interaction between increased curiosity and learning gains.

9.5 Fringe benefits from the interaction

We asked the parent "do you think your child learned something from this interaction?" While originally designed to see whether the parent noticed their child learning new words, the parents' answers were diverse and insightful. Some parents pointed out that their child learned "listening skills" or "learned to wait". We believe that this was partly due to the fact that the tablet app was disabled while the robot was speaking. This raises the question of whether a social robot can be used to teach and assess a child's listening skills.

One parent pointed out that their child learned "how to help another kid how to learn". Similarly, another parent said their child learned "to be patient, to work at another creature's pace". Our framing of the robot as a younger peer that wants assistance in learning how to read was perceived by the parent and child as an opportunity to practice and foster empathy and responsibility over another less capable social agent. This raises the question of whether a social robot can be designed and programmed to improve empathy and consideration of children, and whether that is transferable to other children or adults.

10. CONCLUSIONS AND FUTURE WORK

We have studied the effects an autonomous social robot's curiosity driven behavior has on a child' curiosity. The robot's behavior exhibited several aspects of curiosity and the child was assessed on different aspects as well. We have shown that a fully autonomous robot can be modeled as a peer that impacts curiosity behaviors in children. Moreover, we have shown that only those curiosity aspects which we manipulated increased in children.

These results suggest that manipulating subtle social interaction utterances and expressions can impact children's curiosity. We suggest that other educational HRI studies incorporate these and thus may gain additional positive influence on children learning.

In future work we intend to further study the specificity and generalization of our results, namely, can we manipulate each aspect of curiosity independently and whether those aspects carry to other activities, have long lasting effects and can change the child's mindset. Including a more diverse set of curiosity measures and more model-based curiosity behaviors of the robot will enable the development of a theoretical framework of children's curiosity manipulation as well as a personalized and more social curious companion for children.

11. ACKNOWLEDGMENTS

The authors acknowledge help and support of Jacqueline Kory in the development of the experimental setup. G.G. was supported by the Fulbright commission for Israel, the United States-Israel Educational Foundation. This research was supported by the National Science Foundation (NSF) under Grants CCF-1138986. Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not represent the views of the NSF.

12. REFERENCES

[1] M. P. Arnone, R. V. Small, S. A. Chauncey, and H. P. McKenna. Curiosity, interest and engagement in technology-pervasive learning environments: a new research agenda. *Educational Technology Research and Development*, 59(2):181–198, 2011.

[2] W. A. Bainbridge, J. W. Hart, E. S. Kim, and B. Scassellati. The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics*, 3(1):41–52, 2011.

[3] L. S. Blackwell, K. H. Trzesniewski, and C. S. Dweck. Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development*, 78(1):246–263, 2007.

[4] C. Dweck. *Mindset: The new psychology of success.* Random House LLC, 2006.

[5] R. C. Endsley and S. A. Clarey. Answering young children's questions as a determinant of their subsequent question-asking behavior. *Developmental Psychology*, 11(6):863, 1975.

[6] S. Engel. Children's need to know: curiosity in schools. *Harvard Educational Review*, 81(4):625–645, 2011.

[7] S. Engel. The case for curiosity. *Educational Leadership*, 70(5):36–40, 2013.

[8] A. Gatt and E. Reiter. Simplenlg: A realisation engine for practical applications. In *Proc. of the 12th European Workshop on Natural Language Generation*, pages 90–93. Association for Computational Linguistics, 2009.

[9] K. Haimovitz, S. V. Wormington, and J. H. Corpus. Dangerous mindsets: How beliefs about intelligence predict motivational change. *Learning and Individual Differences*, 21(6):747–752, 2011.

[10] P. L. Harris. *Trusting what you're told: How children learn from others.* Harvard University Press, 2012.

[11] B. Henderson and S. G. Moore. Children's responses to objects differing in novelty in relation to level of curiosity and adult behavior. *Child development*, pages 457–465, 1980.

[12] B. B. Henderson. Parents and exploration: The effect of context on individual differences in exploratory behavior. *Child Development*, pages 1237–1245, 1984.

[13] J. Jirout and D. Klahr. Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review*, 32(2):125–160, 2012.

[14] T. B. Kashdan and M. Yuen. Whether highly curious students thrive academically depends on perceptions about the school learning environment: A study of hong kong adolescents. *Motivation and Emotion*, 31(4):260–270, 2007.

[15] R. Langevin. Is curiosity a unitary construct? *Canadian Journal of Psychology/Revue canadienne de psychologie*, 25(4):360, 1971.

[16] D. Leyzberg, S. Spaulding, M. Toneva, and B. Scassellati. The physical presence of a robot tutor increases cognitive learning gains. In *Proc. of the 34th Annual Conf. of the Cognitive Science Society. Austin, TX*, 2012.

[17] A. N. Meltzoff, P. K. Kuhl, J. Movellan, and T. J. Sejnowski. Foundations for a new science of learning. *science*, 325(5938):284–288, 2009.

[18] J. Movellan, M. Eckhardt, M. Virnes, and A. Rodriguez. Sociable robot improves toddler vocabulary skills. In *Proc. of the 4th ACM/IEEE Int. Conf. on Human robot interaction*, pages 307–308, 2009.

[19] L. Schulz. The origins of inquiry: inductive inference and exploration in early childhood. *Trends in cognitive sciences*, 16(7):382–389, 2012.

[20] A. Setapen. *Creating Robotic Characters for Long-term Interaction.* Thesis, 2012.

[21] F. Tanaka and S. Matsuzoe. 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.

[22] J. Torgesen, R. Wagner, and C. Rashotte. Test of word reading efficiency (towre). austin, texas: Pro-ed, 1999.

[23] F. Wardle. Getting back to the basics of childrens play. *Child Care Information Exchange*, 57:27–30, 1987.