A is for Artificial Intelligence

The Impact of Artificial Intelligence Activities on Young Children's Perceptions of Robots

Randi Williams
MIT Media Lab
Cambridge, Massachusetts
randiw12@media.mit.edu

Hae Won Park
MIT Media Lab
Cambridge, Massachusetts
haewon@media.mit.edu

Cynthia Breazeal
MIT Media Lab
Cambridge, Massachusetts
cynthiab@media.mit.edu

ABSTRACT

We developed a novel early childhood artificial intelligence (AI) platform, *PopBots*, where preschool children train and interact with social robots to learn three AI concepts: knowledge-based systems, supervised machine learning, and generative AI. We evaluated how much children learned by using AI assessments we developed for each activity. The median score on the cumulative assessment was 70% and children understood knowledge-based systems the best. Then, we analyzed the impact of the activities on children's perceptions of robots. Younger children came to see robots as toys that were smarter than them, but their older counterparts saw them more as people that were not as smart as them. Children who performed worse on the AI assessments believed that robots were like toys that were not as smart as them, however children who did better on the assessments saw robots as people who were smarter than them. We believe early AI education can empower children to understand the AI devices that are increasingly in their lives.

CCS CONCEPTS

• Social and professional topics → Computational thinking; **K-12 education**; **Children**; **• Computing methodologies → Artificial intelligence**.

KEYWORDS

AI education, early childhood education, child-robot interaction, social robots

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300677

ACM Reference Format:
Randi Williams, Hae Won Park, and Cynthia Breazeal. 2019. A is for Artificial Intelligence: The Impact of Artificial Intelligence Activities on Young Children's Perceptions of Robots. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, Article 4, 11 pages. https://doi.org/10.1145/3290605.3300677

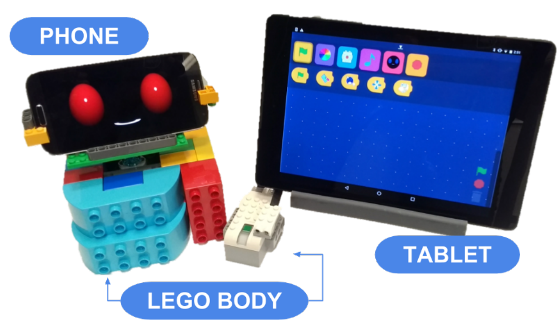

Figure 1: The *PopBots* platform consists of a smartphone-based social robot, LEGO blocks, LEGO WeDo motors, and a block-based programming interface on a tablet.

1 INTRODUCTION

Artificial intelligence, or AI, will have a huge impact on our society and experts say that now is the time to prepare for a rapidly-changing, AI-powered economy [4]. Today's interactive devices are far more advanced than those from even a decade ago. These devices behave more like socially interactive beings than machines – communicating via spoken language, recognizing faces, learning users' preferences, acquiring new skills over time, and more. Even for society's youngest members, AI has begun to impact the ways that many children live, learn, and play [6, 13]. All of this sets the stage for a future where children grow up not just as digital natives, but as AI natives who will have fundamentally different relationships with technology than prior generations [6].

An important question is *when should children begin to learn about AI?* Many young children already interact with devices, such as smart speakers and smart toys, on a daily basis which has led researchers to explore privacy and safety concerns related to children and AI [13]. In our own work, we have found that children eagerly engage with AI, but have not been taught how it works which leads to faulty assumptions about what the technology can do [6, 7]. Therefore, we advocate for intervention as early as possible, even in preschool classrooms (ages 4 to 6).

Previously, we introduced *PopBots*, an early childhood AI and robotics platform that allows preschool through second-grade children (ages 4-7) to explore AI topics by building, programming, and training social robots [22]. In this paper, we sought to understand **how did developmental factors, like perspective taking skills, impact what children could learn about AI?** and **how did children's perceptions of "thinking machines" change after they engaged in educational AI activities?**

2 BACKGROUND

Children's Perceptions of Robots

Social robots interact with people in more human-like ways: speech, non-verbal cues, gaze, emotive expressions, etc. Prior work has shown that young children perceive and respond in rich ways to the non-verbal cues of social robots including back-channeling, attentive behavior, and vocal expressivity [11, 15, 21]. In general, children tend to ascribe emotional states, mental states, intentionality, and morality to intelligent social artifacts [2, 9, 18]. Prior findings suggest that children's psychological attributions and emotional connection to social robots are robust; they persist even after children are shown how they are built and programmed [19].

However, younger children are still developing their understanding of "what it means to be alive" making their attributions less consistent across the population than their older counterparts [9, 16]. The reasons for these differences were not always immediately apparent, but some tended to be dependent on children's individual interactions with the robots (e.g., whether or not it says their name correctly) as well as social and cultural factors (e.g., how other people talk about the robot or smart device) [6, 7, 18].

Mioduser and Levy explored the role of experience in children's understanding of robots by conducting a study where they asked Kindergarten children to describe the actions of a vehicle-like robot as it did increasingly complex tasks [14]. When the robot did simple tasks, children used more mechanical descriptions of the robot such as "It's driving forward, then turning left." However, as the tasks became more complex, children used more psychological descriptions, "It *wants* to go to the light." In a follow-up study, Levy and Mioduser let children program robots before asking children to describe the robot's behavior [12]. They saw that even as behaviors became more complex, children maintained giving technical descriptions of the robot's behavior. However, when then asked to describe the behavior of robots they had not programmed, children reverted to using more psychological descriptions. Still, the experience of programming robots changed the way that children thought about the robots (mechanically versus psychologically), even if the differences did not transfer to new robots. In this work, we expected that at first, young children will relate to robots as social and intellectual beings. Then, after learning more about how a robot's "mind" works, young children will better able to articulate how machine intelligence compares to other people.

Children's Theory of Mind Skills

Perspective-taking, or the ability to view the world from another's vantage point is rooted in one's Theory of Mind skills [1]. Children around the age of 6 or 7 are still developing the final stage of Theory of Mind skills, cognitive perspective taking, or "the understanding of false belief and predicting actions on the basis of beliefs that are false rather than true." [1]. In the *PopBots* activities, children relate to the robot as a social being to understand the algorithms that are running in its "mind". Children are asked to understand the robot's perspective making it likely that their comprehension of AI will be dependent on their Theory of Mind skills.

Wellman and Liu compiled multiple studies to develop a series of tasks to assess children's Theory of Mind skills [20]. The tasks use storytelling and targeted questions to measure children's understanding of diverse desires, diverse beliefs, knowledge access, content false belief, explicit false belief, belief emotion, and real-apparent emotion. Key Theory of Mind abilities that children need to understand AI in *PopBots* activities are knowledge access understanding – that a person may not know something that you know, content false-belief understanding – that another person may believe something incorrect that will impact their behavior, and explicit false belief understanding – knowing how a character will behave given its knowledge state. Wellman and Liu saw that 73% of 3 to 5-year-old children could correctly answer a knowledge access question, that 59% of children could correctly answer a content false belief one, and 57% of children could correctly do explicit false belief [20]. We hypothesize that Theory of Mind development will be a significant factor in children's understanding of AI with *PopBots*.

3 *POPBOTS*: EARLY CHILDHOOD AI ACTIVITIES

To teach children about AI concepts, we developed the Preschool-Oriented Programming (*PopBots*) Platform [22]. Some of our primary design considerations were to help children draw connections between the activities and their own experiences, to appeal to children with varied backgrounds and interests, and to empower children to reflect on and discuss AI.

The platform consists of a social robot toolkit, a programming interface on a tablet computer, and three hands-on activities with assessments for young children to explore machine learning and reasoning algorithms (Figure 1). We developed the activities and assessments through a series of pilot tests and workshops with 26 children. We tested a number of different AI topics, including deep learning and style transfer. For the evaluative study in classrooms, we settled on the 3 topics that were easiest for children to grasp and that had meaningful connections to children's experiences with smart devices. The activities allow children to guide themselves through developing each algorithm from scratch.

PopBots Robot

Children build their own LEGO robot characters using regular LEGO and LEGO DUPLO blocks. The robot is programmable but also has autonomous functionality as it plays an active role. As children go through each hands-on activity, the social robot talks to them – explaining the algorithm logic and encouraging students to try new things.

PopBots Activities

Children use the tablet with the blocks-based interface to program and train the robot for each AI activity. The programming and training interface is entirely picture-based, built on Scratch blocks, to accommodate children who cannot yet read [10]. We used *PopBots* to teach children about three AI concepts, each packaged in a child-friendly activity: Knowledge-Based Systems, Supervised Machine Learning, and Generative AI. We chose these topics because they are relevant to the kinds of AI algorithms that children are exposed to through smart toys and entertainment apps.

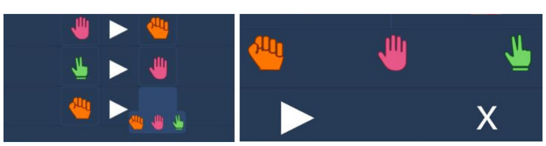

Figure 2: Screenshots of Knowledge-Based Systems activity: *Rock-Paper-Scissors*. The left is the interface for entering the rules of *Rock-Paper-Scissors*. The top line reads "paper beats rock." The right is for playing *Rock-Paper-Scissors* against the robot.

In Knowledge-Based Systems, computers store expert information about how to solve problems then use that information to make decisions. Games are a great example where the system needs to choose the next move depending on what the system predicts the human player might do next. Children explored this topic by entering the rules of *Rock-Paper-Scissors* and then playing against the robot (Figure 2). As the child plays, the robot uses a state probability transition matrix to predict the child's next move. If the robot's guess for the child's next move is greater than chance (33%), the robot says "I think you will put X, so I will put Y because Y beats X." Otherwise, the robot says "I'm not sure what you'll do next, I'll just guess."

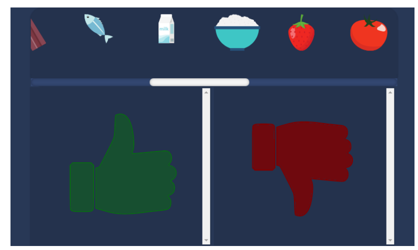

Figure 3: Screenshot of Supervised Machine Learning activity: *Food Classification*. Children label foods as good (healthy) or bad (unhealthy) by dragging them to the appropriate box. Children can ask the robot to guess where a food goes by clicking on it.

Supervised Machine Learning is another common AI algorithm that involves acquiring a knowledge base from examples. Children encounter Supervised Machine Learning systems when they use personalized recommender systems for media streaming applications such as *YouTube Kids*. In our activity, children are introduced to training sets and how robots learn patterns by labelling a subset of foods as healthy or unhealthy (Figure 3); these foods form the training set. Then, children can touch any of the unlabeled foods, the test set, to hear the robot's prediction for how it should be classified given its similarities to foods in the training set. The robot has a database of 20 foods where each food is labeled with its name, color(s), food group(s), and amount of calories and sugar in a 100 gram serving. When children touch foods in the test set, the robot predicts its label by comparing it to known foods in the training set: "Bananas are a lot like corn and corn is healthy. So, I guess, bananas are healthy too." Children experiment to see how different training sets impact the robot's accuracy on the test set.

Finally, Generative AI is very different from the other two topics. We chose this activity to show children that robots can be creative, in their own way. Rather than learning rules, the system generates a new output based on children's input song, within the bounds of parameters the child gives it. Real world examples include camera apps that use style transfer filters to remix photographs. In *PopBots*, children explore how a robot creates a new melody by changing the tempo

Figure 4: Screenshot of Generative AI activity: *Music Remix*. The left interface is defining the tempo and chord progression parameters for musical emotions. The right, for creating songs for the robot to remix.

and progressions of an input melody by the child (Figure 4). For example, happy songs have a faster tempo with an ascending melody in a major key. Children associate different combinations of tempo and progressions with emotions (excited, happy, sad, surprised) to teach musical emotions to the robot. Then, children take turns demonstrating a melody to the robot and listening to how the robot transforms it based on these “emotion filters" to produce a new melody.

PopBots Assessments

We developed assessments for each activity as a set of multiple choice questions to probe what children understood about each AI concept. We verified the assessments in the pilot tests, comparing children's performance on them with their demonstrated understanding in semi-structured interviews [22]. Each question has an associated picture and children select the multiple choice response they think is best. For each topic, we probed children's understanding with 3 or 4 questions about the algorithm's basic functionality, edge cases, and initialization.

Knowledge-Based Systems (KBS)

(1) We teach the robot the normal rules of *Rock-Paper-Scissors*. Then, Sally plays rock and the robot plays paper, who does the robot think has won? Sally or the robot? (Robot. Tests if children understand how the robot uses rules to decide who wins.)
(2) Sally plays paper five times. What does the robot think she will play next? Rock, paper, or scissors? (Paper. Tests if children understand how robot uses its opponent's past behavior to predict future behavior.)
(3) The robot thinks that Sally will play paper next. What will the robot play so that it can beat Sally? Rock, paper, or scissors? (Scissors. Tests if children understand how robot uses its predictions to decide what to do next. Requires an understanding of false belief.)
(4) We changed the rules so that they are all opposite rules (paper beats scissors). Sally plays scissors and the robot plays paper. Who does the robot think has won? Sally or the robot? (Robot. Tests if children understand that the robot's knowledge is limited to the rules it is taught, even if they are wrong.)

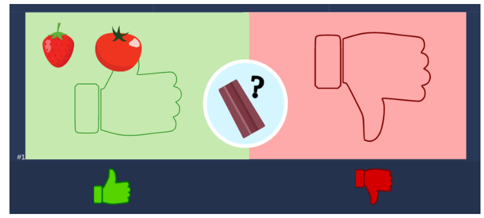

Figure 5: Interface for a question on the Supervised Machine Learning (SML) AI assessment. SML #1: *You teach the robot that strawberries and tomatoes go in the good group. Now, which group will the robot think chocolate goes in? The good group or the bad group?*

Supervised Machine Learning (SML)

(1) You teach the robot that strawberries and tomatoes into the good group, and do not teach it about any other foods. Which group will the robot think chocolate goes in? The good group or the bad group? (Good. Tests if children understand how the algorithm is initialized and that a robot with only good or bad examples will put everything in one category.)
(2) What food does the robot think is most like a tomato? Strawberry, banana, or milk? (Strawberry. Tests if children understand how the robot uses multiple features of foods to determine which is most similar.)
(3) You put ice cream in the good (healthy) category and bananas in the bad (unhealthy) category. What category will the robot put corn in? The good category or the bad category? (Bad. Tests if children understand how the robot puts similar foods in the same category. Requires knowledge access and false belief understanding.)

Generative AI (GAI)

(1) Priya asks the robot to play back a song and she sets the (parameter) bars in the middle. Does the robot play the same song or a different song? (Same. Tests if children understand the robot's initialization where it does not change the tempo or notes of its input song.)
(2) Priya asks the robot to play back a song and she puts the bars to the right. Does the robot play the same song or a different song? (Different. Tests if children understand that changing the parameters given to the robot will result in the robot changing the tempo and notes of the input song.)
(3) When the bars are to the right, does the robot's song still have to have some of the same notes as Priya's input? (Yes. Tests if children understand that the output song is based on the input the robot is given.)

4 MATERIALS

Theory of Mind Assessment

Figure 6: Interface for the Explicit False Belief question on the *Theory of Mind Assessment* question. *The boy is looking for his mittens. He believes that they are in his closet, however, they are really in his backpack. Where will Scott look for his mittens first?* Children select the location where they think the boy will look first.

We converted Wellman and Liu's Theory of Mind assessments [20] into a multiple choice question interface on the tablet or on paper, similar to the one in Figure 6. Given that Wellman and Liu's tasks used stories, we created colorful scenes that would play as a movie on the tablet or that the researcher could read aloud to the students from a book. From Wellman and Liu's original set of tasks, we used the three that were relevant to *PopBots* [20].

Knowledge Access: "Child sees what is in a box and judges (yes - no) the knowledge of another person who does not see what is in a box."

Content False Belief: "Child judges another person's false belief about what is in a distinctive container when the child knows what it is in the container."

Explicit False Belief: "Child judges how someone will search given that person's mistaken belief."

Perception of Robots Questionnaire

Children completed a questionnaire about their perceptions of robots and AI either on a tablet computer or on paper. The questionnaire was in the format of the Monster Game from a previous study (see Figure 7) [6]. In this format, two on-screen characters offer differing (often opposite) opinions about robots. The child decides which character they agree with more, or if their opinion is somewhere between the two. The first question concerns whether the robot has the agency to choose to disobey rules. The next two questions explore how children perceive the robot's intelligence – if robots are intelligent is their intelligence static or can it grow? The last two questions examine if children saw the robot as an intelligent social being and what level of maturity they would assign to it. The bold statements are the target statements that we used in our analysis.

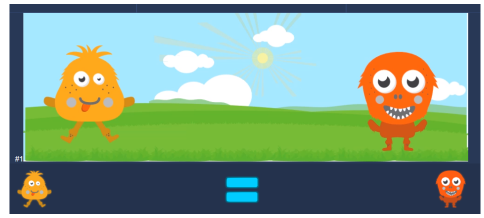

Figure 7: Interface for the *Perception of Robots Questionnaire.* The researcher read the opinions of the two monsters and then children selected which monster they agreed with more or an equal sign if they were somewhere in the middle.

Do you agree with either view or are you in the middle?

(1) **Robots follow rules** / robots do not follow rules.
(2) I am smarter than robots / **robots are smarter.**
(3) Robots can't learn new things / **robots can learn.**
(4) **Robots are like toys** / robots are like people.
(5) **Robots are like children** / robots are like adults.

Participants

Our study was comprised of a series of *PopBots* Workshops that we conducted during Spring 2018 with children from four schools in the Greater Boston Area. We ran the *PopBots* activities in 5 classrooms and collected data from 80 four to six-year-old children. Three schools were public schools where our workshops were carried out during their after school program (Table 1 schools A, B, C). Another was a private school, and our workshop was conducted in the classroom as part of a unit on robots (D, E). All of the children were in preschool, meaning that they had not yet begun formal grade school. Pre-Kindergarten (or Pre-K) children were ages 4–5 and Kindergarten children were ages 5–6. This separation of classrooms by ages was standard for the schools we worked with. Classroom E contained only Pre-K children, Classrooms A, C, and D contained only Kindergarten children, and classroom B was mixed with Pre-K and Kindergarten. Our sample of students included children with a range of backgrounds and socioeconomic statuses, as detailed in Table 1. All students in the classroom were welcome to participate, however, we only collected data from children whose parents completed a consent form.

5 PROCEDURE

First, children completed the *Perception of Robots Questionnaire* and *Theory of Mind Assessment.* The researcher would read each question, then children would respond on their individual tablets or papers. Although the assessments were done with everyone in the classroom, children did not share or discuss their answers until after everyone had answered.

	ED[a]	ELL[b]	Avg. Age	Gender (% Fem.)	N
A	51.4%	32.9%	5.50	33.3%	6
B	38.9%	48.4%	6.00	37.5%	16
C	14.3%	2.2%	5.37	42.1%	19
D	N/A[c]	<1%	5.05	50.0%	22
E	N/A[c]	<1%	4.59	47.1%	17
			5.18	48.8%	80

Table 1: School and Participant Data [5]. (*a*) Percent of students who are economically disadvantaged as defined by the household income of their parent/guardian(s). (*b*) Percent of students who are English language learners. (*c*) Private school classrooms did not share socioeconomic status data.

Then, children went through the *PopBots* AI activities. The order of the activities remained consistent across the different classrooms:

(1) Introduction to Programming with the *PopBots*
(2) Knowledge-Based Systems with *Rock-Paper-Scissors*
(3) Supervised Machine Learning with *Food Classification*
(4) Generative AI with *Music Remix*

Classrooms completed each AI activity in 10-15 minutes. Next children completed the AI assessment to measure the extent to which they understood different topics. After all the AI activities and assessments, children completed the post-test *Perception of Robots Questionnaire*.

Data and Analysis

We collected quantitative data about children's responses from the questionnaires. To support these data, we also recorded observations while children completed the activities. In particular, we noted children's questions and observations, their understanding of the activities as they went through them, and how they would talk through their reasoning about their responses. Each session was also video recorded for later analysis.

6 RESULTS

Performance on Theory of Mind Assessment

Children's scores were graded in the same way as Wellman and Liu, where a question is only correct if a child gets both the target and control question correct [20]. Our sample of 4 to 6-year-old children was quite consistent with Wellman and Liu's data for 3 to 5-year-olds (Figure 8). The knowledge access question was the easiest for children to get correct, at 75%. The false belief questions proved to be more difficult, with both at 55% of correct answers. The similarity of our results to prior work suggests that the *Theory of Mind Assessment* delivered on the tablet is comparable to the original tasks.

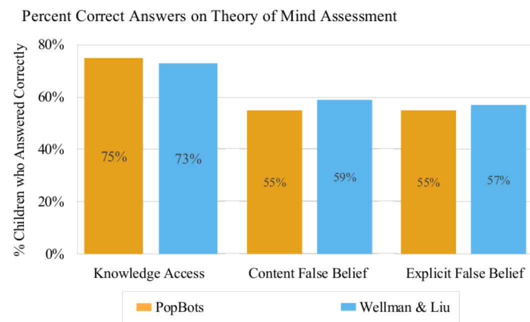

Figure 8: A comparison of children's scores on the *Theory of Mind Assessment* compared to those found in Wellman and Liu [20]. Our results closely match theirs.

Performance on PopBots AI Assessments

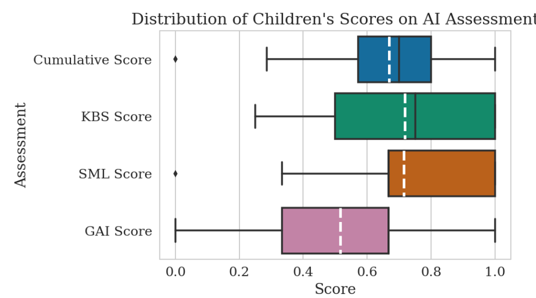

Figure 9: Children's scores on the *PopBots* assessments. The average cumulative was 66.8%. Children best understood Knowledge-Based Systems ($\bar{x}$ = 71.8%), then Supervised Machine Learning ($\bar{x}$ = 71.5%), then Generative AI ($\bar{x}$ = 51.5%).

As reflected by their performance on the AI assessment, children seemed to grasp the AI concepts well. The median score was 70% ($IQR = 57.1\% - 80.0\%$) with an average score of 66.8% (see Figure 9). Kindergarten children performed a bit better than Pre-K children ($Mdn = 70.0\%$ vs $Mdn = 63.3\%$). Children best understood the Knowledge-Based Systems (KBS) activity with *Rock-Paper-Scissors* ($Mdn = 75.5\%, \bar{x} = 71.8\%$), next Supervised Machine Learning (SML) with *Food Classification* ($Mdn = 66.7\%, \bar{x} = 71.5\%$), and finally Generative AI (GAI) with *Music Remix* ($Mdn = 66.7\%, \bar{x} = 51.5\%$). In a previous paper, we found that differences in children's understanding sometimes depended on their age, grade, or the extent to which they explored the activities (especially for GAI) [22].

We expected to see that Theory of Mind skills would impact children's understanding and ability to correctly answer assessments. The third question in the KBS assessment, *KBS #3: "The robot thinks that Sally will play paper next. What will the robot play so that it can beat Sally? Rock, paper, or* *scissors?"* strongly resembles the explicit false belief question *If the boy thinks that his mittens are in the closet, where will he look for them first?*. Only 55% of children could correctly answer the explicit false belief question. We were surprised to find that of the 39 children who initially got the explicit false belief question wrong, 28 of them were able to get KBS #3 correct. Out of all children, 76.6% got that question correct. Furthermore, later children encountered the third SML question *SML #3: "You put ice cream in the good (healthy) category and bananas in the bad (unhealthy) category. What category will the robot put corn in?"*. This question required all three Theory of Mind skills: a child must reflect on the fact that the robot has no prior knowledge of whether corn is healthy or not, recognize that it had a false belief about which foods are healthy, and then follow the logic of Supervised Machine Learning to come to the conclusion that the robot's false belief would cause it to put corn in the unhealthy group with the bananas. On this question, 85.5% of children answered correctly. Reframing this in terms of the order of the activities (Figure 10), we saw that children who incorrectly answered Theory of Mind related questions in the beginning were capable of answering related questions correctly when they came up in the activities. We hypothesized that framing AI concepts through the robot's cognition would help children understand AI algorithms, but we also saw that children's understanding of AI algorithms boosted their Theory of Mind reasoning.

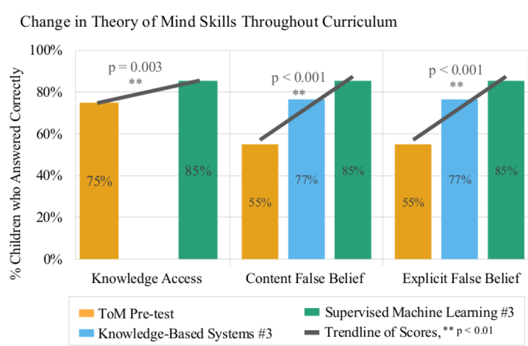

Figure 10: Change in proportion of children who correctly answered Theory of Mind related questions throughout the experiment. We saw an unexpected increase in the number of children who correctly answered Theory of Mind related questions throughout the curriculum.

Perception of Robots Questionnaire

In the results of *Perception of Robots Questionnaire*, a 3-by-3 Chi-square goodness of fit test found that, on every question, the distribution of children's responses differed significantly from an equal distribution, as shown in Figure 11. Therefore, we used 2-by-2 Chi-square goodness of fit tests with adjusted

Pre-test:
Perception of Robots Questionnaire

Statement	Proportion of Responses: Disagree	Neutral	Agree	χ^2 (*df*=2)	*p*
Robots can learn	15%	19%	66%	32.67	< 1e-4**
Robots always follow the rules	3%	35%	62%	30.23	< 1e-4**
Robots are more like children than adults	10%	45%	45%	23.24	< 1e-4**
Robots are more like people than toys	11%	60%	29%	22.29	< 1e-4**
Robots are smarter than me	21%	62%	17%	15.61	< 1e-4**

Figure 11: Results of the *Perception of Robots Questionnaire* pre-test. We used a Chi-square test to see if the observed frequencies differ greatly from the expected frequency of children being equally likely to choose either response.

$\alpha = 0.017$ to determine if any response was significantly likely or unlikely. Most children felt that robots can learn (66%; Agree vs. Disagree $\chi^2 = 19.22, p < 0.0001^{**}$; Agree vs. Neutral $\chi^2 = 14.8, p = 0.0001^{**}$) and very few children disagreed that robots always follow the rules (3%; Agree vs. Disagree $\chi^2 = 31.6, p < 0.0001^{**}$, Neutral vs. Disagree $\chi^2 = 15.04, p = 0.0001^{**}$). These results suggest that children believed that while robots are able to become smarter, they are still not able to think freely outside of the rules.

Most children were unsure of whether or not robots are smarter than themselves (62%; Agree vs. Neutral $\chi^2 = 14.58, p = 0.0001^{**}$, Neutral vs. Disagree $\chi^2 = 12.05, p = 0.0005^{**}$) and whether they were more like people or toys (60%; Agree vs. Neutral $\chi^2 = 5.9, p = 0.015^{*}$, Neutral vs. Disagree $\chi^2 = 19.12, p < 0.0001^{**}$). Children's indecisiveness about whether robots were smarter or not was also seen in an earlier study where younger children (4-6-years old) were less sure about the robot's intelligence than their 7-10-year-old counterparts [6]. Finally, we saw that very few children disagreed that robots are more like children than adults (10%, ; Agree vs. Disagree $\chi^2 = 12.98, p = 0.0003^{**}$, Neutral vs. Disagree $\chi^2 = 12.98, p = 0.0003^{**}$). When reasoning about their answers, children often referred to their previous attributions. Some children believed the robot was not like an adult because adults are usually smarter than they are. Others who said the robot was like an adult, reasoned that robots must always follow rules because adults did.

In separating children by grade, we saw that Pre-K children were significantly more likely to agree that robots are more like children than adults (77%) compared to Kindergarteners (28%; $\chi^2 = 18.16, p < 0.0001^{**}$). This was driven by the fact that all of the four-year-olds said robots were like toys and almost all thought that robots were like children. Children's reasons for their answer on this question did not make it apparent why these differences existed. One possible explanation is that, as found in previous studies, older children are more sensitive to an object's intentionality and draw bigger distinctions between a device that is acting of its own volition versus being propelled or controlled externally [7, 8]. Therefore, perhaps older children had a higher regard for the autonomy of the robot and saw robots as more similar to people, while younger children saw robots as toys and associated them more with children. However, on both of these questions, half of the children answered neutrally so perhaps most children would be unsure unless some other experience caused them to lean in a particular direction.

Post-test:
Perception of Robots Questionnaire

Statement	Proportion of Responses (Disagree / Neutral / Agree)	χ^2 (df=2)	p
Robots can learn	13% / 12% / 75%	40.65	< 1e-4**
Robots always follow the rules	14% / 39% / 47%	10.21	6.1e-2**
Robots are more like people than toys	31% / 33% / 35%	0.04	0.98
Robots are smarter than me	36% / 38% / 26%	1.48	0.48
Robots are more like children than adults	21% / 49% / 25%	2.98	0.23

Figure 12: Results of the *Perception of Robots Questionnaire* post-test. We used a Chi-square test to see if the observed frequencies differ greatly from the expected frequency of children being equally likely to choose either response.

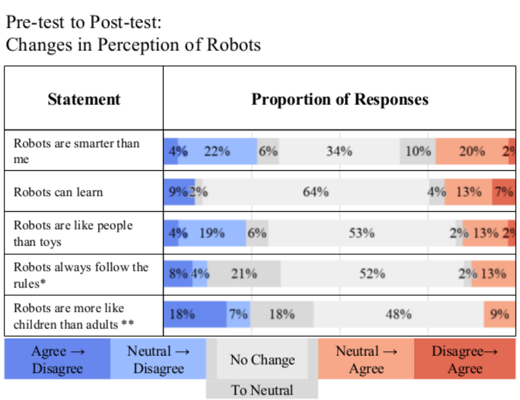

Figure 13: Changes in responses from pre-test to post-test on the *Perception of Robots Questionnaire*. Blue represents the number of children who moved to disagree. Red is the number of children who moved to agree. Gray represents the children who moved to neutral (dark gray) or did not change their answer at all. We used a Wilcoxon signed rank test to determine if the amount of change was significant.

In the post-test, more children agreed that robots could learn (75%) than in the pre-test. This makes sense given that robot "learned" the rules of *Rock-Paper-Scissors*, healthy and unhealthy food classifications, and emotional music parameters from the children in the activities. On the other questions, we saw that fewer children gave neutral responses. This was driven by the fact that groups of children who initially responded with 'Neutral' changed their answers to 'Agree' or 'Disagree' in the post-test. This supports our hypothesis that after learning about AI, children were better equipped to express their understanding of AI.

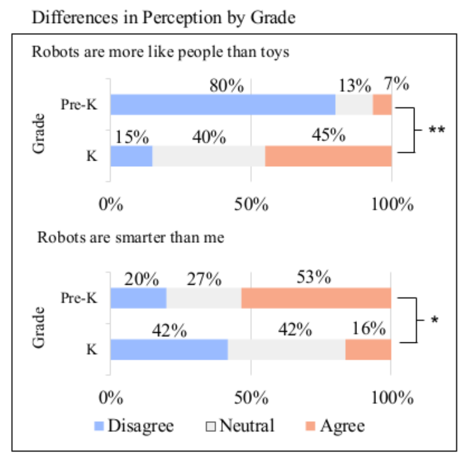

Figure 14: In the *Perception of Robots* post-test, Pre-K children felt much more strongly that robots were like toys and smarter than them compared to the Kindergarten children.

Figure 13 shows how children's responses changed from pre-test to post-test. According to a Wilcoxon signed rank test, there was a significant amount of change in children's answers on two questions: *Robots always follow the rules* ($Z = -2.09, p = 0.039^{*}$) and *Robots are more like children than adults* ($Z = 3.22, p = 0.0010^{**}$). A lot of children who initially said that robots always followed the rules changed their minds, likely influenced by the Generative AI activity where the robots demonstrated creativity. Then, on the other question, a lot of children decided that robots are more like adults than children.

Some of the changes in perceptions of robots were dependent on the grade of the child. On the pre-test, many Pre-K children were not sure if robots were smarter or like people than toys, but after the *PopBots* activities they were more decisive. More than half of the Pre-K children said that robots were smarter than them, compared to only 16% of Kindergarten children ($\chi^2 = 8.05, p = 0.0178^{*}$) (Figure 14). Interestingly, Pre-K children also came to see robots more like toys than people. On the post-test 80% of Pre-K children said the robots were more like toys, compared to 15% of Kindergarten children ($\chi^2 = 21.1, p < 0.0001^{**}$). Pre-K children did not see a contradiction in a toy being more intelligent than themselves.

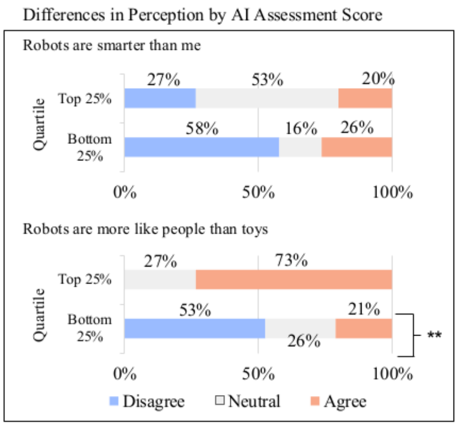

Figure 15: In the *Perception of Robots* post-test, children whose AI assessment scores were in the bottom quartile were more likely to disagree that robots were smarter than them. Children in the top quartile were much more likely to agree that robots were more like people than toys.

Finally, we saw differences in the *Perception of Robots Questionnaire* between children who performed the best in the *PopBots* AI assessments and those who performed the worst (Figure 15). Of those in the bottom quartile (Cumulative AI assessment score $\bar{x} = 46.5\%$), 58% believed that robots were not as smart ($\chi^2 = 5.65, p = 0.059$) as they were and 53% of them believed that robots were more like toys than people. Comparatively, many more of the children in the top quartile (Cumulative AI assessment score $\bar{x} = 92.5\%$) saw robots more as people than toys (73%; $\chi^2 = 13.1, p = 0.001^{**}$). Children's responses in these questions depended on their ability to understand robots as intellectual others. Children in the bottom quartile might not have been able to grasp the robot's cognition as readily as those in the top quartile which is most likely the cause of this large difference in perception.

We saw that before and after the *PopBots* activities, children's previous experiences with robots strongly impacted their perceptions. Here is a portion of a discussion between Jane (6-years-old) and Robert (5-years-old) about two robots that they had seen before. Robert had played on a robot-shaped jungle gym and Jane had seen a glass robot in a movie. Although both children had already completed the *PopBots* activities at this point, some of their ideas about robots were still colored by these experiences. In Robert and Jane's case, they both thought that *PopBots* were like toys, but they had prior experience with larger robots that they considered real.

Researcher: *Are robots toys or are they like people?*
Robert: *Both because sometimes you can go in it.*
Jane: *You can't even go in it.*
Robert: *Sometimes there's a ladder and you can climb and go in it.*
Jane: *No because it's made out of glass. And the robot's not even real. So you can't go in it. They're just toys and you can only play with them. The only real robots are in movies.*
Researcher: *Well, we have robots in my lab.*
Jane: *But they're not real robots. Robots like tell you what to do and [say] whatever you say.*

7 DISCUSSION

We felt that it was important to evaluate the *PopBots* curriculum in classrooms, however, this led to some limitations with the interpretation of the results. The duration of the experiment was limited due to the needs of teachers. Children spent at most 15 minutes with each activity before doing the AI assessments. Another limitation was that children completed the activities in groups rather than individually. In the *PopBots* pilot studies children spent an hour or more working with each activity and were evaluated individually, while likely led to a deeper understanding of the material. In the classroom setting, children's understanding and perceptions of AI were impacted by who was in their group and what was going on in the classroom. Working in groups allowed children to learn and draw conclusions together, but limits the generalizations we can make about our results.

How did developmental factors, like perspective taking skills, impact what children could learn about AI? By taking concepts out of an abstract, mathematical realm and placing them into hands-on activities that rely more on social cognition, *PopBots* was able to help children gain an understanding of AI algorithms. We saw that most children understood presented AI concepts as assessed by the *PopBots* assessments, with a median score of 70%. There was a slight, but not significant, difference in age – the median assessment score for Pre-K children was 63.3% while for Kindergarteners it was 70.0%.

Given that children needed to develop a mental model of the robot's knowledge to understand the *PopBots* activities, we hypothesized that children's perspective-taking skills would limit how much they understood. We used a *Theory of Mind Assessment* similar to that developed by Wellman and Liu to assess children's understanding of knowledge access, content false belief, and explicit false belief [20]. Although many children had not yet fully developed these Theory of Mind skills, when children were challenged to leverage these skills in the *PopBots* assessment, many still answered correctly. This was surprising given that children's Theory of Mind skills naturally develop with age [20] and we did not design *PopBots* activities to help children improve these skills. Therefore, we believe that children's understanding of AI, through the mind of the robot, boosted their Theory of Mind skills.

How did children's perceptions of "thinking machines" change after they engaged in educational AI activities? In a recent study with smart toys and children aged 4–6, we found that young children were more indecisive than their older counterparts when deciding whether AI-enabled devices are intelligent, human-like, or adult-like [6]. In the *Perception of Robots Questionnaire* pre-test, we saw similar results – children often responded with 'Neutral' or 'I'm not sure.' Spending time engaging with the *PopBots* activities led to less indecision and more differences between children. We found, as in previous studies, a relationship between age, how much children understood about robots, and how much they anthropomorphized them [14]. Pre-K children, unlike Kindergartners, were more likely to see the robots as toys. Children in the bottom quartile of performance on the AI assessments, opposite of those in the top, saw the robots more like toys and did not consider them very intelligent.

Before and after the *PopBots* activities, children's reasoning about their answers often referenced fictional robots or things they "just knew" about robots. These findings are in line with our previous study on children's intuitive attributions of intelligence, in which we found that children's understandings of robots are strongly impacted by external influences like their parent's perspectives [7]. After the *PopBots* activities, children's previous experiences with robots still had a powerful hold on their perceptions, emphasizing the importance of early childhood AI education as a means to help next generation AI citizens to form unbiased beliefs about AI technology.

8 DESIGN RECOMMENDATIONS

Earlier exposure to "Technology and Engineering". In the movement towards STEAM in early classroom education, we should not forget to teach children about technology and engineering. Before the *PopBots* activities, a total of 23 children out of 80 had experience with computational thinking, and these children mostly came from one classroom. After the activities, all children were able to program and build *PopBots* themselves, expressing an interest in wanting to learn more. Innovative toolkits, such as *PopBots*, that demystify the technology that young children encounter in daily life are important.

Bring algorithms down to children's eye level. We see that children benefit from seeing and tinkering with real examples of AI that manifest themselves as functional, entertaining, creative, and assistive devices. By opening up the black box of AI and turning abstract ideas into hands-on activities, very young children were able to understand AI. We should build future AI devices that are more transparent and trainable so that children can more easily relate algorithms to their own cognition. These kinds of changes will give children more agency when interacting with smart devices.

Early intervention with sufficient duration. We saw that young children had fluid ideas about objects that cross ontological categories (e.g., alive/not alive). Children's prior experiences and opinions create strong attitudes that can persist even after brief exposure to the *PopBots* platform, as we found in our work. Therefore, it is important to understand each child's unique perspectives and expose children to more information about AI so they can develop an informed understanding. This will require children to be able to explore AI not only in schools, for a few hours, but also in their homes. Platforms like ScratchJr. are a good example of how to bridge this gap [3].

9 CONCLUSION AND FUTURE WORK

In an increasingly AI-powered society, it is important to consider citizen's AI literacy – how much do people really understand AI? Experts have voiced concerns about a global AI skills gap crisis [17]. In order to lessen the gap, it is important to democratize who can access and create with AI. We advocate that now is the time to actively work toward early childhood, inclusive AI education. This work is a novel step toward *Early AI Literacy* that proposes hands-on activities, workshop design, supporting toolkits, and a set of assessments on perceptions of AI and AI learning outcomes. We found that preschool-aged children can learn about AI concepts through appropriately framed content. Growing up with such sense of empowerment about AI concepts and technologies is crucial even for young children as they are already starting to interact with smart toys and smart speakers at home.

In the future, we will explore and develop new activities to expand children's understanding of AI. For example, designing the *PopBots* to interactively assess children's learning and ask questions, and having children program and train robots with different forms. We also think there is an important opportunity to introduce children to the ethical design of AI, understanding how to design and train systems to address bias and promote fairness. Our hope is that this work is useful to educators and parents, and well as for companies building AI-enabled products for use by families and children.

ACKNOWLEDGMENTS

This work was sponsored by Samsung Electronics NBNL, National Science Foundation Grants CCF-1138986 and Graduate Research Fellowship No. 1745302. We would also like to thank the teachers, parents, and students who participated in the experiment and undergraduate students who helped with development and analysis.

REFERENCES

[1] Yvonne Barnes-Holmes, Louise McHugh, and Dermot Barnes-Holmes. 2004. Perspective-taking and Theory of Mind: A relational frame account. *The Behavior Analyst Today* 5, 1 (2004), 15.

[2] Debra Bernstein and Kevin Crowley. 2008. Searching for signs of intelligent life: An investigation of young children's beliefs about robot intelligence. *The Journal of the Learning Sciences* 17, 2 (2008), 225–247.

[3] Marina Umaschi Bers. 2017. *Coding as a playground: Programming and computational thinking in the early childhood classroom.* Routledge.

[4] Jacques Bughin, E Hazan, S Ramaswamy, M Chui, T Allas, P Dahlström, N Henke, and M Trench. 2017. Artificial intelligence–the next digital frontier. *McKinsey Glob Institute URL https://www. mckinsey. de/files/170620_studie_ai. pdf* (2017).

[5] Massachusetts DOE. 2018. School and District Profiles. http://profiles.doe.mass.edu/

[6] Stefania Druga, Randi Williams, Cynthia Breazeal, and Mitchel Resnick. 2017. Hey Google is it OK if I eat you?: Initial explorations in child-agent interaction. In *Proceedings of the 2017 Conference on Interaction Design and Children.* ACM, 595–600.

[7] Stefania Druga, Randi Williams, Hae Won Park, and Cynthia Breazeal. 2018. How smart are the smart toys?: Children and parents' agent interaction and intelligence attribution. In *Proceedings of the 17th ACM Conference on Interaction Design and Children.* ACM, 231–240.

[8] Jennifer L Jipson and Susan A Gelman. 2007. Robots and rodents: Children's inferences about living and nonliving kinds. *Child development* 78, 6 (2007), 1675–1688.

[9] Peter H Kahn Jr, Takayuki Kanda, Hiroshi Ishiguro, Nathan G Freier, Rachel L Severson, Brian T Gill, Jolina H Ruckert, and Solace Shen. 2012. "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. *Developmental psychology* 48, 2 (2012), 303.

[10] MIT Media Lab Lifelong Kindergarten. 2018. Scratch Blocks. https://github.com/LLK/scratch-blocks

[11] Jin Joo Lee, Cynthia Breazeal, and David DeSteno. 2017. Role of speaker cues in attention inference. *Frontiers in Robotics and AI* 4 (2017), 47. https://doi.org/10.3389/frobt.2017.00047

[12] Sharona T Levy and David Mioduser. 2008. Does it "want" or "was it programmed to"? Kindergarten children's explanations of an autonomous robot's adaptive functioning. *International Journal of Technology and Design Education* 18, 4 (2008), 337–359.

[13] Emily McReynolds, Sarah Hubbard, Timothy Lau, Aditya Saraf, Maya Cakmak, and Franziska Roesner. 2017. Toys that listen: A study of parents, children, and internet-connected toys. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.* ACM, 5197–5207.

[14] David Mioduser and Sharona T. Levy. 2010. Making sense by building sense: Kindergarten children's construction and understanding of adaptive robot behaviors. *International Journal of Computers for Mathematical Learning* 15, 2 (01 Jul 2010), 99–127. https://doi.org/10.1007/s10758-010-9163-9

[15] Hae Won Park, Mirko Gelsomini, Jin Joo Lee, and Cynthia Breazeal. 2017. Telling stories to robots: The effect of backchanneling on a child's storytelling. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17).* ACM, New York, NY, USA, 100–108. https://doi.org/10.1145/2909824.3020245

[16] Michael Scaife and Mike van Duuren. 1995. Do computers have brains? What children believe about intelligent artifacts. *British Journal of Developmental Psychology* 13, 4 (1995), 367–377.

[17] Tencent. 2017. 2017 Global Artificial Intelligence White paper. http://www.tisi.org/Public/Uploads/file/20171201/20171201151555_24517.pdf

[18] Sherry Turkle. 1984. *The second self: Computers and the human spirit.* Simon & Schuster, Inc., New York, NY, USA.

[19] Sherry Turkle, Cynthia Breazeal, Olivia Dasté, and Brian Scassellati. 2006. Encounters with kismet and cog: Children respond to relational artifacts. *Digital media: Transformations in human communication* 120 (2006).

[20] Henry M Wellman and David Liu. 2004. Scaling of theory-of-mind tasks. *Child development* 75, 2 (2004), 523–541.

[21] Jacqueline Marie Kory Westlund, Sooyeon Jeong, Hae Won 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. 11 (06 2017).

[22] Randi Williams. 2018. *Leveraging social robots to aid preschool children's artificial intelligence education.* Master's thesis. MIT. https://www.media.mit.edu/publications/pop-study-2018/