A Model-free Affective Reinforcement Learning Approach to Personalization of an Autonomous Social Robot Companion for Early Literacy Education

Hae Won Park[1], **Ishaan Grover**[1], **Samuel Spaulding**[1], **Louis Gomez**[1,2], **and Cynthia Breazeal**[1]
[1]Personal Robots Group, MIT Media Lab
20 Ames Street, Cambridge, Massachusetts, 02142
{haewon,igrover,samuelsp,cynthiab}@media.mit.edu
[2]Electrical Engineering & Computer Science, Wichita State University
1845 Fairmount, 249 Jabara Hall, Wichita, KS 67260
lagomez@shockers.wichita.edu

Abstract

Personalized education technologies capable of delivering adaptive interventions could play an important role in addressing the needs of diverse young learners at a critical time of school readiness. We present an innovative personalized social robot learning companion system that utilizes children's verbal and nonverbal affective cues to modulate their engagement and maximize their long-term learning gains. We propose an affective reinforcement learning approach to train a personalized policy for each student during an educational activity where a child and a robot tell stories to each other. Using the personalized policy, the robot selects stories that are optimized for each child's engagement and linguistic skill progression. We recruited 67 bilingual and English language learners between the ages of 4–6 years old to participate in a between-subjects study to evaluate our system. Over a three-month deployment in schools, a unique storytelling policy was trained to deliver a personalized story curriculum for each child in the Personalized group. We compared their engagement and learning outcomes to a Non-personalized group with a fixed curriculum robot, and a baseline group that had no robot intervention. In the Personalization condition, our results show that the affective policy successfully personalized to each child to boost their engagement and outcomes with respect to learning and retaining more target words as well as using more target syntax structures as compared to children in the other groups.

Introduction

Early literacy and language skills are a significant precursor to children's later educational success (Páez, Tabors, and López 2007; Hart and Risley 1995). Quality preschool programs support the development of key pre-literacy skills such as phonological awareness, alphabetic knowledge, and core vocabulary. These foundations support the development of literacy skills in later grades and can help prevent academic failure (Hart and Risley 1995; Fish and Pinkerman 2003; Páez, Tabors, and López 2007; Snow et al. 2007). Yet, only about 32.7% of eligible 4-year-olds attend preschool (National Institute of Early Education Research 2018). It is noteworthy that in 2015, only 35% of 4th, 36% of 8th and 38% of 12th grade students scored at reading proficiency levels on the National Assessment of Educational Progress (NAEP) (National Center for Educational Statistics 2016). Many of these students do not have learning disabilities but still struggle to comprehend complex text.

School systems are under great pressure to provide instruction in kindergarten that remedies areas of early literacy and English language weakness before children enter primary school. In early childhood, there is a close dependency between oral language development and learning to read (Snow 1991). Language skill development, in turn, requires sufficient exposure to a rich variety of vocabulary in context and spoken language with others (Asaridou et al. 2016). Simply hearing language through passive listening is not enough. Young children need to actively use language while being emotionally and physically engaged in communication to maximize their learning gains (Romeo et al. 2018; Wells 2000).

Unfortunately, a "participation gap" has been identified to exist between families coming from different socioeconomic status (SES) backgrounds with respect to the amount of daily parent-child interactions and active parental involvement in their child's language and early literacy development at home (Romeo et al. 2018; Neuman and Celano 2012). As a result, many preschool age children from low SES families have significantly smaller vocabularies and less-developed vocalizations than their high SES counterparts (Gilkerson et al. 2017). These differences often become magnified over time (Hart and Risley 1995).

In at-risk communities, it is very difficult for a kindergarten teacher to offer a curriculum that addresses the wide diversity of cognitive and pre-literacy starting points at which children enter school. When a child enters kindergarten, she is a unique distribution of the various cognitive, visual, social and linguistic skills needed to be a successful reader (Wolf and Gottwald 2016; Dehaene 2009). Young children would clearly benefit from personalized instruction and active language-based interaction that can measure and adapt to the many intersecting domains of skills and abilities that support the process of learning to read. This reality motivates the development of AI technologies that can continuously assess and effectively personalize to meet individual children's diverse needs. Prior success with middle and high school age students has shown that intelligent tutoring systems (ITS) can automatically assess and adapt to student skill levels to positively impact student learning gains (Corbett and Anderson 1994; Desmarais and Baker 2012; Yudelson, Koedinger, and Gordon 2013). However, to support long-term personalized interaction with preschool-age children, a more engaging, age-appropriate, and autonomous assessment and intervention should be implemented (Woolf 2010).

Social-robot learning companions hold great promise for augmenting the efforts of parents and teachers to promote learning, academic knowledge and positive learning attitudes (Belpaeme et al. 2018). Social robots can physically, socially, and emotionally play and engage with children in the real world with verbal and non-verbal speech acts that resemble those between peers or adults. They can be designed to interact with children in a collaborative, peer-like way during playful educational activities (Michaelis and Mutlu 2018; Park et al. 2017c; Baxter et al. 2017; Park and Howard 2015). Social robots can offer a unique opportunity to personalize social interactions to promote areas of language development important for early literacy skills, learning to read, and academic success.

Though the development of personalized robot tutors has gained increased attention (Leyzberg, Spaulding, and Scassellati 2014; Baxter et al. 2017; Kory 2014), the development and assessment of a fully autonomous, personalized learning companion robot that improves engagement and learning outcomes for young children over months remains a challenge. In this paper, we present algorithmic methods and tools, developed and evaluated in real world contexts, to advance the development of social robots that can personalize to children and sustain engagement over months to foster early literacy and language skills of young English language learners (ELLs). Children engage with social robots in an emotive and relational way, and the affective cues that social robots can elicit from children provides an opportunity to better assess their engagement states. Our affective reinforcement learning personalization policy takes full advantage of these engagement cues. Our robot was able to collect a unique corpus of verbal and non-verbal behaviors as children engaged in a dialogic storytelling task with it. The robot employed a personalization policy trained over months of interaction with Q-learning, a model-free reinforcement learning approach. We compare children's engagement and learning outcomes with a personalized learning companion robot to a non-personalized version (following a fixed curriculum) and a baseline group (with no robot). Our results show that the learned policy effectively personalized to each child and served to boost their engagement and language skills.

Interaction Design

Dialogic storytelling is one of the most important educational activities in kindergarten. Rather than simply reading or telling stories to children, the educator actively involves children in the story by asking questions, and actively listening and responding to children's contributions. We developed a fully autonomous social robot that can engage children between the ages for 4–6 years old in dialogic storytelling activities. The robot records each session and is able to automatically assess children's nonverbal affect cues, as well as analyze their speech samples to assess their lexical and syntax skills.

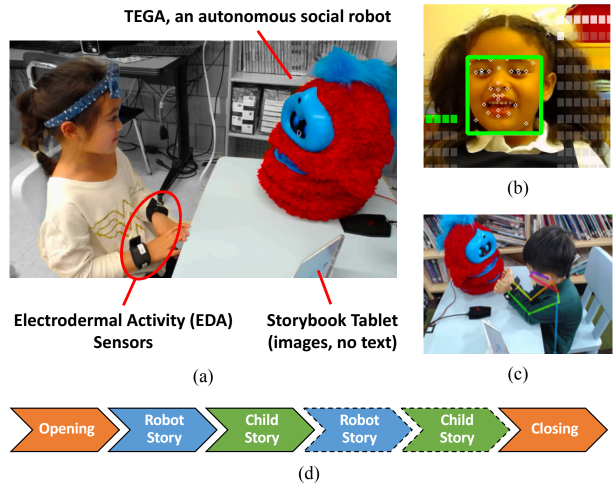

Figure 1: Study setup and phases. (a) The Tega robot and child interact face-to-face with an accompanying digital storybook on a tablet also facing the child. Electrodermal activity (EDA) data is collected by E4 sensors worn on each wrist. (b) A front-facing camera view is used track and recognize children's facial affect. (c) A bird's-eye camera view is used for 2D body-pose tracking. (d) Interaction phases in our study protocol.

Social Robot Platform: The Tega robot is an expressive social robot designed to appeal to and engage with young children as a playful learning companion. The robot supports long-term, real-world deployments in various educational settings such as homes, schools, and therapeutic centers. We designed the robot to interact with children as a peer given the known social influence that children of similar ages can have on each other's language, behavior, and attitudes. Our robot's peer-like attributes include a child-like voice, emotionally expressive body movements, animated facial expressions, and non-verbal cues such as backchanneling to signal active listening while attending to the child (Park et al. 2017b; 2017a; 2017c).

Interaction Phases: Our 12-week study protocol was comprised of several phases. First, we administered pretests to assess children's vocabulary and oral syntax skills. This data was also used to counterbalance children across the three conditions. Children then interacted with the Tega robot once per week over 6–8 sessions in a dialogic storytelling activity (Figure 1). In the *Opening* phase, the child wakes the robot up and the robot greets the child and engages in brief social pleasantries, asking the child about school life, favorite stories, etc. In the *Robot Story* phase, the robot selects and presents a story from its library (comprised of 81 children's picture books) to the child. As the robot tells the story and asks the child questions, the artwork of the story is shown on the tablet. Only the illustrations are shown to isolate the effect of the robot's oral storytelling on the child's learning and to avoid any effects of textual prompts in case the child can read. The types of questions the robot asks are lexical, factual, inferential, and emotional in nature. By asking these questions, the robot has the opportunity to assess the child's engagement as well as her comprehension of the story content. After telling its story, the robot invites the child to retell the story, using the story illustrations on the tablet as a guide if desired (*Child Story* phase). Each story sample from the child provides new observations to measure her lexical and syntactic skill growth. When the child tells the robot that she is finished with storytelling, the robot provides some brief comments, says farewell, and goes back to sleep (*Closing* phase). Depending on the robot's story length, two short stories were exchanged in some sessions instead of one longer story. Once all storytelling sessions were completed, we administered a final post-assessment.

Data Collection: In addition to the recorded dialogues and task state data, children's nonverbal cues were recorded throughout each session (Figure 1). Their facial affect features were extracted from the front-view camera using Affdex (McDuff et al. 2016). We used the arousal metric as a state space feature in our reinforcement learning algorithm for personalization. We also collected electrodermal activity (EDA) data from sensors worn on both children's wrists, starting at least 5 minutes before the *Opening* phase to allow the sensor values to stabilize. EDA is reported to correlate with user engagement state and has been tested on young children for its efficacy (Hernandez et al. 2014). We also post-analyzed the birdeye-view camera frames for children's body pose using OpenPose (Cao et al. 2017). We found that both EDA and leaning-forward body pose data were strongly correlated with engagement (see Results section).

Language Skill Assessments: In the pre/post-test sessions, children's vocabulary and syntactic skills were assessed using a clinically evaluated vocabulary test, target vocabulary test, and personal narrative collection protocols. The Peabody Picture Vocabulary Test (PPVT) (Dunn and Dunn 2007) was used as a measure of children's vocabulary level, and the same format (four picture-based vocabulary identification assessment) was used for the target vocabulary test. Children's personal narratives were collected using Westerveld and Gillon's language sampling protocol (Westerveld and Gillon 2002). The narrative samples were evaluated using the Index of Productive Syntax (IPSyn) (Scarborough 1990) that evaluates the grammatical complexity of spontaneous language samples. We evaluated children's syntax complexity from collected narrative samples using IPSyn's noun phrase, verb phrase, and sentence structure scales.

Story Corpus: A touchscreen tablet was used to show storybook illustrations to the child without text (Park, Coogle, and Howard 2014). The robot's library of 81 storybooks were curated by early literacy experts (from Tufts University's Reading and Language Research Center) to span across the spectrum of age appropriate lexical and syntactic complexity. All titles are commercially available. Every book in the corpus was analyzed for lexical and syntactic complexity using our automatic assessment tools.

In even-numbered sessions, the robot presented a target storybook to all participants. We replicated the target-word protocol from Collins' work on studying preschool English language learner's vocabulary acquisition from listening to teacher's storytelling, and we used the same 4 books referenced in (Collins 2010). Between 5–8 target words were either inserted at appropriate locations in the story or replaced the original words in the books. Of the 25 target words, 6 were verbs, 12 were nouns, and 7 were adjectives. The target words were selected based on 1) applicability to the story so that the target words made sense when replaced with the original word, such as "put on" $\rightarrow$ "donned" and "noise" $\rightarrow$ "clamor", 2) frequency of occurrence so that the word appears exactly twice in the story with accompanying visual illustration, and 3) unfamiliarity of the word to our target-age audience, determined by oral and written word lists (Chall and Dale 1995; Carroll et al. 1971). We cross-checked to confirm that the target words do not appear in the 'First 1,000 English Words' list (Browne and others 2013). Neither do they appear in any of the other storybooks in our corpus. The target vocabulary pre-test administered to all participants showed that children performed no better than chance on these target words (error rate above 75% for multiple choice with four answer choices).

Affective Personalization Policy and Robot Action Selection

We hypothesize that interaction samples collected through repeated encounters with a given child is sufficient to train a personalized robot-action policy that would improve engagement and learning outcomes for that child. Interactions between our robot each of the participants in the personalized (P) group provided episodes of $< s, a, r, s' >$. The tuple represents that the user was in state s, the robot provided action (i.e., story sentences) a, the robot received reward r for its action, and the users state changed to state s' i response to robot's action. These episodes are used to train each user's personalization policy, $Q[s, a]$. In this section we present our Q-learning approach (Watkins and Dayan 1992)to learn an optimal policy without knowing the model in advance.

For personalized storytelling, the goal is to predict the syntax and lexical complexity levels of the storybook that the robot should present to a given child in order to maximize the child's future engagement and learning gains measured through the child's facial affect expressions, question answering, and story retell narratives. Because the reward for each state-action pair is partly computed using the transcript of the child's story retell, the $Q[s, a]$ table is updated at the end of each child's retell.

State space (S). The state space consists of users' task behavior s_{task} and affective arousal s_{affect}. Task behavior states represent users' question answering behavior during the robot's story narration - $\{$*not answered, answered*$\}$, $\{$*with prompt, without prompt*$\}$, $\{$*length of utterance (LU)* ≤ 2, *LU* $> 2\}$. Since verbal prompts and the length of utterance states only make sense when an answer is given, the total number of s_{task} is $1+2\times 2 = 5$. Affective arousal states represent users' facial muscle activation that illustrates a user's expressiveness and arousal level. The raw value of this metric, an output from Affdex (McDuff et al. 2016), ranges between $[0, 100]$. This range is divided into four states, $\{[0, s_{affect_{q1}}), [s_{affect_{q1}}, s_{affect_{q2}}), [s_{affect_{q2}}, s_{affect_{q3}}), [s_{affect_{q3}}, 100]\}$, where $s_{affect_{qn}}$ is an n-th quartile of the s_{affect} range of an individual participant, measured during the first session before the first training. We implemented this adaptive metric because we discovered from prior experience that young children's affective arousal range can vary widely across individuals. In total, the state space consists of $|s_{task}| \times |s_{affect}| = 5 \times 4 = 20$ states.

Action space (A). The action space of the personalization policy represents the robot's storytelling content, i.e., the lexical and syntactic complexity of a given sentence in a storybook. The telling of each sentence is considered one action. The lexical complexity is determined by whether the sentence has a lemmatized word not in the *known word list (κ)*, $a_{lex} = \{\in \kappa, \notin \kappa\}$. The known word list consists of the first 1,000 English words (Browne and others 2013) plus any words that appeared in a child's narrative or story retell. Thus, for each child we maintained their known word list. The syntactic complexity, $a_{syn} = \{low, med, high\}$, is determined by whether the given sentence is comprised only of syntactic structures that are at a lower level than the child's, $a_{syn} = low$, or has at least one syntactic structure that is of a similar level as the child's, $a_{syn} = med$, or has at least one syntactic structure at a higher level, $a_{syn} = high$. Hence, the action space has $|a_{lex}| \times |a_{syn}| = 2 \times 3 = 6$ actions in total.

Reward (R). The reward function is a weighted sum of engagement and learning, $r = 0.5 \cdot engagement + 0.5 \cdot learning$, $-100 \leq r \leq 100$. The intention behind this function is to reward new lexical and syntax learning while bounding it with engagement so that the algorithm doesn't always select a story with the highest level of linguistic complexity. The engagement and learning rewards are computed as follows:

$$engagement = \begin{Bmatrix} 25 \cdot (n-5), & \text{question not answered} \\ 25 \cdot n, & \text{question answered} \end{Bmatrix}, \tag{1}$$

$$learning = \begin{Bmatrix} -100, & \text{no matching phrase} \\ 0, & \text{phrase} \in [\{\in \kappa, low\}] \\ +50, & \text{phrase} \in [\{\in \kappa, \{med, high\}\}, \{\notin \kappa, low\}] \\ +100, & \text{phrase} \in [\{\notin \kappa, \{med, high\}\}] \end{Bmatrix}. \tag{2}$$

where $n = \{1, 2, 3, 4\}$ is the n-th s_{affect} quartile, and the lexical and syntactic complexity of a matching phrase is computed the same way as the action space.

Storybook selection. Given the updated user's lexical (κ) and syntactic (IPSyn category probability) levels from the child's latest story retell, the action space in each storybook is recomputed. Employing an ϵ-greedy algorithm and setting ϵ to decrease in each successive session, $\epsilon = \{0.7, 0.6, 0.5, 0.3, 0.2, 0.2, 0.2, 0.2\}$, we choose the storybook that best balances exploration and exploitation. The exploration value ξ of a storybook m, $0 < \xi_m \leq 1$, is computed as follows:

$$\xi_\mu = \sum_{a \in A} |\hat{a}_m| \cdot \frac{1}{1 + \sum_{s \in S} visits(s, a)}, \tag{3}$$

where $|\hat{a}_m| = \frac{|a_m|}{\sum_{a \in A} |a_m|}$ is the occurrence frequency of action a in storybook m, and $visits(s, a)$ is the total number of times a state-action pair (s, a) has been visited. The learning rate α at every iteration also decreases as each episode (s, a) is re-visited, preserving a minimum rate of 0.125:

$$\alpha = max\left(\frac{1}{1 + visits(s, a)}, 0.125\right). \tag{4}$$

We did not employ any specific exploration-exploitation strategy, but estimated the exploration probability of storybooks in the corpus at a given iteration. The mixed lexical and syntactic elements in the storybooks provided a natural selection of exploration and exploitation opportunities. It was indeed observed that as the sessions progressed, more exploitation actions were selected.

Experimental Setup

Study Conditions. Based on the pre-assessment results, children were divided into three counterbalanced groups based on age, gender, school, lexical, and syntax scales. In the Personalization (P) group, the robot trained a personalized action policy for each child to select storybooks that were predicted to deliver the best learning and engagement outcomes for that particular child. The robot's action policy for children in the Non-personalization (NP) group followed a fixed curriculum with storybooks sampled from our story corpus with varying lexical/syntactic complexity. The Baseline (B) group took pre- and post-tests, but did not interact with the robot at all.

Experts at the Tufts Reading and Language Research Center authored three different levels of the four target storybooks (in addition to the embedding the target words in them as mentioned earlier). In even numbered sessions when children heard the target stories, the personalization policy guided the selection of the story level for children in the P group. Children in the NP group heard level 1 of the given target storybook in session 2, level 2 on week 4, level 2 on week 6, and level 3 on week 8. Level 3 was the most difficult in terms of lexical and syntactic complexity.

The only difference between the P and NP groups was the selection of storybooks from the corpus based on the robot's action policy. Other relational factors that could make a session feel more personal to a child, such referring to the child by name, opening/closing conversations, or the robot's expressiveness were kept consistent across conditions.

Participants. We recruited 73 English language learners (ELL) and bilingual children between the ages of 4–6 years from local public schools in the Greater Boston area[1]. Afterwards, six children who earned significantly higher vocabulary and syntax scores in the pre-test compared to the

[1]The study protocol was approved by the Institutional Review Board (IRB) and parental consent was collected for children who participated in this study. The use of images of children that appear in this paper was approved by each child's parents or guardians.

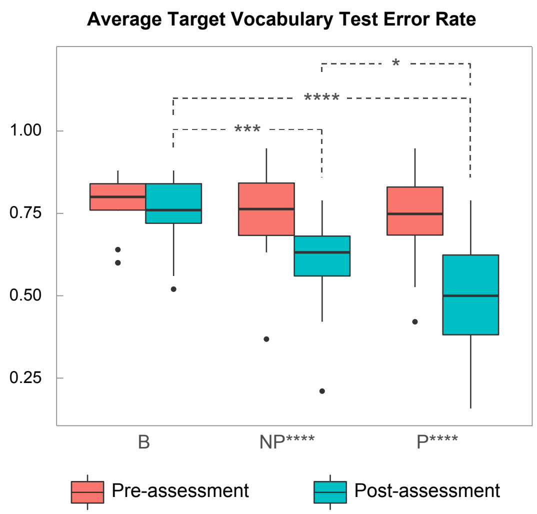

Figure 2: Children learned more vocabulary from interacting with the personalized robot. Children who interacted with the non-personalized robot learned more words than children who did not interact with either robot.

rest of the children ($p = 0.013$) were excluded from the study. In total, 67 children (55% female, age $\mu = 5.36$ years, $\sigma = 0.62$) from 12 classrooms across three schools participated in the study. Later, two children from the NP group withdrew from the study mid-way, due to their family moving out of the country. Over the course of a three-month deployment, a unique storytelling policy was trained to deliver a personalized story curriculum for each child in the Personalized group ($N = 22$). We compared their engagement and learning outcomes to the Non-personalized group ($N = 22$, fixed curriculum robot) and the Baseline group ($N = 23$, no robot intervention).

Results

Children learned more words from a personalized robot peer. Since participants were divided into counterbalanced groups, there was no significant difference in the target vocabulary pre-test scores across conditions (B: $\mu = .79$, $\sigma = .09$; NP: $\mu = .75$, $\sigma = .13$; P: $\mu = .74$, $\sigma = .13$). However, in the post-test, we saw significant effect of the robot's policy (B: $\mu = .75$, $\sigma = .10$; NP: $\mu = .61$, $\sigma = .13$; P: $\mu = .51$, $\sigma = .16$). With paired t-test analysis, both NP and P groups showed a significant interaction effect pre-to-post (NP: $t(19) = 10.02$, $p < .0001$; P: $t(21) = 10.40$, $p < .0001$). Both NP and P groups also showed a significant interaction effect compared to the baseline group. A Welch's independent samples t-test yields (B vs. NP: $t(34.88) = 3.65$, $p < .001$, Cohen's $d = 1.18$; B vs. P: $t(35.77) = 5.65$, $p < .0001$, Cohen's $d = 1.72$). Most notably, we found a significant effect for personalization. With Welch's t-test, (NP vs. P: $t(39.29) = 2.27$, $p = .028$, Cohen's $d = .69$) (Figure 2).

The Storyteller policy effectively personalized to each child. A root mean square (RMS) policy distance between

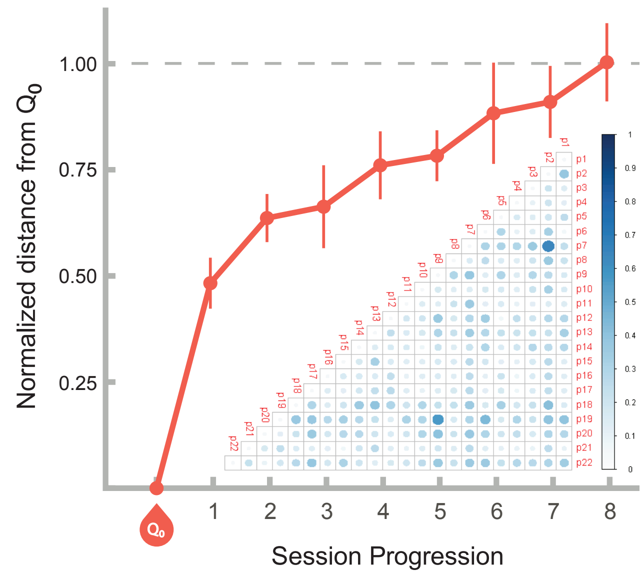

Figure 3: The storyteller policy personalized to each child in the Personalized group. The plot shows the average RMS distance of policies in each session from the initial policy Q_0. The confusion matrix depicts the rate of the number of matching maximum reward state-action pairs between two policies summed after each session over all sessions.

two sessions l and m of user x is computed as,

$$d_x^{l,m} = \sqrt{\frac{\sum_{s \in S} \sum_{a \in A} (Q_x^l[s,a] - Q_x^m[s,a])^2}{|S| \times |A|}}. \quad (5)$$

An average distance between the original policy and policies in each session over all participants shows that the distance increases as the sessions progress, suggesting that the policy evolved over time (Figure 3). The policy did not converge after $N = 8$ sessions. This was expected as there are 120 state-action pairs, and the interaction samples are still quite sparse with each pair being visited less than four times on average. Each policy was trained with approximately 423.32 ± 76.78 episodes per person, 59.21 ± 22.23 per session.

The essence of reinforcement learning is in the policy-guided action selection. We analyzed how the maximum reward-yielding action in each state was similar between any two policies. The number of matching actions $\max_a Q_x[s,a] = \max_a Q_y[s,a]$ for every state was summed after every session, and then was divided by the number of states and number of sessions. When there were multiple actions with a maximum reward, it was counted as a match if at least one action overlapped between policies. The rate of the action-match, $\nu_{x,y} \in [0,1]$, is formulated as,

$$\nu_{x,y} = \frac{1}{|S| \cdot N} \sum_{l=1}^{N} \sum_{s \in S} \Big[\max_a Q_x^l[s,a] = \max_a Q_y^l[s,a]\Big]. \quad (6)$$

The embedded confusion matrix in Figure 3 depicts $\nu_{x,y}$ values for each policy pair. The mean and standard devia-

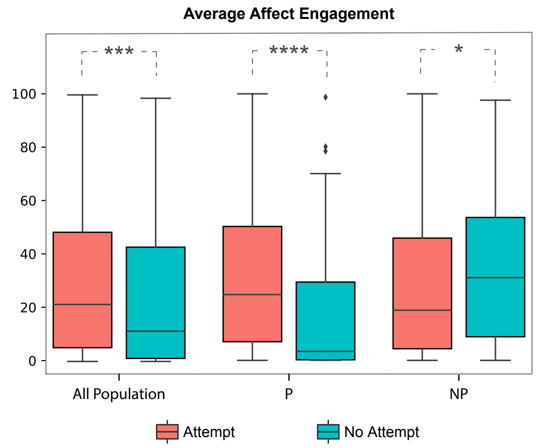

Figure 4: Children learned more vocabulary from interacting with the personalized robot. Children who interacted with the non-personalized robot learned more words than children who did not interact with either robot.

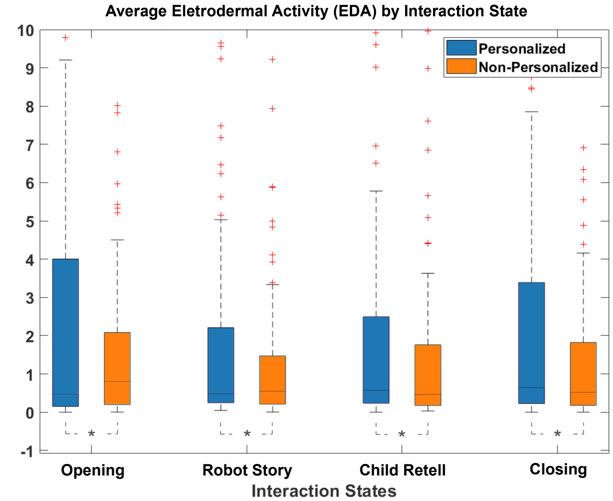

Figure 5: Children in the Personalized condition demonstrated significantly higher Electrodermal Activity (EDA) during all phases of the interaction.

tion of $\nu_{x,y}$ is 0.21 ± 0.095, with maximum value of 0.65 and minimum value of 0.07. This analysis strongly suggests that each policy's action selection strategy evolved quite differently, personalizing to each user over time.

Children's verbal and nonverbal engagement cues are significantly correlated. The state space of the personalization policy consists of user's behavioral engagement cues (i.e., children's verbal responses to the robot's questions and children's facial affect while listening to a story). While the verbal behavior might be regarded as a more direct measure of engagement, we hypothesize that the facial affect cues hold as much information as the verbal cues. We analyzed the correlation between the two signals in the P and the NP conditions.

For all participants in both NP and P groups, when children attempted to answer the robot's question we observed significantly higher affective engagement as compared to when children did not attempt to answer the robot's question (Kruskal-Wallis: $\chi^2(2) = 12.62$, $p < .001$). Children's affective engagement was measured while they listened to the part of the story that held the information relevant to the robot's question (i.e., sentences that just preceded the question). A post-hoc test using Mann-Whitney tests with Bonferroni correction also showed a significant difference between children's attempt to answer and no-attempt ($p < .001$, effect size $r = .095$).

As depicted in Figure 4, it was also observed that this trend was mainly driven by the Personalized group. When children in the P group attempted to answer a question, they showed significantly higher affective engagement compared to when they did not attempt to answer ($\chi^2(2) = 15.32$, $p < .0001$). A post-hoc test also showed a significant difference between attempt and no-attempt ($p < .0001$, $r = .240$).

Children's nonverbal & physiological cues show higher attention and engagement in Personalized condition. Children in the P group showed higher Electrodermal activity (EDA) in all phases of the interaction compared to children in the NP group (*Opening*: $p < .05$, *Robot Story*: $p < .05$, *Child Retell*: $p < .05$, *Closing* $p < .05$). See Figure 5.

We used OpenPose (Cao et al. 2017), an open-source tool for 2D body-pose estimation, to analyze the dynamics of children leaning either toward or away from the robot. Video recordings of each session were programmatically divided into the three interaction phases: *Opening*, *Robot Story* and *Child Retell*.

Our primary measurement was the horizontal slope between the estimated location of the child's neck joint and right or left hip joints to detect when children either leaned toward or leaned away. After this calculation, the slope measurements were smoothed by a median filter of 30 frames to help reduce noise. For each session, we normalized the slope values into the range $[0, 1]$ and binned the values into quartiles. We calculated the percentage of frames belonging to each quartile as an aggregate metric for understanding children's body pose over many sessions. We analyzed the number of transitions across quartiles as a way to approximate short-term leaning behavior. Using the smoothed quartile data, we counted any instance of moving across quartiles as a *Transition*, instances of moving up two quartiles as a *LeanForward* and instances of moving down two quartiles as a *LeanBack*.

Analyzing counts of these short-term pose behaviors reveals significant differences between the P and NP groups (Figure 6). The results show that children in the P condition exhibited significantly more *LeanForward*, *LeanBack*, and overall *Transitions* than children in the NP condition. However this effect is present *only* during the robot's story and child's retell phases of the interaction (*Robot Story*: $p < .05$, *Child Retell*: $p < .05$).

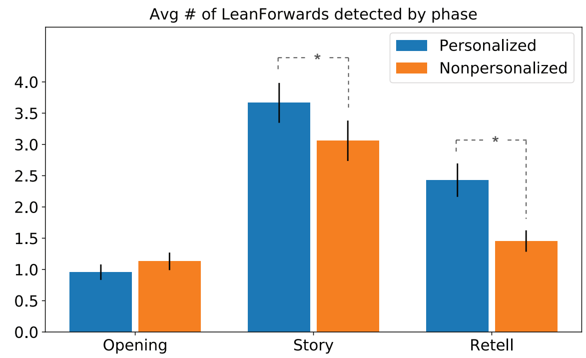

Figure 6: Children in the Personalized condition demonstrated significantly more *LeanForward* behavior during the robot's story and child's retell phases.

Discussion and Conclusion

Our results suggest that all children acquired new vocabulary after storytelling interactions with our robot. While both Personalized and Non-personalized groups performed significantly better than the Baseline group, children who interacted with a robot that learned to personalize to them showed significantly higher engagement and learning outcomes.

Our vocabulary learning results aligns with those reported in (Collins 2010) where a group of ELL children were read eight target stories by a person three times each, over a 3-week period. Children were also given the similar format quizzes before and after. Children in Collins' experimental group were given rich explanations of the target words, whereas in the control condition they received none. They report that children on average scored 0.53 errors in the experimental group, an average of 0.68 errors in the control group, and an average of 0.75 errors in a no-story baseline group. Collins' results are comparable to the results we observed in our study with a robot. This suggests that a storytelling interaction with a personalized and socially relatable robot-peer could be very effective. It further motivates the use of interactive, social, emotive, and relational AI technology to assist teachers and parents to improve children's education.

Our results also suggest that our reinforcement learning (RL) approach to training a personalized interaction policy was effective. We found that our approach yielded personalized policies for each child. This is supported by the accumulated differences between policies as well as differences in the maximum-reward yielding actions in each state. We further provide evidence that verbal and nonverbal engagement cues (i.e., the affective arousal cue and children's question answering behavior) are highly correlated. This supports our decision to use these features in the design of the state space in our RL approach. In order to speed up the convergence of the policy, we may leverage the likely similarity of neighboring states to account for the sparsity of the Q table. We performed further analysis on the affective engagement state transitions and found that an engagement level rarely jumps from extreme to extreme, but rather gradually transitions between states. Using this analysis and modeling a state-transition cost function, each interaction episode should more efficiently train the personalization policy.

Lastly, we evaluated the potential of using other types of nonverbal and physiological data (i.e., body pose and EDA data) to infer children's engagement. We found that children in the Personalized group show higher EDA and more leaning forward behavior while interacting with the robot, signaling higher attention and engagement levels. These results will assist in future development of personalized engagement models.

Our next research goal is to install our robots in schools and homes in areas with high concentration of low-SES families and study their long-term effect. The price of commercialized social robots is dropping at a rapid pace, lowering the adoption barrier of these new technologies.

Acknowledgments

This work was supported by a National Science Foundation grant, IIS-1523118. We sincerely thank Huili Chen, Randi Williams, Pedro Reynolds-Cuellar, Nikhita Singh, David Cruz, Meng Xi, Sooyeon Jeong, and Jacqueline Kory Westlund for helping to develop the storybook assets, conducting user studies, communicating with the schools, and maintaining our ten Tega robots. We also thank Prof. Stephanie Gottwald, Niermala Mohan, Elizabeth Dielentheis, and Kevin Poisson at Tufts University's Reading and Language Research Center for their input and feedback on developing language skill assessments and selecting the storybooks for this study.

References

Asaridou, S. S.; Demir-Lira, Ö. E.; Goldin-Meadow, S.; and Small, S. L. 2016. The pace of vocabulary growth during preschool predicts cortical structure at school age. *Neuropsychologia*.

Baxter, P.; Ashurst, E.; Read, R.; Kennedy, J.; and Belpaeme, T. 2017. Robot education peers in a situated primary school study: Personalisation promotes child learning. *PloS one* 12(5):e0178126.

Belpaeme, T.; Kennedy, J.; Ramachandran, A.; Scassellati, B.; and Tanaka, F. 2018. Social robots for education: A review. *Science Robotics* 3(21):eaat5954.

Browne, C., et al. 2013. The new general service list: Celebrating 60 years of vocabulary learning. *The Language Teacher* 37(4):13–16.

Cao, Z.; Simon, T.; Wei, S.-E.; and Sheikh, Y. 2017. Realtime multi-person 2d pose estimation using part affinity fields. In *CVPR*.

Carroll, J. B.; Davies, P.; Richman, B.; and Davies, P. 1971. *The American Heritage word frequency book*. Houghton Mifflin Boston.

Chall, J. S., and Dale, E. 1995. *Readability revisited: The new Dale-Chall readability formula*. Brookline Books.

Collins, M. F. 2010. Ell preschoolers english vocabulary acquisition from storybook reading. *Early Childhood Research Quarterly* 25(1):84–97.

Corbett, A. T., and Anderson, J. R. 1994. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction* 4(4):253–278.

Dehaene, S. 2009. *Reading in the brain: The new science of how we read*. New York: Penguin Books.

Desmarais, M. C., and Baker, R. S. 2012. A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction* 22(1-2):9–38.

Dunn, L. M., and Dunn, D. M. 2007. *PPVT-4: Peabody picture vocabulary test*. Pearson Assessments.

Fish, M., and Pinkerman, B. 2003. Language skills in low-ses rural appalachian children: Normative development and individual differences, infancy to preschool. *Journal of Applied Developmental Psychology* 23(5):539–565.

Gilkerson, J.; Richards, J. A.; Warren, S. F.; Montgomery, J. K.; Greenwood, C. R.; Oller, D. K.; Hansen, J. H.; and Paul, T. D. 2017. Mapping the early language environment using all-day recordings and automated analysis. *American journal of speech-language pathology* 26(2):248–265.

Hart, B., and Risley, T. R. 1995. *Meaningful differences in the everyday experience of young American children.* Paul H Brookes Publishing.

Hernandez, J.; Riobo, I.; Rozga, A.; Abowd, G. D.; and Picard, R. W. 2014. Using electrodermal activity to recognize ease of engagement in children during social interactions. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 307–317. ACM.

Kory, J. J. M. 2014. *Storytelling with robots: Effects of robot language level on children's language learning*. Ph.D. Dissertation, Massachusetts Institute of Technology.

Leyzberg, D.; Spaulding, S.; and Scassellati, B. 2014. Personalizing robot tutors to individuals' learning differences. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 423–430. ACM.

McDuff, D.; Mahmoud, A.; Mavadati, M.; Amr, M.; Turcot, J.; and Kaliouby, R. e. 2016. Affdex sdk: A cross-platform real-time multi-face expression recognition toolkit. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 3723–3726. ACM.

Michaelis, J. E., and Mutlu, B. 2018. Reading socially: Transforming the in-home reading experience with a learning-companion robot. *Science Robotics* 3(21):eaat5999.

National Center for Educational Statistics. 2016. The condition of education 2015 (nces 2015-144). National Center for Education Statistics. Washington, DC.

National Institute of Early Education Research. 2018. Percent of 4-year-olds enrolled in state pre-k. state of preschool yearbooks 2017.

Neuman, S. B., and Celano, D. 2012. *Giving our children a fighting chance: Poverty, literacy, and the development of information capital*. Teachers College Press.

Páez, M. M.; Tabors, P. O.; and López, L. M. 2007. Dual language and literacy development of spanish-speaking preschool children. *Journal of applied developmental psychology* 28(2):85–102.

Park, H. W., and Howard, A. M. 2015. Retrieving experience: Interactive instance-based learning methods for building robot companions. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 6140–6145. IEEE.

Park, H. W.; Gelsomini, M.; Lee, J. J.; Zhu, T.; and Breazeal, C. 2017a. Backchannel opportunity prediction for social robot listeners. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2308–2314.

Park, H. W.; Gelsomini, M.; Lee, J. J.; and Breazeal, C. 2017b. 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, 100–108. ACM.

Park, H. W.; Rosenberg-Kima, R.; Rosenberg, M.; Gordon, G.; and Breazeal, C. 2017c. Growing growth mindset with a social robot peer. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '17, 137–145. New York, NY, USA: ACM.

Park, H. W.; Coogle, R. A.; and Howard, A. 2014. Using a shared tablet workspace for interactive demonstrations during human-robot learning scenarios. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2713–2719. IEEE.

Romeo, R. R.; Segaran, J.; Leonard, J. A.; Robinson, S. T.; West, M. R.; Mackey, A. P.; Yendiki, A.; Rowe, M. L.; and Gabrieli, J. D. 2018. Language exposure relates to structural neural connectivity in childhood. *Journal of Neuroscience* 38(36):7870–7877.

Scarborough, H. S. 1990. Index of productive syntax. *Applied psycholinguistics* 11(1):1–22.

Snow, C. E.; Porche, M. V.; Tabors, P. O.; and Harris, S. R. 2007. Is literacy enough? pathways to academic success for adolescents. *Brookes Publishing Company*.

Snow, C. E. 1991. The theoretical basis for relationships between language and literacy in development. *Journal of Research in Childhood Education* 6(1):5–10.

Watkins, C. J., and Dayan, P. 1992. Q-learning. *Machine learning* 8(3-4):279–292.

Wells, G. 2000. Dialogic inquiry in education. *Vygotskian perspectives on literacy research* 51–85.

Westerveld, M., and Gillon, G. 2002. Language sampling protocol. *Christchurch: University of Canterbury*.

Wolf, M., and Gottwald, S. 2016. *Tales of Literacy in the 21st Century*. Literary Agenda Series. Oxford University Press.

Woolf, B. P. 2010. *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann.

Yudelson, M. V.; Koedinger, K. R.; and Gordon, G. J. 2013. Individualized bayesian knowledge tracing models. In *International Conference on Artificial Intelligence in Education*, 171–180. Springer.