



Designing Personalized Pedagogical AI Agents to Support Children's Exploratory Learning

Can a social robot support children to become better learners? How personalizing a social robot's behavior can encourage learner exploration.

By Xiajie Zhang

DOI: 10.1145/3611680

OPEN ACCESS

Exploration is a key component in children's development and learning. Children gather information from their environment to gain familiarity with the stimuli through exploration. Exploration can be characterized as being curiosity-driven or initiated by an attention-grabbing stimulus. An example is a child being intrigued by a strange word and wanting to know more or a free exploration of an environment until something catches the eye. These two types of exploration can have a mutual influence on each other. From a pedagogical perspective exploration has been predominantly discussed in two ways: free and guided approaches. With the former, learners are granted full control

inside of the learning environment, and the results of their learning are evaluated as a pedagogy. In the latter, an adult or expert is accompanied in the learning environment to give appropriate guidance. For instance, "guided play" is one contemporary method of providing learners with autonomy while still implementing guidance from adults based on the learner's needs. Researchers found guided exploration has the potential to lead to more efficient problem-solving and social-emotional benefits for the learner,

such as a higher sense of self-competency after learning compared to free exploration [1]. Though promising, the dynamic and interactive nature of the dual-person in guided instruction poses challenges in theorizing and algorithmic formation, making it challenging to implement for each learner [2]. As a potential solution, personalized educational technology may offer the opportunity to collect large amounts of interaction data for data mining and computationally model personal differences in pedagogy reception.

PERSONALIZED AI AGENTS FOR LEARNER EXPLORATION

Since the early work of Reeves and Nass on the media equation (1996), numerous researchers have investigated how humans treat computing-based media as social actors. According to Reeves and Nass, human's interaction with computers and other types of media is fundamentally social. We respond to and react to them as if we are interacting with another human. In child-robot interaction research, previous work has demon-



strated the potential of social robots in peer-influencing through social learning [3, 4]. For instance, researchers have explored how children emulate creativity from a creative peer robot in a collaborative drawing task [5]. Those studies suggest promising research where social robots are designed to promote better learning behaviors among the learner-peer. Taking the insights from the previous pedagogical research with expert guidance, we wish to answer the following three questions: (1) Can social robots be a good educational support in guided pedagogy and designed to support children's exploratory learning?, (2) Does children's exploratory behavior change over time with a social robot's guidance?, and (3) Can children's exploration help their vocabulary learning?

JIBO, A SOCIAL ROBOT FOR EDUCATIONAL SUPPORT AND EXPLORATORY LEARNING

In the summer of 2022, 42 children from 40 families were recruited from different states across the U.S. for

The larger research synthesis of this project is to build pedagogical support to help children become better learners.

our study. After the study enrollment and group randomization, the participants received a robot station and an informative manual to help them set it up at home. The robot station involves three main components, a social robot, a tablet with interactive storybooks for the child to read, and a mini-computer inside of the station to handle the communications between three parties. The particular robot in the station is Jibo, an animated robot with three degrees-of-freedom.

Figure 1. A child using the literacy robotic station for vocabulary learning tasks.



It is designed to engage and interact with people of all ages in their home environment. After the station setup, families lived with the robot for 2 to 4 weeks to complete eight stories with the robot. To ensure children's motivation to read in the interaction, the families could decide when to use the station based on the child's interests. Thus, the stations were fully autonomous in the deployment. At the end of the study, 31 children completed the sessions on time and were included in the study analysis. Families dropped out or were excluded in the analysis because of family emergency, early termination due to time conflicts, unresolvable network and technical issues, or consistent disengagement from the child. Out of the 31 children, 16 were from the treatment group, where the robot applies explorative guidance in the interaction, and 15 were from the control group, where the child has full autonomy.

We first designed robot behaviors to guide children's exploration in literacy learning. Based on the literature review in pedagogical research for guided play and exploration, we designed Jibo's explorative guidance with three components. In the first component, we designed two social macro-actions for the robot to facilitate exploration in a peer child: exploration demonstration and dialogic curiosity. In the case of the exploration demonstration, the robot carries out a sequence of steps to demonstrate how to leverage resources in the storybook to learn about keywords. The sequence of steps is: (a) verbally expressing the intent of learning, (b) pointing out the resources on the platform that can provide information, (c) executing the action to explore the stimuli and seeking the information, and (d) displaying the confirmation of learning. The design of this macro-action is to leverage social emulation theory and show children the process of targeted exploration. In the macro-action for dialogic curiosity, the robot displays its curiosity by asking a storybook-related question. It is a two-turn dialogic interaction that ends with the social robot Jibo verbally responding to the children's answer. The dialogic

curiosity is designed to spark interest and curiosity about the storybook's content (see Figure 1).

The second component of the robot's explorative guidance is a personalization policy. In guided-play pedagogy, one of the characteristics of the expert's guidance is the personalization based on the learner's social-emotional state. Researchers have identified the dual-person social paradigm as one of the key challenges in theorizing and building computational models for effective guidance strategies in guided-play pedagogy [6]. In our work, we formulated learner personalization with a reinforcement learning (RL) framework, a widely used method in intelligent tutoring system (ITS) research. To personalize the pedagogical agent's action based on the learner's emotional and cognitive traits, we constructed a Markov decision process (MDP) with a three-dimensional state space that encompasses children's affective and behavioral states. The three dimensions are:

1. Children's engagement levels on the current page are measured by affect recognition tools from children's facial expressions. The engagement levels are categorized into low, medium, and high levels. The exact bounding box for each level is personalized to each child's affect range (we use children's affect range in a pre-study session for the bounds).

2. Children's prior exploratory behavior in the last round.

3. Keywords unknown to children in the current page's literary content.

The three variables in the state space form a total of 12 states in the MDP. In order to learn a policy that promotes children's exploratory behaviors, we constructed the reward function based on children's exploratory behaviors after the robot's action (as shown in Figure 2). We quantified children's exploration in the digital storybook app as children's self-initiated interactions with different tappable components on the storybook page. The number of children's self-exploration and diversity of exploration stimulus are accounted for with a logarithmic function in the reward so that the first explorations

Figure 2. The Markov decision process (MDP) model for the personalization of agent's behavior.

At each time step t in an interaction, children's current states are used to select the agent's action a_t . Afterward, children's self-initiated exploration before $t + 1$ is used to calculate the immediate rewards for the current action.

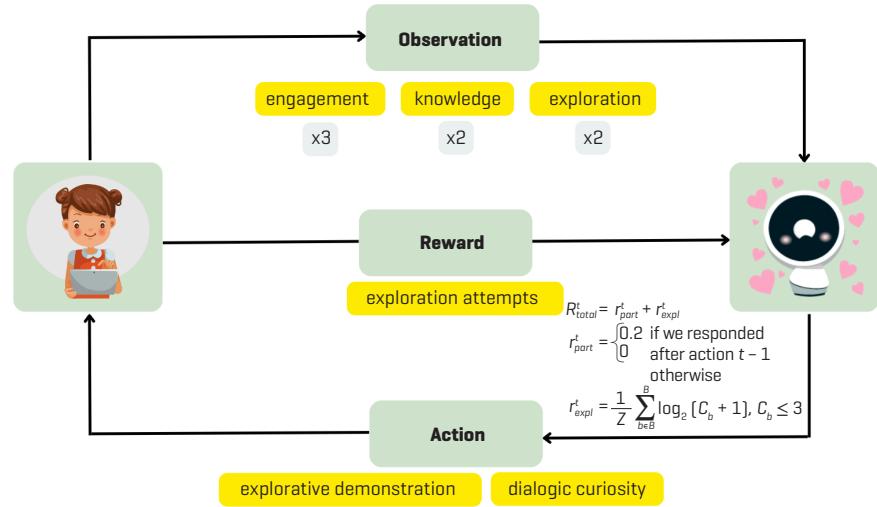
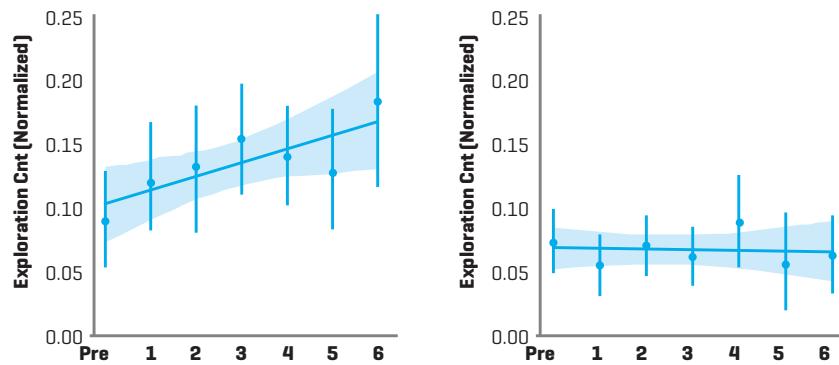


Figure 3. Linear regression in treatment [left] and control [right] group. The x-axis is the session number. The y-axis is the normalized count of exploratory behaviors (by the total count of exploration stimuli).



with different stimulus yields high reward gains. With this reward design, the robot's action is rewarded based on how much children explored and how many different resources they used.

The last component of the robot's explorative guidance is the strategy for proactivity. In theory, the proactivity and timing of guidance should be part of the personalization policy, and the frequency of guidance should depend on the learner's exact situation. However, in this preliminary study to evaluate the efficacy of a

robot's explorative guidance, we applied a simplified proactivity strategy where the robot selects a macro action based on the learner's states at a fixed interval (once after reading one page).

To evaluate the robot's explorative guidance paradigm, we designed a between-subject study with participants. The robot's explorative guidance condition is compared to a control condition where the child is given the full agency and autonomy to explore in a story-reading interaction. In the control condition, the robot's



Advertise with ACM!

Reach the innovators and thought leaders working at the cutting edge of computing and information technology through ACM's magazines, websites and newsletters.



Request a media kit with specifications and pricing:

Ilia Rodriguez

+1 212-626-0686

acmmediasales@acm.org



behaviors are reactive and initiated by a robot button on the tablet (i.e., the child chooses to interact with the robot when they wish), and the robot's behavior is limited to dialogic curiosity to avoid strong social emulation effects from the exploration demonstration in the control group. This study design is to evaluate the robot's personalized guidance with a child-centric learning paradigm.

THE INFLUENCE OF A SOCIAL ROBOT'S EXPLORATIVE GUIDANCE

To observe children's exploration progression on the time scale, we performed linear regression on the number of exploratory behaviors and session progression. The treatment group's regression showed a statistically significant increasing trend ($R^2 = 0.051$, $F(1, 97) = 5.16$, $p = .025$) while no trend was found in the control group (the visualization of the linear trend is shown in Figure 3). Moreover, we found session number significantly predicted children's exploration ($\beta = 0.01$, $p = .025$). This result shows children's exploration increased over time when they interacted with the guided-exploration robot. It confirms the results from previous human robot interaction research that with properly designed robot behavior, humans adapt and emulate behaviors from the robot after interacting with them [3, 4, 5]. Further, explorative behaviors are possible to be simulated by children from a peer-like companion.

GUIDED EXPLORATION AND VOCABULARY ACQUISITION

Children's vocabulary knowledge was assessed with a PPVT¹-style assessment tool before and after each session. The target vocabulary in each assessment are the keywords from the corresponding session. Through this, we can capture the change in children's knowledge of the target words after they read each story. We quantified children's vocabulary growth in each session based on the difference in assessment results.

¹ The Peabody Picture Vocabulary Test is an assessment product offered by Pearson to measure receptive vocabulary acquisition.

Spearman's rank correlation was performed for the correlation between children's exploration and learning gain. The treatment group had a positive correlation between exploration and learning ($r = 0.22$, $p = .046$). Still, no correlation was found in the control group ($r = -0.01$, $p = .92$). We found a correlation between children's vocabulary knowledge and their exploratory behaviors in the guided exploration group. This correlation suggests children's exploration with the guided exploration robot is related to their learning but not in the control group.

The result validates the feasibility of designing social robots to emulate exploratory behaviors in child peers. What's more interesting is that children's emulated exploratory behavior is associated with their vocabulary learning gain. This correlation suggests supporting child-centric learning with personalized guidance is a promising research direction that might have the potential to improve children's learning efficacy. As the first study to personalize a robot's exploration scaffolding, this study is meant to funnel more pedagogical agent research into supporting children's learning through behavioral adaptation and meta-cognitive skill training, instead of performance-driven optimization.

LEARNER EFFICACY AND POST-INTERVENTION EFFECT

The larger research synthesis of this project is to build pedagogical support to help children become better learners. By giving children the tools

Pedagogical studies in guided play/exploration show that when learners are given appropriate amounts of support, they yield better self-competency.

and helping them efficiently use them, children can discover their interests and curiosity and have a sustainable model of learning and exploration. In addition, pedagogical studies in guided play/exploration show that when learners are given appropriate amounts of support, they yield better self-competency. As pointed out by researchers, one big challenge of the guided approach is the complexity of the guidance model for the learner. One critical question the current work did not address is the amount of guidance needed. Our work applied a singular strategy for the timing and frequency of guidance for all learners, while in reality, it might be factorized by the learner's traits. Another question that the current research cannot answer is whether this improvement has a long-term impact on children's learning gain compared to a performance-driven pedagogical approach. For instance, does children's exploratory learning persist after the interaction when the robot's guidance is removed? Moreover, how does the relationship between children's exploration and learning progress after the removal?

REINFORCEMENT LEARNING FOR PERSONALIZED PEDAGOGICAL AGENTS

Decades of research in RL for education have shown fruitful results in the computational approach to personalized learning. A review shows much of the previous effort in RL for education focuses on cognitive gain as the performance drive [2]. However, recent psychological perspectives argue for the importance of social-emotional aspects of learning. Our RL model is an effort to make this shift from cognitive performance-driven optimization to behavioral-driven personalization. In the current study, we weren't able to train seed policy before online learning due to the limited research on statistical models that approximate learner's behaviors. Building those approximated simulation environments is critical for faster learning and convergence in online adaptation. A promising future research direction is to better bridge psychological research with compu-

With properly designed robot behavior, humans adapt and emulate behaviors from the robot after interacting with them.

tational models of learner's behaviors, thus, enabling more popularized training environments for personalized pedagogical agents. Further, the current study still uses a handcrafted reward function with feature selections, which is laborious. Given a dataset of learner behaviors in a specific context, how to effectively inverse a reward function will be fruitful.

MACRO ACTIONS AND MICRO ACTIONS IN AGENT DESIGN

In the current study, the agent's behavioral design is inspired by psychological and pedagogical research in expert-child interaction. Within the study scope, we experiment with two designed behaviors to support children's exploration, each involving micro-action steps. The agent's behavioral design is a process that requires extensive research and effort and limits the flexibility in deployments. An important area for pedagogical agent research and social psychology is how to expedite the process of discovering and validating effective agent actions in learning interventions.

In conclusion, the topic of personalization has always been a hot topic in pedagogical AI research. The effect of personalizing a sequence of instructions or agent actions to maximize the learner's cognitive growth (learning gain, performance scores, etc.) is well-studied through decades of research. With development and research in learner-centered pedagogy, one could imagine AI tutors moving away from traditional instruction-oriented teaching to pro-

vide personalized support tailored to each student's learning. Might that support be guided exploration, growth mindset, pep talks, or emotional regulation, AI agents have the potential to personalize beyond just task difficulty but personalize for a better long-term learning experience through cultivating good learning abilities and skills among learners. However, personalizing the agent's actions for children's behavioral change or meta-cognitive skill adaptation suffers great complexity and data scarcity issues. Thus, a joint effort from the pedagogical research, psychology, and machine learning communities is required to advance the development of social robots that support a diverse group of learner's flourishing and growth.

References

- [1] Yu, Y. et al. The theoretical and methodological opportunities afforded by guided play with young children. *Frontiers in Psychology* 9 (2018). DOI:10.3389/fpsyg.2018.01152.
- [2] Doroudi, S., Aleven, V., and Brunskill, E. Where's the reward? A review of reinforcement learning for instructional sequencing. *International Journal of Artificial Intelligence in Education* 29, 4 (2019), 568–620.
- [3] Gordon G., Breazeal C., and Engel S. Can children catch curiosity from a social robot? *ACM/IEEE International Conference on Human-Robot Interaction*, (2015), 91–98.
- [4] Park, H.W. et al. Growing growth mindset with a social robot peer. *ACM/IEEE International Conference on Human-Robot Interaction* (2017), 137–145.
- [5] Ali, S. et al. Social robots as creativity eliciting agents. *Frontiers in Robotics and AI* 8 (2021), 673730.
- [6] Debowski, S., Wood, R.E., and Bandura, A. Impact of guided exploration and enactive exploration on self-regulatory mechanisms and information acquisition through electronic search. *Journal of Applied Psychology* 86, 6 (2001) 1129–41. DOI: 10.1037/0021-9010.86.6.1129. PMID: 11768056.

Biography

Xiajie Zhang is a Ph.D. student at MIT, advised by Prof. Cynthia Breazeal. His research interests focus on designing and evaluating AI agents that promote well-being and human flourishing in diverse social contexts that supplement our social roles.

Copyright is held by the author.
1528-4972/23/09 \$15.00



This work is licensed under a Creative Commons Attribution International 4.0 License.