Long-term Child-Robot Tutoring Interactions: Lessons Learned

Aditi Ramachandran and Brian Scassellati
Department of Computer Science
Yale University
New Haven, CT 06511
{aditi.ramachandran, brian.scassellati}@yale.edu

***Abstract*—As social robots become more widely used as educational tutoring agents, it is important to study how children interact with these systems, and how effective they are as assessed by learning gains, sustained engagement, and perceptions of the robot tutoring system as a whole. In this paper, we summarize our prior work involving a long-term child-robot interaction study and outline important lessons learned regarding individual differences in children. We then discuss how these lessons inform future research in child-robot interaction.**

I. Introduction

Due to the well-established benefits of one-on-one human tutoring on the learning performance of students, the education community is still searching for methods of instruction that can provide similar advantages for individual students [1], [2]. Social robots are a natural option to explore as effective tutoring agents. Research in the human-robot interaction domain has demonstrated that physically present robots can increase cognitive learning gains, enjoyment, and compliance [3]–[5]. Furthermore, a robot that employs socially supportive behavior can also impact learning [6]. Social robots also have the capability to provide highly personalized tutoring interactions by monitoring cognitive and affective states of individuals, which can potentially lead to the robot learning what actions to take for a given individual over time [7]. This provides a foundation to explore whether social robots can be used to shape behavior, sustain engagement and foster long-term learning gains, particularly as tutoring agents during learning interactions.

One key aspect of tutoring interactions is how students seek help from the tutor. Research in the intelligent tutoring systems community has found that students often engage in unproductive help-seeking behavior such as help overuse (rapid hint requests, for example) or help aversion (lack of use of available help) and this may negatively impact learning [8]. Therefore, it is critical to understand the role of help-seeking behaviors in robot tutoring, and whether these behaviors can be shaped by the social robot tutor. This provided the motivation for us to design and run a human-robot interaction study aimed at answering some of these questions.

II. School Deployment

Our prior work involved conducting a long-term child-robot school deployment in which children interacted with a robot in a one-on-one tutoring context [9]. This study was focused on assessing the help-seeking behaviors of the children in the study, as well as understanding the effectiveness of the robot tutoring system that employed mechanisms to shape these behaviors.

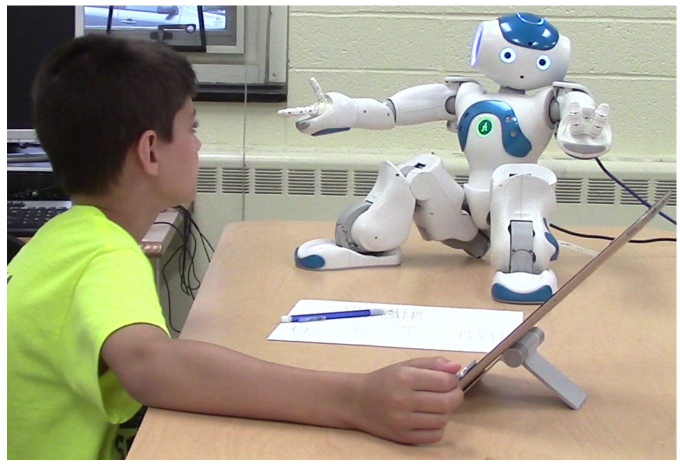

Fig. 1. Child interacting with the NAO robot in a math tutoring scenario.

A. Participants and Setup

The study had 29 children (fifth and sixth grade students), and each child interacted with the robot for four separate sessions over approximately two weeks. The interaction was a math-based interaction in which the students completed a series of fractions problems with the robot during each session. A picture of our robot-child tutoring setup can be seen in Figure 1.

For each math problem, there was exactly three hints, which the child could request from the robot. This is called on-demand help, which refers to help given by the learning environment that must be explicitly requested by the learner [8]. The hints had to be requested in order, and each successive hint contained more information relevant to solving the problem at hand.

B. Experimental Design

Our study had two conditions, which we referred to as the control condition and the adaptive condition. In the control

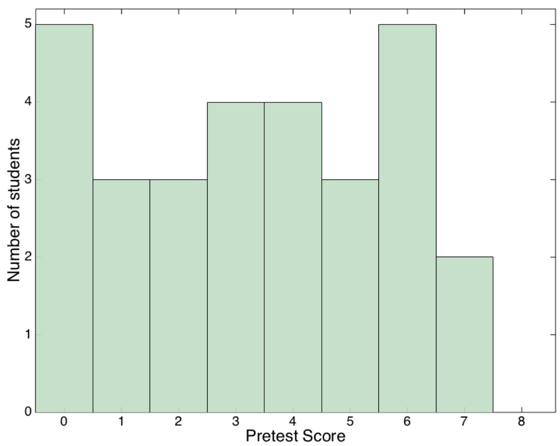

Fig. 2. Histogram of incoming knowledge level for participants. Prior knowledge level was measured by pretest score, which was eight questions total.

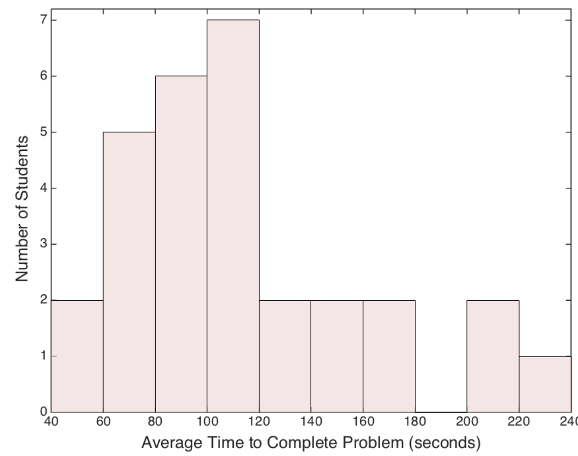

Fig. 3. Histogram of average problem length for participants. This was averaged across all four tutoring sessions.

condition, participants had access to the on-demand help features of the tutoring system. In the adaptive condition, participants also had access to on-demand help, however, the robot employed two simple strategies to regulate the use of the help features. The strategies involved the robot automatically providing a hint when too little help was requested, or denying a hint request and requiring an attempt when too much help was requested.

The two groups were compared by assessing help-seeking behavior change and learning gains. Behavior change was measured by the difference in the number of "suboptimal" help-seeking behaviors that occurred in session one and in session four. The students also took a pretest before session one and a posttest after session four, which was used to measure normalized learning gains over all four tutoring sessions.

C. Results

The full paper describing the methodology and results in greater detail can be found in [9]. The main results of the paper were that the group of children who received the behavior-shaping strategies from the robot improved both their help-seeking behavior and learning gains statistically significantly more than the control group. In addition to these results, by collecting this data over several weeks at multiple sites and observing the tutoring sessions ourselves, we also gained some insights that are of relevance to the child-robot interaction community. We will outline and discuss these "lessons learned" in the following section of this paper.

III. Lessons Learned

By observing the child-robot tutoring sessions, we were able to quickly perceive that there were many individual differences across the students that seemed to dictate how the tutoring session progressed, especially prior knowledge level, and overall level of sustained attention during the session.

A. Incoming Knowledge Level

Despite the fact that the students were all approximately the same age and came from the same state, it was clear that there was a vast range of incoming knowledge levels at play. We measured this by administering a pretest before the child completed the first tutoring session with the robot. As we were very interested in assessing whether the students learned, it turned out to be crucial to administer this pretest. Because of the range of scores present (see Figure 2), it became clear that rather than comparing aggregated test scores across groups, it is more relevant and useful to understand how much a child improved relative to their original score. In future long-term interactions, it is important to build in ways to measure baseline values in order to compare performance within an individual. This is necessary due to the vast and often unexpected differences between each child.

B. Level of Focus/Attention

Each of the four sessions contained the same number of questions, but students could take as long as they needed to complete each session. One of the salient differences between tutoring sessions of different children was the time taken to complete each problem during each session (see Figure 3). This was somewhat unanticipated due to the fact that all children were from a similar geographical area and received the identical content during the sessions. These differences can likely be explained in part by prior knowledge level, as those who had weaker incoming knowledge certainly took longer to work through each of the exercises. However, from observing the sessions in person, we noted that the student's ability to maintain focus on the math task and continue working diligently throughout the entire session also played a role in the distribution of total length of the session across students.

The reasons for when children were able to maintain focus on the math task are likely tied to attributes and internal states that are extremely difficult to measure reliably, such as disengagement, frustration, boredom, confusion, and motivation level. Despite the complexity of these emotional states, loss of focus or attention also correlates to more observable behaviors within the tutoring environment, such as large changes in timing to answer questions, or drops in accuracy. Because the tutoring system had no mechanism by which to handle these behaviors, students who stopped paying attention (or experienced boredom, frustration, etc) were always left to return to the task completely independently and consequently took longer to complete the problems. These observations regarding the crucial role of student attention during a learning interaction have inspired us to introduce mechanisms to address these individual differences in future robot-child tutoring studies. Furthermore, we plan to specifically investigate when to use non-task activities as breaks to sustain overall engagement during learning interactions.

IV. Future Work

Due to the differences observed between children regarding their attention during learning, we believe it is important to study techniques of keeping students engaged throughout the course of a tutoring interaction. The attention spans of students listening to teachers can be as short as five minutes [10]. Accommodating the short attention spans of students, specifically younger students, during learning interactions is therefore a crucial aspect of an autonomous robot tutoring system. One method of doing this involves having the robot use social support and allow the students to take breaks using non-task activities (such as a simple game, or a stretch break). Using off-task breaks during a traditional learning interaction has the potential to improve child-robot interactions, especially for younger students who may have trouble maintaining attention over an extended period of time.

Social robotics research has demonstrated that personalized robot behavior can be more effective than static behavior applied across all users [11]. In addition, robot behavior contingent on the user's behavior can improve evaluations of the robot as an interaction partner [12]. Therefore, regarding non-task activities during a learning task, it is our aim to understand whether non-task breaks that are contingent upon the user's actions are more effective than fixed-timing breaks on learning outcomes, perceptions of the robot, and user experience. Furthermore, these non-task breaks during a learning interaction can be administered in many different ways. How to use breaks in this context to maintain student engagement and potentially motivation is not yet understood.

We propose a study in which we compare three groups of children, who each interact with the robot in a one-on-one tutoring session. One group will have robot behavior that provides non-task breaks on a fixed schedule. The other two groups will receive breaks contingent on the user's behavior, allowing us to evaluate the effectiveness of the contingent breaks. Each of the two contingent conditions will utilize different user behavior on which the breaks will be contingent. One will focus on administering the breaks as a reward, where the student receives non-task activities for improving their performance (accuracy or timing). The other group will have robot behavior that triggers a break whenever the student is performing poorly, which aims to use the breaks as a tool to combat frustration or disengagement. By designing and comparing two conditions in which the user actions that trigger the breaks differ, we can better understand which triggering mechanisms are more effective during a learning interaction.

V. Conclusion

In this paper, we outlined the promise of using social robots as tutoring agents and summarized a field deployment in which students interacted with a robot one-on-one for multiple tutoring sessions over the course of a few weeks. We presented some important lessons learned from completing this long-term interaction study, specifically focusing on how individual differences across children affect these learning interactions. Lastly, we proposed future research we plan to conduct, which we derived as a natural consequence of the lessons learned from the completed school deployment.

Acknowledgment

This work was supported by the National Science Foundation award #1139078 (Socially Assistive Robots).

References

[1] B. S. Bloom, "The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring," *Educational researcher*, vol. 13, no. 6, pp. 4–16, 1984.

[2] K. VanLehn, "The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems," *Educational Psychologist*, vol. 46, no. 4, pp. 197–221, 2011.

[3] D. Leyzberg, S. Spaulding, M. Toneva, and B. Scassellati, "The physical presence of a robot tutor increases cognitive learning gains," in *34th Annual Conference of the Cognitive Science Society (Cog. Sci.'12)*, 2010.

[4] A. Pereira, C. Martinho, I. Leite, and A. Paiva, "iCat, the chess player: the influence of embodiment in the enjoyment of a game," in *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 3*. International Foundation for Autonomous Agents and Multiagent Systems, 2008, pp. 1253–1256.

[5] W. A. Bainbridge, J. Hart, E. S. Kim, and B. Scassellati, "The effect of presence on human-robot interaction," in *The 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*. IEEE, 2008, pp. 701–706.

[6] M. Saerbeck, T. Schut, C. Bartneck, and M. D. Janse, "Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor," in *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2010, pp. 1613–1622.

[7] A. Ramachandran and B. Scassellati, "Adapting difficulty levels in personalized robot-child tutoring interactions," in *Proceedings of the Workshops at the 28th AAAI Conference on Artificial Intelligence*, 2014.

[8] V. Aleven, E. Stahl, S. Schworm, F. Fischer, and R. Wallace, "Help seeking and help design in interactive learning environments," *Review of Educational Research*, vol. 73, no. 3, pp. 277–320, 2003.

[9] A. Ramachandran, A. Litoiu, and B. Scassellati, "Shaping productive help-seeking behavior during robot-child tutoring interactions," in *The Eleventh ACM/IEEE International Conference on Human Robot Interation*. IEEE Press, 2016, pp. 247–254.

[10] G. Petty, *Teaching today: A practical guide*. Nelson Thornes, 2009.

[11] D. Leyzberg, S. Spaulding, and B. Scassellati, "Personalizing robot tutors to individuals' learning differences," in *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2014, pp. 423–430.

[12] M. P. Michalowski, S. Sabanovic, and P. Michel, "Roillo: Creating a social robot for playrooms," in *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006)*, 2006, pp. 587–592.