Robots, Young Children, & Alternative Input Methods

Yanghee Kim and Lixiao Huang

January 2021

Rapid Community Reports

Workshop Outcomes

Digital Promise®
Accelerating Innovation in Education

ISLS International Society of the Learning Sciences

CIRCLS

Robots, Young Children, & Alternative Input Methods

Social robots show promise for student learning yet analyzing children's behavior and applying learning theories remains challenging.

Authors

Yanghee Kim, Northern Illinois University, ykim9@niu.edu
Lixiao Huang, Arizona State University, lixiao.huang@asu.edu

Abstract

Research and development of social robotic applications for young children have been growing rapidly since 2010. Considerable research has demonstrated the positive impacts of social robots on children's learning and development in various domains. Nonetheless, our knowledge about child/robot interaction is still fragmented across several disciplines. This workshop aimed to elucidate the current status of designing and evaluating the efficacy of child/robot collaborative systems and prioritize research challenges. The workshop confirmed the educational potential of embodied sociable robots for physical, socio-emotional, and intellectual development of young children. There was a consensus among the participants that research on child robot interaction necessitated multidisciplinary collaboration to develop pedagogically and technologically sound applications. This report intends to inform interested researchers from learning and computer sciences about future research in child/robot collaborative systems.

Keywords

Social robotics, social robots, child/robot interaction, child development, human-computer interaction, multimodal learning analytics, interaction design

Suggested Citation: Kim, Y. & Huang, L. (2021). Robots, Young Children, & Alternative Input Methods. In Rapid Community Report Series.

Introduction

Social (or sociable) robots are embodied humanoid robots designed to interact with people in an interpersonal manner (Breazeal et al., 2016). Research and development of child/robot interaction (CRI) is a subset of social robotics that investigates the design and development of social robots in various sectors such as health, education, and industry. CRI research highlights the use of social robots for the purpose of developing young children not only intellectually but also socially, emotionally, and even physically (Belpaeme et al., 2018). For example, prior to the workshop researchers in the fields of engineering and computer science were already developing social robots as personalized learning companions for early literacy, observing affective relationships between children and robots (e.g., hugging), and using the robots to support children with autism.

This workshop on Robots, Young Children, & Alternative Input Methods brought together relevant researchers and developers and co-constructed shared knowledge about the current status of and the challenges in designing and evaluating the efficacy of child/robot collaborative systems. The workshop organizers acknowledged that there was a growing body of research on social robotics for young children, but that investigators from various disciplines needed time together to articulate a stronger research agenda and better approaches. Fifteen selected investigators were invited from diverse fields including learning sciences, computer science, engineering, psychology, education, and communication. The participants brought relevant expertise and current research to inform each other and engaged in discussion collectively, seeking to find answers to these five core questions:

- What is the current status of research and development efforts in child/robot interaction?
- What are the theoretical perspectives that may guide research on developing child/robot collaborative systems?
- What are important research issues in designing and engineering assistance

for developing child/robot interactions?
- What technologies are available to design child/robot interaction and to collect data to assess efficacy?
- What are the challenges and opportunities in developing such technologies and research programs?

More broadly, a goal was to explore how this work aligns with Federal research funding goals, across multiple program areas—and thereby to energize and empower a community of investigators to identify means of supporting the needed research.

Workshop attendees and structure

The workshop was held at Northern Illinois University on January 25-26, 2018. Participants were twenty-two researchers, thirteen doctoral students, one postdoctoral researcher, and four private industry professionals. The participants' disciplinary backgrounds included engineering, computer science (vision technology, automatic speech recognition, and computational linguistics), learning sciences, psychology, and communication (See Appendix for a list of participants).

The workshop was organized into two major activities, each a day long. The first day was for invited participants to present current multidisciplinary research on child/robot interaction and to discuss theoretical and technical aspects that can support this research. With the goal of planning future work, the second day began with a summary of the questions raised at the end of day one, followed by three breakout groups that discussed future work, as well as the identification of opportunities for collaboration and alignment with funding priorities.

Day 1 Sessions included:

- Current Status of Research on Child/Robot Interaction
- Theoretical Frameworks for Research on Child/Robot Interaction
- Visual Image Processing and Embodiment
- Automatic Speech Recognition and Dialogue Generation
- Ethnographic Observation of Children

Day 2 started with an overview of known funding priorities (particularly, for the National Science Foundation [NSF]). Then participants joined cross-cutting breakout groups:

- Robot 101: Platform & User Experiences

- Further Discussion on Speech Technology
- Content Authoring for Science Learning

After the workshop, activity continued through:

- Attending a summit of Cyberlearning workshops and the Cyberlearning '19 conference
- Posting content and presentations to the CREATE Center website (see Additional Resources)
- Organizing related symposia at AERA 2019 and the 27th IEEE International Conference on Robot and Human Interactive Communication
- Submitting collaborative research proposals to NSF.

Figure 1: Workshop Attendees

Key issues

Following the organization of Day 1, we reported out on key issues by sessions. Overall, designing and studying sociable robots for young children raises challenges that deserve attention by computer scientists, learning scientists, engineers, and educational researchers, in general, committed to better understanding the future of learning in this truly interdisciplinary area of research.

Findings from early projects

Child/robot interaction research to date confirms that children demonstrate sustained task engagement when learning with social robots, often leading to the successful achievement of intended goals. The research also poses great challenges including (1) the development of natural interpersonal communication between children and robots and (2) the valid assessments of children's multisensory and multimodal learning behaviors.

Dr. Cynthia Breazeal (Associate Professor, MIT Media Lab) introduced her research on personal robots and advocated for AI-enabled social robots that engaged and supported young children holistically, which includes emotion, cognition, social interactions, and physical interactions. Some preliminary findings include (1) children express more joy, attention, and relatedness toward robots as pediatric companions than avatars or plush companions; (2) children retain language better with expressive robots than with expressionless robots; (3) children self-report a higher growth mindset after interacting with a robot designed for this purpose.

Dr. Yanghee Kim (Professor at Northern Illinois University) presented a social robotic app for 3-7-year-old ESL children to improve their English language and literacy skills. The robotic interaction design strategies included (1) inviting children into conversations repeatedly, (2) allowing children to speak and engage in activities in either their native or second language, and (3) always demonstrating empathy with children. The main findings include (1) children are highly engaged and develop affectionate relationships with the robot, (2) children interact with the robot like they would with a friend, and (3) children are very forgiving of the robot's mistakes.

Dr. Lixiao Huang (Postdoctoral Fellow at Duke University) introduced representative child/robot interaction (CRI) research conducted by the international CRI community: e.g., robots to treat autism or stuttering. The research highlights the need for a human mediator in the child/robot interactions because (1) even state-of-the-art robots are not robust enough to meet children's varied needs, (2) humans have the basic psychological need for relatedness to other humans that robots cannot replace, and (3) adding humans in the loop seems to support children's real-world adaptability.

Theories of learning

Applicable learning theories to guide the design and evaluation of child/robot interaction can include embodied cognition, group cognition, social-emotional learning, and domain specific learning & design theories (e.g., for science learning). Researchers in this area informed the participants about current theories and relevant tools.

Dr. Insook Han (Assistant Professor at Temple University) introduced embodied cognition, emphasizing perceptual and physical experiences in human learning. Embodied cognition takes into account the perception-action link grounded in a physical and social environment. It can guide the examination of young children's physical, socio-cognitive, and emotional interactions with a robot that can automatically capture their movements, gestures, and verbal exchanges.

Dr. Ying Xie and Dr. Kyung Kim (Assistant Professors at Northern Illinois University) gave a presentation on group cognition as the examination of socio-linguistic interactions among individuals while they produce cognitive artifacts. Research on children's dialogue in formal and informal settings has been very limited. The presenters introduced a tool that visualized written dialogue into network graphs and showed how the tool could be used to explore the interaction between children and robots.

Dr. Vinci Daro (Director of STEM Learning at Stanford University) introduced the social emotional learning (SEL) theory and discussed challenges in conducting SEL research in the conventional context: (1) a tendency to isolate SEL skills from content learning, (2) the difficulty of measuring SEL competencies, and (3) teacher/peer biases in interactions with students from diverse backgrounds. Dr. Daro acknowledged the potential of social robots for addressing these challenges. That is, all children (regardless of their backgrounds) can develop content knowledge and socio-emotional relationships while interacting with bias-free robots which can automate the measurement of real-time multimodal behaviors.

Dr. Jiyoon Yoon (Associate Professor at University of Texas Arlington) gave a presentation on early science learning research and emphasized that children should be able to "DO" science so as to enhance their acquisition of scientific concepts and facts. Dr. Yoon introduced three approaches to doing science: (1) developmentally appropriate practice – enabling a sense of agency, multisensory science experiences, and social/cultural relevance; (2) the 5E learning model (Engage, Explore, Explain, Elaborate, & Evaluate), and scaffolding children's questioning.

Visual image processing

A notable characteristic of child/robot interaction (CRI) is physical and expressive manifestations of the children's engagement. Very often, children embody and express their intent before they talk or sometimes without talking at all. Such bodily actions and expressions could be credible indicators of engagement and learning. AI-enabled vision technology is a handy tool to advance CRI research.

Dr. Xiaojun Qi (Professor at Utah State University) presented current techniques for tracking facial expressions to recognize children's emotions while children interacted with a robot. The first method is called structured multitask multi-view sparse tracker, which casts face tracking as a sparse approximation problem in a particle filter framework to track one face. The second method is called the multi-Bernoulli filtering technique, which applies the random finite set multi-target multi-Bernoulli filter to detect and track multiple faces simultaneously and without explicit detection. She further presented a deep neural network (a custom version of the VGG13 network), which is focused on the facial expression recognition (FER+) database, which recognizes children's facial expressions to classify them into two emotional states (i.e., happy and neutral). She concluded her talk with several challenges she observed in tracking and recognizing emotions.

Dr. Aaron Kline (Engineer, Pediatric Research Lab at Stanford Medical School) presented a system developed by his research group that seamlessly integrates sensors, real-time social cues, and feedback in behavioral therapy. He described the approaches his research group uses to help reinforce emotional awareness for children with autism, including face tracking and emotion recognition. He emphasized that the inclusion of learners in the design process would improve their engagement in the learning experience.

Dr. Karthik Ramani (Professor at Purdue University) presented Ziro, a prototype of design-build-play robots for kids in STEM learning. He demonstrated that children could learn through design and making. Ziro has vision components and is integrated with Amazon Echo (for voice integration) to do a variety of tasks. He explained the motion flow system for gesture recognition and concluded that multimodal sensing, particularly human emotions, can allow new forms of AI-based interactions.

Speech recognition and dialogue generation

The language and literacy of young children are still developing and vary widely among children in proficiency. Developing a robot to converse with children and understand children's speech therefore is crucial for the efficacy of child/robot interaction. This group presented the current technological and methodological statuses and great challenges in advancing this line of research due to the lack of children's voice data.

Chad Dorsey (President and CEO, the Concord Consortium) introduced a few instances of using automatic speech recognition (ASR) in learning sciences research and the promises of ASR to advance learning sciences. There are several challenges in CRI research Linguistic variation is high across ages. Error rates are high because young children's speech is still developing: approx. 60% accuracy for children vs. 95% for adults in a lab setting. Speech patterns in a natural environment are unexplored and seem very challenging due to unexpected background noise. Children's voice data are largely lacking, thereby limiting the advance of ASR for children.

Dr. Abeer Alwan (Professor at the University of California-Los Angeles) discussed advanced ASR techniques. The challenges her group faces with children's speech include: (1) a lack of large databases of children's speech and significant intra- and inter-speaker variability, (2) significant variability in pronunciation due to different linguistic backgrounds and misarticulations, (3) low signal-to-noise ratio in the natural environment (e.g., classroom), and (4) distinguishing reading errors from pronunciation differences.

Ajith Alexander (President, Oxford Wave Research USA) performed an analysis of sample audio from Dr. Yanghee Kim's child-robot interaction recordings and was able to diarize[1] the voices of the robot, a facilitator, and two children. The major takeaways include (1) diarization works well for two speakers or one-on-one interactions of a child with a robot -- accuracy on more than two speakers tends to be low, (2) ambient noise in a classroom poses challenges; (3) gender-based separation is weak for children, and (4) the corpus of child vocal data is limited; diarization of children's speech as a field is largely nascent. Recommendations for future data collection for optimal diarization include: (a) constraining the recording environment to as few speakers as possible, (b) using an individual microphone for each child during data collection to reduce difficult post processing problem, (c) using an ambient microphone to pick up background noises, (d) getting children to speak longer phrases at least a few times while recording rather than yes/no responses.

Dr. Marilyn Walker (Professor at University of California-Santa Cruz) introduced her team's work on open domain dialog with Slugbot. Challenges include personalization, scaling conversational interaction, adapting to new domains, and multi-domain and multi-modal dialog systems. Currently, Dr. Walker's research lab has useful resources for interacting with children about the content of a story. She identified a need for better dialogue management strategies for conversation with children, controlling the nonverbal behaviors of the robot to demonstrate personality and make the robot engaging to the children.

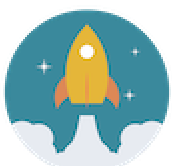
Tony Zhao (a doctoral student at CMU) introduced DialPort, a dialog system that produces natural dialog in prescribed topics such as weather and restaurants.

[1] Speaker diarization: the process of partitioning a conversational audio stream according to individual speakers.

Ethnographic methods

Dr. Laura Johnson (Associate Professor at Northern Illinois University) introduced using qualitative ethnographic observations to study child/robot interactions. She discussed Spradley's matrix for descriptive observations and how these could help researchers pay attention to many elements within observations, including spaces, actors, activities, objects, goals, time, events, and feelings. Regarding specific theoretical and methodological approaches to observing children engaged in communication, she introduced Corsaro's (2012) work on interpretive reproduction and peer cultures and Hymes' (1974) work on the ethnography of communication. Questions were raised about the reliability of observations across observers.

Recommendations for future work



The workshop successfully achieved its goal of bringing together interested researchers from relevant disciplines—computer science, learning science, and engineering. The participants informed each other about their research and reached a consensus on the current obstacles and opportunities in this area of research and development. Two key recommendations called for collaboration among investigators and across projects to (a) advance speech and vision recognition technologies for this population, necessitating a better corpus of shared data and (b) refine learning theories and research methods to better address the issues associated with young children learning with social robots.

Further recommendations were generated from the breakout groups on the second day, who selected three topical issues and conducted focused discussions on each topic.

Designing a powerful educational robotics platform is an unsolved problem that requires multidisciplinary collaboration.

With regard to **platforms and user experiences,** participants agreed that there are no satisfying educational robot platforms. The group discussed the desired features (e.g., physical embodiment, socio-emotional interaction, and speech recognition & generation) and proposed to design an ideal robot platform for education in classrooms and at home. To achieve this long-term goal, there is a great need for diverse expertise from a variety of disciplines including human factors, curriculum design, speech recognition, computer vision, pedagogy, and personalization. A place to start is a survey of existing educational robots and users' feedback on their experiences (benchmark) in order to identify critical user needs in terms of robots' functionality and affordances.

A recommended mid-term goal is to form a community for design, which incorporates (1) advancement of all aspects of the robot: speech, vision, personality, interface, activity, and curriculum design, (2) the best way to tell a story, (3) best practices for speech recognition, (4) the best technique to detect children's emotions, (5) how to personalize the robot in terms of appearance and characters, and (6) how to generate robot dialogs and behaviors based on a chosen personality.

With regard to **speech technology,** there is a great need for a children's speech database. This group discussed specifications for such a database, which include metadata on age, gender, language, grade level, region (zip code of residence), and speech and hearing disabilities. The speech data for individuals (4-8 years old) include commonly used elements (e.g., digits, alphabet), a set of sentences representing full (ideally) phonetic coverage, spontaneous interaction data consisting of 20-30 minutes per individual, with 1:1 boy/girl ratio, 200-300 children per age band, 100 individuals within each major dialect/accent variation, including individuals with speech challenges (i.e., autism and stuttering), and with geographical/socio-political dialect variations. As a short-term goal, the team will check existing corpora that match the characteristics listed and search for communities underrepresented in current data sets.

Once a database exists, research should be focused on unsolved problems in recognizing children's speech, and research in this area could become very productive.

With regard to **authoring science learning activities**, the group sees social robots as powerful for engaging young children in science. In a scientific inquiry process, generating scientific questions in each inquiry cycle is important, and a robot can help children to generate proper questions. A robot can also help children to evaluate their understanding through the questions and answers generated by children.

To achieve a vision of supporting inquiry learning, challenging research goals must be addressed. A key topic is better ways to author and generate dialogue. Hard issues include automating the authoring process, supporting multi-party dialogues, generating questions and responding to scientific content, as well as classifying children's responses in ways that are relevant to the learning goals of scientific activity.

A recommended emphasis is on improving robot dialog and robot behaviors for a specific subject matter, like science learning. To do this, expertise is required in dialogue generation, science education, teacher education, HCI research and design, and learning sciences.

Once better authoring supports are available, additional recommendations focus on formative evaluation and iterative refinement. A key recommendation is to support research on the design and early-stage evaluation. Many iterations and improvements are likely to be needed before social robots for science learning are ready for larger field tests and summative evaluations.

To conclude, the role of AI and robots in education as well as other sectors of life has been well-acknowledged in recent years, and the development of this topic takes into account a range of technological and human factors. The workshop participants expressed strong interest in working together as a community and willingness to participate in similar workshops like this one. Federal funding opportunities will be crucial to make progress in multidisciplinary community building and knowledge development in CRI research. Some agencies like NSF have already emphasized multidisciplinary collaboration, but robotics-related funding programs are still focused largely on engineering and computer scientific aspects of development. Educational robotics programs that weigh learning and developmental issues equally with technological development could advance our deep understanding of CRI and lead to substantial and sustained robotic applications for education.

For further information about this workshop and its discussion, please read the full white paper and view the presentation videos.

References

Admoni, H., & Scassellati, B. (2017). Social eye gaze in human-robot interaction: a review. *Journal of Human Robot Interaction, 6*(1), 25-63. doi:10.5898/JHRI.6.1.Admoni

Breazeal, C. L. (2002). *Designing sociable robots*. Cambridge, MA: The MIT Press.

Breazeal, C., Dautenhahn, K., & Kanda, T. (2016). Social robotics. In K. O. Siciliano B. (Ed.), *Springer handbook of robotics* (pp. 1935-1972): Springer, Cham.

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics, 3*(1), 1-17. doi:10.1126/scirobotics.aat5954

Kim, Y., Butail, S., Tscholl, M., Liu, L., & Wang, Y. (2020, March 23-27). *An exploratory approach to measuring collaborative engagement in child robot interaction.* Paper presented at the Learning Analytics and Knowledge (LAK20), Frankfurt, Germany.

Kim, Y., & Smith, D. (2017). Pedagogical and technological augmentation of mobile learning for young children. *Interactive Learning Environments, 25*(1), 4-16. doi:10.1080/10494820.2015.1087411

Kory-Westlund, J. M., & Breazeal, C. (2019). Exploring the effects of a social robot's speech entertainment and backstory on young children's emotion, rapport, relationship, and learning. *Frontiers in Robotics and AI*. doi:10.3389/frobt.2019.00054

Randall, N. (2019). A survey of robot-assisted language learning (RALL). *ACM Transactions on Human Robot Interaction, 9*(1). doi:10.1145/3345506

Zhang, Y., Song, W., Tanb, Z., Zhu, H., Wang, Y., Lam, C. M., . . . Yia, L. (2019). Could social robots facilitate children with autism spectrum disorders in learning distrust and deception? *Journal of Psychiatric Research, 98*, 140-149. doi:10.1016/j.chb.2019.04.008

Resources

MIT Media Lab, Personal Robots:
https://www.media.mit.edu/groups/personal-robots

Social and Intelligent Robotics Research Lab:
https://uwaterloo.ca/social-intelligent-robotics-research-lab/robots

Social Robotics Lab: https://scazlab.yale.edu

Center for Cross-disciplinary Research on Engaging Advanced Technology for Education: http://createcenter.net

ROBOKIND (ROBOT 4 AUTISM): https://robokind.com

Acknowledgements

The research projects presented in the workshop are sponsored by private and public agencies including NSF (CCF-1138986, EAGER-174805, IIP-1632154, IIS-1623561, IIS-1704167, NRI-1734380, etc.) and SK-Telecom Corp.

Rapid Community Reports are supported by NSF grant 2021159. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Appendix

Workshop Participants:

Ajith Alexander (Oxford Wave Research). Abeer Alwab, Ph.D. (University of California Los Angeles), Cynthia Breazeal Ph.D. (MIT Media Lab), Vinci Daro Ph.D. (Stanford University (SCALE)), Chad Dorsey (The Concord Consortium), Insook Han Ph.D. (Temple University), Lixiao Huang, Ph.D. (Duke University), Laura Johnson, Ph.D. (Northern Illinois University), Kyung Kim, Ph.D. (Northern Illinois University), Yanghee Kim, Ph.D. (Northern Illinois University), Xiaojun Qi, Ph.D. (Utah State University), Karthik Ramani, Ph.D. (Purdue University), Marilyn Walker, Ph.D. (University of California Santa Cruz), Ying Xie, Ph.D. (Northern Illinois University), Ji-yoon Yoon, Ph.D. (University of Texas, Arlington), Afshan Amber (University of California Los Angeles), Amir Hossein Farzaneh (Utah State University), Yan Chen (Northern Illinois University), Chris Kraner (Northern Illinois University), Kevin Bowden (University of California Santa Cruz), Javanmardi Mohammadreza (Utah State University), Sam Spaulding (MIT Media Lab), Saurav Mukhopodhyay (Northern Illinois University), Sooyeon Jeong (MIT Media Lab), Tracy Rogers-Tryba (Northern Illinois University), Jiaqi Wu (University of California Santa Cruz), Mengxi Zhou (Northern Illinois University), Tiancheng Zhao (Carnegie Mellon University).