Final Workshop Report

CIRCL Workshop:
Robots, Young Children, & Alternative Input Methods

January, 25 - 26, 2018

Organizing Committee

Yanghee Kim, Chair, Northern Illinois University
Vinci Daro, Stanford University
Insook Han, Temple University
Lixiao Huang, Duke University
Laura Johnson, Northern Illinois University
Xiaojun Qi, Utah State University
Ying Xie, Northern Illinois University
Jiyoon Yoon, University of Texas at Arlington

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This two-day workshop on Robots, Young Children, & Alternative Input Methods, was held at the Northern Illinois University on January 25-26, 2018. The workshop brought together a pool of the world's leading investigators interested in interaction design and technology development in this topic. The work goals included (i) to explore the current status of research on children-robot interaction, (ii) to discuss theoretical and technical aspects that can support the research, and (iii) to explore the potential for future research in the area from the social, emotional and cognitive, and educational perspectives. The goals required the assembly of distinguished researchers from a diverse array of interrelated fields, including learning sciences, computer science, engineering, psychology, and others. The participants were selected as those who understand the necessity of cross-disciplinary collaboration for inquiring into the development of a socio-technical partnerships between an embodied, humanoid robot and young children in ways that promote children's intellectual, affective and social development.

During the workshop, the researchers shared their respective expertise and on-going research to inform each other and discuss core research questions in designing and evaluating the efficacy of the child/robot collaborative system:

- What are the current statuses of research and development efforts in child/robot interaction?
- What are theoretical perspectives that may guide research on developing child/robot collaborative systems?
- What are important research issues in engineering assistance for child development with a robot?
- What technologies are available to design child/robot interaction and collect data to assess the efficacy?
- What are the challenges and opportunities in developing such technologies and research programs?
- In what way are the research issues aligned with the NSF goal of Broadening Participation in STEM education and STEM workforce (particularly, the NSF initiative INCLUDES, Human Technology Partnership)?

Current Status of Research on Child/Robot Interaction:

The workshop highlighted some of cutting-edge developments in children-robot interaction and social robotics. Both the workshop presenters and discussions agreed that they have to tackle hard problems in many dimensions, including robot design and functionalities, robot cognition and intelligence, and last but not least, the human dimension of how people perceive robots and how we relate to robots. Presenters made a few points on what robots are, and what they are not, and emphasize useful roles they can play in supporting children's learning and development. Particularly promising are areas where robots are not meant to replace humans, but carry out tasks that they are

specifically good at. Presenters illustrated those points with examples of research projects they have been involved in, and highlighted lessons that they have learned over the years in quest to develop robot companions that provide useful assistance and are socially acceptable for children. These examples include work on robotic mediators for children with autism and special needs, including research on robot-assisted therapy for children with autism and robot-mediated literacy, as well as work with AI-enabled social robots which are meant to engage and support children holistically. It is concluded that robots can play a valuable role in human society, e.g. as helpful tools, mediators, assistants and companions, if we exploit the strengths of both robots and humans, considering the specific requirements in real-world application areas, while preserving human dignity, privacy and independence.

Theoretical and Technical Aspects of Research on Child/Robot Interaction:

If we are to build human-like robots that can interact naturally with people, our robots must know not only about the properties of objects but also the properties of animate agents in the world. One of the fundamental social skills for humans is the attribution of beliefs, goals, and desires to other people. This set of skills has often been called a "theory of mind." The presenters presented some of theories on the development of theory of mind in child/robot interaction, including embodied cognition, group cognition, social emotional learning, and early science learning, and discusses the potential application of these theories to building robots with similar capabilities. For example, research needs are identified in applying embodied cognition into younger learners, since most of embodied cognition research in educational contexts have been mainly conducted with adults and upper-level elementary learners. Along with the theory, this workshop also had in-depth discussion on technology for the construction of humanoid robots that engage in natural human social dynamics (such as finding faces and eyes, distinguishing animate from inanimate stimuli, and speech and gesture recognition are introduced). The presentations described their iterative experimental methodology of testing in and out of the technologies (lab based and real world), and outlined discoveries that were made doing so.

Future Work: Breakout Group Recommendations

Participants in the breakout discussion focused on these three priority areas where further research is needed. These key areas are discussed in the sections below.

Robot Platform: The group went through a list of robots that have been used in educational research and agreed that there is no satisfying educational robot platform. This group discussed the desired features and proposed short-, mid-, and long-term goals to design an ideal robot platform for education in classrooms and at home.

Speech Technique: The group agreed on further development and collaboration of speech related techniques. Due to a great need for a children's speech database, this group discussed plans to build such a database, including commonly used elements

(e.g., digits, alphabet), set of sentences representing full (ideally) phonetic coverage, spontaneous interaction data, individuals within each major dialect/accent variation, etc. They will check existing corpora that match the characteristics listed and search for communities underrepresented in current data sets.

Content Authoring: The group discussed the possibilities of using social robots in science learning. In scientific inquiry process, generating scientific questions in each inquiry cycle is important and a robot can help students to generate proper questions as well as evaluate their understanding through the questions and answers generated by students. To enable this support by a robot, they discussed what is the current state of dialogues generations and what needs to be further developed for supporting better automation, multi-party dialogues (speaker ID, localization), automated question generation (exploration/content), automated classification of responses. They set a three-year project timeline for this project.

In summary, the overall consensus of the group was that, despite our professional diversity and the short time available for us to work together, we achieved substantial coherence, common language and integrity. The communication and mutual interest in the work of each others' specialized disciplines, and the wish to embrace the complementary frameworks were impressive. The final sense of the group was that this workshop represented a worthwhile opportunity to derive the maximum benefit from the deep need of child/robot research to enhance the learning capabilities and cognitive skills of young learners at all aspects of society. The workshop successfully brought together an interdisciplinary group of leading thinkers in order to explore current status of research on child/robot interaction, to identify possible theories and cutting edge technologies that can guide and support for the design and development of child/robot collaborative system, and to discuss appropriate directions and methods for future work. We hope that the workshop, this report, and subsequent efforts will lead to and support improvements in interaction design and technology development, particularly for young learners.

PARTICIPANTS

The workshop participants included twenty-two researchers, thirteen doctoral students, one postdoctoral researcher, and four private industry professionals. Participants' disciplinary backgrounds ranged broadly from Engineering, Computer Science (vision technology, automatic speech recognition, and computational linguistics), Learning sciences, Communication, and Psychology. More details about the participants are presented in Appendix 1.

Presenters:

Aaron, Kline, Stanford University
Alexander, Ajith, Oxford Wave Research
Alwan, Abeer, UCLA
Breazeal, Cynthia, MIT
Dorsey, Chad, Concord Consortium
Han, Insook, Temple University
Huang, Lixiao, Duke University
Johnson, Laura, Northern Illinois University
Kim, Kyung, Northern Illinois University
Kim, Yanghee, Northern Illinois University
Qi, Xiaojun, Utah State University
Ramani, Karthik, Purdue University
Walker, Marilyn, UC Santa Cruz
Xie, Ying, Northern Illinois University
Yoon, Jiyoon, University of Texas at Arlington

WORKSHOP SESSIONS AND SCHEDULE

Day One (Thursday): Learning and Sharing

The first day had two goals: (1) to present current multidisciplinary research on children-robot interaction and (2) to discuss theoretical and technical aspects that can support the research. Day one included an overview session on Outcome of Social, Educational Robots (session 1), followed by sessions on Child Development and Learning Theories (session 2), Vision Processing (session 3), Speech Recognition (session 4), and Ethnographic Observations (session 5).

Time	**Event**
7:30-9 a.m.	Breakfast and registration
9-9:30 a.m.	Greetings and introduction
9:30-11 a.m.	Sociable, educational robots for young children (Lead: Y. Kim) Presented by Breazeal, Y. Kim and Lixiao
11-11:15 a.m.	Break (coffee and snack)
11:15 a.m. -12:45 p.m.	Child development and learning theories(Lead: Han) Embodied cognition (Han) Group cognition (Xie and K. Kim) Social, emotional learning (Daro) Early science learning (Yoon)
12:45-1 p.m.	Break
1-2 p.m.	**Lunch**
2-3:30 p.m.	Visual processing and embodiment (Lead: Qi) Presented by Qi, Kline and Ramani)
3:30-3:45 p.m.	Break (coffee and snack)
3:45-5:15 p.m.	Speech and dialogue (Lead: Alexander) Learning sciences research on speech and dialog (Dorsey) Speech recognition (Alwan) Speech diarization (Alexander) Dialog generation (Walker and Zhao)
5:15-5:30 p.m.	Break
5:30-6 p.m.	Ethnographic observations (Johnson)
6:30-8:30 p.m.	**Dinner**

Day Two (Friday): Planning Future Work

The second day began with a summary of the questions raised at the end of day one, followed by five breakout group sessions discussing future work, including the identification of opportunities for collaboration and alignment with funding priorities.

Time	**Event**
7:30-9 a.m.	Breakfast and networking
9-9:30 a.m.	NSF Funding priorities Presented by D'Angelo
9:30 a.m.-noon	Breakout groups by topic/problem areas (Lead: Johnson) Goals/objectives Major activities and outcomes Collaboration plan (expertise and timeline) Each group needs a scribe, discussion leader and spokesperson Please submit group discussion notes to the workshop committee at lixiao.huang@duke.edu.
noon-1 p.m.	**Lunch**
1 -2 p.m.	Breakout group reports and wrap-up (Lead: Johnson)
2-2:30 p.m.	Post-workshop survey
2:30-4 p.m.	Organizing committee meeting

1. Current Status of Research on Child/Robot Interaction
(Discussion lead: Dr. Yanghee Kim, Northern Illinois University)

What are the current status of research and development efforts in child and robot interaction? Speakers in the session 1, *Outcome of Social, Educational Robots***,** shared their on-going social-educational robot research projects, including research methods, research challenges and research plans.

Dr. Cynthia Breazeal (associate professor at MIT Media Lab) presented social robots as personalized learning companions for early literacy. As AI becomes increasingly essential, Dr. Breazeal's team strives to design humanistic AI-enabled technologies (social robots) that empower all of us to become the people and society we aspire to be as part of daily life. She advocated that AI-enabled educational technologies should engage and support learners holistically, including emotion, cognition, social interactions, and physical interactions. Learning with personalized peer-like social robots provides friendly companionship, collaboration, interpersonal interaction and social cues, perspective-taking, social modeling, and emotional

engagement. When including parents in the child-robot interactions, the robot fosters socially inclusive interaction with others, facilitates adult participation and scaffolding, and provides important opportunities for parent engagement at home. Some preliminary findings include the following results: (1) children express more joy, attention, and relational touch toward robots as pediatric companions than avatar and plush companions for in-patient care; (2) children retain more similar phrases and words, as well as longer storylines, with expressive robots than with flat robots; (3) children self-report having higher growth mindset after interacting with growth mindset robot; (4) children are more emotionally expressive with robots than with tablets; and (5) emotional data improves prediction of word-reading skill. So far their team has a dataset for 218 hours of time-synchronized multi-modal interaction, including interaction data, surveys, sensor data, and assessments. Dr. Breazeal plans to continue with research and to design personalized learning companions for early literacy.

Dr. Yanghee Kim (professor at Northern Illinois University) presented her research on educational affordance of embodied robots. Project DEAR aims to design an engaging and affable robot for 3-7-year-old ESL children to improve their early English literacy skills. Project IDEAl is an inclusive design that mediates children's collaboration. Based on the literature, Dr. Kim's team developed a theoretical framework for crafting interactions between robot and children. Their goals of robot design are three-fold. They wanted the robot to (1) consistently invite children into conversations, (2) provide opportunities for children to speak and engage in activities in either their native, second, or developing language, and (3) always demonstrate empathy with children. To do this job, they used human mediator sessions to develop scripts for robot mediation. They used the Wizard of Oz method, explicit and repeated invitation, opportunities for children to explore, and limited interaction sessions up to 20 minutes. Her team found the following findings: (1) children developed affectionate relationships and were very engaged with the robot, (2) children interacted with the robot like they would with a friend, (3) children were very forgiving of the robot's mistakes, and (4) children gradually learned to work with their peers, taking turns and listening. Dr. Kim also introduced more evidence of an affectionate relationship with Skusie, including physical (touching and hugging), words, and unwillingness to leave. The goal of the project was to design the robot to take over the mediation role from humans. She plans to integrate curricular content and artifacts into the child-robot interactions. To evaluate changes in children's engagement and collaboration skills, she needs computational models to design and assess children's behaviors.

Dr. Lixiao Huang (postdoctoral fellow at Duke University) presented more examples from the US and internationally to complement Cynthia and Yanghee's research areas on ESL and emotional support. Dr. Huang showed one example of applying a social-educational robot in treating autism, and pointed out the problem of evaluating the effect of child-robot interaction in an isolated environment without testing it in a human-human interaction later. She also showed an example of applying a social-educational robot in treating stuttering, in which the human tutor played an enormous role in the facilitating process. In this stuttering treatment case, the learning

activities and tutor's facilitation becomes more important than the robot itself. Dr. Huang used the examples to illustrate the importance of including other humans in the children's interactions with robots for three reasons. First, even the state-of-art robots are not robust enough to meet children's varied needs. Second, humans have the basic psychological need for relatedness to other humans that robots cannot replace. Third, the goal of child-robot research is to increase children's adaptability in the real world, so adding humans in the loop is ultimately good for children's well-being. Based on the issues existing in the examples, Dr. Huang raised three big questions all researchers need to consider in all child-robot interaction research, including 1) what specific areas (i.e., storytelling, moral training, emotional support in health care, playmate interaction, etc.) do children need help? 2) How do we design both robots and activities to include humans in the loop? 3) How do we measure the results effectively considering its transfer in real-life scenarios? Dr. Huang concluded with a big picture of various human-robot interaction applications in a lifespan (from infant to older adults; from personal usage to public use) that will all need to be informed by these three questions. She suggested human factors' methods such as task analysis to identify the critical issues in each application, and improve the design of child-robot interaction.

2. Theoretical Framework of Research on Child/Robot Interaction
(Discussion lead: Dr. Insook Han, Temple University)

What are theoretical perspectives that may guide research on developing child/robot collaborative systems? Speakers in the session 2, *Child Development and Learning Theories*, suggested some of the theoretical framework of child development and learning theories that can support the research, including embodied cognition, group cognition, social emotional learning, and early science learning.

Embodied Cognition: Dr. Insook Han (assistant professor at Temple University) presented embodied cognition. Embodied cognition is a relatively new theoretical perspective to emphasize perceptual and physical experiences in human learning. Compared to the traditional view of considering knowledge or conception as purely mental and independent of our abilities to perceive, cognitive linguists and cognitive psychologists have tried to examine human cognition from this new perspective. Based on embodied cognition, many educational interventions have been designed and implemented that involve multisensory experiences, perceptual/bodily interaction with physical worlds. Embodied cognition research in educational contexts has mainly focused on math, science and language learning as well as many adults and elementary students. Considering that young learners' learning is inherently perceptual and multi-modal, there is room for researching with younger learners with the embodied cognition framework. A social robot can be a good tool to facilitate young learner's physical interaction as well as to capture their movement, gestures, and verbal exchanges.

Group Cognition: Dr. Ying Xie and Dr. Kyung Kim (assistant professors at Northern Illinois University) presented group cognition. Gerry Stahl defined group cognition as the

study of group interaction. After discussing the three levels of group cognition and why this topic is important to the community, they introduced their cutting edge text-based dialogue analysis tool, Graphical Interface of Knowledge Structure (GIKS), that can visualize written dialogue into network graphs. They showed how this tool can be used for exploring the interaction between robot and children. Previous analysis of group cognition mainly used college students and adults' written text. The database of children's writing in formal and informal settings is still lacking.

Social Emotion Learning: Dr. Vinci Daro (researcher at Stanford University and Director of STEM learning at Envision Learning Partners) presented "social emotional learning" as a lens for exploring ways that social robots might be used in research and in interventions to support children in school settings. Dr. Daro presented a brief rationale for focusing on social emotional learning (SEL), and then provided a brief overview of the constructs that have become dominant in the field of SEL research and development. She identified three challenges in conducting SEL research for which social robots may be relevant: a tendency in research designs (and intervention designs) to isolate the instruction or assessment of SEL skills from academic content instruction; the difficulty of measuring SEL competencies, and growth in these competencies; the significant role of teacher and researcher biases in interactions with students from different backgrounds, including linguistically diverse students. The rest of Dr. Daro's presentation described the general outlines of how SEL frameworks might help situate research and interventions involving social robots, with a focus on students' identity development as learners, language development for multilingual students, and the equity considerations important in these lines of research.

Early Science Learning: Dr. Jiyoon Yoon (associate professor at University of Texas Arlington) presented early science learning and emphasized that children need to improve their abilities to "DO" science to enhance children's acquisition of scientific concepts and facts. There are three approaches for doing science: Developmentally Appropriate Practice (DAP), 5E Instructional Model, and Questioning.

3. Technology for Research on Child/Robot Interaction

What technologies are available to design child/robot interaction and collect data to assess the efficacy? What are the challenges and opportunities in developing such technologies and research programs? Speakers in the session 3 (Vision processing), 4 (Speech recognition), and 5 (Ethnographic observation) introduced some of promising technology that can be used in robot/children interaction.

Vision Processing and Embodiment (Lead: Dr. Xiaojun Qi, Utah State University)

The visual processing and embodiment session consisted of three presentations: face tracking and emotion recognition in robot/children interaction (Dr. Xiaojun Qi), autism glass project: expression recognition glasses for autism therapy (Dr. Aaron

Kline), and children-robot interaction (CRI) for engaged learning through design and making (Dr. Karthik Ramani).

Dr. Xiaojun Qi (professor at Utah State University) presented a talk to explain the current techniques for tracking faces and recognizing emotion in robot and children interaction. She explained two tracking methods that have been developed by her Computer Vision Research Laboratory. The first method is called structured multi-task 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 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 (e.g., custom version of the VGG13 network), which is trained on the facial expression recognition (FER+) database, to recognize two facial expressions (i.e., happiness and neutral) of the kids in robot and child interaction. She finally concluded her talk by presenting several challenges in tracking and emotion recognition.

Dr. Aaron Kline (professor at Stanford University) 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 also emphasized the importance of including learners directly in the design process to improve their engagement in the learning experience. Finally, he mentioned the same or similar technologies could be leveraged to help measure engagement during child–robot interaction.

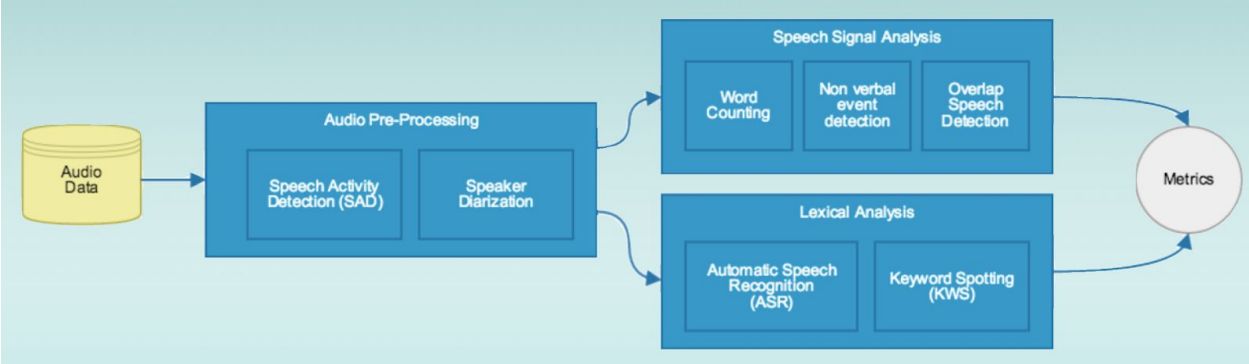
Dr. Karthik Ramani (professor at Purdue University) presented a talk to explain the latest in Ziro, which is a prototype of design-build-play robots for kids to motivate them in STEM learning. He demonstrated that kids can learn through design and making. Ziro has some vision components and is integrated with Amazon Echo (has voice integration) to do a variety of tasks. Finally, 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 (Lead: Dr. Ajith Alexander, Oxford Wave Research*)*

This session focused on the current techniques for speech recognition and had four speakers introducing learning sciences research on speech and dialog (Dr. Chad Dorsey), speech recognition techniques (Dr. Abeer Alwan), speech diarization (Dr. Ajith Alexander), and dialog generation (Dr. Marilyn Walker and Mr. Tony Zhao).

Dr. Chad Dorsey (president and CEO at the Concord Consortium) presented an overview, possibilities, and barriers for learning sciences research. Speech technologies are important for learning science research, and it can be broken down into the figure below. Audio data collection uses microphones, LENA device, Microphone arrays, and

beamforming. When it comes to preprocessing, speaker diarization technology can help identify and separate individual speakers in a single audio track, while speech activity detection separates speech from the acoustic background. When it comes to speech signal analysis, word counting, turn counts, sharing, nonverbal sentiment detection, social signals detection (laughter, filled vs. unfilled pauses, and overlapped speech), and stress detection are all helpful. When it comes to lexical analysis, automatic speech recognition and keyword spotting are the two main techniques. Then use fusion method to sum the parts. These analysis methods apply to education research areas such as collaboration, argumentation and reasoning, teacher questioning, facilitation, and classroom ecology, student motivation and engagement. The method provides feedback for teaching, enables insider data collection, auto-extraction of argumentation instances, and enables longitudinal analysis of learning at scale. Finally, a few barriers include the following: error rates are still high for automated speech recognition of speech signals; youth speech development is an issue; linguistic variation is highly significant, both from youth to adults and across ages; naturalistic speech patterns are unexplored; naturalistic acoustic environments are challenging; few datasets of child speech exist.



Dr. Abeer Alwan (Professor at University of California-Los Angeles) presented speech processing techniques. Dr. Alwan argued that the variability in the way humans produce speech due to, for example, gender, accent, age, and emotion necessitates data-driven approaches to capture significant trends and behavior in the data. The same variability, however, may not be modeled adequately by such systems, especially if data are limited and corrupted by noise. Challenges in automatic speech recognition of children's' speech include the following aspects: (1) lack of large databases of children's speech Significant intra- and inter-speaker variability, (2) significant variability in pronunciations due to different linguistic backgrounds, and misarticulations, (3) low signal-to-noise ratio in the classroom, and (4) distinguishing reading errors from pronunciation differences. Due to these challenges, the processing of kids speech has significantly higher error rates than that of adult speech. Dr. Alwan developed pronunciation modeling and hypotheses to deal with children's speech processing. Some preliminary findings showed characteristics of vowels and constants. Dr. Alwan identified the need for an "expert human" (Golden Standard) in the loop for designing and evaluating the system, as well as the importance of professional development and

attention to ESL. Dr. Alwan plans to collect more data of both reading, story retelling, and other tasks. He is interested in looking into shared datasets. He also plans to measure in a longitudinal fashion how kid's speech changes with age, especially at the early ages (5-8). He will also look into challenging cases, such as Autism and ESL learners, which might help develop better systems for all (the Edison example).

Dr. Ajith Alexander (Oxford Wave Research) presented speech diarization. Dr. Alexander's team specializes in research and bespoke product development, speaker recognition, speaker diarization, mobile audio and voice analysis, media similarity and time-synchronisation, and forensic audio enhancement. He performed an analysis of sample audio from Dr. Yanghee Kim's child-robot interaction classroom recordings and was able to diarize the robot, facilitator, and two children. The major takeaways include the following aspects: (1) diarization works well for two speakers or one on one interactions of a child with a robot; (2) accuracy on multiple speakers tends to be low; (3) strong results distinguishing between "children's voices" versus robot and facilitator voices; (4) separation across each of the children yields weaker results; (5) ambient noise, including hum from robot movement, in a classroom environment poses challenges; (6) benefit of gender-based separation is weak when children are subjects; and (7) corpus of data is limited, diarization of children's speech as a field is largely nascent. Based on the analysis, Dr. Alexander proposed recommendations on future data collection for best diarization results: (a) constrain the recording environment to as few speakers as possible, (b) mic children individually during data collection phase to reduce post activity processing because post-processing is a harder problem, (c) use a room microphone to pick up background noises so that it can be subtracted from speech, (d) get children to speak longer phrases at least a few times in a recording versus getting yes/no answers because short clusters are harder to work with.

Dr. Marilyn Walker (professor at University of California-Santa Cruz) and Tony Zhao (doctoral student at CMU) presented dialog generation agents for children. Dr. Walker introduced her team's work on open domain dialog with Slugbot as part of the Amazon Alexa Prize Challenge. There are many challenges with building dialog agents for children still to be addressed; these include personalization, scaling conversational interaction, adapting to new domains, and multi-domain and multi-modal dialog systems. Because personalization is impossible with scripted behaviors, the goal is to develop technology that supports dialog interaction around any content, and to handle content that is narratively structured or expository, in addition to the most typical case of content from a structured database. They are also doing work on producing expressive linguistic and nonverbal behaviors. Currently, the NLDS lab at UCSC resources that can be useful for interacting with children around the content of a story, such as the DramaBank corpus of Story Intention Graphs (SIG) for Aesop's Fables, a corpus of gesture annotated stories, software for expressive personality generation of both verbal and nonverbal behaviors, and software for converting monologic story tellings to first person dialogic tellings of stories. They plan to specifically work on representations and dialogue management strategies for conversation with children using a robot in the context of the exhibits at UCSC's Seymour Marine Center,

controlling the nonverbal behaviors of the robot to demonstrate personality, and make the robot engaging to the children. They also are working on an NSF Cyberlearning Grant for literacy that involves an animated agent interacting with children to improve their narrative comprehension and social language skills, with an initial focus on stories such as Aesop's Fables, but possibly expanding to other children's stories.

Tony Zhao introduced DialPort, a dialog system that produces natural dialog in prescribed subjects, such as weather, restaurants. They propose that one way to collect data of interacting with children would be to involve their mothers. In the past, this clever solution provided good quality audio and facilitated IRB approval. Maxine Eskenazi at CMU who is the PI on DialPort would be interested in collaborating on an extended version of DialPort that could help with the paucity of databases of children's speech. They would like to discuss this possibility further as an extension to their current NSF CRI grant in collaboration with others at the workshop.

Ethnographic Observations (Lead: Dr. Laure Johnson, Northern Illinois University)

Dr. Laura Johnson (associate professor at Northern Illinois University), provided a brief presentation on conducting qualitative ethnographic observations. She began with an overview of ethnographic fieldwork and a discussion of the levels of participation a researcher might take on within fieldwork. She also outlined different types of notes a researcher might utilize within fieldwork and how these can be used to record descriptive and reflective information and evidence about the setting and participants. In particular, she highlighted Spradley's matrix for descriptive observations and how these help researchers pay attention to many elements within observations, including spaces, actors, activities, objects, goals, time, events, and feelings.

Dr. Johnson also discussed the use of particular theoretical and methodological approaches to observing children engaged in communication, such as Corsaro's (2012) work on interpretive reproduction and peer cultures, and the work of Hymes (1974) in the ethnography of communication. Questions after this presentation focused on how to ensure that observations are reliable and consistent across researchers/observers. Dr. Johnson emphasized the importance of researchers acknowledging their distinct lenses and perspectives, as informed by their disciplines and theoretical positions, and how these diverse perspectives can actually strengthen and enhance findings, providing more nuanced and complex explanations of particular phenomena and practices. Some researchers might also make use of procedures and processes, such as calculating inter-rater reliability, which helps a team achieve a degree of consensus regarding observations.

4. Future Work: Specific Breakout Group Recommendations

Robot 101 Platform & User experiences group

The group went through a list of robots that have been used in educational research. However, there is no satisfying educational robot platform. The group

discussed the desired features and proposed to design an ideal robot platform for education in classrooms and at home. To achieve that long-term goal, we need expertise from human factors, curriculum design, speech recognition, computer vision, pedagogy, and personalization.

The mid-term goal is to form a community for design, incorporating the following aspects: 1) advancement of each aspect of the robot: speech, vision, personality, interface, activity, curriculum design; 2) the best way to tell a story; 3) the best practice of speech recognition; 4) the best technique to detect children's emotions; 5) personalize the robot in terms of appearance and characters; and 6) generate robot dialogs and behaviors based on a chosen personality.

The short-term goal is to survey the existing educational robots and users' feedback on their experiences (benchmark) and identify users critical needs of robots' functionality and affordances.

Speech technique group

The speech technique breakout group focused on further development and collaboration of speech related techniques. Due to a great need for a children's speech database, this group discussed plans to build such a database. The parameters of metadata include 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), set of sentences representing full (ideally) phonetic coverage, spontaneous interaction data, 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. They will check existing corpora that match the characteristics listed and search for communities underrepresented in current data sets. The specific plan is to work toward an INCLUDES EAGER proposal.

Content Authoring group

The content authoring breakout group discussed the possibilities of using social robots in science learning. In scientific inquiry process, generating scientific questions in each inquiry cycle is important and a robot can help students to generate proper questions as well as evaluate their understanding through the questions and answers generated by students. To enable this support by a robot, we discussed what is the current state of dialogues generations and what needs to be further developed for supporting better automation, multi-party dialogues (speaker ID, localization), automated question generation (exploration/content), automated classification of responses. Discussions further focused on programming, improving students' performance, dialogue system, question pools, and evaluative tools based on students' responses. The emphasis was on better design of robot dialog and robot behaviors.

Expertise is required in the following areas: dialogue generation, science education, teacher education, HCI research and design, and learning sciences.

The project timeline will take at least three years. Year one is for developing content, data collection, and in parallel investigating adapting current dialogue management strategies to multi-party. Year two is to evaluate how well question generation works with initial data, reassess what extra data or functionality is needed, and refine multi-party systems. Refined components will be added, as well as qualitative observation studies. Year three is to synthesis the work and deploy field tests.

NSF Funding, Priorities, and CIRCL (Dr. Cynthia D'Angelo)

Dr. Cynthia D'Anglelo (senior researcher specializing in science education, simulations, games, and learning in SRI International's Center for Technology in Learning) introduced the 10 big ideas revealed by NSF in 2016 that represents the cutting-edge, long-term research topics in science and engineering. The ones that are most relevant to the workshop include the following items: The Future of Work at the Human-Technology Frontier; Harnessing Data for 21st Century Science and Engineering; Growing Convergent Research at NSF; and NSF INCLUDES (Inclusion across the Nation of Communities of Learners of Underrepresented Discoverers in Engineering and Science): Enhancing STEM through Diversity and Inclusion.

Dr. Cynthia D'Anglelo emphasized that cyberlearning is a cross-directorate effort and embraces several categories: Computer and Information Science and Engineering (CISE), Education and Human Resources (EHR), Engineering (ENG), and Social, Behavioral and Economic Sciences (SBE). A new program called Cyberlearning for Work at the Human-Technology Frontier has such interdisciplinary feature. The new program requires innovation in BOTH learning and computer science/tech. Lastly, D'Anglelo stated the general guidelines for different types of NSF funding grants and answered people's questions.

After Workshop Survey

Return rate. 25 people filled out the survey. The return rate is 62.5%.

1. *Overall impression of the workshop*

Six people reported moderately satisfied, and nineteen people reported satisfied.

2. *Experiences of communicating with workshop organizers*

22 people reported satisfied, two people reported moderately satisfied, and one person reported neutral.

3. *Whether the workshop met individuals' expectations*

20 people reported agree, three people reported moderately agree, and two people reported neutral.

4. *Whether the workshop sessions were useful and led participants to find new ideas to advance my research*

18 people reported agree, and five people reported moderately agree, and two people reported neutral.

5. *Whether the workshop overall helped participants better understand the issues around robot/child interaction design*

20 people reported agree, four people reported moderately agree, and one person reported neutral.

6. *Whether the participants find anyone they would like to collaborate in this topic area*

23 people reported yes.

7. *"what session/aspect of the workshop was most helpful for you? Please explain".*

14 participants reported the breakout discussions being most helpful, as well as other small discussions after the speakers' presentations. Five participants reported the technical presentations being most helpful, specifically, they mentioned the vision processing and speech recognition. Three participants reported the opening session being the most helpful with getting familiar with the research areas. The rest 3-4 people liked the overall arrangement, the hosting of the workshop, the funding priority talk, and the rich variety of topics.

8. *"what session/aspect of the workshop was least helpful for you? Please explain."*

23 people responded to this question. Eight people did not report any less helpful sessions. Then four people reported the talks not having enough interaction time, such as panel discussion, increasing discussion time after each talk, and adding wiki or small discussion groups on the first day as well. Another four people reported not being able to connect with the child development and learning theories session.

9. *Will you be interested in attending a similar workshop like this in the future?"*

All 25 people reported yes.

10. *"How would you improve the workshop in the future?"*

The improvement feedback included several aspects. First, people suggested more time for breakout discussion and spread out the first day talks. They also suggested increasing the number of breakout group topics and reduce the size of the groups. Based on the feedback, what might be better is to add a breakout session on the first day in the afternoon or evening, so we have a breakout session on both days

and people can participate in multiple groups. If possible, it is suggested to post the breakout groups earlier for consideration.

Second, there are three follow up recommendations: (a) make the workshop as an annual or biannual event; (b) build a resource website for recommended papers and educational robot materials; and (c) increasing participants' follow-up connections.

Third, logistics issues include having power strips at the tables on the first day, hosting the breakout groups in seperate rooms to reduce noise distraction, and using a different method for virtual participants.

Fourth, two participants suggested adding more technical researchers from Engineering and Computer Sciences, in addition to speech recognition. One concern is user study researchers and technical researchers may not find each other' topics engaging.

Dissemination of the Results of the Workshop and Next Step

Bringing the group of scientists and educational experts together for the workshop in this cross-disciplinary field will itself provide an impetus for enhanced research in the area. The participants will be encouraged to develop courses and projects exploring various aspects of this novel area. It is expected that the student participants will be inspired to conduct research studies in the ensuing topics. The workshop culminated in a Final Report with the recommendations [……]. Reports of the meeting will be disseminated by its publication in [somewhere].

This workshop was well received in that many participants talked about having it again annually or biannually. One possibility is to host the workshop domestically at NIU again next year, with the benefit of low cost and a continuation of participants. The other possibility is to host the workshop at other related conferences, with the benefit of having a more diverse audience, but many previous participants might not be able to attend international conferences.

One such related conference is International Conference of the Learning Sciences (ICLS). "ICLS is a major international event, organized biennially by ISLS, which gathers people involved in all aspects of the field of the learning sciences, including empirical, conceptual, theoretical, design-based, practitioner and policy perspectives. The conference theme for ICLS 2018 is "Rethinking learning in the digital age: making the Learning Sciences count,". Now more than ever, the learning sciences have a key role to play in unpacking the complexity of the teaching and learning process. AI and Automation in the workplace, including within education, will alter what we need to learn and how we need to teach it. Therefore, as scientists and educators, we need to explore learning in real-world settings in an interdisciplinary manner to understand how learning may be facilitated both with and without technology. In addition, there is now an additional imperative to guide the commercial development of Educational Technologies to ensure that they are pedagogically sound."

The second conference options is RO-MAN conference. “The 27th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2018, will be held in Nanjing and Tai'an, China, from 27 to 31 of August 2018. This leading forum covers state-of-the-art innovative results, the latest developments cover a wide range of topics related to Robot and Human Interactive Communication. The theme of the conference is social intelligence in interactive robots. Original contributions including basic findings, multi-disciplinary approaches towards friendly, open and useful robots in practical/real-life applications such as healthcare, industry edutainment, etc., are highly encouraged.”

Conclusion

The workshop successfully achieved its goal of forming a diverse community of children-robot interaction research by gathering lead researchers from a variety of backgrounds to update each other on the status of critical aspects of current research on this topic. The workshop not only educated participants on a wide range of areas within this field, but produced several tangible collaborative ideas. Participants expressed strong interest in working together as a community in the future.

Appendix 1 Participants

Name	Short Bio
Aaron, Kline	Lead Mobile Engineer Stanford University School of Medicine
Alexander, Ajith	Ajith Alexander represents the North America operations for Oxford Wave Research, a company focused on audio acquisition, processing and pattern recognition. Ajith is a Microsoft veteran, where he deployed user identity solutions at Fortune 500 companies including children's brands such as Lego. Ajith graduated from the University of Illinois at Urbana-Champaign in 2008 with a Bachelors and Masters in Electrical Engineering.
Alwan, Abeer	Abeer Alwan received her Ph.D. in EECS from MIT in 1992. Since then, she has been with the EE department at UCLA where she is now a Full Professor and Vice Chair of Undergraduate Affairs. She established the Speech Processing and Auditory Perception Laboratory at UCLA. Dr. Alwan is a recipient of several awards including: the NSF Research Initiation Award, the NIH FIRST Award, the UCLA-TRW Excellence in Teaching Award, Okawa Foundation Award in Telecommunication, and the Engineer Council Educator Award. She is a Fellow of the Acoustical Society of America, IEEE, and International Speech Communication Association (ISCA). She was a Fellow at the Radcliffe Institute, Harvard University, co-Editor in Chief of Speech Communication, Chair of the IEEE Flanagan Committee and Vice Chair of the IEEE Awards committee. She was recently elected to the Board of Governors of the IEEE Signal Processing Society.
Breazeal, Cynthia	Our research group develops social robots, deploys them, and studies their impact on human behavior and outcomes. For the purposes of this gathering, we have developed numerous technologies and run a number of longitudinal studies of social robots for personalized early childhood literacy skill development. We've been more recently exploring how younger children understand machine intelligence.

Daro, Vinci	Vinci Daro is Director of STEM Learning at Envision Learning Partners, and Research Partner at the Stanford Center for Assessment, Learning, and Equity. She has a PhD in cultural anthropology, from UNC Chapel Hill, and her postdoc was in Mathematics education at UC Berkeley. Her dissertation focused on the development activist identities in the context of transnational social movement networks and the “edge effects” of protest events. She has a professional background in mathematics curriculum development and teacher professional development, and her current research focuses on social emotional learning, student identity and agency in the context of mathematics learning, mathematical concepts learning in children, learning as a social process, and language development as part of conceptual learning in mathematics.
Dorsey, Chad	Chad Dorsey is President and CEO of the Concord Consortium, which has been an innovation leader in researching and developing STEM educational technology for more than two decades. Chad's experience ranges across the fields of science, education, and technology. Chad has led teacher professional development workshops as a member of the Maine Mathematics and Science Alliance, taught science in classrooms from middle schools through college, and guided educational reform efforts at the district and school levels. While earning his B.A. in physics at St. Olaf College and his M.A. in physics at the University of Oregon, Chad conducted experimental fluid mechanics research, built software models of Antarctic ice streams, and dragged a radar sled by hand across South Cascade Glacier.
Han, Insook	Insook Han is an Assistant Professor in the Department of Teaching and Learning at Temple University. Her research interest focuses on emerging technologies in teaching and learning, embodied cognition, and preservice teacher education.
Huang, Lixiao	Dr. Lixiao Huang is a Postdoctoral Associate at the Humans and Autonomy Lab at Duke University. She received her Ph.D. degree in Human Factors and Applied Cognition in Psychology from NC State University. Her research interests include 1) humans emotional, cognitive, and behavioral responses to robots (especially emotional attachment, intrinsic motivation, and trust in autonomy); 2) application of human factors inhuman robot interface design, and 3) the effectiveness of robotics education. Her recent work investigates humans' trust in autonomy through the development of the Human Autonomy Interface for Exploration of Risks (HAIER).
Laura, Johnson	Laura Ruth Johnson is an Associate Professor in the Department of Educational Technology, Research and Assessment at Northern

	Illinois University. She teaches classes in qualitative research methods, including courses in ethnographic research, interview methods, and community-based/participatory action research. Her recent book is entitled Community-based Qualitative Research: Approaches for Education and the Social Sciences.
Kim, Kyung	Dr. Kyung Kim is interested in examining the role of visualization technology in supporting science literacy of liguistically diverse learners (e.g., English language learners), for visually impaired learners, and for learners at risk of reading and writing (e.g., dyslexics).
Kim, Yanghee	Yanghee Kim, Ph.D, is the Morgridge Endowed Chair of Teacher Education and Professor of Educational Technology, Research, and Assessment at Northern Illinois University. Her team design human/computer interactions in educational applications, particularly using anthropomorphic interfaces, like virtual agents and embodied robots, for culturally and linguistically diverse groups of learners. Using up-to-date technology, they inquire into ways to enhance positive affect and identity, engagement, and learning of challenging topics.
Qi, Xiaojun	Dr. Xiaojun Qi is a Professor, the interim department head of Computer Science, and the director of Computer Vision Laboratory at Utah State University. She has expertise in artificial intelligence focusing on machine learning and computer vision. She has established and sustained a nationally recognized research program to solve challenging research problems. She has acquired competitive funding from reputable agencies and has published 97 high-quality manuscripts. She has worked as a PI on more than 20 research projects and has supervised 82 students. She has also gained her professional leadership experience by serving on technical program committees, NSF panels, and as a reviewer.
Ramani, Karthik	Ramani, Karthik is a professor in Mechanical Engineering and also Electrical and Computer Engineering. My research interests for this workshop in design based education with a hands-on build-play focus such as by using new robotic architectures as a means to learn by tinkering and making. I have been teaching a toy design class and conducting design workshops for children as well as collaborating with learning scientists.

Walker, Marilyn	Marilyn Walker, is a Professor of Computer Science at UC Santa Cruz, and a fellow of the Association for Computational Linguistics (ACL), in recognition of her for fundamental contributions to statistical methods for dialog optimization, to centering theory, and to expressive generation for dialog. Her current research includes work on computational models of dialogue interaction and conversational agents, analysis of affect, sarcasm and other social phenomena in social media dialogue, acquiring causal knowledge from text, conversational summarization, interactive story and narrative generation, and statistical methods for training the dialogue manager and the language generation engine for dialogue systems. Before coming to Santa Cruz in 2009, Walker was a professor of computer science at the University of Sheffield. From 1996 to 2003, she was a principal member of the research staff at AT&T Bell Labs and AT&T Research, where she worked on the AT&T Communicator project, developing a new architecture for spoken dialogue systems and statistical methods for dialogue management and generation. Walker has published more than 200 papers and has 10 U.S. patents granted or pending. She earned a B.A. in computer and information science at UC Santa Cruz, M.S. in computer science at Stanford University, and M.A. in linguistics and Ph.D. in computer science at the University of Pennsylvania.
Xie, Ying	Xie Ying is an Assistant Professor at Northern Illinois University. She earned her Ph.D. in Instructional Systems from Penn State University. Her current research focuses on the integration of emerging technologies and cognitive tools to promote a) reflective thinking, higher-order learning, knowledge construction and group cognition, b) second language acquisition, and c) mathematical problem posing.
Yoon, Ji-yoon	Jiyoon Yoon is Associate Professor of Science Education at the University of Texas at Arlington. She received her Ph.D. from Indiana University in Bloomington and her research has been focused on teacher education (Early Childhood through Grade 6) by integrating science with educational technology, like robots, for better understanding of science concepts.