

Designing Robots for Children: Multisensory Interactions and Multimodal Assessments

Abstract

In this symposium, a panel of researchers with multidisciplinary backgrounds will introduce the AERA community to the fast-growing research on educational robots for children. The panel will present the current status of educational robotic research, relevant theoretical frameworks, and technological advances involved in the design and assessment of child robot interactions. The audience will be invited to discuss with the panel designing robotic interventions that can support teachers and learners in classrooms and that can provide inclusive contexts to overcome the real-world constraints. We will also invite discussion on ways to facilitate cross-disciplinary research collaboration among educational researchers, engineers, and computer scientists for sustainable robotic research geared toward promoting children's intellectual, affective and social development.

Symposium Summary

Objective

The development of humanoid robots has been growing fast over the last decade, and education is recognized as one of the most promising areas for application of humanoid robotics. The objective of this symposium is to introduce the AERA community to this fast-growing research area and invite the audience to examine the promise and limitations of educational robots through productive discussions geared toward answering six questions:

- Q1. What is the current status of research and development on child/robot interaction (CRI)?
- Q2. What are the cognitive and social theoretical perspectives that may guide research on CRI?
- Q3. In what ways can CRI be designed to be culturally sustainable for all children?

Kim, Y., Han, I., Daro, V., Kline, A., D'Angelo, C. M., & Johnson, L. R. (2019). Symposium on designing robots for children: Multisensory interactions and multimodal assessments. *The Annual Conference of American Educational Research Association (AERA)*: Toronto, Canada. April 5 - 9.

Q4. What is the potential of robots to augment the capabilities of teachers and students in classrooms?

Q5. What technologies are available to design and evaluate CRI?

Q6. What are the opportunities and challenges in this line of research?

Overview and Structure

The first sixty minutes will be devoted to the presentations by a panel of researchers with multidisciplinary backgrounds. Presentation 1 will introduce three representative lines of research where educational robots have been applied. Next, embodied cognition (Presentation 2) and social emotional learning (Presentation 3) will be discussed as theoretical foundations for the design and research on child/robot interaction. Following that, two cutting-edge technologies will be introduced as unobtrusive measures of the efficacy of CRI: speech recognition and processing (Presentation 4) and computer vision and image recognition (Presentation 5). Presentation 6 will discuss how ethnography can complement and validate the information gathered from the technology-based measurements. The remaining thirty minutes will be open for audience participation, devoted to joint discussion between the panel and the audience, focused on Q3 and Q4. The audience will be invited to deliberate on how humanoid robots can be designed to promote children's intellectual, affective and social development and also encouraged to bring in thought-provoking questions and constructive criticism.

Significance

Compared to other electronic devices, humanoid robots offer unique social affordances (Breazeal, 2002) and are often designed as playmates. The embodiment and mobility of a robot can make CRI more natural and social, adequately supporting children's holistic development (i.e., both psycho-motor and socio-cognitive development). The various sensors embedded in the

Kim, Y., Han, I., Daro, V., Kline, A., D'Angelo, C. M., & Johnson, L. R. (2019). Symposium on designing robots for children: Multisensory interactions and multimodal assessments. *The Annual Conference of American Educational Research Association (AERA)*: Toronto, Canada. April 5 - 9.

robot's body can capture real-time multimodal information of CRI unobtrusively, enabling the assessment of robotic interventions to be comprehensive and ecologically valid. Importantly, when designed carefully, robots can facilitate inclusive learning contexts since robots, speaking in any language, are relatively free from social and cultural biases prevalent in humans. The effectiveness of educational robots to date has been promising, but educational robotic research has been limited. This research has been growing mainly in engineering and computer science, and is gaining attention from policy makers worldwide. The research is often missing solid learning-theoretical grounding. The research questions fail to reflect current educational needs. Through this symposium, we hope to invite the interest of educational researchers in educational robotic research and facilitate cross-disciplinary efforts for sustainable educational robotic research.

Presentation 1: Current Research on Sociable Educational Robots

This presentation will address *Q1. What is the current status of research and development on child/robot interaction?* The presentation will introduce three lines of educational robot research for young children: 1) a robot as a personalized learning companion for early literacy, 2) a robot as a mediator for children's collaboration, and 3) the role of a robot and a human facilitator to support ASD children's emotional development.

The Personal Robots Group at the MIT Media Lab pioneered sociable robots for young children (<https://www.media.mit.edu/groups/personal-robots/overview>). The personalized peer-like sociable robots provide friendly companionship, collaboration, perspective-taking, social modeling, and emotional engagement. This research has found that (1) children are more engaged and become more relational toward robots as companions than they did to an avatar; (2) children retain more phrases and words used by the robot, as well as telling longer storylines

Kim, Y., Han, I., Daro, V., Kline, A., D'Angelo, C. M., & Johnson, L. R. (2019). Symposium on designing robots for children: Multisensory interactions and multimodal assessments. *The Annual Conference of American Educational Research Association (AERA)*: Toronto, Canada. April 5 - 9.

with expressive robots than with non-expressive robots; (3) children demonstrate a growth mindset after interacting with a growth mindset robot; (4) children become more emotionally expressive with a robot than with a tablet; and (5) the emotional data extracted from 218 hours of time-synchronized multimodal interaction have improved the prediction of children's word-reading skill.

Second, project IDEAL (<http://www.createresearch.net/projects.html>) is an inclusive design that uses a robot to mediate collaboration among children coming from different linguistic and cultural backgrounds, grounded in multicultural education and intercultural communication (Kim, et al., 2018). The robot-mediation strategies are threefold. The robot (1) consistently invites children into conversations, (2) provides opportunities for children to speak and engage in activities in either their native or a second language, and (3) always demonstrates empathy with children. This research to date shows that (1) children develop affectionate relationships and are very engaged with the robot, (2) children interact with the robot as they would with a friend, (3) children are very forgiving of the robot's mistakes, and (4) children gradually learn to work with their peers, by taking turns and listening.

Third, another popular line of research deals with robotic intervention for children with Autism Spectrum Disorder (Shamsuddin, et al., 2012; Kim, et al., 2013), examining the inclusion of human facilitators in the children's interactions with a robot. This examination is critical in that i) even the state-of-art robots are not robust enough to meet children's varied needs, ii) humans have the basic psychological need for relatedness to other humans that robots cannot replace, and iii) the goal of the child-robot research is to increase children's adaptability in the real world, so adding humans into the loop ultimately provides a seamless transition from child/robot interaction to human-to-human interaction, eventually leading to children's well-

Kim, Y., Han, I., Daro, V., Kline, A., D'Angelo, C. M., & Johnson, L. R. (2019). Symposium on designing robots for children: Multisensory interactions and multimodal assessments. *The Annual Conference of American Educational Research Association (AERA)*: Toronto, Canada. April 5 - 9.

being. In an example of applying a robot to treat stuttering, the learning activities and tutor's facilitation becomes more important than the robot itself (<http://www.design4hri.net/research-projects.html>).

Presentation 2: Application of Embodied Cognition to Research on Child Robot Interaction

This presentation will discuss the theory of embodied cognition as a theoretical framework for CRI research and design, addressing *Q2. What are the cognitive and social theoretical perspectives that may guide research on child/robot interaction (CRI)?* A recent view of cognition emphasizes that our perceptual experiences, such as seeing, hearing and feeling, are vital elements to reach conceptual understanding because our thoughts and knowledge emerge from dynamic interactions between our body and the physical world. The primary lines of research have discussed embodied cognition for many years. Cognitive linguistics, which explains the cognitive process of understanding linguistic expressions with metaphorical mapping process (e.g., Lakoff & Johnson, 1999), Cognitive psychology explains how people understand a new concept with perceptual simulation based on prior multisensory experiences (e.g., Barsalou, 2008). Based on this emerging paradigm of understanding human cognition created by merging these two lines of research, educational researchers have investigated the effectiveness of interventions that provide perceptual and embodied experiences in learning. Examples of embodied interventions include the use of gestures (Hu, Ginns, & Bobis, 2015) and bodily action (Sidhu & Pexman, 2016), which were often enhanced by technologies such as multimodal simulations including haptic sensations (Han & Black, 2011), touch screens (Agostinho et al., 2015), or augmented reality (Lindgren & Johnson-Glenberg, 2013). In addition to these technologies, educational robots have gained scholarly attention as emerging technology that has a potential to facilitate young children's learning. While previous studies on embodied

cognition can provide implications for designing and developing educational robots as well as child/robot interactions, little research has been done to connect these two evolving areas.

Young children's interaction with robots using gestures and bodily actions is one area to be explored based on embodied cognition. Considering that the majority of recent studies have examined the effectiveness of embodiment in mathematics, science and language learning, educational robots can also be applied to conceptual learning. For example, the use of gestures has been researched as evidence that the body is involved in thinking (Alibali & Nathan, 2012). In mathematics learning, in particular, an educational intervention that fosters the use gestures that correspond to mental representations required to solve problems can facilitate creation and elaboration of new ideas (Goldin-Meadow, Cook, & Mitchell, 2009). Also, bodily actions and behavioral performances enhance learning in language acquisition (Lan, Fang, Legault, & Li, 2015) and science (Lindgren & Johnson-Glenberg, 2013). Thus, robots can be used as a learning companion to facilitate learners' cognitive processes by modeling gestures/actions, evaluating learners' gestures/actions to assess their level of conceptual understanding, providing proper gestural feedback, or recording gestural data.

Considering young learners tend to learn better with multisensory experiences and perceptual/bodily activities and are motivated by the use of educational robots, research exploring the intersection between embodied cognition and robots will help us design and develop appropriate interventions for young learners.

Presentation 3: Social and Emotional Learning through Child Robot Interaction

This presentation will address the cognitive and social theoretical perspectives that may guide research on child/robot interaction (CRI). A focus in the past 15 years among educators, policy-makers, and researchers on "social emotional learning" provides another lens for

understanding affordances and limitation of humanoid robots in contexts of schooling and learning. The purpose of this presentation is to bring social practice theory (Holland and Lave 2009; Bourdieu 1990) to bear on questions about possible roles for humanoid robots in learner identity development processes.

Social emotional learning (SEL) is a category of competencies that have been identified by education reformers and researchers as valued outcomes of school that are beyond mastery of narrowly defined disciplinary skills and academic content learning. Focusing on this 'beyond' – also referred to as “non-cognitive skills,” “metacognition,” and “soft skills” – supports research on collective and individual identity processes, particularly insofar as SEL frameworks call for attention to constructs such as ‘student agency’ and ‘collaboration skills’ and ‘self-management’.

Three ongoing challenges in conducting research on social emotional learning in school contexts are (1) a tendency in research designs (and intervention designs) to isolate the instruction or assessment of SEL skills from academic content instruction; (2) the difficulty of *measuring* SEL competencies, and measuring *growth* in these competencies; and (3) the significant role of teacher and researcher biases in interactions with students from different backgrounds, including linguistically diverse students. Humanoid robots offer promise to researchers wanting to side-step these challenges, while simultaneously facilitating the very competencies and positive learner identity processes associated with valued SEL outcomes.

Humanoid robots offer opportunities to experiment with ways of supporting students' sense of belonging (Goode and Inzlicht 2006) and sense of agency in school learning contexts. Giving students “voice and choice” during learning tasks, including opportunities to make decisions, revise their thinking and their work, and interpret and critique others' ideas are known approaches toward more equitable school experiences and outcomes.

Because identity development is related to interaction – and both individual action and collective action – and both are deeply social, one’s identity, or ‘sense of self’, develops over time in many different contexts of social and cultural experience. Humanoid robots offer novel ways of cultivating a “modicum of agency” (Holland et. al. 1996) by which a person can define themselves, rather than being thoroughly and continually defined and determined by the social and cultural forces in their immediate surroundings and historically-marked social position. This presentation will provide an analysis of video of student-robot interaction that illustrates how social practice theory can be applied to better understand students’ identity development as learners, including students’ use and development of language, when mediated by social robots.

Presentation 4: Computer Vision Affordances for Child Robot Interaction

Recent advancements in computer vision are providing exciting new opportunities to develop child/robot learning experiences that both respond to participant involvement and engagement interactively, and measure learning intervention efficacy. Addressing the question *Q5*, this presentation will discuss an array of these computer vision techniques and their applications. The presentation will also introduce a case study of an interactive experience for reinforcing emotion recognition for children with autism (Washington, 2017). Addressing the question *Q3*, we will also discuss key challenges in the field, particularly the impact of diversity on facilitating equitable learning contexts.

Face tracking, emotion recognition and pose estimation are areas of computer vision that can provide educators with important insight into the reactions, body language and overall engagement of the participants in a learning experience. Recent advancements allow for these techniques to be employed in real-time via increasingly affordable technology. One example of

Kim, Y., Han, I., Daro, V., Kline, A., D'Angelo, C. M., & Johnson, L. R. (2019). Symposium on designing robots for children: Multisensory interactions and multimodal assessments. *The Annual Conference of American Educational Research Association (AERA)*: Toronto, Canada. April 5 - 9.

this, “Superpower Glass” (Washington, 2017) leverages these techniques and wearable and mobile technologies to reinforce emotion recognition for children with autism. While a child socially interacts with others, the system tracks the emotional responses of the participants and provides reinforcing cues for the child as to the emotions they may be seeing. A session review app allows the child and caretakers/educators to replay the video of an interaction and the emotions that were detected, fostering further discussion. Initial studies have found that children with autism strongly connect with the technology, show increased socialization skills, and exhibit them in contexts outside of the learning sessions themselves (Washington, 2016; Voss, 2016; Washington, 2017). The session videos are analyzed for engagement and efficacy using the same face tracking and emotion recognition capabilities used for the sessions themselves – demonstrating how such techniques can serve to both create interactive learning experiences and assess their impact.

Other recent advancements in computer vision extend these capabilities beyond learners to the spaces and objects around them. Object detection and semantic segmentation focus on providing insight into the overall understanding of a scene – not just tracking the people and objects within view, but also determining the relationships between them. Equipping child/robot interactions with this level of contextual awareness opens exciting opportunities to foster collaboration and track the progression of complex multi-step learning processes.

However, in order for these capabilities to translate into experiences that are equitable for all learners, diversity in the field of artificial intelligence, as with STEM more broadly, still remains a key challenge. Ethnic and gender underrepresentation seen in academic programs and in the workplace are also reflected in the massive datasets used to drive these technological advancements. As a result, we risk providing educators focusing on increasing engagement of

underrepresented groups with tools that do not perform as effectively for those very learners.

Thus, bridging these diversity gaps is critical for the creation of equitable learning contexts.

Addressing these challenges will then allow these advancements in computer vision, along with speech recognition and similar technologies, to fully enable exciting new multimodal opportunities to create interactive, collaborative child/robot experiences for all learners.

Presentation 5: Speech Recognition and Processing

This presentation will include an overview of speech data collection and speech processing procedures and methods, addressing the question, *Q5. What technologies are available to design and evaluate CRI?* It will discuss advances and issues surrounding automated speech recognition, speaker diarization (trying to distinguish between multiple speakers in one audio stream), and microphone options.

Speech-related technologies that most consumers are familiar with (year and making their way into more and more household devices (e.g., Alexa, Siri, Hey Google, etc.) involve many different technologies stacked together to produce a service, like a personalized assistant. Developing interactions and rules for a robot to listen and/or talk to a child involves many of these same technologies, the most important of which is probably automated speech recognition (ASR). While algorithms and advances in ASR for adult speech have progressed significantly over the past decade, these advances have not translated to successful ASR for younger speakers (Russell, 2007; Gerosa 2009). Young children have vocal tracts that are significantly smaller than those of adults, causing speech to occur at higher frequencies and children's speech contains much higher variability than adult speech (Russell, 2007; Gerosa, 2009). Due to the dramatic differences in their speech, when recognition systems trained on adult speech are applied to speech from younger children, error rates can easily double or triple (Russell, 2007).

ASR is an important part of many systems that are attempting to determine how well children can read or trying to engage in conversations, but there are other types of analyses from speech data that are relevant for other kinds of applications. Speech contains more information than just what words might have been said. Speech features such as prosody (tonality), rate of speech, loudness contours, and intonation patterns (e.g., questioning vs. statement), as well as more complex utterance styles such as sarcasm, reflect more than just propositional meaning. Also, patterns of interaction between or among speakers—such as who speaks more, latencies between conversational turns, disfluencies and discourse markers, or how often individual speakers overlap with each other or negotiate who gets the floor—provide rich information about individual speakers and about groups as a whole (e.g., Ford & Couper-Kuhlen, 2004; Sacks, Schegloff, & Jefferson, 1974). Many of these types of speech analyses are more successful with younger children since they do not require ASR techniques, however they are still not up to the same accuracy levels as adult speakers.

The presentation will also discuss different projects that have made advances in conversational analysis or tools that use speech processing and related technologies. There are many different applications for speech technologies within conversational agents. These systems are text-based, like DialPort and SlugBot that can help answer specific questions. Nonetheless, interactions between humans and robots is one of the main areas where much research is focused right now.

Presentation 6: Ethnographic Observation of Child Robot Interaction

This presentation will provide an overview of ethnographic methods of participant observation and discuss their use in studies that involve communication and collaboration between children and humanoid robots. The presentation will i) review key principles and

features of ethnographic participant observation, ii) discuss key theoretical frames that can be used to interpret communicative interactions within robot-mediated communicative activities involving young children, and iii) provide analytical examples from an on-going study of how humanoid robots can facilitate cross-cultural and bi-lingual collaborative interactions between kindergarten aged-children from Spanish and English dominant families.

Participant observation is an essential approach in ethnographic field work, involving immersive participation and fine-grained observation of particular settings in order to better understand beliefs, practices, behaviors, and interactions within specific contexts (Spradley, 1980). The focus in these observations is on participant meanings and local knowledge—what ethnographers have termed *emic*—and “thick description” (Geertz, 1973); researchers must set aside their own beliefs and assumptions about what is “going on” in setting and avoiding foisting external meanings on the context. Ethnographic observations are largely inductive; the goal is not to test hypotheses developed *a priori*, but allow theories and hypotheses to emerge from the setting, which will become working hypotheses that will be developed and refined through further observations (Erickson, 1986).

Participant observation occurs along a continuum, and researchers may decide to take a more passive role in settings as they observe—acting essentially as a bystander at a scene—or a more active one, where they are seeking to understand cultural rules and norms of behavior by participating in activities of setting, (Spradley, 1980). Most classroom studies involve researchers as moderate participants, where they are peripherally involved and maintaining a balance between roles as participant and observer. Researchers may audio or video record observations—and this would be essential in studies examining in-depth communication and dialogue—along with preparing descriptive field notes from observations, which would include

Kim, Y., Han, I., Daro, V., Kline, A., D'Angelo, C. M., & Johnson, L. R. (2019). Symposium on designing robots for children: Multisensory interactions and multimodal assessments. *The Annual Conference of American Educational Research Association (AERA)*: Toronto, Canada. April 5 - 9.

portraits of participants, descriptions of setting, and accounts of particular events (Emerson, Fretz, & Shaw, 2011). Initial observations are highly descriptive, whereas later observations become more focused. Researchers may also prepare reflective, methodological, and analytical notes and memos.

A useful interpretive frame for making sense of interactions among humanoid robots and children include the ethnography of communication, which views language as occurring within socially contextualized speech communities and settings (Hymes, 1974); the focus of analysis of ethnographic data is on speech situations, events, and acts, which can help researchers identify roles of interlocutors, types of and norms for communication, and how certain communicative practices function within the setting. Observations of children also need to be mindful of issues of power—among children, between children and adults, and among adults—and reflect on one's own positionality as an “adult” among children (Corsaro, 2012; Eder & Corsaro, 1999; Thorne, 1993). Researchers must authentically look and listen to children and recognize children's active production of knowledge and peer culture (see Corsaro's theory of interpretive reproduction, Corsaro, 2012).

References

- Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the learning sciences, 21*(2), 247-286.
- Agostinho, S., Tindall-Ford, S., Ginns, P., Howard, S. J., Leahy, W., & Paas, F. (2015). Giving learning a helping hand: finger tracing of temperature graphs on an iPad. *Educational Psychology Review, 27*(3), 427-443.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645.

- Kim, Y., Han, I., Daro, V., Kline, A., D'Angelo, C. M., & Johnson, L. R. (2019). Symposium on designing robots for children: Multisensory interactions and multimodal assessments. *The Annual Conference of American Educational Research Association (AERA)*: Toronto, Canada. April 5 - 9.
- Bourdieu, P. (1977). *Outline of a Theory of Practice*. New York: Cambridge University Press.
- Bourdieu, P. (1990). *The Logic of Practice*. Palo Alto: Stanford University Press.
- Corsaro, W.A. (2012) Interpretive reproduction in children's play. *American Journal of Play*, 4(4), 488-504.
- Eder, D. & Corsaro, W. (1999). Ethnographic studies of children and youth: Theoretical and ethical issues. *Journal of Contemporary Ethnography*, 28(5), 520-531.
- Emerson, R.M., Fretz, R.I. & Shaw, L.L. (2011). *Writing ethnographic field notes*. Chicago, IL: University of Chicago Press.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. Wittrock (Ed.) , *Handbook of research on teaching* (pp. 119-161). New York, NY: McMillan.
- Ford, C. E., & Couper-Kuhlen, E. (2004). Conversation and phonetics: Essential connections. *Sound Patterns in Interaction. Cross-Linguistic Studies from Conversation. Amsterdam: Benjamins*, 3–25.
- Geertz, C. (1973). Thick description: Towards an interpretive theory of culture. In *The interpretation of cultures*. New York, NY: Basic Books.
- Gerosa, M., Giuliani, D., Narayanan, S., & Potamianos, A. (2009). A review of ASR technologies for children's speech. In *Proceedings of the 2nd Workshop on Child, Computer and Interaction - WOCCI '09* (pp. 1–8). New York, New York, USA: ACM Press. doi:10.1145/1640377.1640384
- Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, 20(3), 267-272.

- Kim, Y., Han, I., Daro, V., Kline, A., D'Angelo, C. M., & Johnson, L. R. (2019). Symposium on designing robots for children: Multisensory interactions and multimodal assessments. *The Annual Conference of American Educational Research Association (AERA)*: Toronto, Canada. April 5 - 9.
- Good, C., & Inzlicht, M. (2006). How environments can threaten academic performance, self-knowledge, and sense of belonging. In *Stigma and group inequality: Social Psychological Approaches*. Shana Levin and Colette Van Laar, eds. Mahwah, NJ: Erlbaum
- Han, I., & Black, J. B. (2011). Incorporating haptic feedback in simulation for learning physics. *Computers & Education*, 57(4), 2281-2290.
- Holland, D., Lachicotte, W., Jr., Debra Skinner, D., & Cain, C. (1998). *Identity and Agency in Cultural Worlds*. Cambridge, Mass: Harvard University Press.
- Holland, D., & Lave, J. (2001). *History in Person: Enduring struggles and identities in practice*. Santa Fe: School of American Research Press.
- Holland, D., & Lave, J. (2009). Social practice theory and the historical production of persons. *In Actio: An International Journal of Human Activity Theory* (2), 1-15.
- Hu, F. T., Ginns, P., & Bobis, J. (2015). Getting the point: tracing worked examples enhances learning. *Learning and Instruction*, 35, 85-93.
- Hymes, D. (1974). *Foundations in sociolinguistics: An ethnographic approach*. Philadelphia, PA: University of Pennsylvania Press.
- Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., & Scassellati, B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders*, 43(5), 1038-1049. doi:10.1007/s10803-012-1645-2
- Kim, Y., Marx, S., Pham, H., & Nguyen, T. (2018). *Designing technology as a cultural broker for young children: Challenges and opportunities*. The 13th International Conference of the Learning Sciences (ICLS): London, UK. June 23 - 27, 2018
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh* (Vol. 4). New York: Basic Books.

Kim, Y., Han, I., Daro, V., Kline, A., D'Angelo, C. M., & Johnson, L. R. (2019). Symposium on designing robots for children: Multisensory interactions and multimodal assessments. *The Annual Conference of American Educational Research Association (AERA)*: Toronto, Canada. April 5 - 9.

Lan, Y. J., Fang, S. Y., Legault, J., & Li, P. (2015). Second language acquisition of Mandarin Chinese vocabulary: Context of learning effects. *Educational Technology Research and Development*, 63(5), 671-690.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher*, 42(8), 445-452.

Penuel, W. R., DiGiacomo, D. K., van Horne, K., & Kirshner, B. (2016). A social practice theory of learning and becoming across contexts and time. *Frontline Learning Research* 4(4), 30-38.

Russell, M. J., & D'Arcy, S. M. (2007). Challenges for computer recognition of children's speech. In *Proceedings of the Speech and Language Technology in Education Workshop (SLaTE2007)* (pp. 108–111). Retrieved from http://www.isca-speech.org/archive/slate_2007/sle7_108.html

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 696–735.

Shamsuddin, S., Yussof, H., Ismail, L. I., Mohamed, S., Hanapiah, F. A., & Zahari, N. I. (2012). Initial response in HRI-a case study on evaluation of child with autism spectrum disorders interacting with a humanoid robot Nao. *Procedia Engineering*, 41, 1448-1455.

Sidhu, D. M., & Pexman, P. M. (2016). Is moving more memorable than proving? Effects of embodiment and imagined enactment on verb memory. *Frontiers in psychology*, 7, 1010.

Spradley, J.P. (1980). *Participant observation*. Belmont, CA: Wadsworth.

Thorne, B. (1993). *Gender play: Girls and boys in school*. Rutgers University Press

- Kim, Y., Han, I., Daro, V., Kline, A., D'Angelo, C. M., & Johnson, L. R. (2019). Symposium on designing robots for children: Multisensory interactions and multimodal assessments. *The Annual Conference of American Educational Research Association (AERA)*: Toronto, Canada. April 5 - 9.
- Voss, C., Washington, P., Haber, N., Kline, A., Daniels, J., Fazel, A., De, T., McCarthy, B., Feinstein, C., Winograd, T., & Wall, D. (2016). Superpower glass: delivering unobtrusive real-time social cues in wearable systems. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*, 1218-1226
- Vygotsky (1978). *Mind in Society*. Cambridge, Mass.: Harvard University Press. Chap. 7, "The Role of Play in Development": 92-104.
- Washington, P, Voss, C., Haber, N., Tanaka, S., Daniels, J., Feinstein, C., Winograd, T., & Wall, D. (2016). 2016. A wearable social interaction aid for children with autism. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI '16)*, 2348-2354
- Washington, P., Voss, C., Kline, A., Haber, N., Daniels, J., Fazel, A., De, T., Feinstein, C., Winograd, T., & Wall, D. (2017). Superpower glass: A wearable aid for the at-home therapy of children with autism. In *Proceedings of the 2016 ACM International Joint Conference on Mobile Wearable Ubiquitous Technology* 1, 3, Article 112 (September 2017), 22 pages. doi: [10.1145/3130977](https://doi.org/10.1145/3130977)
- Zwiers, J. (2017). Developing Oral Language to Foster Students' Academic Literacy: Cultivating Students' Inner Language of Comprehending Through Classroom Conversation. Chapter 8 of the *Handbook of Research on Teaching the English Language Arts, 4th Edition*. Diane Lapp and Douglas Fisher, eds. International Reading Association and the National Council of Teachers of English.
- Zwiers, J. (2011). *Academic Conversations: Classroom talk that fosters critical thinking and content understandings*. Portsmouth, NH: Stenhouse.