Published in final edited form as:

J Autism Dev Disord. 2016 November ; 46(11): 3615–3621. doi:10.1007/s10803-016-2896-0.

Evaluation of an Intelligent Learning Environment for Young Children with Autism Spectrum Disorder

Zhi Zheng¹, Zachary Warren^{2,3,4,5}, Amy Weitlauf^{2,5}, Qiang Fu¹, Huan Zhao¹, Amy Swanson⁵, and Nilanjan Sarkar^{1,5,6}

¹Departments of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN

²Department of Pediatrics, Vanderbilt University, Nashville TN

³Department of Psychiatry, Vanderbilt University, Nashville TN

⁴Department of Special Education, Vanderbilt University, Nashville TN

⁵Vanderbilt Kennedy Center, Treatment and Research Institute of Autism Spectrum Disorders, Vanderbilt University, Nashville, TN

⁶Department of Mechanical Engineering, Vanderbilt University, Nashville, TN

Abstract

Researchers are increasingly attempting to develop and apply innovative technological platforms for early detection and intervention of Autism Spectrum Disorder (ASD). This pilot study designed and evaluated a novel technologically-mediated intelligent learning environment with relevance to early social orienting skills. The environment was endowed with the capacity to administer social orienting cues and adaptively respond to autonomous real-time measurement of performance (i.e., non-contact gaze measurement). We evaluated the system with both toddlers with ASD (n=8) as well as typically developing infants (n=8). Children in both groups were able to ultimately respond accurately to social prompts delivered by the technological system. Results also indicated that the system was capable of attracting and pushing toward correct performance autonomously without user intervention.

Keywords

autism spectrum disorders; social communication; adaptive systems; early identification

Correspondence concerning this article should be addressed to Zachary Warren, Vanderbilt University, 230 Appleton Pl., PMB #74, Nashville TN 37203, (615) 936-7159, phone: 615-601-9253, fax: 615-343-6687. zachary.warren@vanderbilt.edu.
Zhi Zheng, Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, USA. Zachary Warren, Department of Pediatrics and Psychiatry and Special Education, Vanderbilt Kennedy Center/Treatment and Research Institute for Autism Spectrum Disorders, Vanderbilt University, Nashville, TN, USA. Amy Weitlauf, Department of Pediatrics, Vanderbilt Kennedy Center/Treatment and Research Institute for Autism Spectrum Disorders, Vanderbilt University, Nashville, TN, USA. Qiang Fu, Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, USA. Huan Zhao, Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, USA. Amy Swanson, Vanderbilt Kennedy Center/Treatment and Research Institute for Autism Spectrum Disorders, Vanderbilt University, Nashville, TN, USA. Nilanjan Sarkar, Departments of Electrical Engineering and Computer Science and Mechanical Engineering, Vanderbilt University, Nashville, TN, USA.

As we unravel the complex neurogenetic and environmental influences of Autism Spectrum Disorder (ASD), there are increasing opportunities and needs for the development of sophisticated technological tools that help us better understand very early symptom profiles for detection and potentially intervention/prevention. With recent advances in our understanding of the earliest neurobehavioral vulnerabilities associated with autism spectrum disorder (ASD) in risk samples, there is more capacity than ever to conceptualize the emergence of developmental vulnerabilities at extremely early points in development (i.e., infancy) (Veenstra-VanderWeele and Warren 2015). Adaptive technological systems capable of both quantifying and responding in real-time to behaviors highly relevant to very early social communication development with precision beyond common human observation may represent a promising opportunity in this regard. In the current work we designed and evaluated the feasibility/tolerability of an adaptive, intelligent learning environment capable of (1) utilizing non-contact eye-gaze tracking to index visual attention and (2) adaptively responding to this gaze measurement in a manner that might promote social orienting over time. Specifically, we developed and tested a system with infants and young toddlers that was capable of (1) calling the child's name, (2) autonomously detecting the child's response, and (3) adapting the system to promote correct orientation to subsequent name calls. We evaluated how toddlers with ASD responded to the system and how infants and younger toddlers would respond as proof-of-concept of feasibility and potential clinical value of future sophisticated technological applications with ASD risk populations.

Multiple studies have used eye tracking methodologies to assess deficits in social orienting and responsiveness in young children with and at risk of ASD (Bedford et al. 2012; Bekele et al. 2014; Chevallier et al. 2015; Falck-Ytter et al. 2012; Guillon et al. 2014; Jones and Klin 2013; Magrelli et al. 2013; Noris et al. 2012; Rehg et al. 2013; Shic et al. 2013). However, these systems have not typically utilized eye tracking information to alter feedback in responsive, real-time settings. Such systems hold appeal in that they reduce the need for human participation or guidance, operating in relative independence to scaffold prompts for certain behaviors and, in theory, optimize participants' engagement with defined targets within system. Further, evidence exists that children with ASD may initially respond more readily to robotic or technological systems as compared to some aspects of early social orienting (Klin et al. 2009; Warren et al. 2014). Given the limited availability of human therapists in many communities, adaptive technological systems may hold future promise for supplementing (but not replacing) evidence-based one-on-one care. The concept of portable, accessible systems that respond to children's gaze or physiological signals to optimize interactions, although purely experimental at this point, is an important avenue that we explore within the current work. Previous work (Bekele et al. 2014; Warren et al. 2014; Zheng et al. 2013) has examined the use of adaptive intervention platforms for children and adolescents with ASD using both robotic and Virtual Reality platforms, with physiologically responsive systems showing promise for enhancing social engagement (Lahiri et al. 2015) in adolescents. However, this study represents a novel downward extension of these principles to an autonomous system for use with infants and young toddlers.

To establish the feasibility and proof-of-concept of such an approach, we present a pilot study of a system that integrates video stimuli with real-time, responsive, non-invasive eye tracking technology to measure children's responses to their name being called within the

technological environment itself. A poor response to one's name being called, in children with adequate hearing, has consistently been identified as an early red-flag for ASD in infancy and toddlerhood (Gammer et al. 2015; Nadig et al. 2007; Stenberg et al. 2014) and is often included as part of standardized, widespread screening checklists and diagnostic instruments. In the current work we evaluated the system with toddlers with ASD, as well as infants with unknown risk status to demonstrate the system's usability, tolerability, and applicability to these respective groups.

Methods

Participants

Sixteen children participated in this pilot study, eight toddlers with a clinical diagnosis of ASD (average age 2.19 years, range: 23–30 months) and eight infants and toddlers without developmental concerns (TD: average age 1.33 years, range: 9–22 months). Participants with ASD were recruited through an existing university clinical research registry. The registry includes individuals who received a clinical diagnosis of ASD from a licensed clinical psychologist and scored at or above the clinical cutoff on the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord et al. 2012). The registry also contains information regarding cognitive abilities (Mullen Scales of Early Learning, 1995). All toddlers with ASD received a score of 2 or above on the ADOS-2 Response to Name item (6 children's scores = 3; 2 children's scores = 2). Such item scores indicate that during this assessment the children demonstrated challenges responding to their name when called by a human administrator and their caregiver.

Participant characteristics are listed in Table 1. We intentionally designed the user study with two separate populations of interest: (1) toddlers with ASD around age 2, since this represents an age at which young children can be reliably diagnosed (Corsello et al. 2013), (2) TD infants and young toddlers recruited as a proxy for demonstrating potential use with infants who may be deemed at-risk (i.e., young infants not routinely diagnosed with ASD). Given fundamental differences of these groups in inclusion, analyses did not directly compare groups. These groups were not intended as controls for one another. Nor was the study intended to contrast TD infants from toddlers with ASD.

System Design and Procedures

The system architecture was comprised of four distinct elements which interacted via a network in real-time (see Zheng et al. 2015). Elements included (1) a name prompting sub-system utilizing five mounted monitors (Samsung 32 inch flat TVs); (2) an attention tracking sub-system made up of four spatially distributed cameras (Logitech 930e web cameras; the resolution was 720p); (3) an animated attractor, utilized to help shift the child's gaze from current to target location; and (4) a feedback mechanism used for reinforcement. An autonomous, within-system supervisory controller (see Figure 1) was used to direct the interaction of these elements by reporting results and responding to incoming commands. Specifically, in the current system, it directed the global execution logic by prompting the sub-system and activating the name prompting or attractors when needed. The supervisory controller used the attention tracking sub-system to track the participant's gaze. The

attention tracking sub-system detected the children's respond using a camera array-based algorithm. This algorithm first detects the participant's head pose, and then approximate the gaze direction of the participant using the head pose. If the participant's gaze direction was pointing inside the region of a monitor, then we considered the participant responded to this monitor. The supervisory controller then simultaneously used this gaze information to make real-time decisions on how the prompting sub-system should respond. That is, the system's ability to track where the participant looked allowed it to independently move through different prompts.

In order to conduct an initial validation and understanding of error rate for the system, we asked 7 TD children and 3 adults to randomly look at a monitor and look away. In 100 trials, this system has an error rate of 5% in detecting whether the gaze of the participant was on one monitor or not. This error rate was tested based on large monitor regions, i.e., whether the gaze direction was inside or outside the region of a particular monitor, since where the participant was looking at exactly in one monitor was not important in this study. However, if the interactive objects are smaller objects rather than large monitors, then this system may have a higher error rate.

The experiment room layout is seen in Figure 2. The name prompting sub-system, made up of monitors 1–5, covered a large part of the room. It generated the name calling prompts and provided animated attention attractors (e.g., a bouncing ball) that could be activated where the child was calculated to be looking and subsequently move across the system to the target (i.e., children following the bouncing ball would be directed to the social target). Cameras 1–4 (i.e., attention tracking sub-system) tracked the participant's gaze. The child was seated in a small Rift on chair in the center of the circle. By arranging the cameras in front of the child and along a half circle concentric with the monitor half circle, children's frontal facial views could be captured by at least one of the cameras when he looked at any of the surrounding monitors. Seated children could then turn toward any of the monitors in front of or beside them while the system tracked their attention based on gaze.

During the protocol's response to name trials (RTN), when a child did not respond to his or her name being called on the target monitor (e.g., pre-recorded video of research assistant calling child's specific name), a hierarchical prompt model based upon System of Least Prompts (Doyle et al. 1988) was utilized. This model started with the least intrusive prompt at Level 1. If the child did not respond successfully, we added more intrusive information (e.g., different forms of attention attractors that moved or changed appearance) at each subsequent prompt level to heighten the attractor's effect. These escalated until the children looked at the target monitor.

This hierarchical sequence of enhanced prompts is listed in Table 2.

Level 1 prompts consisted simply of a prerecorded video of a trained female examiner calling the name of the child twice (each name calling last for about 2 seconds). If the participant did not look at the target monitor during Level 1, Level 2 was given. At Level 2, the attractor appeared, immediately bouncing from the monitor closest to the child's gaze to the target monitor (where the examiner continued to call child's name. If still no response, at

Level 3, the attractor bounced continuously for 2 seconds on the monitor closest to the child's current gaze in order to help capture the child's attention, then bounced in the direction of the target monitor. At Level 4, sound effects were added to Level 3 prompts. There was a 1 sec lag between different prompt levels. After the initial 2 name callings in Level 1, the name calling was repeated until the end of the trial (child looked at the target screen or all four level of prompts were finished). For example, if the trial lasted for 10 seconds, since each name calling last for about 2 seconds, there would be 5 times of name calls.

At Level 2 to Level 4, the attention attractor bounced at a speed of 38.4 cm per second in horizontal, and the distance between the centers of two adjacent monitors was 86.4cm. If the child looked at the target screen (target hit) in response to his name, whether immediately or at any point during the four levels of prompts, the attention tracking sub-system recognized the success, terminated the name calling immediately, and delivered praise (examiner's face responding "Good job! You found me!") with a firework animation. Otherwise, Level 1 would be finished after the 2 name callings (about 4 seconds), and Level 2 to 4 would be ended when the attention attractor bounced (from the child's gaze direction at the beginning of the Level) to the center of the target monitor. Therefore, the length of Level 2–4 depended on where the participant was looking at the beginning of the Level. The system was programmed to ultimately activate the praise module when a child did not correctly find target after a pre-specified time sequence.

Each participant completed a single experimental session consisting of the following steps. First, the child watched a short pre-recorded video of a confederate (e.g., same research assistant used in name trials) welcoming the child. They then watched a child-friendly video (e.g., a Sesame Street clip) that was displayed on one monitor at a time, until all 5 monitors displayed a section of the clip consecutively. This was intended to help participants understand that a video could appear on any of the 5 mounted screens. RTN trials were then presented, starting with the filmed face of the confederate appearing on screen only calling the child's name (e.g., "Sarah!") and progressing through the hierarchy in Table 2. After five RTN trials, a second fun video was played as a "break", followed by five final RTN trials and a "Good-bye video" of the confederate praising the child's work and saying good-bye.

The name calling happened on Monitor 5, 1, 4, 2, 3 in the first five trials and Monitor 1, 5, 2, 4, 3 in the last five trials. There was a 3 sec lag time between different trials. Monitor 3 was right in front of the participant, Monitor 1 and 5 were 90 degrees away on the left and right side of the participants, and Monitor 2 and 4 was 45 degrees away from the participants on the left and right sides, respectively. Therefore, a name call could occur in a location at, near, or far away from the child's current eye gaze, depending on where the participant was looking at dynamically.

Results

The primary objective of the current work was to assess the feasibility, tolerability, and potential application of a novel, autonomous social communication relevant technological system for (1) toddlers diagnosed with ASD and (2) much younger children without ASD.

For each group, we analyzed two aspects of participant responses: (1) the prompt level required for participants to respond to name, and (2) the length of time from first prompt to when a child looked at the target monitor. Children in the TD and ASD groups were analyzed separately given age, developmental, and clinical differences.

Regarding tolerability of the system, all participants tolerated the system well in spite of their young ages. No sessions were terminated due to participant distress or engagement challenges. On average, toddlers with ASD completed the sessions in 5 min, 25 sec (SD = 10 sec, Min = 5 min, 13 sec; Max = 5 min, 43 sec), and the TD infants completed the sessions in 5 min, 23 sec (SD = 7 sec; Min = 5 min 13 sec; Max = 5 min, 33 sec), respectively.

Regarding within system performance, all participants in the ASD group eventually hit the target (i.e., turned toward their name) across all trials. Most of the target hits ($n = 66$) were at prompt Level 1, name calling only on the target monitor. The rest were divided between prompt Level 2 ($n = 11$), name calling with a bouncing attractor, prompt Level 3 ($n=2$), and prompt Level 4 ($n=1$). On average, children in the ASD group required a prompt level of 1.23 to hit the target, taking an average of 3.07 seconds between initial prompt and success. TD infants, an average of 10 months younger than the toddlers with ASD, performed similarly to those in the ASD group by ultimately hitting the target across all trials ($n = 68$ on Level 1, $n = 11$ on Level 2, and one child missing the target for one of the five trials). The average time required for TD infants to hit the target was 1.19 seconds. The standard deviation listed in Table 3 show large variability in the response time in both groups. This could be mainly due to the heterogeneous development trajectory and behavioral pattern of children, different participant would have different capability in response to name. Based on our observation, for example, some participants were actively searching among the monitors, and thus could hit a target monitor quickly. However, there were other participants who tended to wait for a monitor to show the name calling for a while before turn to another monitor, and thus needed longer time to hit a target other than this particular monitor.

We also calculated the prompt level needed by the participants to hit a target on different directions, and did not find significant differences among them. From the observation, we found that these children in most cases were turning around to search for different monitors. We did not find they have particular preferences on any of the monitors.

Discussion

In the current pilot feasibility study, we studied the development and application of an innovative closed-loop adaptive technological system with potential relevance to a core area of vulnerability (i.e., social orienting) related to ASD in infancy and toddlerhood. The ultimate objective of this study was to empirically test the feasibility and usability of an adaptive technological learning environment, capable of intelligently administering early social orienting prompts and adaptively responding based on within system measurements of performance. We also conducted a preliminary examination of child performance, both in a toddler with ASD sample and TD infant sample, not to directly compare performance but rather to examine system functionality in these separate target populations.

The system was extremely well tolerated in this sample. No sessions were suspended or interrupted due to distress, non-compliance, disengagement, or parental request. This is in stark contrast to previous works and technological systems requiring wearable devices or significant constraints in motion (see Bekele, 2013 for example) and stresses the need to develop applied technologies capable of indexing child behaviors in more naturalistic settings without physical constraint.

Both toddlers with ASD as well as typically developing infants and younger toddlers were able to ultimately respond accurately to prompts delivered by the technological system within the standardized protocol. Further, the system was capable of attracting and pushing toward correct performance autonomously without user intervention. This finding is particularly important in the context of noting that all of the children within our clinical ASD sample were demonstrating potent challenges responding to their name when formally assessed for this skill on the ADOS-2, suggesting the potential benefits of utilizing technology to facilitate and push social orienting over time.

Collectively these preliminary findings are promising in both supporting system capabilities and potential relevance of future technological applications very early in development. Technological assessment and learning environments endowed with enhancements for successfully assessing and shaping more sophisticated early joint attention and social orienting skills, might be capable of rapidly indexing these skills from an evaluation and monitoring perspective and potentially taking advantage of baseline enhancements in non-social attention preference (Klin et al. 2009; Annaz et al. 2012) in order to meaningfully enhance skills related to coordinated attention over time (i.e., intervention-like technology). In this capacity, creating technological systems that can augment or supplement the work of a human therapist could increase children's access to intervention strategies while simultaneously detecting subtle social-communication differences that are difficult to detect without computer assistance.

While our data provides preliminary evidence that intelligent, adaptive technological environments and non-contact measurements of gaze may be useful tools for future assessment and potential intervention format, ultimately we present only feasibility and tolerability data. The question of the impact of such systems in realistically advancing detection and intervention paradigms remains open. Intelligent technologies will likely necessitate much more sophisticated paradigms and approaches that specifically target, enhance, and accelerate skills for meaningful impact on this population. Yet autonomous technologies (Feil-Seifer and Mataric 2011; Liu et al. 2008) that harness powerful differences in attention to technological stimuli, such as the current system may hold great promise in this regard in the future.

There are also several methodological limitations of the current study that are important to highlight. The small sample size examined and the limited time frame of interaction are the most powerful limits of the current study. As such, while we are left with data suggesting the potential of the technological learning environment, the utilized methodology significantly restricts our ability to realistically comment on the value and ultimate clinical utility of this system as applied to very young children with and without ASD, or in comparison to one

another. Eventual success and clinical utility of intelligent technologically-mediated systems hinges upon their ability to detect and potentially promote meaningful change in core skills that are tied to dynamic neurodevelopmentally appropriate learning across environments. We did not systematically intend to assess learning within this system; rather we indexed simple initial behavioral responses within system application. As such, questions regarding whether such a system could constitute a true assessment or skill intervention paradigm remain open. Finally, although we made attempts to ensure that toddlers with ASD had received evaluations with gold-standard assessment tools (e.g., ADOS, clinician diagnosis), we did not have rigorous assessment data on the comparison sample on these same instruments. As such, our ability to comment on the specific clinical characteristics matched with performance differences regarding this technology is limited.

Despite limitations, this work is the first to our knowledge to design and empirically evaluate the usability, feasibility, and preliminary efficacy of an autonomous, three-dimensional, technological learning environment capable of modifying response based on within system measurements of gaze and social orienting via non-contact gaze detection. Few other existing intelligent technological and robotic systems (Feil-Seifer and Mataric 2011; Liu et al. 2008) have specifically addressed how to detect and flexibly respond to individually derived, socially and disorder relevant behavioral cues within an intelligent paradigms of relevance for very young children with ASD. Movement in this direction introduces the possibility of realized technological assessment and potentially intervention tools that are not simple response systems, but systems that are capable of necessary and more sophisticated adaptations. Systems capable of such adaptation may ultimately be utilized to understand and promote meaningful change related to the complex and important social communication impairments of the disorder itself.

Acknowledgments

This study was supported by in part by the National Institute of Health under Grants 1R01MH091102-01A1 and R21 MH103518. Work also includes core support from EKS NICHD of the NIH under Award U54HD083211 and by CTSA award UL1TR000445.

References

- Annaz D, Campbell R, Coleman M, Milne E, Sweetenham J. Young children with autism spectrum disorder do not preferentially attend to biological motion. *Journal of Autism and Developmental Disorders*. 2012; 42:401–408. [PubMed: 21547412]
- Bekele E, Crittendon J, Zheng Z, Swanson A, Warren ZE, Sarkar N. Assessing the utility of a virtual environment for enhancing facial affect recognition in adolescents with autism. *Journal of Autism and Developmental Disorders*. 2014; 44(7):1641–50. [PubMed: 24419871]
- Bedford R, Elsabbagh M, Gliga T, Pickles A, Senju A, Charman T, et al. Precursors to social and communication difficulties in infants at-risk for autism: gaze following and attentional engagement. *Journal of Autism and Developmental Disorders*. 2012; 42(10):2208–18. [PubMed: 22278030]
- Chevallier C, Parish-Morris J, McVey A, Rump K, Sasson N, Herrington J, et al. Measuring social attention and motivation in autism spectrum disorder using eye-tracking: stimulus type matters. *Autism Research*. 2015; doi: 10.1002/aur.1479
- Corsello CM, Akshoomoff N, Stahmer AC. Diagnosis of autism spectrum disorders in 2 year-olds: a study of community practice. *Journal of Child Psychology and Psychiatry*. 2013; 54(2):178–185. [PubMed: 22905987]

- Doyle PM, Wolery M, Ault MJ, Gast DL. System of least prompts: A literature review of procedural parameters. *Research and Practice for Persons with Severe Disabilities*. 1988; 13(1):28–40.
- Falck-Ytter T, Fernell E, Hedvall A, von Hofsten C, Gillberg C. Gaze performance in children with autism spectrum disorder when observing communicative actions. *Journal of Autism and Developmental Disorders*. 2012; 42(10):2236–45. [PubMed: 22354708]
- Feil-Seifer, D.; Mataric, C. Proceedings of the 6th International Conference on Human-Robot Interaction. New York, NY: ACM Press; 2011. Automated detection and classification of positive vs. negative robot interactions with children with autism using distance-based features; p. 323-330.
- Gammer I, Bedford R, Elsabbagh M, Garwood H, Pasco G, Tucker L, et al. Behavioral markers for autism in infancy: scores on the autism observation scale for infants in a prospective study of at-risk siblings. *Infant Behavioral Development*. 2015; 38:107–115.
- Guillon Q, Hadjikhani N, Baduel S, Rogé B. Visual social attention in autism spectrum disorders: insights from eye tracking studies. *Neuroscience & Behavioral Reviews*. 2014; doi: 10.1016/j.neubiorev.2014.03.013
- Hertz-Picciotto, I. Environmental Risk Factors in Autism: Results from Large-Scale Epidemiologic Studies. In: Amaral, D.; Geschwind, D.; Dawson, G., editors. *Autism Spectrum Disorders*. New York, NY: Oxford University Press; 2011. p. 827-862.
- Jones W, Klin A. Attention to eyes is present but in decline in 2–6 month-olds later diagnosed with autism. *Nature*. 2013; 504(7480):427–431. [PubMed: 24196715]
- Klin A, Lin DJ, Gorrindo P, Ramsay G, Jones W. Two-year-olds with autism orient to nonsocial contingencies rather than biological motion. *Nature*. 2009; 459(7244):257–261. [PubMed: 19329996]
- Krakowiak P, Walker CK, Bremer AA, Baker A, Ozonoff S, Hansen RL, Hertz Picciotto I. Maternal metabolic conditions and risk for autism and other neurodevelopmental disorders. *Pediatrics*. 2012; 129(5)doi: 10.1542/peds.2011-2583
- Lahiri U, Bekele E, Dohrmann E, Warren ZE, Sarkar N. A physiologically informed virtual reality based social communication system for individuals with autism. *Journal of Autism and Developmental Disorders*. 2015; 45(4):919–931. [PubMed: 25261247]
- Liu C, Conn K, Sarkar N, Stone W. Online affect detection and robot behavioral adaptation for intervention of children with autism. *IEEE Transactions on Robotics*. 2008; 24(4):883–896.
- Lord C, Risi S, Lambrecht L, Cook E, Leventhal B, DiLavore P, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with spectrum of autism. *Journal of Autism and Developmental Disorders*. 2000; 30(3):205–223. [PubMed: 11055457]
- Lord, C.; Rutter, M.; DiLavore, PC.; Risi, S.; Gotham, K.; Bishop, S. *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2)*. Western Psychological Services; Torrance, CA: 2012.
- Magrelli S, Jermann P, Noris B, Ansermet F, Hentsch F, Nadel J, et al. Social orienting of children with autism to facial expressions and speech: a study with a wearable eye-tracker in naturalistic settings. *Frontiers in Psychology*. 2013; doi: 10.3389/fpsyg.2013.00840
- Mullen, EM. *Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance Services Inc.; 1995.
- Nadig A, Ozonoff S, Young G, Rozga A, Sigman M, Rogers SJ. A prospective study of response to name in infants at risk for autism. *Archives of Pediatric and Adolescent Medicine*. 2007; 161(4): 378–83.
- Noris B, Nadal J, Barker M, Hadjikhani N, Billard A. Investigating gaze of children with ASD in naturalistic settings. *PLOS One*. 2012; doi: 10.1371/journal.pone.0044144
- Rehg J, Abowd G, Rozga A, Romero M, Clements M, Sclaroff S, Rao H. Decoding children's social behavior. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2013:3414–3421.
- Shic F, Macari S, Chawarska K. Speech disturbs face scanning in 6-month olds who develop autism spectrum disorder. *Biological Psychiatry*. 2013; doi: 10.1016/j.biopsych.2013.07.009
- State MW, Levitt P. The conundrums of understanding genetic risks for autism spectrum disorders. *Nature Neuroscience*. 2011; 14doi: 10.1038/nn.2924

- Stenberg N, Bresnahan M, Gunnes N, Hirtz D, Hornig M, Lie K, et al. Identifying children with autism spectrum disorder at 18 months in a general population sample. *Paediatric and Perinatal Epidemiology*. 2014; 28(3):255–262. [PubMed: 24547686]
- Veenstra-VanderWeele J, Warren ZE. Intervention in the context of development: Pathways toward new treatments. *Neuropsychopharmacology*. 2015; 40:225–237. [PubMed: 25182180]
- Warren ZE, McPheeters ML, Sathe N, Foss-Feig JH, Glasser A, Veenstra-VanderWeele J. A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics*. 2011; 127(5):e1303–1311. [PubMed: 21464190]
- Warren ZE, Zheng Z, Shuvajit D, Young E, Swanson A, Weitlauf A, Sarkar N. Development of a robotic intervention platform for young children with ASD. *Journal of Autism and Developmental Disorders: Special Issue on Technology*. 2014; doi: 10.1007/s10803-014-2334-0
- Zheng, Z.; Zhang, L.; Bekele, E.; Swanson, A.; Crittendon, J.; Warren, ZE.; Sarkar, N. Impact of robot-mediated interaction system on joint attention skills for children with autism. 13th International Conference on Rehabilitation Robotics (ICORR) 2013; Seattle Washington. June 24–26, 2013; 2013.
- Zheng, Z.; Fu, Q.; Zhao, H.; Swanson, A.; Weitlauf, A.; Warren, ZE.; Sarkar, N. Universal Access in Human-Computer Interaction Access to Learning, Health and Well-Being. Switzerland: Springer International Publishing; 2015. Design of a computer-assisted system for teaching attentional skills to toddlers with ASD; p. 721-730.

Figure 1.
Global system architecture.

Figure 2.
Experiment room layout.

Table 1

Participant characteristics

Group	Age in Years (SD)	ADOS Raw Score	IQ
ASD	2.19 (0.16)	23.75 (2.31)	55.00 (8.49)
TD	1.33 (0.40)	N/A	N/A

Table 2

Name Prompting Levels

<i>Prompt level</i>	<i>Content</i>
Level 1	Name calling on the target monitor
Level 2	Level 1 + Attractor bouncing from the gaze direction to the target monitor.
Level 3	Level 1 + Attractor <i>first bouncing at the gaze direction for 2 seconds</i> , and then bouncing from the gaze direction to the target monitor.
Level 4	Level 1 + Attractor first bouncing at the gaze direction, and then bouncing from the gaze direction to the target monitor <i>with special sound effect</i> .

Table 3

Average prompt levels and time required for ASD and TD groups to respond to name

	Prompt level (1–4)	Time to hit target (in seconds)
ASD mean (SD)	1.23 (0.55)	3.07 (3.32)
TD mean (SD)	1.19 (0.55)	2.66 (2.25)