Eye Tracking as a Measure of Receptive Vocabulary in Children with Autism Spectrum Disorders

Nancy C. Brady, **Christa J. Anderson**, **Laura J. Hahn**, **Sara M. Obermeier**, and **Leah L. Kapa**
University of Kansas

Abstract

This study examined the utility of eye tracking research technology to measure speech comprehension in 14 young boys with autism spectrum disorders (ASD) and 15 developmentally matched boys with typical development. Using eye tracking research technology, children were tested on individualized sets of known and unknown words, identified based on their performance on the Peabody Picture Vocabulary Test. Children in both groups spent a significantly longer amount of time looking at the target picture when previous testing indicated the word was known (known condition). Children with ASD spent similar amounts of time looking at the target and non-target pictures when previous testing indicated the word was unknown (unknown condition). However, children with typical development looked longer at the target pictures in the unknown condition as well, potentially suggesting emergent vocabulary knowledge.

Keywords

Speech comprehension; Receptive language; Autism spectrum disorders

Speech comprehension refers to the ability to understand the meaning of spoken words. The ability to comprehend a variety of words develops over time, and many studies have documented how spoken word comprehension progresses throughout development (Facon, Facon-Bollengier, & Grubar, 2002; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Zampini & D'Odorico, 2009). The comprehension of speech is an important component of overall language comprehension, because individuals who are able to understand more spoken words will be more able to understand the information conveyed to them with speech. In addition, scores on speech comprehension measures tend to correlate with other aspects of language development (e.g., commenting, requesting, joint attention, inventory of words, etc.; McDuffie, Yoder, & Stone, 2005; Watt, Wetherby, & Shumway, 2006); nonverbal cognitive development (Brady, Thiemann-Bourque, Fleming, & Matthews, in press; Cunningham, Glenn, Wilkinson, & Sloper, 1985); and other variables, including maladaptive behaviors (i.e., negative correlations with aberrant behavior, stereotypies, and

Correspondence concerning this manuscript should be directed to Nancy C. Brady, University of Kansas, Speech-Language-Hearing: Sciences & Disorders, 1000 Sunnyside Ave., Room 3008, Lawrence KS 66045; nbrady@ku.edu.
Nancy C. Brady, Speech-Language-Hearing: Sciences & Disorders; Christa J. Anderson, Life Span Institute; Laura J. Hahn, Life Span Institute; Sara M. Obermeier, Schiefelbusch Institute for Life Span Studies; Leah L. Kapa; Schiefelbusch Institute for Life Span Studies

restricted and repetitive behaviors; McClintock, Hall, & Oliver, 2003; Ray-Subramanian & Ellis Weismer, 2012).

Profiles comparing speech comprehension scores to other aspects of language and cognitive development have been used to describe the behavioral phenotypes associated with specific disorders such as Williams syndrome, Down syndrome (Mervis, 2004), and autism spectrum disorders (ASD; Ellis Weismer, Lord, & Esler, 2010). Specifically, in young children with ASD, speech comprehension may be one of the best predictors of long-term language outcomes (Bopp & Mirenda, 2011; McDuffie et al., 2005; Smith, Mirenda, & Zaidman-Zait, 2007). In addition, children with ASD have impaired comprehension skills compared to children with typical development (Charman, Drew, Baird, & Baird, 2003). Luyster, Qiu, Lopez, and Lord (2007) found that children with ASD have similarly delayed receptive language abilities as children with other developmental disabilities when nonverbal mental age is considered. However, the amount of delay in receptive language, relative to productive language, appears to differ for children with ASD. Whereas the profile for language of children with typical development shows relatively advanced comprehension compared to production (Bornstein & Hendricks, 2012; Fenson et al., 1994), this pattern is reversed in children with ASD. For example, Ellis-Weismer et al. (2010) evaluated the profiles of receptive and expressive language abilities in 257 toddlers with ASD between 24 and 36 months of age across multiple language measures, and found that they had more severe receptive language delays than expressive language delays. Similar findings were reported in two other studies of children with ASD (Hudry et al., 2010; Maljaars, Noens, Scholte, & Berckelaer-Onnes, 2012).

Receptive language skills are particularly important for children with ASD who communicate with augmentative and alternative communication (AAC). Children with more advanced receptive language skills learn more words across multiple modes over time (Brady et al, in press; Drager et al., 2006; Romski & Sevcik, 1996). Individuals learning to use graphic symbol representations appear to use their existing speech-referent knowledge to map onto new symbol-referent relationships (Sevick & Romski, 2002). Additionally, it has been proposed that greater comprehension of spoken words can be used to access working memory processes while learning AAC (Thistle & Wilkinson, 2013). Thus, receptive language appears to facilitate AAC learning, but additional research indicates that learning AAC also strengthens receptive language skills. For example, Brady (2000) demonstrated that comprehension for object names improved after learning to use speech-generating devices for two children, one of whom had ASD. Similarly, Preis (2006) found that the use of AAC symbols improved the comprehension of verbal commands for 4 out of 5 participants with ASD.

Measuring receptive language in children with ASD who are learning AAC poses a challenge. Many standardized measures of early language development, including speech comprehension, have been described as unsuitable for young children with ASD (Volden et al., 2011). For example, when Volden and colleagues examined scores on the Preschool Language Scales 4th edition (PLS-4; Zimmerman, Steiner, & Pond, 2002) from 294 children with ASD, they found that almost 30% of the participants performed at the floor level (i.e., less than a standard score of 50). It may be that difficult response demands, required by

traditional vocabulary assessments, may lead to underestimates of vocabulary knowledge in children with ASD. For example, assessments such as the Peabody Picture Vocabulary Test 4 (PPVT-4; Dunn & Dunn, 2007) require that the child point to a picture that corresponds with a referent after it is named by the test administrator (e.g., *Point to ball*). This task may be difficult for children with limited behavioral repertoires and motor delays. Similarly, Brady and colleagues (in press) recently found that nearly half of their participants, who were between 4 and 7 years of age, with minimal verbal skills, were unable to pass training items on the PPVT-4.

Problems with assessing vocabulary in children with ASD may be somewhat alleviated by altering standardized tasks; however, this would limit the interpretability of results and invalidate comparisons to norm-referenced data. Beyond problems with comparing standardized test performance of children with ASD to peers with typical development, children's limited ability to respond appropriately during standardized testing leads to ambiguity in interpreting their performance. Thus, poor performance on standardized language assessments among children with ASD may stem either from true underlying deficits in language knowledge or inability to employ task appropriate responding, independent of their language knowledge. A significant problem in the area of language development in children with ASD is accurately assessing language knowledge independent of the child's potentially limited ability to provide appropriate task responses. Therefore, it is important to have accurate measures of speech comprehension that allow researchers and clinicians to correctly assess the language knowledge of children with ASD.

To address this problem, parent/caregiver reports of speech comprehension can be used. Instruments such as the MacArthur-Bates Communicative Development Inventory (CDI; Fenson, Marchman, Thal, Reznick, & Bates, 2006) ask parents to report on their child's vocabulary comprehension. However, the validity of parent reports of speech comprehension depends on the parent's ability to infer lexical comprehension on the basis of their child's behavior. Parents are typically not provided with any specific instructions about how to determine if a child understands a word and different interpretations of these instructions have led to variations in outcomes (Tomasello & Mervis, 1994). Thus, while parent report measures are found to be stable over time (Yoder, Warren, & Biggar, 1997), the accuracy of parents inferring comprehension for specific vocabulary items has been called into question (Styles & Plunkett, 2009).

Implicit Learning Paradigms

Implicit learning assessment paradigms, previously used in infant studies, are one method that has proved useful for examining underlying vocabulary knowledge. Two of the most common implicit learning methods are (a) habituation and (b) intermodal preferential looking paradigm (IPLP; Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987), neither of which require speech production or advanced motor skills. In the habituation paradigm, the progression of looking time is assessed during the presentation of a stimulus during repeated exposure, then a novel stimulus is introduced alongside the repeated stimulus and looking time is assessed. Increased looking to the novel stimulus is indicative of habituation to the

repeated stimulus and suggests the novel stimulus is processed as different from previously presented stimuli (e.g., Colombo, 2001).

In contrast, IPLP relies on looking or gaze patterns to assess comprehension, typically by including an auditory stimulus. Specifically, IPLP involves simultaneously presenting two visual stimuli, in the left and right field of vision (spaced far enough apart so that the location of the look can be distinguished), along with an auditory stimulus that corresponds with one visual stimulus (Golinkoff et al., 1987). The infants' responses are videotaped and the proportion of time spent looking at each stimulus is determined using a frame-by-frame analysis procedure. Comprehension is defined by a larger proportion of looking time to the matching visual stimulus, confirming implicit understanding of the word. Using IPLP, receptive language abilities of pre-verbal infants have been measured as early as 6 months of age (Tincoff & Jusczyk, 1999). In addition, the IPLP has been validated as a method for word comprehension by showing that infants look to words their parents have previously indicated their child knew (Behrend, 1990; Houston-Price, Mather, & Sakkalou, 2007; Robinson, Shore, Hull Smith, & Martinelli, 2000). Furthermore, these studies suggest that IPLP may be a sensitive measure of emergent word comprehension, based on the fact that during IPLP tasks infants looked to target words that had been reported by parents to be "heard, but not understood" (Houston-Price et al., 2007; Robinson et al., 2000).

IPLP has also been used to assess language among children and adults with severe cognitive, language, and motor deficits, such as cerebral palsy (Cauley, Golinkoff, Hirsh-Pasek, & Gordon, 1989; Geytenbeek, Heim, Vermeulen, & Oostrom, 2010), and in children with ASD (Bebko, Weiss, Demark, & Gomez, 2006; Goodwin, Fein, & Naigles, 2012; Swensen, Kelley, Fein, & Naigles, 2007). Specifically, using IPLP allowed Swensen and colleagues (2007) to explore key aspects of language comprehension at an earlier age in children with ASD than previous studies – all children with ASD in this study had mental ages of less than 2.5 years. The results of this study indicated that, among other findings, children with ASD displayed comprehension of subject-verb-object word order before they produced this sentence structure. This pattern is similar to what is observed in peers with typical development, but unlike the patterns that have been reported in individuals with ASD using behavioral measures (e.g., Hudry et al., 2010; Ray-Subramanian & Ellis Weismer, 2012). Thus, IPLP demonstrated emergent understanding that had not yet been demonstrated productively.

Although the results of the study by Swensen and colleagues (2007) suggested similar linguistic processing in children with and without ASD, other studies have used IPLP to identify differences in language processing by children with ASD. Specifically, Bebko and colleagues (2006) used IPLP to assess looking time differences for synchronous and asynchronous cues in 4-to 6-year-old children with ASD. Their hypothesis was that children without ASD would look longer at linguistic stimuli that were presented with synchronous audio and visual components than to stimuli that were temporally asynchronous; whereas children with ASD would not show this pattern. Results confirmed their hypothesis, suggesting that children with ASD are less sensitive to linguistic stimuli. Finally, using IPLP, Goodwin and colleagues (2012) found that there is delayed comprehension of object-wh and subject-wh questions in children with ASD, compared to chronological age-matched

children with typical development. However, this delay was no longer observed when children with ASD were matched to children with typical development on overall language ability (Goodwin et al., 2012).

In all of these studies, the use of IPLP allowed for the examination of language comprehension, but with fewer demands placed on coordinating motor movements (i.e., pointing to a picture corresponding with the auditory word) and working memory than standardized measures of language comprehension (Goodwin et al., 2012). While IPLP has its advantages, it is a laborious assessment method due to the frame-by-frame video coding of the participant's looking behavior. Nonetheless, these studies demonstrate the benefit of using looking time to assess comprehension in children with ASD.

Eye-tracking as an Objective Measure of Implicit Understanding

Recent advances in technology indicate that the use of eye tracking research technology could greatly facilitate the measurement of preferential looking. Eye tracking research technology systems and related software, like GazeTracker[TM] (GT; Eye-Gaze Response Interface Computer Aid [ERICA], 2001), provide data about overt looking behaviors such as the total and average amount of time spent looking at a specified target stimulus, along with the lag-time until fixation on the target stimulus. All of this information can be obtained and analyzed much more efficiently and with greater precision than the information gained from the traditional frame-by-frame video coding employed in IPLP. Specifically, infrared technology is used to measure pupil size and corneal reflection to compute line of gaze, allowing the exact location of the look to be measured, providing a more precise and objective measure of looking than IPLP. Therefore, just as Golinkoff and colleagues (1987) suggested that the IPLP took Spelke's intermodal perception paradigm (1976, 1979) to a new level, the use of eye tracking research technology systems may be the next logical advancement of IPLP.

Eye-tracking Research with Children with ASD

Eye tracking research technology is emerging as a valuable technique to examine looking behaviors under a variety of task conditions in children with ASD. For example, researchers have examined preferential looking to social versus non-social images in young children with ASD (Anderson, Colombo, & Shaddy, 2006; Pierce, Conant, Hazin, Stoner, & Desmond, 2011), face and emotion recognition (Johnson, Gillis, & Romanczyk, 2012), and the ability of children and adolescents with ASD to predict and correctly interpret the intentional actions of others (Vivanti et al., 2011; Vivanti, Nadig, Ozonoff, & Rogers, 2008). Differences in patterns of visual attention and visual scanning have been identified in young children with ASD (Anderson et al., 2006; Mercadante, Macedo, Baptista, Paula, & Schwartzman, 2006) and it has been suggested that these differences may have diagnostic value.

Eye tracking research technology has been used to expand our knowledge of social deficits that are the primary characteristic of individuals with ASD. In comparison to age-matched peers, young children with ASD (2 to 5 years of age) have similar scanning times to socially relevant stimuli (e.g., Anderson et al., 2006; Speer, Cook, McMahon, & Clark, 2007), while

older children and adults with ASD tend to have decreased scanning to socially relevant stimuli (e.g., Riby & Hancock, 2009a, 2009b; Sasson, Turner-Brown, Holtzclaw, Lam, & Bodfish, 2008), indicating a possible developmental progression of social avoidance. In sum, eye tracking research technology has been demonstrated as useful for examining processing in children with autism, but has not yet been applied to receptive language.

Purpose of the Present Study

The present study is a preliminary investigation of the feasibility of eye tracking research technology as a possible method for assessing speech comprehension in children with ASD. Our primary purpose was to document that the behaviors recorded with eye tracking research technology map onto overt conventional behaviors, such as pointing to pictures, in a sample of young boys with ASD and a developmentally matched sample of young boys with typical development. Thus, the present study was designed to lay the groundwork for future studies to advance the measurement of speech comprehension in children with ASD who may be learning AAC. All children were first tested with the PPVT-4 in order to establish a corpus of words that were known and unknown. Based on these results from the PPVT-4, we hypothesized that children in both groups would spend a significantly longer amount of time looking at target pictures than non-target pictures, when they had previously demonstrated comprehension for the target words from the PPVT-4 (i.e., known condition). We also hypothesized that there would not be any significant differences in looking time to target and non-target pictures for words that they had not indicated comprehension from the PPVT-4 (i.e., unknown condition).

Method

Participants

A total of 29 boys between the ages of 42 and 82 months were recruited for this study, 14 children with ASD and 15 children with typical development. All of the children with ASD had a formal diagnosis of either Autistic Disorder (AD; $n = 10$) or Pervasive-Developmental Disorder-Not Otherwise Specified (PDD-NOS; $n = 4$) by established diagnostic clinics in the Kansas City area, using DSM-IV criteria. These clinics include teams of licensed professionals who have been formally trained and certified to provide an official diagnosis along the ASD spectrum. Average IQ was calculated using the Early Learning Composite of the Mullen Scales of Early Learning (MSEL; Mullen, 1995). The mean IQ for children with ASD was 88.1 ($SD = 18.86$, range 51–112) and 110.60 ($SD = 15.73$, range 81–135) for children with typical development. Group differences in IQ were significant, $F(1,27) = 12.265$, $p = .002$, $^2 = .31$. Although children differed in IQ, they were similar in terms of developmental levels, as can be seen in Table 1. As would be expected, however, there was greater variability within the group of children with ASD. The average mental age-equivalent (MA) scores from the four subscales of MSEL --Verbal Reception, Fine Motor, Receptive Language, and Expressive Language -- as well as the scores of each subscale, were similar for children with and without ASD. None of the children with typical development had a biological parent or sibling with a diagnosis of ASD or other mental

health issues such as schizophrenia and/or cognitive/language delays. See Table 1 for age, demographic information, MSEL scores, and ADOS-G scores.

Exclusion/attrition criteria—Children were excluded from participation if their parents reported impairments in vision and/or motor functioning that could significantly impede their ability to participate in the testing session (i.e., severe vision loss or inability to sit upright without assistance). With the exception of the ASD diagnosis, none of the children had a history of chronic illness. In addition, none of the children who participated in the current study had symptoms of acute illness (e.g., allergies, cold/flu, ear infections, etc.) at the time of the testing sessions. All exclusion criteria were based on parent report.

Autism diagnosis—For children with ASD, autism diagnosis was verified by a trained member of our laboratory team through administration of the Autism Diagnostic Observation Schedule-Generic (ADOS-G; Lord, Rutter, & DiLavore, 1997). There are four modules of the ADOS-G, and the administrator chooses which module to administer based on the CA of the subject and level of language. For the current study, only Modules 1 ($n = 3$) and 2 ($n = 11$) were used due to the age-range of the participants. Module 1 is for children who have minimal language such as single words; Module 2 is for children who have phrase speech. Modules 3 and 4 are for older children and adults and thus are not applicable to the current study. The administration of this test is concise and it yields scores on three difference components: Communication, Qualitative Impairments in Reciprocal Social Interaction, and Stereotyped Behaviors and Restricted Interests. An overall score can also be calculated, which is comprised of the Communication and Qualitative Impairments in Reciprocal Social Interaction scores. For Module 1 and 2, scores of 12 and above are in the autism range. For Module 1, scores between 7 and 11 are in the autism spectrum range; while scores between 8 and 11 are in the autism spectrum range for Module 2. Scores on the ADOS-G reflect severity of impairment, with higher scores indicating more severe impairment. The means and ranges of overall scores are provided in Table 1.

Procedure

Children were recruited from a variety of developmental disability organizations in metropolitan and suburban areas of Kansas City, Kansas and Missouri, and through a pre-established commercial list of families interested in participating in research activities. Each child was scheduled and seen for two testing sessions, approximately one week apart, at our Lawrence laboratory. Participants were given $40 for attending each laboratory session ($80 total) to compensate for time and travel.

Session 1: Standardized testing—During Session 1, all of the children were administered the Peabody Picture Vocabulary Test 4, Form A (PPVT-4; Dunn & Dunn, 2007; Smith, 1997), along with the MSEL (Mullen, 1995). Parents were asked to either not be in the room with the child during PPVT-4 administration or to turn their back to the child to ensure that they were unaware of which words their child could receptively identify. This allowed us to ensure that the parents did not teach their child these words during the week between Sessions 1 and 2; parents were also verbally informed that any teaching of incorrectly identified words on the PPVT-4 between sessions would invalidate the study.

Session 2: Eye tracking—Results from the PPVT-4, obtained during Session 1, were used to create individualized stimulus sets for each child that began with the four practice words from PPVT-4 Form A (*dog*, *ball*, *bike* and *chair*), followed by a randomized presentation of the first 12 PPVT-4 words that the child correctly identified (known) and the last 12 PPVT-4 words that the child incorrectly identified (unknown). The individualized stimulus sets were presented to each child using an eye tracking research technology system and interface program (see section on eye tracking research technology system). During the eye tracking protocol, if parents remained in the room with the child, they were asked to remain out of the child's line of sight and to not interact with their child during the session, in order to minimize the potential for parental influence on the child's looking behaviors. Following the administration of the eye tracking protocol, the children were administered any unfinished subscale of the MSEL from the first testing session (e.g., child became too fatigued during session one). In addition, children with ASD were also administered the appropriate Module (1 or 2) of the ADOS-G.

Stimuli—To represent administration of the PPVT-4 using the eye tracking research technology, static color pictures that corresponded with vocabulary items in the PPVT-4 were selected from available clip art programs. Actual PPVT-4 pictures were not used in order to protect copyright agreements; however, the images were chosen to look as similar as possible to the PPVT-4 target and non-target pictures. As in the PPVT-4, each stimulus consisted of four pictures in a grid format with one target and three non-target pictures. These pictures were presented on a white background, which was split into four equal quadrants. Each quadrant measured 8 in × 6 in (20.32 cm × 15.24 cm), and was presented full-screen on a 16 in × 12 in (40.64 cm × 30.48 cm) monitor. The four images within each stimulus were matched for size; each picture was confined to a space of 4 in × 4 in (10.16 cm × 10.16 cm). Each picture was centered within its quadrant, leaving a 4-inch (10.16 cm) gap between each picture horizontally, and a 2-inch (5.08 cm) gap between each picture vertically.

The eye tracking research technology interface program was used to present each individualized stimulus set on a centrally located computer monitor. Each child was first presented with the four practice stimuli successively. Each practice stimulus was presented for 3s and was followed by an auditory directive, *Look at* (target word), and a 2s pause, allowing the child to make a looking response, then the directive was repeated and followed by another 2s pause. This was followed by the presentation of the child's individualized stimulus set. Twelve known and 12 unknown stimulus sets were randomly presented. For each stimulus set, 3s of visual presentation was followed by a directive that consisted of the oral presentation of the target word alone (i.e., *belt*), and a 2s pause that enabled the child to make a looking response; then the oral directive and 2s pause were repeated. Before and after each stimulus, a 2s interstimulus slide was presented, which consisted of a low-level central fixation cross hair presented at the child's midline. In addition, to keep the child's attention on the computer screen, three attention-grabber videos (presented for 30s each) were randomly interspersed throughout the presentation set. These were followed by the presentation of a calibration array to ensure accuracy of gaze measurement.

Eye-tracking research technology system—The pictures were presented on a 40.6 cm computer monitor, which subtended a 21.6° visual angle at the viewing distance. Looking responses were recorded using an Applied Science Laboratory (ASL) E6 eye-tracking system, Model 504 (ASL, 2008) with the GazeTracker interface program (ERICA, 2001), setup in a darkened interior room. The pan/tilt module, a component of the ASL system, uses near infrared technology to illuminate the eye and telephoto an image of the eye onto an eye camera. The E6 control unit then extracts the pupil and reflection of the light source on the cornea to compute gaze location at a 60 Hz sampling rate.

Each child was secured in a child-sized booster seat with over the shoulder straps and a seatbelt to ensure the child's safety and to minimize movement. We told children that they were going to sit in a seat with straps *like a go-cart*. The booster seat was secured onto a hydraulic chair, which enabled adjustment of the child's eye height to be approximately centered with the mid-point of the stimulus monitor (124.5 cm). The booster seat was slightly reclined to minimize head movement.

Prior to the start of the experimental session, a 9-point standard calibration was presented, which consisted of dynamic and multimodal cartoons presented individually at each of the 9 calibration target points; this was designed to naturally draw the child's attention to each target point without verbal instruction. Once an accurate calibration was achieved, the experimental stimulus set was presented. Calibration was monitored throughout the session, using three calibration checks placed randomly throughout the stimulus presentation. These checks were similar to the initial calibration and consisted of a 9-point calibration array of multimodal cartoons. As with the initial calibration, the experimenter visually monitored the calibration to ensure accuracy. Calibration was deemed accurate if pupil and corneal reflection cross-hairs still lined up through the center of the calibration point when it was presented. If calibration was inaccurate, the session was paused, calibration was re-established, and the session was resumed. A final check of calibration accuracy was included following the final slide to provide another check of gaze accuracy throughout the session. Thus, calibration was monitored during the session by the experimenter and then verified during data coding, by ensuring that the gaze location was on the actual calibration target. Only data that occurred while calibration was deemed accurate were used in the final analysis; however, because calibration was monitored and adjusted throughout the session, none were eliminated because of calibration.

Data Extraction and Reduction

The GazeTracker interface program was used to extract line of gaze from the ASL eye-tracking system. This program allows for analysis of looking time based on specific regions of interest, which were created for each stimulus slide. Thus, both practice, known and unknown stimuli were divided into two regions of interest: (a) the target pictures (pictures that corresponded with the verbal directive), and (b) non-target pictures (the three pictures that did not correspond with the verbal directive). Using these defined regions of interest we examined the looking in each of the four quadrants during the following periods (a) Pre-command, 3s prior to the first directive, (b) Post command 1, second 1, during the first second after the first directive; (c) Post command 1, second 2, during second 2 after the first

directive; (d) Post-command 2, second 1, during the first second after the second directive; and (e) Post command 2, second 2, second 2 after the second directive.

Dependent Measures

The dependent measures were the amount of looking time in seconds spent fixated on the target area of interest versus the time spent on the non-target area of interest. Looking to the target was determined by the number of milliseconds spent on the target area of interest. For looking time to the non-target area of interest, however, we accounted for the relative difference in space on the screen devoted to target (1 of 4 pictures) vs. non-target (3 of 4 pictures area), by performing an additional adjustment to the non-target looking time. Specifically, we divided the looking time to non-target by three to get an average looking time, across the three non-target pictures. Thus, if a participant fixated on any of the three pictures in the non-target area of interest for .45s, this looking time was adjusted for size, that is, .45/3 = .15s. Time spent looking off screen, blinking, or other causes of momentary calibration loss were not included; thus, the amount of looking time spent on target and non-target areas of interest do not typically sum to 1 second.

Results

Preliminary Analyses

Prior to the examination of proportion of looking to known and unknown pictures, we conducted preliminary analyses to examine preferential looking and the time course of looking-based responses.

Preferential looking—To ensure that looking was due to the verbal directive and not a preference for a particular picture, we first examined proportion of looking during the precommand period for all stimuli and groups. Trials were not used in final analysis if the participant spent > 25% of the total looking time (exclusive of off-looking time) during the precommand period in any one quadrant, indicating that the child may have a preference for that particular image, which could interfere with subsequent responses. That is, looking more than 25% time to one quadrant could indicate a bias that would interfere with determining looking to known vs. unknown pictures. This criterion was established based on the presumption that looking times should be comparable across the four quadrants prior to the presentation of the directive. Using this criterion, an average of 1.6 trials were eliminated per participant (range = 0 – 4) for the known/unknown conditions, and an average of 0.5 trials were eliminated per participant (range = 0 – 2) for the practice condition. We also calculated the time spent looking away from the screen or blinking (off-looking time) for each participant. During the precommand, off-looking time for children with ASD averaged .11s (*SD* = .24). Off-looking time for children with typical development averaged .08s (*SD* = .19). The amount of off-looking time during the known/unknown condition trials for each participant is included in Table 2. All of the remaining calculations based on looking times to target and non-target pictures excluded off-looking time.

Time period with greatest response discrepancy—We determined which looking period provided the most accurate and disparate responses between targets and non-targets

based on the mean differences between proportion of looking to targets versus non-targets during the four practice time periods: (a) Command 1, second 1; (b) Command 1, second 2; (c) Command 2, second 1; and (d) Command 2, second 2. Paired-sample t-tests indicated a statistically significant difference between the proportion of looking time to targets versus non-targets following Command 1, second 1, $t(28) = 3.77$, $p = .001$. No statistically significant differences were found between targets and non-targets during Command 1, second 2, $t(28) = 1.61$, $p = .12$); Command 2, second 1, $t(28) = .67$, $p = .51$); or Command 2, second 2, $t(28) = .91$, $p = .37$). Responses that occurred 1s following the first command showed the greatest difference between target and non-target responses ($d = .735$). The mean difference in proportion was 2% during this 1-s period. In comparison, the mean differences during the other three time periods were 1% ($d = .321$), .5% ($d = .134$), and .7% ($d = .172$), respectively. These differences were similar for children with and without autism; that is, there were not significant effects for group. To ensure brevity, only responses that occurred during Command 1, second 1 were used in the following analyses.

Responses to Target Words in Known and Unknown Conditions

To determine if eye tracking research technology responses corresponded with behavioral responses (pointing) on the PPVT-4, we examined the time spent looking at the target word during known and unknown conditions versus the mean time spent looking at non-target words in these same conditions (after subtracting off-looking time as explained previously). Preliminary Independent Samples t tests indicated that there were no statistically significant differences between children with ASD and children with typical development on the time spent looking at target and non-target words in the known condition, $t(27) = −1.84$, $p = .08$, $d = −.76$; $t(27) = 1.13$, $p = .27$, $d = .25$; or in the unknown condition, $t(27) = −1.94$, $p = .06$, $d = −.74$; $t(27) = .02$, $p = .99$, $d = 0$). Because of this we did not make comparisons between the two groups on looking time to target and non-target pictures. Instead, we examined separate within group comparisons—one for children with ASD and one for children with typical development.

Within group analysis—To examine within group differences, paired samples t tests were performed for the children with ASD and children with typical development on the looking time to target and non-target words in each condition (known vs. unknown). For this analysis, we adjusted the p value to .025 to account for two t-tests per participant group. The mean looking times for children with ASD are shown in Figure 1; Figure 2 shows mean looking times for children with typical development. Children with ASD spent significantly more time looking at the target pictures ($M = .24$, $SD = .07$) as compared to non-target pictures ($M = .13$, $SD = .04$) in the known condition, $t(13) = 5.37$, $p < .001$, $d = 1.54$. Similarly, children with typical development spent significantly more time looking at the target pictures ($M = .32$, $SD = .14$), compared to non-target pictures ($M = .12$, $SD = .04$) in the known condition, $t(14) = 5.40$, $p < .001$, $d = .35$. For children with ASD, there were no statistically significant differences between looking time to the target pictures ($M = .14$, $SD = .09$) and non-target pictures ($M = .15$, $SD = .05$) in the unknown condition, $t(13) = −.42$, $p = .68$, $d = −.11$. However, for children with typical development, the difference in looking times to the target pictures ($M = .21$, $SD = .10$) and non-target pictures ($M = .15$, $SD = .05$)

in the unknown condition was significant, $t(14) = 2.27$, $p = .04$, $d = .68$. This last finding went against our hypothesis and we will discuss this finding more in the Discussion.

Eventually, it is hoped that this line of research will determine how looking times for individual items reflect emergent word knowledge for individual children. Therefore, we looked at the degree to which the group results reflected the performances of individual children, and the degree to which the reported mean looking times for words within the sets of known and unknown words for children with ASD were representative of their looking times to individual items. The standard errors for the means for proportions of looking time to both known and unknown target pictures were .004, indicating little variability across stimuli. In addition, we looked at each participant's mean looking time to known and unknown pictures, and his or her r off-looking times, and these are presented in Table 2. As is apparent from the data in Table 2, the group results reflected the performance of many of the children with ASD. Using a criterion of at least a .10 s difference, ASD Participants 3, 4, 5, 6, 10, 12, and 14 all showed large differences in looking time to the targets compared to the non-targets when words were known (i.e., in the known condition as determined by their responses on the PPVT-4). Only ASD Participant 4 showed this pattern in the unknown condition.

For children with typical development, large differences in individual looking times within the known condition were shown for Participants 1, 2, 3, 4, 8, 9, 11, 12, 13, and 15. However, large differences in looking times between the target and the non-targets within the unknown condition were also detected for five participants with typical development (3, 7, 9, 10, and 13). In addition, smaller differences in this same direction were also observed for Participants 2, 5, and 8 with typical development.

Discussion

This study represents a preliminary investigation aimed at determining the feasibility of using eye tracking research technology as a method for measuring spoken word comprehension in children with ASD who are minimally verbal. In order to do this, we first needed to verify that eye gaze behaviors measured with an eye tracking research technology system mapped onto other behavioral indicators of receptive word understanding. This was accomplished by comparing eye gaze behaviors to traditional pointing responses in a group of children with ASD and a group of developmentally age-matched children with typical development who were all able to participate in a traditional *Point to* __ type of assessment (e.g., the PPVT-4).

Our hypothesis was that children would look longer at pictures for which they had previously indicated understanding, based on standard administration of the PPVT-4. This hypothesis was supported by the group data for children with ASD. Specifically, within the known condition, children, as a group, looked longer at target pictures than at non-target pictures. For children with ASD, as a group, there were no significant differences in looking at target vs. non-target pictures in the unknown condition (see example of fixation patterns shown in Figure 3). This pattern was observed to hold for the majority of individual children with ASD as well. For children with typical development, however, only the responses in

the known condition supported our hypotheses. Counter to our hypothesis, within the unknown condition significant differences were observed for looking to target vs. non-target pictures for children with typical development; further examination of the individual data for these children indicated that the significant differences in the unknown condition were largely accounted for by five participants' data. The eye tracking data from these five participants suggests that they may potentially have known at least some words in the unknown condition, despite having missed them on the PPVT-4.

In fact, similar discrepant results have been observed when examining word comprehension in infants using the IPLP, based on parental report (e.g., Houston-Price et al., 2007; Robinson et al., 2000). Specifically, using IPLP, 22-month-old infants indicated comprehension of words that parental report indicated had been heard by the infant, but were not yet understood (Robinson et al., 2000). Furthermore, Houston-Price et al. (2007) found that 15-, 18-, and 21-month-old infants would look at target images for words parents had reported as unknown during the IPLP task. These authors and others have viewed such responses as indicative of emergent language comprehension (see Golinkoff, Ma, Song, & Hirsh-Pasek, 2013 for a review). That is, a child may look at a named item even if they are not sure enough of its meaning to point to it.

It is also possible, however, that eye gaze responses that do not match behavioral responses reflect problems in the procedures. In our study it is possible that children with typical development failed to point to some known target pictures during PPVT-4 testing because of fatigue or acts of noncompliance (Golinkoff et al. 2013). Other possible explanations include procedural variables implemented while using the eye tracking research technology. For example, the pictures used during eye tracking were not identical to the pictures presented in the PPVT-4; therefore, it is possible that some of the pictures presented using the eye tracking technology were more salient representations than those presented during PPVT-4 testing. Past research using IPLP has suggested that the image used as stimuli matters, and this is most likely also true when using eye tracking research technology; therefore, future studies using this technology to measure receptive vocabulary should include salience trials to examine whether the visual stimuli are similar in terms of attractiveness and quality (Golinkoff et al., 2013).

The present study is an attempt to lay the groundwork for addressing the prevalent problem of assessing speech comprehension in children who do not readily participate in other forms of standardized test administration. This is a significant problem for children with ASD, who are often reported to be untestable with traditional paradigms/assessments. The current study is a first step in a line of research aimed at providing researchers and clinicians with a method to utilize alternative responses, such as eye gaze, to indicate implicit understanding by some children with ASD, as well as children with other types of disabilities. The findings here contribute to the growing research base demonstrating the validity of alternative test response modes (e.g., McDougall, Vessoyan, & Duncan, 2012).

Developing better language comprehension measures is an important area for further research for individuals who use AAC. Accurate information about an individual's comprehension would facilitate the development of educational and communication

programs that more accurately match his or her underlying abilities. For example, decisions about vocabulary selection for AAC devices could be influenced by a more accurate description of comprehension levels. It is also feasible that measuring comprehension of other content, outside of vocabulary, could be improved with strategies such as those explored in the current investigation. However, more research is needed before eye tracking research technology can be useful in typical clinical applications.

Limitations and Future Directions

One desirable outcome for future studies would be determining an objective criterion for looking time differences that could be interpreted as reflecting differential responses. For example, in the current study we observed that mean looking time to the target picture across all of the children was .28 seconds for the 1-s post-command, based on a four choice array; however, there was a range of responses as indicated in Table 2. Future research is needed to identify the number of looking responses for known words that meet or exceed an identified minimum difference in looking time in order to develop a method for identifying an individualized objective criterion for looking time that could be interpreted as a known response. In addition, future research is needed to establish for whom this is a reliable and valid measure of receptive vocabulary.

Although all of the children with ASD showed a difference in looking times to target pictures in known conditions, some differences were small and one was only .01s. In addition to variance across participants, minimal variance was also noted across items in each stimulus set. Additional review of the data across participants indicated minimal differences in looking times across items in the known condition for each participant, that is, the overall mean looking time scores were representative of responses to most items and were not attributable to a small subset of items. We also reviewed the data to see if some items were particularly salient and responded to with longer looking times across participants. However, we were not able to answer this question because each stimulus set was individualized and there were few common pictures across sets. Future research studies with larger numbers of participants responding to the same items would allow for item level analyses, which are necessary in order to identify response patterns that could be associated with characteristics of the pictures themselves, such as color or shape.

The current study focused on comparing proportions of looking time, but eye tracking research technology data are incredibly rich, and other variables should also be investigated in future studies. For example, in the IPLP paradigms, latency to looking to named words has been a robust outcome variable used to indicate speed of linguistic processing (Fernald, Zangl, Portillo, & Marchman, 2008; Golinkoff et al., 1987). Therefore, latency measures might also be valuable for eye tracking research on vocabulary comprehension.

Further replication and extension of this research is essential to ensure that this methodology can be applied to individuals with various levels of cognitive and language ability and to identify if meaningful results can be obtained from these populations. The current study included a small sample size and only children who were testable on a traditional PPVT-4. Also, we relied on parent report to describe vision and hearing status, and future studies should confirm these reports with assessments or additional documentation. Additional

research using this methodology needs to be applied to larger samples, including the target population of children who are typically untestable when presented with traditional assessments. It is possible that only a subset of these children can also be successfully assessed using the current methodology. Therefore, an important goal of future research will be to fine tune and individualize the current methodology for successful use with populations that have more significant challenges associated with typical testing paradigms.

In the current study, we relied on behavioral pointing data from administration of the PPVT-4 to indicate the known and unknown words administered during the eye tracking research technology session. However, the PPVT-4 was presented according to the instructions in the manual and each item was presented once; and the response array consisted of the four-choice grid. Therefore, there is a 25% chance that the participant may have correctly identified the right word by chance. Additional steps could be added in future studies to help ensure identification of known words such as additional presentations with different placements of the stimuli within the grids or obtaining additional validation information from parents or other knowledgeable individuals. In addition, we were not able to use identical pictures from the PPVT-4 due to copyright restrictions, and future studies may wish to verify the equivalence of the pictures. We compared responses on the PPVT-4 to responses using eye tracking research technology for individual vocabulary items; however, the PPVT-4 is designed to measure global understanding of single words. Therefore, in future studies, researchers may wish to investigate the comparability in overall scores obtained with eye tracking and other vocabulary measures..

The stationary nature of the equipment imposes another limitation, as participants must come to our research laboratory for testing. This may increase stress for the child and is inconvenient for families. We are currently investigating the feasibility of using a mobile eye tracking research technology system to address this limitation. The use of a mobile system might be especially helpful for children who are difficult to test using the traditional method because it will be possible to follow their eye-gaze during administration of the actual PPVT-4 and other assessments. Furthermore, it will be easier to ensure that the child is looking at the stimulus before presenting the directive, to individualize the redirection of their attention, and to conduct the assessments in classrooms with typical lighting (i.e., not in a darkened room).

Finally, the use of eye tracking research technology as an assessment tool is clearly more expensive and potentially more time consuming than behavioral forms of assessment. In addition, specific training is required to use the technology. Thus, the benefits of the results obtained must be evaluated in light of these costs. However, obtaining an accurate assessment is important and the use of this technology may be warranted for children who are difficult to test; development of less expensive equipment and more clinically friendly procedures could help alleviate some of these concerns.

In conclusion, eye tracking research technology may hold promise for increasing the accuracy of spoken word comprehension measures for individuals who are difficult to test. The results from the present study indicate that this technology might be a valid measure for speech comprehension for some children with ASD. Future studies are needed to

demonstrate successful testing with children who have severe disabilities and are difficult to test, to identify characteristics of successful participants, to assess item level differences, and to address the limitations imposed by the apparatus.

Acknowledgments

References

Anderson CJ, Colombo J, Shaddy DJ. Visual scanning and pupillary responses in young children with autism spectrum disorder. Journal of Clinical and Experimental Neuropsychology. 2006; 28:1238–1256.10.1080/13803390500376790 [PubMed: 16840248]

ASL. ASL EYE-TRACK 6 eye-tracking system with optics. Bedford, MA: ASL; 2008.

Bebko J, Weiss J, Demark J, Gomez P. Discrimination of temporal synchrony in intermodal events by children with autism and children with developmental disabilities without autism. Journal of Child Psychology and Psychiatry. 2006; 47:88–98. [PubMed: 16405645]

Behrend DA. The development of verb concepts: Children's use of verbs to label familiar and novel events. Child Development. 1990; 61:681–696. [PubMed: 2364743]

Bopp KD, Mirenda P. Prelinguistic predictors of language development in children with autism spectrum disorders over four–five years. Journal of Child Language. 2011; 38:485–503.10.1017/S0305000910000140 [PubMed: 20609280]

Bornstein MH, Hendricks C. Basic language comprehension and production in >100,000 young children from sixteen developing nations. Journal of Child Language. 2012; 39:899–918.10.1017/S0305000911000407 [PubMed: 22129486]

Brady N. Improved comprehension of object names following voice output communication aid use: Two case studies. Augmentative and Alternative Communication. 2000; 16:197–204.

Brady N, Thiemann-Bourque K, Fleming K, Matthews K. Predicting language outcomes for children learning AAC: Child and environmental factors. Journal of Speech, Language, and Hearing Research. in press.

Cauley K, Golinkoff RM, Hirsh-Pasek K, Gordon L. Revealing hidden competencies: A new method for studying language comprehension in children with motor impairments. American Journal on Mental Retardation. 1989; 94:53–63. [PubMed: 2751891]

Charman T, Drew A, Baird C, Baird G. Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). Journal of Child Language. 2003; 30:213–236. [PubMed: 12718299]

Colombo J. The development of visual attention in infancy. Annual Review of Psychology. 2001; 52:337–367.10.1146/annurev.psych.52.1.337

Cunningham C, Glenn S, Wilkinson P, Sloper P. Mental ability, symbolic play and receptive expressive language of young children with Down's syndrome. Journal of Child Psychology and Psychiatry. 1985; 26(2):255–265. [PubMed: 2579964]

Drager KDR, Postal VJ, Carrolus L, Castellano M, Gagliano C, Glynn J. The effect of aided language modeling on symbol comprehension and production in 2 preschoolers with autism. American Journal of Speech-Language Pathology. 2006; 15:112–125.10.1044/1058-0360(2006/012) [PubMed: 16782684]

Dunn, L.; Dunn, D. The Peabody Picture Vocabulary Test. 4. San Antonio, TX: Psychological Corporation; 2007. (PPVT-4)

Ellis Weismer S, Lord C, Esler A. Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. Journal of Autism and Developmental Disorders. 2010; 40:1259–1273.10.1007/s10803-010-0983-1 [PubMed: 20195735]

Eye-Gaze Response Interface Computer Aid (ERICA), Inc. GazeTracker [Computer software and manual]. Charlottesville, VA: 2001.

Facon B, Facon-Bollengier T, Grubar J. Chronological age, receptive vocabulary, and syntax comprehension in children and adolescents with mental retardation. American Journal on Mental Retardation. 2002; 107:91–98. [PubMed: 11853526]

Fenson L, Dale P, Reznick JS, Bates E, Thal DJ, Pethick SJ. Variability in early communicative development. Monographs of the Society for Research in Child Development. 1994; 59(5):1–173. [PubMed: 7845413]

Fenson, L.; Marchman, VA.; Thal, DJ.; Reznick, S.; Bates, E. The MacArthur-Bates Communicative Development Inventories: User's guide and technical manual. 2. Baltimore: Brookes Publishing; 2006.

Fernald, A.; Zangl, R.; Portillo, AL.; Marchman, VA. Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In: Sekerina, IA.; Fernandez, EM.; Clahsen, H., editors. Developmental psycholinguistics: On-line methods in children's language processing. Philadelphia, PA: John Benjamins; 2008. p. 97-135.

Geytenbeek JJM, Heim MMJ, Vermeulen RJ, Oostrom KJ. Assessing comprehension of spoken language in nonspeaking children with cerebral palsy: Application of a newly developed computer-based instrument. Augmentative and Alterative Communication. 2010; 26:97–107.

Golinkoff RM, Hirsch-Pasek K, Cauley KM, Gordon L. The eyes have it: Lexical and syntactic comprehension in a new paradigm. Journal of Child Language. 1987; 14:23–46. [PubMed: 3558524]

Golinkoff RM, Ma W, Song L, Hirsh-Pasek K. Twenty-Five Years Using the Intermodal Preferential Looking Paradigm to Study Language Acquisition: What Have We Learned? Perspectives on Psychological Science. 2013; 8:316–339.10.1177/1745691613484936

Goodwin A, Fein D, Naigles L. Comprehension of wh-questions precedes their production in typical development and autism spectrum disorders. Autism Research. 2012; 5:109–123. [PubMed: 22359403]

Houston-Price C, Mather E, Sakkalou E. Discrepancy between parental reports of infants' receptive vocabulary and infants' behavior in a preferential looking task. Journal of Child Language. 2007; 34:701–724. [PubMed: 18062356]

Hudry K, Leadbitter K, Temple K, Slonims V, McConachie H, Aldred C, Charman T. Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. International Journal of Language & Communication Disorders. 2010; 45:681–690.10.3109/13682820903461493 [PubMed: 20102259]

Huttenlocher J, Haight W, Bryk A, Seltzer M, Lyons T. Early vocabulary growth: Relation to language input and gender. Developmental Psychology. 1991; 27:236–248.

Johnson AL, Gillis JM, Romanczyk RG. A brief report: Quantifying and correlating social behaviors in children with autism spectrum disorders. Research in Autism Spectrum Disorders. 2012; 6:1053–1060.10.1016/j.rasd.2012.01.004

Lord, C.; Rutter, M.; DiLavore, PS. Autism Diagnostic Observation Schedule-Generic. Los Angeles, CA: Western Psychological Services; 1997. Manual and testing form

Luyster R, Qiu S, Lopez K, Lord C. Predicting outcomes of children referred for autism using the MacArthur-Bates Communicative Development Inventory. Journal of Speech, Language, and Hearing Research. 2007; 50:667–681.

Maljaars J, Noens I, Scholte E, Berckelaer-Onnes I. Language in low-functioning children with autistic disorder: differences between receptive and expressive skills and concurrent predictors of language. Journal of Autism and Developmental Disorders. 2012; 42(10):2181–2191.10.1007/s10803-012-1476-1 [PubMed: 22350453]

Mercadante MT, Macedo EC, Baptista PM, Paula CS, Schwartzman JS. Saccadic movements using eye-tracking technology in individuals with autism spectrum disorders. Arquivos De Neuro-Psiquiatria. 2006; 64:559–562. [PubMed: 17119790]

McClintock K, Hall S, Oliver C. Risk markers associated with challenging behaviours in people with intellectual disabilities: a meta-analytic study. Journal of Intellectual Disability Research. 2003; 47:405–416.10.1046/j.1365-2788.2003.00517.x [PubMed: 12919191]

McDuffie A, Yoder P, Stone W. Prelinguistic predictors of vocabulary in young children with autism spectrum disorders. Journal of Speech, Language, and Hearing Research. 2005; 48:1080–1097.

Mervis, C. Cross-etiology comparisons of cognitive and language development. In: Rice, ML.; Warren, SF., editors. Developmental language disorders: From phenotypes to etiologies. Mahway, NJ: Lawrence Erlbaum Associates; 2004. p. 153-188.

Mullen, EM. Mullen Scales of Early Learning, American Guidance Service (AGS) Edition. Circle Pines, MN: American Guidance Service, Inc; 1995.

Pierce K, Conant D, Hazin R, Stoner R, Desmond J. Preference for geometric patterns early in life as a risk factor for autism. Archives of General Psychiatry. 2011; 68:101–109.10.1001/archgenpsychiatry.2010.113 [PubMed: 20819977]

Preis J. The effect of picture communication symbols on the verbal comprehension of commands by young children with autism. Focus on Autism and Other Developmental Disabilities. 2006; 21:194–210.10.1177/10883576060210040101

Ray-Subramanian C, Ellis Weismer S. Receptive and expressive language as predictors of restricted and repetitive behaviors in young children with autism spectrum disorders. Journal of Autism and Developmental Disorders. 2012; 42:2113–2120.10.1007/s10803-012-1463-6 [PubMed: 22350337]

Riby D, Hancock PJ. Do faces capture the attention of individuals with Williams syndrome or autism? Evidence from tracking eye movements. Journal of Autism and Developmental Disorders. 2009a; 39:421– 431.10.1007/s10803-008-0641z [PubMed: 18787936]

Riby D, Hancock PJ. Looking at movies and cartoons. Eye-tracking evidence from Williams syndrome and autism. Journal of Intellectual Disability Research. 2009b; 53:169– 181.10.1111/j.1365-2788.2008.01142x [PubMed: 19192099]

Robinson, CW.; Shore, WJ.; Hull Smith, P.; Martinelli, L. Developmental differences in language comprehension: What 22-month-olds know when their parents are not sure. Poster presented at the International Conference on Infant Studies; Brighton, England. Jul. 2000

Romski, M.; Sevcik, R. Breaking the speech barrier: Language development through augmented means. Baltimore, MD: Paul H. Brookes Publishing; 1996.

Sasson NJ, Turner-Brown LM, Holtzclaw TN, Lam KSL, Bodfish JW. Children with autism demonstrate circumscribed attention during passive viewing of complex social and nonsocial picture arrays. Autism Research. 2008; 1:31–42.10.1002/aur.4 [PubMed: 19360648]

Sevcik, R.; Romski, M. The role of language comprehension in establishing early augmented conversations. In: Reichle, J.; Beukelman, D.; Light, J., editors. Exemplary practices for beginning communicators: Implications for AAC. Baltimore, MD: Paul H. Brookes Publishing; 2002.

Smith A. Development and course of receptive and expressive vocabulary from infancy to old age: Administrations of the Peabody Picture Vocabulary Test, Third Edition, and the Expressive Vocabulary Test to the same standardization population of 2725 subjects. The International Journal of Neuroscience. 1997; 92:73–78. [PubMed: 9522257]

Smith V, Mirenda P, Zaidman-Zait A. Predictors of expressive vocabulary growth in children with autism. Journal of Speech, Language, and Hearing Research. 2007; 50:149–160.

Speer LL, Cook AE, McMahon WM, Clark E. Face processing in children with autism: Effects of stimulus contents and type. Autism. 2007; 11:265– 277. [PubMed: 17478579]

Spelke ES. Infants' intermodal perception of events. Cognitive Psychology. 1976; 8:553–560.

Spelke ES. Perceiving bimodally specified events in infancy. Developmental Psychology. 1979; 15:626–636.

Styles S, Plunkett K. What is 'word understanding' for the parent of a one-year-old? Matching the difficulty of a lexical comprehension task to parental CDI report. Journal of Child Language. 2009; 36:895–908.10.1017/S0305000908009264 [PubMed: 19079845]

Swensen LD, Kelley E, Fein D, Naigles LR. Processes of language acquisition in children with autism: evidence from preferential looking. Child Development. 2007; 78:542–557.10.1111/j.1467-8624.2007.01022.x [PubMed: 17381789]

Thistle JJ, Wilkinson KM. Working Memory Demands of Aided Augmentative and Alternative Communication for Individuals with Developmental Disabilities. Augmentative and Alternative Communication. 2013; 29:235–245. [PubMed: 23902430]

Tincoff R, Jusczyk PW. Some beginnings of word comprehension in 6-month-olds. Psychological Science. 1999; 10:172–175.

Tomasello M, Mervis C. The instrument is great, but measuring comprehension is still a problem. Monographs of the Society for Research in child Development. 1994; 59:174–179.

Vivanti G, McCormick C, Young GS, Abucayan F, Hatt N, Nadig A, Rogers SJ. Intact and impaired mechanisms of action understanding in autism. Developmental Psychology. 2011; 47:841–856.10.1037/a0023105 [PubMed: 21401220]

Vivanti G, Nadig A, Ozonoff S, Rogers SJ. What do children with autism attend to during imitation tasks? Journal of Experimental Child Psychology. 2008; 101:186–205.10.1016/j.jecp.2008.04.008 [PubMed: 18582895]

Volden J, Smith IM, Szatmari P, Bryson S, Fombonne E, Mirenda P, Thompson A. Using the Preschool Language Scale, Fourth Edition to characterize language in preschoolers with autism spectrum disorders. American Journal of Speech-Language Pathology. 2011; 20:200–208.10.1044/1058-0360(2011/10-0035) [PubMed: 21478278]

Watt N, Wetherby A, Shumway S. Prelinguistic predictors of language outcome at 3 years of age. Journal of Speech, Language, and Hearing Research. 2006; 49:1224–1237.10.1044/1092-4388(2006/088)

Yoder PJ, Warren SF, Biggar HA. Stability of maternal reports of lexical comprehension in very young children with developmental delays. American Journal of Speech-Language Pathology. 1997; 6:59–64.

Zampini L, D'Odorico L. Communicative gestures and vocabulary development in 36-month-old children with Down's syndrome. International Journal of Language & Communication Disorders. 2009; 44:1063–1073.10.1080/13682820802398288 [PubMed: 19105070]

Zimmerman, IL.; Steiner, VG.; Pond, RE. Preschool Language Scale. 4. San Antonio, TX: The Psychological Corporation; 2002.
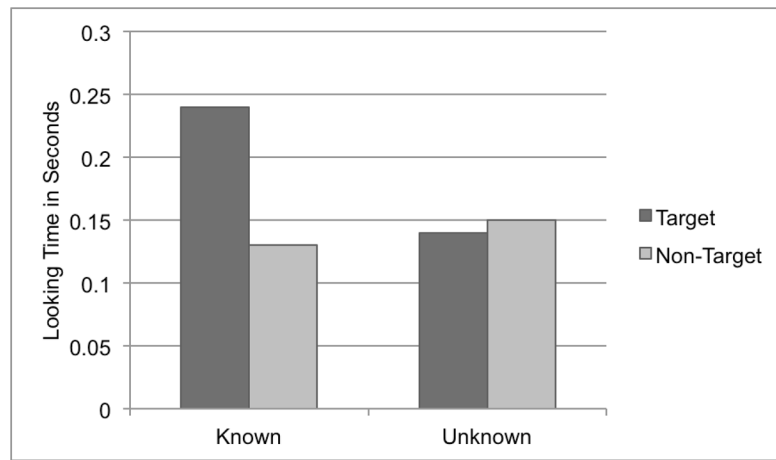
Figure 1.
Mean looking time in seconds to target and non-target regions of interest for both known and unknown conditions for children with ASD. Looking to non-target regions were adjusted for relative size.
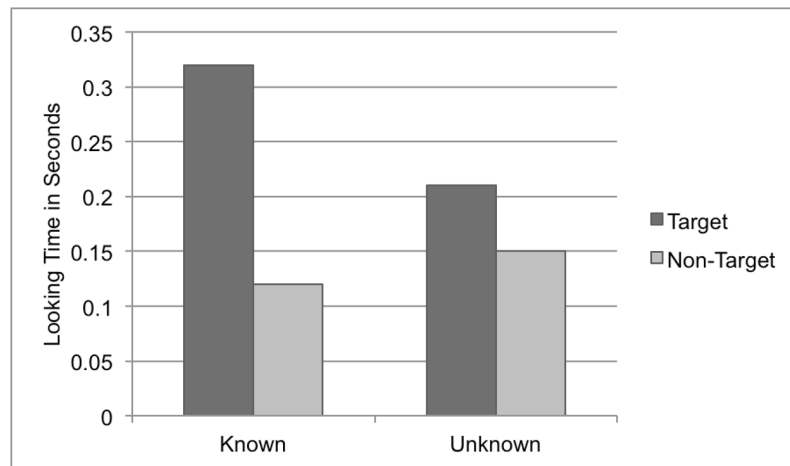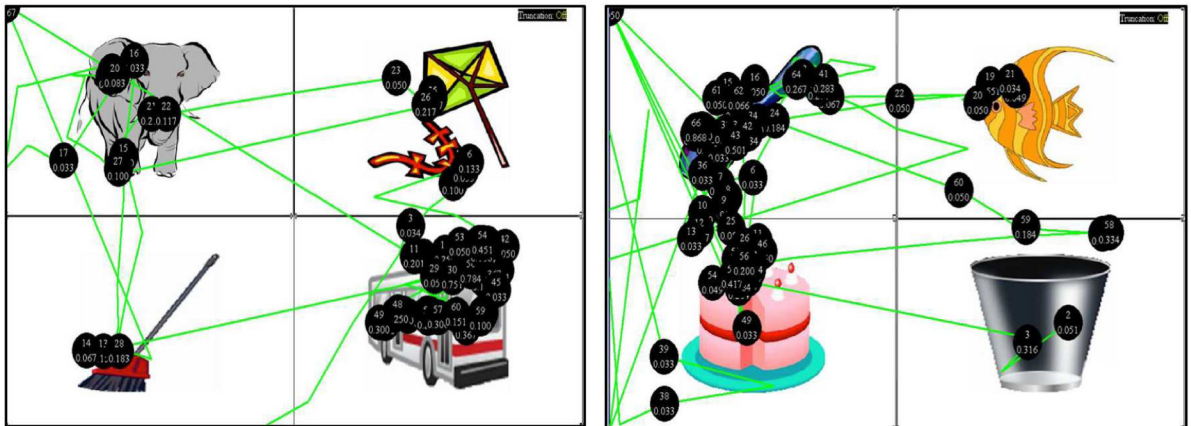
Figure 2.
Mean looking time in seconds to target and non-target regions of interest for both known and unknown conditions for children with typical development. Looking to non-target regions were adjusted for relative size.

Known

Bus

Spoon

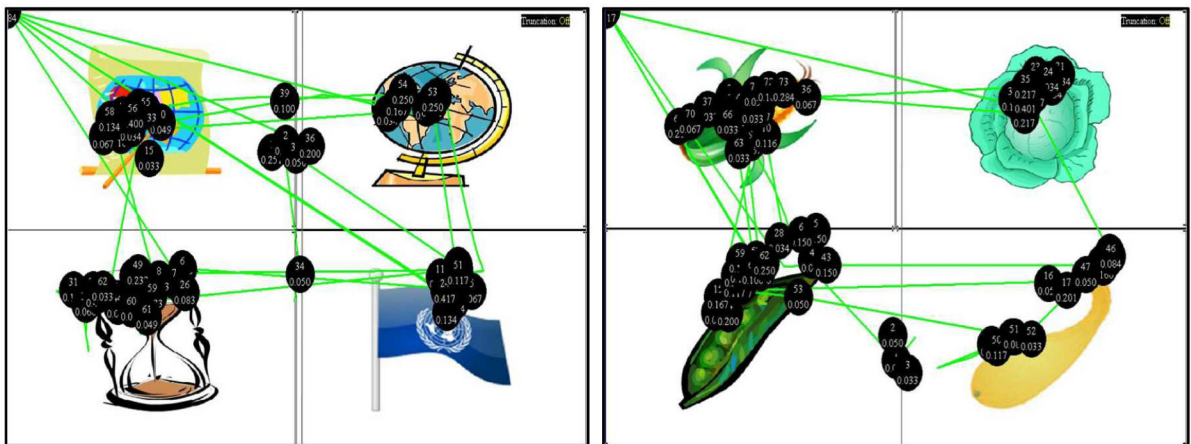# Unknown

Globe

Squash

Figure 3.
Example of gaze patterns to known and unknown words.

Table 1

Participant Demographic Information

Characteristic	Children with ASD ($n = 14$)		Children with typical development ($n = 15$)	
	M	Range	M	Range
Age [a]	61.07	46–82	56.47	42–70
Mullen Scales of Early Learning				
Early Learning Composite [**]	88.1	51–112	110.6	81–135
Overall Mental-Age[a]	55	29.5–66.5	59.63	43.25–67.25
Visual Reception T-Score	42.86	22–59	51.13	20–70
Visual Reception Age Equivalence [a]	55.07	20–69	56.6	45–69
Fine Motor T-Score	44.71	20–60	55.40	42–72
Fine Motor Age Equivalence[a]	56.29	30–68	58.40	42–68
Receptive T-Score	44.64	20–62	58.73	43–90
Receptive Age Equivalence [a]	56.43	27–69	61.33	37–69
Expressive T-Score	41.5	20–62	60.40	41–96
Expressive Age Equivalence [a]	52.07	18–70	62.20	40–70
ADOS=G: Modules 1 and 2				
Social	9.04	4.5–13	0	
Communication	7.14	2–22	0	
Behavior	3.92	1–7	0	
Total	16.18	6.5–30	0	
Parent Education Level [b]	3.85	0–7.5	4.77	2–8.5

Note.

[**] Significant difference between children with ASD and children with typical development, ($p < .01$).

[a] Presented in months.

[b] Years beyond high school averaged for both parents.

Social = Qualitative impairments in reciprocal social interaction; Behavior = Stereotyped behaviors and restricted interests; Total = total ADOS score.

Table 2

Looking Time in Seconds to Target and Average Looking Time to Non-Target Regions of Interest and Off-looking times for Known and Unknown Conditions for Participants with ASD and Those with Typical Development

Participant	Known condition			Unknown condition		
	Target	Non-target[a]	Off-looking time[2]	Target	Non-target[1]	Off-looking time[b]
ASD 1	.17	.13	.41	.09	.19	.33
ASD 2	.11	.08	.66	.03	.07	.76
ASD 3	.23	.13	.39	.17	.10	.51
ASD 4	.36	.14	.22	.38	.15	.18
ASD 5	.26	.15	.28	.12	.15	.42
ASD 6	.25	.04	.64	.09	.10	.61
ASD 7	.22	.17	.28	.19	.15	.37
ASD 8	.21	.15	.35	.19	.16	.33
ASD 9	.24	.15	.32	.08	.16	.43
ASD 10	.31	.17	.17	.24	.16	.27
ASD 11	.17	.13	.45	.03	.19	.39
ASD 12	.31	.11	.36	.07	.11	.60
ASD 13	.20	.19	.22	.11	.15	.45
ASD 14	.36	.13	.24	.16	.26	.04
ASD mean	.24	.13	.36	.14	.15	.41
TD 1	.43	.11	.22	.08	.18	.39
TD 2	.26	.17	.22	.20	.16	.32
TD 3	.38	.15	.16	.28	.17	.22
TD 4	.54	.12	.11	.20	.20	.20
TD 5	.15	.11	.52	.24	.16	.26
TD 6	.12	.06	.70	.04	.05	.83
TD 7	.22	.15	.33	.40	.14	.19
TD 8	.37	.14	.21	.26	.18	.21
TD 9	.32	.02	.62	.20	.05	.65

Participant	Known condition			Unknown condition		
	Target	Non-target[a]	Off-looking time[2]	Target	Non-target[1]	Off-looking time[b]
TD 10	.16	.16	.36	.34	.13	.26
TD 11	.39	.10	.30	.20	.25	.04
TD 12	.52	.09	.21	.17	.21	.22
TD 13	.45	.12	.21	.32	.16	.19
TD 14	.17	.10	.55	.06	.07	.74
TD 15	.29	.15	.26	.17	.17	.32
TD Mean	.32	.12	.33	.21	.15	.34

Note.

[a] This is the mean looking time in seconds for the three non-target regions.

[b] Looking offscreen = 1s − looking time to target region + total looking time to all three non-target regions.