Do minimally verbal and verbally fluent individuals with autism spectrum disorder differ in their viewing patterns of dynamic social scenes?

Daniela Plesa Skwerer[1] (iD), Briana Brukilacchio[2], Andrea Chu[1],
Brady Eggleston[1], Steven Meyer[1] and Helen Tager-Flusberg[1]

Abstract

Attending preferentially to social information in the environment is important in developing socio-communicative skills and language. Research using eye tracking to explore how individuals with autism spectrum disorder deploy visual attention has increased exponentially in the past decade; however, studies have typically not included minimally verbal participants. In this study, we compared 37 minimally verbal children and adolescents with autism spectrum disorder with 34 age-matched verbally fluent individuals with autism spectrum disorder in how they viewed a brief video in which a young woman, surrounded by interesting objects, engages the viewer, and later reacts with expected or unexpected gaze-shifts toward the objects. While both groups spent comparable amounts of time looking at different parts of the scene and looked longer at the person than at the objects, the minimally verbal autism spectrum disorder group spent significantly less time looking at the person's face during the episodes where gaze following—a precursor of joint attention—was critical for interpreting her behavior. Proportional looking-time toward key areas of interest in some episodes correlated with receptive language measures. These findings underscore the connections between social attention and the development of communicative abilities in autism spectrum disorder.

Keywords

dynamic scene, eye tracking, minimally verbal autism spectrum disorder, visual social attention

Investigations of visual social attention in autism spectrum disorders (ASD) have surged in the past decade, especially after unobtrusive eye-tracking technology became widely available for research. Enthusiasm for this topic was incited, in part, by the intriguing findings reported by Klin and colleagues in 2002 (Klin, Jones, Schultz, Volkmar, & Cohen, 2002a, 2002b), who examined the visual fixation patterns of adolescents and adults watching highly emotionally charged scenes from the 1967 movie version of Edward Albee's play *Who's afraid of Virginia Woolf?* These authors found that, in contrast to the visual scanning patterns showed by neurotypical peers, who consistently focused on the protagonists' faces, in particular the eye region, the participants with ASD looked significantly less at the eyes and more at the protagonists' mouth, body, or various objects in the scenes. Since this seminal study, research using eye tracking to explore how individuals with ASD orient to and engage attention toward social and nonsocial stimuli has increased rapidly, but findings of

atypicalities in visual attention deployment remain mixed (see Frazier et al., 2017; Guillon, Hadjikhani, Baduel, & Rogé, 2014; Papagiannopoulou, Chitty, Hermens, Hickie, & Lagopoulos, 2014 for recent reviews and meta-analyses of eye-tracking studies).

Eye movements have been studied as measures of attention monitoring, interest, problem-solving, and language comprehension in older verbal individuals with ASD (e.g. Bavin et al., 2014; Klin et al., 2002b; Sasson, Turner-Brown, Holtzclaw, Lam, & Bodfish, 2008; Venker, Eernisse, Saffran, & Weismer, 2013) and, more recently, in infants and toddlers (e.g. Chawarska, Macari, & Shic,

[1]Boston University, USA
[2]The University of Texas at Austin, USA

Corresponding author:
Daniela Plesa Skwerer, Department of Psychological & Brain Sciences, Boston University, 100 Cummington Mall, Boston, MA 02215, USA.
Email: dplesas@bu.edu

2013; Elsabbagh et al., 2012; Jones & Klin, 2013; Pierce et al., 2016). Much research has been conducted on the deployment of attention to faces as potential windows into the mechanisms underlying the social impairments found in ASD (Dawson, Webb, & McPartland, 2005; Sasson, 2006; Schultz, 2005; Weigelt, Koldewyn, & Kanwisher, 2012). Difficulty processing information from faces early in development has been linked to socio-cognitive limitations that hinder the acquisition of language, a process heavily dependent on social interactive processes, such as initiating and responding to episodes of joint attention, which involve gaze monitoring (Bedford et al., 2012; Chawarska, Macari, & Shic, 2012; Chawarska & Shic, 2009; Mundy, Sigman, & Kasari, 1990). The ability to follow a person's gaze is an important prerequisite for joint attention (Butler, Caron, & Brooks, 2009; Carpenter, Nagell, & Tomasello, 1998; Shepherd, 2010), which plays a significant role in the development of communication abilities and language in both typical development (e.g. Baldwin, 1995; Moore & Dunham, 1995; Tomasello & Farrar, 1986) and in autism (Adamson, Bakeman, Deckner, & Romski, 2009; Akechi et al., 2011; Baron-Cohen, Baldwin, & Crowson, 1997; Charman, 2003; Leekam, Lopez, & Moore, 2000; Loveland & Landry,1986; Mundy, Sigman, & Kasari, 1994; Toth, Munson, Meltzoff, & Dawson, 2006). Therefore, it is not surprising that an extensive body of research has examined this foundational ability in young children with ASD or in infants at risk for ASD, compared with those developing typically. A majority of these studies concluded that sensitivity to eye gaze is atypical in ASD, as shown by children's difficulties with spontaneously following another person's eye gaze to share attention (Bedford et al., 2012; Gillespie-Lynch, Elias, Escudero, Hutman, & Johnson, 2013). However, evidence for typical attentional cueing from eye-gaze direction has also emerged, especially when evaluated using experimental tasks (see Chawarska, Klin, & Volkmar, 2003; Falck-Ytter & von Hofsten, 2011; Nation & Penny, 2008 for reviews). Leekam, Hunnisett, and Moore (1998) found that differences between school-aged children with ASD in their ability to orient spontaneously to another person's head turn depended on their verbal mental ages, reporting that mainly children with mental ages below 48 months had difficulties with spontaneous gaze following. Research with older verbal individuals with ASD, using more complex stimuli, such as brief videos of social scenes, commonly focused on allocation of social attention during *free viewing* of the images/videos. *Social attention* in this context refers to the process of directing attention to aspects of people in a scene (Chevallier et al., 2015). Studies using eye-tracking technology usually compared looking-time at people/faces versus at nonsocial information (objects, background), and yielded mixed results across studies and tasks: some researchers reported that individuals with ASD without intellectual disabilities showed a reduced likelihood to follow a protagonist's gaze spontaneously when viewing a social scene (Fletcher-Watson, Leekam, Benson, Frank, & Findlay, 2009; Norbury et al., 2009; Riby & Hancock, 2008; Riby, Hancock, Jones, & Hanley, 2013); in contrast, others have reported typical patterns of looking behavior in response to gaze cueing in participants with ASD who have IQ within normal range (Freeth, Chapman, Ropar, & Mitchell, 2010). Examining visual attention to social scenes in teenagers, Norbury and colleagues (2009) found differences in viewing patterns related to participants' language status (e.g. between those with and without language impairments), while Rice and colleagues (Rice, Moriuchi, Jones, & Klin, 2012) reported significant variation in children's visual scanning of complex social scenes based on four distinct cognitive profiles among nonintellectually disabled children with ASD.

In sum, numerous studies have documented atypical patterns of social attention orienting in individuals with ASD across a range of experimental paradigms, and in real or simulated social interactions (Caruana, McArthur, Woolgar, & Brock, 2017; Franchini et al., 2017; Shic, Bradshaw, Klin, Scassellati & Chawarska, 2011; Senju, Tojo, Dairoku, & Hasegawa, 2004), but few have focused on individual differences across the wide spectrum of abilities in ASD. So far, eye-tracking studies have shown that findings depended on the tasks and type of stimuli used (isolated faces/objects, complex scenes, static images or dynamic stimuli, cf. Chevallier et al., 2015; Speer, Cook, McMahon, & Clark, 2007), on the context and task demands (experimental, passive viewing, interactive, cf. Freeth, Foulsham, & Kingstone, 2013; Noris, Nadel, Barker, Hadjikhani, & Billard, 2012), as well as on sample characteristics (intellectual functioning, age and verbal mental age, or communication abilities, cf. Leekam et al., 2000; Norbury et al., 2009; Rice et al., 2012). A relatively small sample size and the exclusion of individuals with ASD with more severe intellectual disabilities are common limitations of many of these studies, restricting the generalizability of the findings with respect to the broad autism spectrum. Even when studies included larger, heterogeneous samples of individuals with ASD and focused on patterns of variability in visual social engagement (Rice et al., 2012), participants' average IQ was not in the range of intellectual disability (i.e. standard score below 70).

Only recently have investigators started to focus on associations between eye-movement data and other phenotypic characteristics, besides autism symptom severity such as expressive and receptive language. Findings of these studies generally supported the hypothesis of a significant relationship between social attention and communication ability profiles in both young children and adolescents with ASD (e.g. Chawarska et al., 2012; Murias et al., 2018; Norbury et al., 2009). The span of verbal abilities among individuals with ASD ranges from those who

remain nonverbal into adulthood to those who become highly proficient in their expressive language (Kim, Paul, Tager-Flusberg, & Lord, 2014). Yet the sources of this heterogeneity and their possible links to processes of social attention deployment are not well understood.

As noted, the majority of previous research focused either on young, preverbal infants and toddlers (Chawarska et al., 2013; Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998; Elsabbagh et al., 2013; Jones & Klin, 2013; Klin, Lin, Gorrindo, Ramsay, & Jones, 2009; Swettenham et al., 1998) or on older children, adolescents, and adults who are able to speak (Fletcher-Watson et al., 2009; Klin et al., 2002b; Riby & Hancock, 2008; Riby et al., 2013). To date, it is unknown whether the approximately 30% of individuals with ASD who do not develop functional speech by school age differ in their attention allocation to social and nonsocial information in the environment, or whether their language and communication limitations are related to particular difficulties in attending to and processing socially relevant cues. Because of the challenges in testing this population, they have generally not been included as study participants in earlier research (Tager-Flusberg & Kasari, 2013; Tager-Flusberg et al., 2017).

This study was motivated by two main goals: one was to investigate whether distinctive patterns of visual social attention differentiated minimally verbal (MV-ASD) from verbally fluent (V-ASD) individuals with ASD, when viewing naturalistic dynamic scenes. Given that the ability to follow gaze is an important prerequisite for joint attention, we were interested in examining whether MV-ASD children and adolescents were sensitive to the attentional focus of a protagonist in a naturalistic scenario, as indicated by following the gaze and head turn of a person shown in a brief video. Another goal was to examine whether visual social attention was related to measures of language ability and to diagnostic measures of autism symptomatology. We presented participants with a brief video modeled after a task used by Chawarska and colleagues (2012), which was adapted to make it more appropriate for older children and adolescents. The video depicted a young woman making a snack at a table, surrounded by four interesting objects. In the video, the protagonist addresses the viewer in greeting, comments on her activity, and then reacts to the sudden movement of one of the objects, a mechanical toy spider, by shifting her gaze appropriately toward the moving object. In a later episode when the spider moves again, the woman shifts her gaze unexpectedly, toward an object placed opposite the spider (a static panda). Our primary aim was to explore whether the two groups differed in their allocation of visual attention to the protagonist and the objects in the video as a function of the events presented. More specifically, we hypothesized that the V-ASD participants would pay more attention to the protagonist's face and gaze behavior,

especially in the unexpected gaze-shift episode, when her behavior should surprise typical viewers. We predicted that in the latter episode the V-ASD participants would demonstrate the tendency to spontaneously follow the protagonist's gaze toward the target of her attention (i.e. will follow her gaze/head direction of movement toward the panda), whereas this viewing pattern will be diminished or absent in the MV-ASD group. We also predicted that visual attention toward the protagonist—in particular, her face and direction of gaze, as well as the target of her attention—would be positively related to measures of language ability and negatively related to aspects of autism symptom severity.

Methods

Participants

Participants were 71 individuals with ASD, divided into two groups based on language ability. A total of 37 participants (8 girls) ranging in age between 8.6 and 20.2 years ($M = 13.56$ years, $SD = 3.4$) were described by their parents as having little to no functional speech used in a range of social contexts. Criteria for assignment to the MV-ASD group included lack of spontaneous functional speech or inconsistent simple phrase speech of no more than three units, as defined by the Autism Diagnostic Observation Schedule-Second Edition (ADOS-2; Hus et al., 2011) Module 1. This definition of MV-ASD has been used in the previous literature (Bal, Katz, Bishop, & Krasileva, 2016). The other 34 participants (8 girls), aged between 8.9 and 20.9 years, ($M = 14.97$ years, $SD = 3.4$) were verbally fluent (V-ASD) and used complex phrase speech consistently. Diagnoses of all participants were confirmed using the ADOS-2 and the Autism Diagnostic Interview-Revised (ADI-R; Le Couteur, Lord, & Rutter, 2003). The MV-ASD participants were administered Module 1 of either the ADOS-2 or the Adapted ADOS (A-ADOS; Hus et al., 2011), depending on their age: the MV-ASD participants above 12 years were assessed with the A-ADOS, which uses play materials more appropriate and engaging for adolescents. The V-ASD participants were administered Modules 3 or 4 of the ADOS-2, as appropriate for their age and language level. Social-affective and restrictive and repetitive behavior symptom severity were calculated with the ADOS calibrated symptom severity scores (CSS), which are comparable across ADOS modules (Hus, Gotham, & Lord, 2014). Table 1 summarizes the demographic characteristics of the two groups.

The Peabody Picture Vocabulary Test (PPVT-4; Dunn & Dunn, 2007) was administered to assess receptive word knowledge. Nonverbal IQ (NVIQ) was assessed using the Leiter-3 (Roid, Miller, Pomplun, & Koch, 2013) for the MV-ASD participants, and the WASI-II (Wechsler, 2011) for the V-ASD participants. The Leiter-3 is a test commonly

used with minimally and low-verbal individuals with ASD (Kasari, Brady, Lord, & Tager-Flusberg, 2013) because it does not require verbal instructions or verbal responding, facilitating a reliable assessment of nonverbal reasoning abilities relatively independent of language. The Perceptual Reasoning Index of the WASI-II was used to obtain an estimate of NVIQ for the V-ASD group. In addition to the

Table 1. Demographic Characteristics of Participants.

	MV-ASD (N = 37)		V-ASD (N = 34)	
	M	SD	M	SD
Chronological Age (years)	13.56	3.5	14.97	3.4
Gender:				
Male/Female	29 /8		26 /8	
Race (N):				
African-American	2		2	
Asian	4		2	
White	27		22	
Hispanic	0		3	
Native Hawaiian or other Pacific islander	1		0	
More than one race	3		5	

MV-ASD: minimally verbal children and adolescents with autism spectrum disorder; V-ASD: verbally fluent children and adolescents with autism spectrum disorder; *SD*: Standard deviation.

ADI-R, parents completed the Vineland Adaptive Behavior Scales-2 (VABS-2; Sparrow, Cicchetti, & Balla, 2005), administered in an interview format. Table 2 summarizes the descriptive characteristics of the groups.

All participants had normal or corrected to normal vision and no significant sensory or neurological impairments, according to a brief medical history survey completed by parents. Only participants from predominantly English-speaking homes were included in the study. Informed consent and participant assent were obtained from caregivers and from V-ASD participants, as appropriate. All study procedures were approved by the Institutional Review Board of the university in which the study was conducted.

Procedures

Eye-tracking task. Participants' eye movements were recorded with a TOBII T60 XL eye-tracker run by *Tobii Studio 2.0.3 software* (Tobii Technology AB, Danderyd, Sweden). This system requires no headgear and has relatively high tolerance for head movements. We used a five-point calibration and adapted the choice of calibration method (adult or infant) to each participant. The choice of calibration method was dictated by the need to maximize the likelihood of attracting a fixation with minimum verbal instructions. Five and even two-point calibration methods are commonly used with individuals with severe intellectual disabilities (Wilkinson & Mitchell, 2014).

Table 2. Behavioral characteristics of participants.

	MV-ASD (N = 37)		V-ASD (N = 34)	
	M (SD)	Range	M (SD)	Range
Receptive vocabulary[1]	25.88 (10.4)	20–64	101.53 (25.9)	31–135
Nonverbal reasoning[2]	62.14 (14.7)	30–94	104.15 (23.2)	64–152
VABS-2 Adaptive Behavior Composite	48.94 (8.9)	30–69	76.52 (13.4)	37–104
VABS-2 Communication	47.72 (9.9)	28–70	78.97 (17.8)	42–118
Receptive language[3]	26.88 (14.5)	1–59	102.58 (71.6)	16–216
Expressive language[4]	20.61 (10.8)	3–42	109.24 (81.3)	5–276
ADI-R				
Total A[5]	26.83 (3.4)	15–30	20.58 (5.0)	10–29
Total B (NV)[6]	12.47 (1.8)	9–14		
Total B (V)[7]			15.15 (5.2)	7–25
ADOS severity scores:				
Overall CSS	7.73 (1.5)	6–10	7.58 (2.3)	1–10
Social-Affect-CSS	7.24 (1.5)	5–10	7.27 (2.3)	1–10
Restricted Repetitive Behavior-CSS	8.49 (1.5)	5–10	7.67 (2.0)	1–10

MV-ASD: minimally verbal children and adolescents with autism spectrum disorder; V-ASD: verbally fluent children and adolescents with autism spectrum disorder; *SD*: Standard deviation; VABS: Vineland Adaptive Behavior Scales; ADI-R: Autism Diagnostic Interview-Revised; NV: nonverbal; V: verbal; ADOS: Autism Diagnostic Observation Schedule-Second Edition; CSS: calibrated severity score.
[1]Standard scores from the PPVT-4 assessment.
[2]Standard scores from the Leiter-3 for the MV-ASD group and from the WASI-perceptual reasoning scale for the V-ASD group.
[3,4]Age equivalent scores in months.
[5]ADI total on qualitative abnormalities in reciprocal social interaction.
[6]ADI total on qualitative abnormalities in communication—nonverbal subjects.
[7]ADI total on qualitative abnormalities in communication—verbal subjects.

The eye-tracking task featured a video of a young woman making a snack. The movie display area was a rectangle subtending $35° \times 23.4°$ of visual angle. Four interesting objects were placed surrounding the woman, who was shown in the center of the scene seated at a table facing the camera. The objects (iPad, toy spider, toy-Panda, Jack-in-box toy) were about the same size, subtending $8.9° \times 8.9°$ of visual angle, and were positioned on the table and on top of two boxes placed on the left and the right sides of the protagonist. Other AOIs included the <u>face/head</u> of the protagonist ($8.7° \times 10.4°$ of visual angle) and the <u>hands/activity area</u> ($11.8° \times 7.4°$ of visual angle). The subregions of the <u>eyes</u> and <u>mouth</u> subtended $5.2° \times 2°$ and $4° \times 1.6°$ of visual angle, respectively (see Figure 1). The video was divided into six episodes based on the protagonist's behavior (see Table 3). Three episodes were critical for assessing social visual attention: (a) episode 2: *"Verbal greeting"* in which the protagonist lifts her head, looks toward the camera, and addresses the viewer; (b) episode 3: *"Expected gaze-shift"* showing the mechanical toy spider starting to move on the table and the protagonist's gaze following the spider's movement; and (c) episode 5: *"Unexpected gaze-shift"* depicting the spider moving again, but the protagonist looking toward an unmoving object (a panda bear) placed diagonally opposite from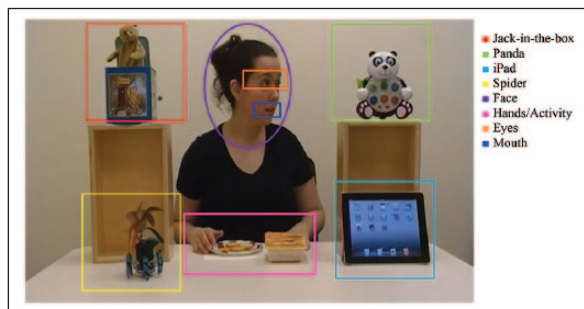 the spider. It should be noted that in the gaze-shift episodes the young woman also turns her head, not just the eyes, toward the object of her attentional focus, so there is no ambiguity about her direction of gaze for the viewer watching her behavior. The total duration of the video is 75 s, and the duration of each episode is listed in Table 3.

Participants were seated approximately 60 cm from the monitor, with eye-level approximately even with the center of the scene. Up to five calibration attempts were conducted with each participant, at successive visits if needed, before the task was administered. After successful calibration, the participants' compliance and interest in watching the movie varied significantly and the amount of valid data contributed by each participant across the video duration, according to the TOBII system, ranged from 1% to 93% in the MV-ASD group ($M=49.5\%$, $SD=29.1$) and from 2% to 99% in the V-ASD group ($M=62.3\%$; $SD=35.1$). We included in the analyses participants with more than 15% valid data across the movie duration, with the additional constraint that they needed to provide data in at least five of the six episodes of the video. Nine MV-ASD participants who had no fixations in two or more episodes or provided less than 15% valid data across the video were excluded from further analyses. Five V-ASD participants were excluded based on these criteria, resulting in 28 MV-ASD participants and 29 V-ASD participants with gaze data included in analyses. Because our main interest was in capturing the characteristics of visual attention allocation to a complex scene by MV-ASD individuals who have ordinarily not been included in eye-tracking studies, we could not afford to employ more stringent gaze data validity criteria without having to exclude a significant number of participants, potentially biasing the characterization of the attentional processes that may be distinctive to this ASD subpopulation. The MV-ASD and V-ASD groups were matched on chronological age, $F(1, 56)=0.25$, $p=0.88$, and on ADOS calibrated severity scores (CSS) (Gotham, Pickles, & Lord, 2009). The excluded participants within each group did not differ on age, receptive language, IQ, or

Figure 1. Composition of the scene—snapshot from episode 5.

Table 3. Description of the movie episodes.

	Duration (s)	Event	Audio-track
Episode 1	22	Protagonist starts preparing a snack; she looks down at the box of crackers and plate she placed on the table	Silent episode
Episode 2	11	Protagonist looks directly into the camera and speaks to the viewer, then resumes preparing the snack	*Oh hi, how are you? It's my snack time*
Episode 3	8	Mechanical spider starts to move across the table; protagonist shifts gaze toward the spider and follows its movement until it stops	Silent episode
Episode 4	15	Protagonist turns her attention back to preparing the snack and speaks to the viewer	*I'm making cheese and crackers*
Episode 5	12	Spider starts moving across the table; Protagonist shifts gaze in opposite direction, looking at the toy panda bear	Silent episode
Episode 6	7	Protagonist resumes preparing the snack and speaks toward the camera	*Now I'm ready to eat. I am hungry*

ADOS symptom severity (based on ADOS CSS) from those who were retained.

Analytic approach. First, we compared the groups in their overall attention across all episodes by calculating proportional looking-time to the video (i.e. their gaze falling within the media frame) relative to the video total duration, to obtain an individual measure of general attention to the dynamic scene. Individual looking-time at the video was used in later analyses to calculate proportional looking-time within each area of interest (AOI). More specifically, all analyses involving within-AOIs visual fixation data were conducted on proportional variables calculated as looking-time within a particular AOI divided by the participant's total looking-time at the scene (i.e. within the media frame), considered both across the movie duration and within the duration of each episode. This approach was intended to mitigate the potential biasing effects of missing data in particular episodes when analyzing participants' attention allocation to predefined AOIs relative to the key video events. Because of the differences in cognitive functioning between the two groups, we covaried NVIQ standard scores in all analyses of proportional looking-time data.

Next, we analyzed participants' distribution of visual attention to the person and the four objects collapsed across all six episodes, to examine whether the salience—as indexed by proportional viewing time—of social (protagonist) and nonsocial (toys) elements of the dynamic scene, differed for the two groups.

The next set of analyses explored attention to specific AOIs that were tied to a priori predictions based on salient events in each of the key episodes. We tested whether AOIs and episodes differentially influenced viewing time in the MV-ASD and the V-ASD groups with a mixed-model analysis of covariance (ANCOVA) and followed main effects and interactions with post hoc comparisons reported by key episode. Because the primary purpose of the study was to determine whether and how MV-ASD individuals differ from V-ASD peers in their visual attention allocation to salient AOIs as a function of the events in the video, we prioritized reporting comparisons between participant groups, within key episodes, for particular AOIs relevant for interpreting the scene: <u>face</u>, in episodes 2 (protagonist addresses the viewer), 3 (protagonist shifts gaze toward the moving spider), and 5 (protagonist shifts gaze toward the stationary panda), <u>spider</u> in episodes 3 and 5 (in which it starts to move unexpectedly), and <u>panda</u> in episode 5 (because it is the target of the protagonist's unexpected gaze shift).

We also compared the two groups in the proportion of individuals who made a responsive fixation toward the targets of the protagonist's gaze after looking at her face in the two gaze-shifting episodes. This additional nonparametric approach was meant to test whether participants in the two groups showed a spontaneous gaze-following

tendency, regardless of the amount of viewing time spent within the relevant AOIs. Participants were categorized into those who did and those who did not make a fixation in the relevant AOIs in the key episodes, and chi-square tests were used to compare the MV-ASD and V-ASD groups based on these categories of responders.

Finally, to determine whether social attention as indexed by looking-time data was related to language abilities and to autism symptom severity, we investigated correlations between proportional looking-time to the specific AOIs listed above and scores on measures of receptive and expressive language, and ratings of autism symptomatology.

Results

Overall viewing of the video

A one-way ANCOVA conducted on looking-time at the scene relative to the total video duration, controlling for NVIQ, yielded a significant group effect, $F(1, 56) = 4.83$, $p = 0.032$, $\eta^2 = 0.081$ showing that the MV-ASD group spent on average less time ($M = 56.5\%$) than did the V-ASD group ($M = 72.2\%$) attending to the video overall. However, the groups did not differ in their initial attention to the video during the first episode, $F(1, 56) = 0.538$, $p = 0.46$. When controlling for individual looking-time at the scene (i.e. within the media frame), the proportional viewing time spent within the six most relevant AOIs (i.e. the sum of looking-time spent within the six nonoverlapping AOIs—face/head, hands/activity area, spider, panda, iPad, and Jack-in-the-box—divided by the individual time spent looking at the entire screen) did not differ by group: $F(1, 56) = 0.353$, $p = 0.56$. Both groups looked at the relevant AOIs on average for over 85% of the time they attended to the screen (85.3% for the MV-ASD and 88% for the V-ASD, respectively). Table 4 presents the proportion of valid looking-time by participant group for each of the three key episodes.

As noted above, analyses of visual attention to particular AOIs were conducted on proportional looking-time data (i.e. variables of interest were standardized by individual looking-time at the scene across or within episodes, respectively). An inspection of these data revealed a positively skewed distribution; therefore, logarithmic transformations were applied to normalize the data distribution. For ease of interpretation, however, Table 4 presents the untransformed percentages of looking-time within AOIs relative to individual time attending to the scene in the three key episodes.

Distribution of overall visual attention between the person and objects

We first compared the groups in their proportional attending with the objects (i.e. the sum of looking at the iPad,

panda, spider, and Jack-in-the-box relative to individual looking at the scene) versus attending to the protagonist (i.e. looking at the face and the hands/activity area, relative to looking at the scene) during the entire duration of the movie. A mixed-model ANCOVA with AOI (person, objects) as the within-subjects factor and *group* (MV-ASD vs V-ASD) as the between-subjects factor on proportional looking-time measured across the movie duration yielded a significant main effect of AOI, $F (1, 50) = 4.13$, $p = 0.04$, $\eta_p{}^2 = 0.076$, but no main effect of group $F (1, 50) = 0.55$, $p = 0.461$ or interaction between group and AOI, $F (1, 50) = 2.29$, $p = 0.14$. Both the groups looked proportionally longer at the person ($M = 54.87\%$, $SD = 29.53$ in the MV-ASD group and $M = 62.52\%$, $SD = 18.92$ in the V-ASD group, respectively) than at the objects ($M = 29.53\%$, $SD = 14.19$ in the MV-ASD group and $M = 25.5\%$, $SD = 11.92$ in the V-ASD group, respectively) across the six episodes.

Next, we examined whether the participants' allocation of attention to the objects and to the protagonist depended on the content of the events viewed, as defined by the protagonist's behavior toward the viewer in episode 2 (verbal greeting), and toward the moving and stationary objects in the scene (in episodes 3 and 5 in which the protagonist shifts her gaze to objects). We conducted analyses of proportional looking-time in each AOI relative to individuals' viewing time within each episode, to minimize potential biasing effects of missing data in particular episodes. All participants retained in analyses provided data in the three key episodes, 2, 3, and 5.

Distribution of attention within each AOI as a function of episode content

An initial mixed-model ANCOVA, with AOI (6) and episode (6) as within-subjects factors and group (2) as the between-subjects factor, covarying NVIQ, yielded a significant main effect of AOI, $F (5, 250) = 6.25$, $p = 0.0001$, $\eta_p{}^2 = 0.11$, and a significant main effect of episode, $F(5, 250) = 2.61$, $p = 0.02$, $\eta_p{}^2 = 0.03$, which were qualified by a significant three-way interaction between AOI, episode, and group, $F (25, 1250) = 1.58$, $p = 0.035$, $\eta_p{}^2 = 0.031$. Following the significant three-way interaction, we analyzed participant group differences in proportional looking-time to predicted AOIs within each key episode (2, 3, and 5). Table 4 presents the untransformed proportional looking-time data for every AOI by key episode and participant group.

Episode 2—Verbal greeting. In this episode, we were primarily interested in whether the protagonist's verbal greeting influenced how the MV-ASD versus V-ASD participants allocated attention to the *face*. Group differences for proportional attending to the face in this episode were not statistically significant, $t (55) = -1.69$, $p = 0.096$, with both the groups spending about a third of their viewing time looking at the young woman's face when she addressed the viewer (see Table 4).

An additional analysis was conducted for this episode involving only the eyes and mouth as AOIs: because the *face* AOI included both the mouth and the eye regions, we further investigated whether the group similarities in proportional viewing time of the face involved a similar or a different distribution of attention between the two facial features—eyes and mouth. A separate group (MV-ASD, V-ASD) X AOI (eyes, mouth) analysis of variance (ANOVA) for proportional looking-time in episode 2 yielded a significant main effect of AOI, $F (1, 55) = 5.24$, $p = 0.026$, $\eta_p{}^2 = 0.088$, but no significant group X AOI interaction, $F(1, 55) = 1.72$, $p = 0.195$, $\eta_p{}^2 = 0.03$: both the groups looked longer at the mouth than at the eyes in this episode (Figure 2 and Table 4).

Episode 3—Expected gaze-shift. The primary comparisons of interest in this episode involved looking at the protagonist's face as she turned her gaze toward a moving spider, and looking at the spider, which was the object of her attentional focus and was unexpectedly moving. Both the groups looked significantly longer at the moving spider than at the protagonist's face during episode 3, $t (27) = -3.66$, $p = 0.001$ in the MV-ASD group and $t (28) = -2.6$, $p = 0.015$ in the V-ASD group. However, the two groups differed significantly in their looking behavior at the face in this episode, as the MV-ASD participants spent on average proportionally less viewing time (10.5%) on the face AOI compared with the V-ASD group, who spent on average over 21% of their looking-time on the protagonist's face, $t (55) = -3.18$, $p = 0.002$. Proportional viewing time at the spider did not differ significantly between the MV-ASD and V-ASD groups in episode 3.

Episode 5—Unexpected gaze-shift. In episode 5, the primary comparisons of interest involved the protagonist's face, the moving spider, and the panda toward which the young woman shifts her gaze unexpectedly. The groups differed significantly in their proportional viewing time for two AOIs: for the face, $t (55) = -3.01$, $p = 0.004$ and for the panda, $t (55) = -2.63$, $p = 0.011$, with the V-ASD participants looking proportionally longer at both these AOIs than the MV-ASD participants did (see Table 4).

Table 4 also presents the percentage of participants who made a responsive fixation to the panda after a fixation on the protagonist's face, in each group. A significantly lower proportion of participants made at least a fixation on the panda among the MV-ASD individuals (21.4%) compared with 55.2% of the V-ASD group, $\chi^2 = 6.84$, $p = 0.009$.

Table 4. Proportional looking-time per AOI and key episode, and percentage of participants who made a responsive fixation to selected AOIs in the episodes involving gaze-sifting.

Episode	MV-ASD		V-ASD		F/χ^2	p value
	(N = 28)		(N = 29)			
	M	(SD)	M	(SD)		
2. *Verbal greeting*						
Valid time (%)	89.83	16.1	91.8	14.75	0.76	ns
Hands/activity (%)	32.33	25.02	40.51	20.74	3.98	ns
Face (%)	32.27	27.94	37.98	18.17	2.09	ns
Mouth (%)	13.91	21.75	17.38	15.93	1.59	ns
Eyes (%)	11.29	20.48	6.49	11.07	0.613	ns
Toy spider (%)	10.86	22.05	6.62	7.17	1.31	ns
iPad (%)	3.24	6.37	1.32	4.16	2.89	ns
To panda (%)	3.59	5.14	1.52	3.36	3.07	.ns
Jack in the box (%)	5.06	10.48	3.85	10.35	0.271	ns
3. *Expected gaze-shift*						
Valid time (%)	85.16	28.98	88.02	22.37	1.64	ns
Hands/activity (%)	14.92	16.69	13.92	12.02	1.57	ns
Face (%)	10.48	15.37	21.17	15.51	10.16	0.002
Mouth (%)	4.53	11.87	7.08	9.32	3.77	.ns
Eyes (%)	1.75	3.76	5.01	10.93	2.37	ns
Toy spider (%)	36.88	30.34	47.03	28.5	3.61	ns
iPad (%)	6.01	19.32	0.79	1.55	2.45	ns
Toy panda (%)	5.12	11.82	1.29	3.22	1.73	.ns
Jack in the box (%)	4.56	9.07	3.89	8.43	0.123	ns
N who made a responsive fixation to spider (%)	75		96.6		5.48	0.019
5. *Unexpected gaze-shift*						
Valid time (%)	76.83	31.74	87.83	22.93	1.86	ns
Hands/activity (%)	35.68	26.76	33.47	18.59	0.630	ns
Face (%)	12.47	14.17	23.99	15.34	9.05	0.004
Mouth (%)	3.42	6.35	9.33	11.52	7.01	0.011
Eyes (%)	4.30	9.08	7.80	10.39	3.73	0.051
Toy spider (%)	16.22	16.68	21.67	14.95	3.28	ns
iPad (%)	4.46	12.08	0.753	1.35	1.16	ns
Toy panda (%)	2.79	8.05	5.56	7.07	6.9	0.011
Jack in the box (%)	5.21	14.97	1.32	2.37	0.78	ns
N who made a responsive fixation to panda (%)	21.4		55.2		6.84	0.009
N who made a responsive fixation to spider (%)	71.4		89.7		3.04	ns

Relations between visual social attention and measures of cognition, language ability, and autism symptomatology across and within episodes

First, we examined the possible relations between proportional looking-time in each relevant AOI, collapsed across episodes, and cognitive functioning (NVIQ), considering significance with Bonferroni correction at $p = 0.008$ (0.05/6). Only the correlation between proportional looking-time at the spider collapsed across episodes and NVIQ was significant, r (53) = 0.375, $p = 0.002$. Proportional looking-time at the spider collapsed across episodes was also correlated with Vineland Adaptive Behavior composite score, r (50) = 0.376, $p = 0.006$, but no other gaze-related variables were significantly correlated with any measures of cognition, communication, adaptive functioning, or autism symptom severity when considered across episodes.

To address specific questions about the possible relationships between gaze-following ability, attending to another person's attentional focus and language-related skills, we conducted correlational analyses separately for the episodes involving the protagonist's gaze shift (3 and 5), controlling for age and IQ. More specifically, we investigated whether looking-time at the AOIs that provided
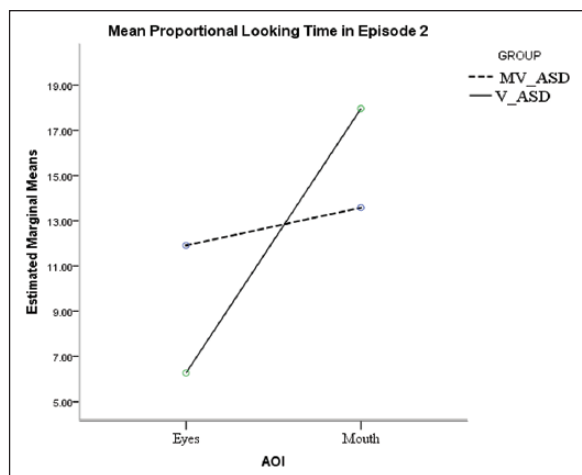
Figure 2. Episode 2—Proportional looking-time (mean %) at the protagonist's eyes and mouth, by group.

cues for interpreting the protagonist's behavior in particular video segments (i.e. the young woman's face and the spider in episodes 3 and 5; the panda in episode 5) correlated with language abilities.

In episode 3 (*Expected gaze-shift*), proportional looking-time at the protagonist's *face* was positively correlated with PPVT-4 scores, after controlling for age and NVIQ, r (49) = 0.442, p = 0.001. Proportional looking-time at the spider, however, was not correlated with language measures in this episode, once NVIQ was partialled out. Proportional looking-time at the face was also positively correlated with PPVT-4 scores in episode 5 (*Unexpected gaze-shift*), r (48) = 0.348, p = 0.014. Interestingly, in episode 5, proportional looking-time at the panda—the object toward which the protagonist unexpectedly shifted her gaze when the spider started to move—was positively correlated with both PPVT-4 scores, r (47) = 0.309, p = 0.01, and with the Vineland Communication Domain score, r (47) = 0.364, p = 0.005, after controlling for age and NVIQ.

We further examined correlations among measures of autism symptom severity obtained from the ADOS and the ADI diagnostic assessments, and proportional looking-time spent on the protagonist's mouth in episode 2, face/eyes and spider in episode 3, and face/eyes and panda in episode 5 (on both the ADOS and the ADI higher scores indicate more impairment). Only two looking-time AOI-related variables showed significant correlations with ASD symptomatology: in episode 3 proportional looking-time at the protagonist's face was negatively related to scores on the ADI for *qualitative abnormalities in reciprocal social interaction, r* (41) = −0.498, p = 0.001. In episode 5, looking-time at the protagonist's eyes was negatively correlated with ADOS overall CSS, r (51) = −0.333, p = 0.007. No significant relationships were found between looking-time variables and ADOS CSS for any other AOIs in any of the episodes.

Discussion

In this study, we compared MV-ASD children and adolescents with age-matched V-ASD participants in their viewing of naturalistic dynamic scenes, focusing on how they distributed attention to areas of the scene that involved social cues, such as a protagonist's face and gaze behavior. The majority of past research using eye-tracking methods to assess social attention in ASD has compared individuals with ASD with neurotypical controls. In this study, we wanted to explore whether investigating similarities and differences between MV-ASD and V-ASD children and adolescents in their spontaneous viewing patterns of a naturalistic video clip could provide insights into the possible connections between social attention atypicalities and failure to acquire spoken language in individuals with autism. We hypothesized that proportional looking-time toward AOIs that provided social cues to interpreting the events in the video, especially in the episode when the protagonist's behavior was unexpected, would be related positively to communication abilities and negatively to scores on ADOS and ADI items targeting joint attention and social reciprocity, an expectation that was largely supported by our findings.

Our results point to several commonalities and differences in how MV-ASD and V-ASD individuals deploy their attention to the components of a naturalistic scene involving a person and a set of interesting objects. Of note, although the MV-ASD participants tended to pay, on average, less attention to the entire video than their V-ASD peers, initial attention to the scene in the first episode was similar and relatively high in both groups, suggesting that they started similarly motivated to attend to the task. Relative to total movie duration, both V-ASD and MV-ASD participants spent proportionally more time looking at the protagonist compared with looking at the interesting objects placed around her in the scene. Consistent with findings reported by Chawarska and colleagues (2012) for toddlers, and Rice et al. (2012) for school-aged children with ASD, our results do not suggest a generalized disinterest in looking at people in a social scene, even in the presence of an intriguing moving mechanical toy. Instead, our findings suggest that MV-ASD participants may be less motivated to attend to and interpret the protagonist's *behavior* in a complex scene. Chawarska and colleagues (2012) found that toddlers with ASD showed diminished attention to an actor's face compared with the comparison groups, particularly in the condition where dyadic bids (child directed speech and eye contact) were present. In our study, we found group differences mainly in the segments that entailed interpreting the actor's gaze shift toward and away from a surprising moving object (episodes 3 and 5): in these episodes, the MV-ASD participants spent proportionally less viewing time on the protagonist's face than did the V-ASD group. Moreover,

significantly fewer participants in the MV-ASD group compared with the V-ASD group followed the protagonist's line of regard to look at the object of her attention when she shifted her gaze unexpectedly toward a static toy. It is notable, however, that even in the V-ASD group there were participants who did not look at the panda in episode 5, despite spending viewing time on the protagonist's face. Just over half of the V-ASD group followed the protagonist's gaze shift to the panda. The tendency to follow spontaneously the direction of another person's gaze toward the target of that person's attention is of particular significance for establishing episodes of joint attention, particularly in an interactive context. In a social context, this tendency could reflect responsiveness to others' bids for joint attention. In a free-viewing passive paradigm, following spontaneously an actor's direction of gaze may indicate the development of a foundational prerequisite for joint attention, although it does not constitute proof of joint attention abilities. Other studies conducted with *verbal* children and adolescents with ASD (Freeth et al., 2010; Riby et al., 2013) have reported subtle differences in gaze following between individuals with ASD and IQ matched typically developing children, or individuals with Williams syndrome. For instance, Riby et al. (2013), requested explicit responses from participants about the target of an actor's gaze in a social scene, after a free-viewing phase. These authors showed that, when cued to follow an actor's gaze in a naturalistic scene, participants with ASD looked more at the face and eyes but did not increase gaze to the correct targets of the actor's attention, continuing to look much longer than their controls at implausible targets. In the spontaneous viewing phase, however, they spent less time on people's faces and eyes than the control groups did. It appears from these results that atypicalities in spontaneous gaze following remain common among individuals with ASD across a wide range of verbal abilities. However, in our study, analyses relating visual social attention variables to language measures indicated that looking-time at the most salient AOIs in the gaze-shifting episodes (i.e. the face in episodes 3 and 5, and the panda in episode 5) was positively correlated with receptive language scores on a vocabulary test (PPVT-4), as well as with a parent report of communication abilities (Vineland Communication domain scores). Thus, participants who allocated more attention to the protagonist's face and to the focus of her attention in the relevant episodes had better communication abilities according to these language assessments. Most significantly, proportional looking-time at the protagonist's face/eye-region in the unexpected gaze-shift episode was positively correlated with standardized measures of receptive vocabulary, suggesting a meaningful relationship between the ability to attend to visual social cues and language comprehension among children and adolescents with ASD. This relationship is particularly salient because the visual attention deployment measures

in our sample were largely independent of NVIQ or overall ASD symptom severity on the ADOS. The lack of sensitivity to the social cue of gaze shifting suggests that the MV-ASD participants may have difficulties understanding the referential nature of looking. In our paradigm, even school-aged MV-ASD children and adolescents showed either a lack of understanding or a lack of interest in interpreting the protagonist's gaze shift, which was surprising in the context shown.

Reports in the literature relating the allocation of visual social attention to communication abilities, or to autism symptomatology, vary widely. While some researchers found direct predictive relations between looking patterns and level of social competence or disability (e.g. Jones, Carr, & Klin 2008; Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Thorup et al., 2018), others have reported no correlations between gaze metrics and autism symptoms (see Guillon et al., 2014 for a review). In our study, we found few and quite specific associations between gaze to the person-related AOIs and ASD symptomatology measured by the ADOS and the ADI: only for the unexpected gaze-shift episode (5), looking at the protagonist's eyes was negatively correlated with the ADOS CSS, while looking at the face in the expected gaze-shift episode was negatively correlated with scores for *qualitative abnormalities in reciprocal social interaction*, on the ADI. Thus, the participants who showed more impairment in social interactive abilities on the two diagnostic assessments of autism were those who tended to look proportionally less at the protagonist's eyes/face in the episodes when these AOIs provided cues for interpreting her behavior in the video. These correlations suggest a possible link between the ability to attend to the subtler social cue of gaze shifting and lower levels of ASD symptom severity.

To our knowledge, this is the first study to use a naturalistic video to directly compare the gaze allocation patterns of minimally verbal and verbally fluent age-matched children and adolescents with ASD. The group differences we found were mainly related to attention toward a protagonist's face, eyes, or target of attention, when these AOIs provided or failed to provide relevant cues for interpreting the actor's behavior in the scene. It is likely that the differences found between the MV-ASD and V-ASD groups in gaze time allocation to particular AOIs reflect decreased attention, among the MV-ASD participants, to behaviors that entail inferring the underlying intentions of the protagonist, suggesting either a lack of understanding, or a lack of interest in trying to interpret other people's actions. The free-viewing paradigm used in our study, while less demanding than protocols that require an explicit response from participants, is not conducive to refuting such alternative explanations. Regardless of the underlying causes, these findings suggest that MV-ASD children may be less able to learn from interactive opportunities involving

people's shifts of attentional focus, which are critical for establishing joint attention; this limitation may further impair their ability to detect and interpret social cues, and may have downstream influences on their development of language and communication abilities.

Study limitations

In a first effort to characterize how MV-ASD children and adolescents deploy their visual attention to a dynamic scene showing a person involved in a routine activity, we started by documenting similarities and differences between individuals with ASD who remained non- or minimally verbal after age 8, and verbally fluent peers with ASD, in their viewing patterns. As designed and conducted, our study did not include a non-ASD control group and does not address larger theoretical questions about the nature of joint attention and gaze-following atypicalities in autism, or their underlying mechanisms and neural underpinnings. We focused on viewing patterns in a passive paradigm to determine whether MV-ASD children and adolescents differ from V-ASD peers in their spontaneous allocation of attention to a protagonist's gaze and target of looking (attentional focus), as gaze-following ability is an important prerequisite of the ability to participate in joint attention episodes. We acknowledge that for probing joint attention abilities, interactive paradigms that involve social partners are more appropriate. Recent research has made tremendous progress using technology to record eye movements during real-life interactions or to employ virtual reality in simulating social interactive contexts, while ensuring tight experimental controls and even recording brain activity during such interactions (see Caruana, McArthur, et al., 2017 for a review of these studies). While watching another person in a video looking at various objects may not capture the essence of this social process, the "third person" perspective involved in passive viewing paradigms is not without any merits. Indeed, these have been used extensively in research on the allocation of social attention, using eye-tracking technology. Our choice of a more "traditional" free-viewing paradigm was motivated by the need to facilitate comparisons between findings from research conducted with cognitively able participants with ASD using similar stimuli, and research with MV-ASD individuals who have usually been excluded from eye-tracking studies. The methodological limitations of our study are directly related to the difficulties of engaging MV-ASD participants in research tasks: for instance, we used only one video clip as stimulus, without a comparable nonsocial viewing condition to match for nonsocial viewing differences; also, we did not fully control for the salience of the particular elements of the scene by using another set of objects and a male protagonist. We also acknowledge as limitations the less stringent inclusion criteria for analyses of looking-time data than those used in studies with cognitively able individuals, and the use of a five-point calibration method instead of a nine-point calibration for eye tracking. These constraining methodological decisions were dictated by the need to reduce testing time and attentional demands for the MV-ASD participants in particular. Caution in the interpretation of our results is also needed: even though we covaried NVIQ in all our analyses, we cannot rule out the possibility that differences in proportional viewing time between the MV-ASD and the V-ASD groups may not be truly independent of nonsocial cognitive processes (e.g. oculomotor control, other aspects of attention or motivation to attend to the task), in the context of scene viewing. In sum, we acknowledge inherent methodological limitations driven by our goal to provide a realistic description of the social attention characteristics of this subpopulation of more severely impaired individuals with ASD, while minimizing task demands.

Conclusion

Our results suggest specific and subtle differences in viewing patterns between MV-ASD and V-ASD children and adolescents that were related to particular aspects of language and communication skills, primarily receptive vocabulary. These findings have important implications for the possibility of training social attention allocation to promote the development of other abilities, including those related to understanding and using language. Future research should explore whether interventions targeting basic social attention processes could improve outcomes in communication skills for MV-ASD children and adolescents. While research on individuals at the "neglected end of the spectrum" (Tager-Flusberg & Kasari, 2013) is slowly emerging, the wealth of phenotyping information provided by these efforts holds promise for better understanding the significant heterogeneity of ASD, as well as for developing ways to improve social functioning in affected individuals.

ORCID iD

Daniela Plesa Skwerer (iD) https://orcid.org/0000-0003-4136-8189

References

Adamson, L. B., Bakeman, R., Deckner, D. F., & Romski, M. A. (2009). Joint engagement and the emergence of language in young children with autism and Down syndrome. *Journal of Autism and Developmental Disorders*, *39*, 84–96.

Akechi, H., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2011). Do children with ASD use referential gaze to learn the name of an object? An eye-tracking study. *Research in Autism Spectrum Disorders*, *5*, 1230–1242.

Bal, V. H., Katz, T., Bishop, S. L., & Krasileva, K. (2016). Understanding definitions of minimally verbal across instruments: Evidence for subgroups within minimally verbal children and adolescents with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *57*, 1424–1433.

Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 131–158). Mahwah, NJ: Lawrence Erlbaum.

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development*, *68*, 48–57.

Bavin, E. L., Kidd, E., Prendergast, L., Baker, E., Dissanayake, C., & Prior, M. (2014). Severity of autism is related to children's language processing. *Autism Research*, *7*, 687–694.

Bedford, R., Elsabbagh, M., Gliga, T., Pickles, A., Senju, A., Charman, T., . . . BASIS Team. (2012). Precursors to social and communication difficulties in infants at-risk for autism: Gaze following and attentional engagement. *Journal of Autism and Developmental Disorders*, *42*, 2208–2218.

Butler, S. C., Caron, A. J., & Brooks, R. (2009). Infant understanding of the referential nature of looking. *Journal of Cognition and Development*, *1*, 359–377.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*, 1–143.

Caruana, N., McArthur, G., Woolgar, A., & Brock, J. (2017). Simulating social interactions for the experimental investigation of joint attention. *Neuroscience & Biobehavioral Reviews*, *74*, 115–125.

Charman, T. (2003). Why is joint attention a pivotal skill in autism? *Philosophical Transactions of the Royal Society*, *358*, 315–324.

Chawarska, K., Klin, A., & Volkmar, F. (2003). Automatic attention cueing through eye movement in 2-year old children with autism. *Child Development*, *74*, 1108–1122.

Chawarska, K., Macari, S., & Shic, F. (2012). Context modulates attention to social scenes in toddlers with autism. *Journal of Child Psychology and Psychiatry*, *53*, 903–913. doi:10.1111/j.1469-7610.2012.02538.x

Chawarska, K., Macari, S., & Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biological Psychiatry*, *74*, 195–203.

Chawarska, K., & Shic, F. (2009). Looking but not seeing: Atypical visual scanning and recognition of faces in 2 and 4-year-old children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *39*, 1663–1672. doi:10.1007/s10803-009-0803-7

Chevallier, C., Parish-Morris, J., McVey, A., Rump, K. M., Sasson, N. J., Herrington, J. D., & Schultz, R. T. (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Research*, *8*, 620–628.

Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, *28*, 479–485.

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. *Development Neuropsychology*, *27*, 403–424.

Dunn, L. M., & Dunn, L. M. (2007). *Peabody picture vocabulary test* (4th ed.). Circle Pines, MN: American Guidance Service.

Elsabbagh, M., Gliga, T., Pickles, A., Hudry, K., Charman, T., & Johnson, M. H., & BASIS Team. (2013). The development of face orienting mechanisms in infants at-risk for autism. *Behavioural Brain Research*, *251*, 147–154.

Elsabbagh, M., Mercure, E., Hudry, K., Chandler, S., Pasco, G., Charman, . . . BASIS Team. (2012). Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. *Current Biology*, *22*, 338–342.

Falck-Ytter, T., & von Hofsten, C. (2011). How special is social looking in ASD: A review. *Progress in Brain Research*, *189*, 209–222. doi:10.1016/B978-0-444-53884-0.00026-9

Fletcher-Watson, S., Leekam, S. R., Benson, V., Frank, M. C., & Findlay, J. M. (2009). Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia*, *47*, 248–257.

Franchini, M., Glaser, B., Wood de Wilde, H., Gentaz, E., Eliez, S., & Schaer, M. (2017). Social orienting and joint attention in preschoolers with autism spectrum disorders. *PLoS ONE*, *12*(6), e0178859. doi:10.1371/journal.pone.0178859

Frazier, T. W., Strauss, M., Klingemier, E. W., Zetzer, E. E., Hardan, A. Y., Eng, C., & Youngstrom, E. A. (2017). A meta-analysis of gaze differences to social and nonsocial information between individuals with and without autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *56*, 546–555.

Freeth, M., Chapman, P., Ropar, D., & Mitchell, P. (2010). Do gaze cues in complex scenes capture and direct attention of high functioning adolescents with ASD process social information in complex scenes. Combining evidence from eye movements and verbal descriptions. *Journal of Autism and Developmental Disorders*, *41*, 364–371.

Freeth, M., Foulsham, T., & Kingstone, A. (2013). What affects social attention? Social presence, eye contact and autistic traits. *PLoS ONE*, *8*(1), e53286.

Gillespie-Lynch, K., Elias, R., Escudero, P., Hutman, T., & Johnson, S. P. (2013). Atypical gaze following in autism: A comparison of three potential mechanisms. *Journal of Autism and Developmental Disorders*, *43*, 2779–2792.

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum

disorders. *Journal of Autism and Developmental Disorders*, *39*, 693–705.

Hus, V., Gotham, K., & Lord, C. (2014). Standardizing ADOS domain scores: Separating severity of social affect and restricted and repetitive behaviors. *Journal of Autism and Developmental Disorders*, *44*, 2400–2412.

Hus, V., Maye, M., Harvey, L., Guthrie, W., Liang, J., & Lord, C. (2011). The adapted ADOS: Preliminary findings using a modified version of the ADOS for adults who are non-verbal or have limited language. In *Poster presented at the International Meeting for Autism Research, San Diego, CA*.

Jones, W., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry*, *65*, 946–954. doi:10.1001/archpsyc.65.8.946

Guillon, Q., Hadjikhani, N., Baduel, S., & Rogé, B. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience & Biobehavioral Reviews*, *42*, 279–297.

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, *504*(7480), Article 427.

Kasari, C., Brady, N., Lord, C., & Tager-Flusberg, H. (2013). Assessing the minimally verbal school-aged child with autism spectrum disorder. *Autism Research*, *6*, 479–493.

Kim, S. H., Paul, R., Tager-Flusberg, H., & Lord, C. (2014). Language and communication in autism. In F. R. Volkmar, R. Paul, S.J. Rogers, & K. A. Pelphrey (Eds.), *Handbook of autism and pervasive developmental disorders diagnosis, development, and brain mechanisms* (4th ed., Vol. 1, pp. 230–262). New York, NY: Wiley.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002a). Defining and quantifying the social phenotype in autism. *The American Journal of Psychiatry*, *159*, 895–908.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002b). Visual fixation patterns during viewing of natu-ralistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, *59*, 809–816.

Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social con-tingencies rather than biological motion. *Nature*, *459*(7244), 257–261.

Le Couteur, A., Lord, C., & Rutter, M. (2003). *Autism diagnostic interview-revised*. Los Angeles, CA: Western Psychological Services.

Leekam, S. R., Hunnisett, E., & Moore, C. (1998). Targets and cues: Gaze following in children with autism. *Journal of Child Psychology and Psychiatry*, *39*, 951–962.

Leekam, S. R., Lopez, B., & Moore, C. (2000). Attention and joint attention in preschool children with autism. *Developmental Psychology*, *36*, 261–273.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule (ADOS)* (2nd ed., Manual: Modules 1–4). Washington, DC: Western Psychological Services.

Loveland, K. A., & Landry, S. H. (1986). Joint attention and lan-guage in autism and developmental language delay. *Journal of Autism and Developmental Disorders*, *163*, 335–349.

Moore, C., & Dunham, P. J. (Eds.) (1995). *Joint attention: Its origins and role in development*. Mahwah, NJ: Lawrence Erlbaum.

Mundy, P., Sigman, M., & Kasari, C. (1990). A longitudinal study of joint attention and language development in autistic children. *Journal of Autism and Developmental Disorders*, *20*, 115–128.

Mundy, P., Sigman, M., & Kasari, C. (1994). Joint attention, developmental level, and symptom presentation in autism. *Development and Psychopathology*, *6*, 389–401.

Murias, M., Major, S., Davlantis, K., Franz, L., Harris, A., Rardin, B., . . . Dawson, G. (2018). Validation of eye-track-ing measures of social attention as a potential biomarker for autism clinical trials. *Autism Research*, *11*, 166–174.

Nation, K., & Penny, S. (2008). Sensitivity to eye gaze in autism: Is it normal? Is it automatic? Is it social? *Development and Psychopathology*, *20*, 79–97.

Norbury, C. F., Brock, J., Cragg, L., Einav, S., Griffiths, H., & Nation, K. (2009). Eye-movement patterns are associated with communicative competence in autistic spectrum dis-orders. *Journal of Child Psychology and Psychiatry*, *50*, 834–842.

Noris, B., Nadel, J., Barker, M., Hadjikhani, N., & Billard, A. (2012). Investigating gaze of children with ASD in natural-istic settings. *PLoS ONE*, *7*(9), e44144.

Papagiannopoulou, E. A., Chitty, K. M., Hermens, D. F., Hickie, I. B., & Lagopoulos, J. (2014). A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. *Social Neuroscience*, *9*, 610–632.

Pierce, K., Marinero, S., Hazin, R., McKenna, B., Barnes, C. C., & Malige, A. (2016). Eye tracking reveals abnormal visual preference for geometric images as an early bio-marker of an autism spectrum disorder subtype associated with increased symptom severity. *Biological Psychiatry*, *79*, 657–666.

Riby, D. M., & Hancock, P. J. B. (2008). Viewing it differently: Social scene perception in Williams syndrome and autism. *Neuropsychologia*, *46*, 2855–2860.

Riby, D. M., Hancock, P. J. B., Jones, N., & Hanley, M. (2013). Spontaneous and cued gaze-following in autism and Williams syndrome. *Journal of Neurodevelopmental Disorders*, *5*, Article 13. doi:10.1186/1866-1955-5-13

Rice, K., Moriuchi, J. M., Jones, W., & Klin, A. (2012). Parsing heterogeneity in autism spectrum disorders: Visual scanning of dynamic social scenes in school-aged children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*, 238–248.

Roid, G. H., Miller, L. J., Pomplun, M., & Koch, C. (2013). *Leiter International Performance Scale* (Leiter 3). Torrance, CA: Western Psychological Services.

Sasson, N. J. (2006). The development of face processing in autism. *Journal of Autism and Developmental Disorders*, *36*, 381–394. doi:101007/s1080300600763

Sasson, N. J., Turner-Brown, L. M., Holtzclaw, T. N., Lam, K. S. L., & Bodfish, J. W. (2008). Children with autism demonstrate circumscribed attention during passive view-ing of complex social and nonsocial picture arrays. *Autism Research*, *1*, 31–42.

Schultz, R. T. (2005). Developmental deficits in social percep-tion in autism: The role of the amygdala and fusiform face

area. *International Journal of Developmental Neuroscience*, *23*, 125–141.

Senju, A., Tojo, Y., Dairoku, H., & Hasegawa, T. (2004). Reflexive orienting in response to eye gaze and an arrow in children with and without autism. *Journal of Child Psychology and Psychiatry*, *45*, 445–458.

Shepherd, S. T. (2010). Following gaze: Gaze-following behavior as a window into social cognition. *Frontiers in Integrative Neuroscience*, *4*, Article 5. doi:10.3389/fnint.2010.00005

Shic, F., Bradshaw, J., Klin, A., Scassellati, B., & Chawarska, K. (2011). Limited activity monitoring in toddlers with autism spectrum disorder. *Brain Research*, *1380*, 246–254. doi:10.1016/j.brainres.2010.11.074

Sparrow, S., Cicchetti, D., & Balla, D. (2005). *Vineland Adaptive Behavior Scales* (2nd ed.). Circle Pines, MN: AGS Publishing.

Speer, L. L., Cook, A. E., McMahon, W. M., & Clark, E. (2007). Face processing in children with autism. *Autism*, *11*, 265–277.

Swettenham, J., Baron-Cohen, S., Charman, T., Cox, A., Baird, G., Drew, A., & . . . Wheelwright, S. (1998). The frequency and distribution of spontaneous attention shifts between social and nonsocial stimuli in autistic, typically developing, and nonautistic developmentally delayed infants. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*, 747–753.

Tager-Flusberg, H., & Kasari, C. (2013). Minimally-verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism Research*, *6*, 468–478.

Tager-Flusberg, H., Plesa Skwerer, D., Joseph, R. M., Brukilacchio, B., Decker, J., Eggleston, B., . . . Yoder, A. (2017). Conducting research with minimally-verbal participants with autism spectrum disorder. *Autism*, *21*, 852–861.

Thorup, E., Nyström, P., Gredebäck, G., Bölte, S., Falck-Ytter, & EASE Team. (2018). Reduced alternating gaze during social interaction in infancy is associated with elevated symptoms of autism in toddlerhood. *Journal of Abnormal Child Psychology*, *46*, 1547–1561.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*, 1454–1463.

Toth, K., Munson, J., Meltzoff, A. N., & Dawson, G. (2006). Early predictors of communication development in young children with autism spectrum disorder: Joint attention, imitation, and toy play. *Journal of Autism and Developmental Disorders*, *36*, 993–1005.

Venker, C. E., Eernisse, E. R., Saffran, J. R., & Weismer, S. E. (2013). Individual differences in the real-time comprehension of children with ASD. *Autism Research*, *6*, 417–432.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of intelligence* (2nd ed.). San Antonio, TX: Pearson.

Weigelt, S., Koldewyn, K., & Kanwisher, N. (2012). Face identity recognition in autism spectrum disorders: A review of behavioral studies. *Neuroscience & Biobehavioral Reviews*, *36*(3), 1060–1084.

Wilkinson, K. M., & Mitchell, T. (2014). Eye tracking research to answer questions about augmentative and alternative communication assessment and intervention. *Augmentative and Alternative Communication*, *30*, 106–119.