Distance from Typical Scan Path When Viewing Complex Stimuli in Children with Autism Spectrum Disorder and its Association with Behavior

Elena J. Tenenbaum, Samantha Major, Kimberly L. H. Carpenter, Jill Howard, Michael Murias & Geraldine Dawson

ONLINE FIRST

Springer

Springer

ORIGINAL PAPER

Distance from Typical Scan Path When Viewing Complex Stimuli in Children with Autism Spectrum Disorder and its Association with Behavior

Elena J. Tenenbaum[1,2,3] · Samantha Major[1] · Kimberly L. H. Carpenter[1,2,3] · Jill Howard[1] · Michael Murias[1,4] · Geraldine Dawson[1,2,3]

Abstract
Eye-tracking is often used to study attention in children with autism spectrum disorder (ASD). Previous research has identified multiple atypical patterns of attention in children with ASD based on areas-of-interest analysis. Fewer studies have investigated gaze path, a measure which is dependent on the dynamic content of the stimulus presented. Here, rather than looking at proportions of looking time to areas of interest, we calculated mean fixations frame-by-frame in a group of typically developing children (36 to 72 months) and determined the distance from those typical fixations for 155 children with ASD (27–95 months). Findings revealed that distance from the typical scan path among the children with ASD was associated with lower communication abilities and greater ASD symptomatology.

Keywords Eye-tracking · Autism · Social attention · Communication

Introduction

For nearly two decades, eye-tracking has been used to distinguish between the gaze patterns of individuals with autism spectrum disorder (ASD) and neurotypical individuals while viewing naturalistic social stimuli (Klin et al. 2002b; Pelphrey et al. 2002). In light of group differences in patterns of social attention, the correspondence between gaze

---

✉ Elena J. Tenenbaum
   elena.tenenbaum@duke.edu

1   Duke Center for Autism and Brain Development, Duke University, Durham, NC 27705, USA

2   Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC 27705, USA

3   Duke Institute for Brain Sciences, Duke University, Durham, NC 27708, USA

4   Present Address: Medical Social Sciences, Northwestern University, Chicago, IL 60622, USA

---

patterns and behavioral measures (Klin et al. 2002a; Murias et al. 2018), and evidence for genetic influence on scan paths (Constantino et al. 2017; Kennedy et al. 2017), gaze patterns are increasingly being considered a biological marker of underlying neurological function (Murias et al. 2018). Many of the studies exploring group differences have relied on proportion of looking time to predefined areas of interest collapsed across the length of the stimuli presented. While overall looking time to specific areas is clearly important, this measure does not capture one of the key features of the eye-tracking technology: time course data. Though attention to facial features such as the eyes or mouth may be the normative pattern among typically developing (TD) individuals, the content of the video matters a great deal and a "typical fixation" can change moment-to-moment depending on the stimuli (Chawarska et al. 2012). Here we attempt to make use of this time-locked data by (1) determining the average gaze path of TD children when they are viewing a well-established set of eye-tracking stimuli (Chawarska et al. 2012), (2) measuring the extent to which each child with ASD in this study showed a different gaze path from the typical group, and (3) assessing the relationship between distance from a typical scan path and behavioral functioning in children with ASD.

Typical Patterns of Attention to Dynamic Social Stimuli

As illustrated in a recent comprehensive review of the eye tracking stimuli used in ASD research between 2005 and 2019 (Mastergeorge et al. 2020), multiple factors have now been shown to influence how children attend to social stimuli (Chita-Tegmark 2016; Falck-Ytter et al. 2013a, b; Guillon et al. 2014). These include age of the participant (Nakano et al. 2010), IQ of the participant (Sargezeh et al. 2019), motivational factors (Schomaker et al. 2017), content of the scene (Hanley et al. 2013; Rice et al. 2012; Sasson et al. 2008), affective content (Wagner et al. 2013), dynamic nature of the stimuli (Chevallier et al. 2015; Hanley et al. 2013; Speer et al. 2007), and social or dyadic nature of the presentation (Chawarska et al. 2012; Shic et al. 2011, 2013). Given that some of these factors, such as the affective content, can change from moment to moment in a dynamic stimulus, measures of overall proportion of looks to a given area averaged over the length of a trial or stimulus set will likely fail to capture valuable information about social attention in ASD.

Since the early eye-tracking studies of social attention in ASD (Klin et al. 2002b; Pelphrey et al. 2002) many groups have used eye-tracking to identify group differences in proportion of looks to areas of interest within social stimuli (Chawarska et al. 2013, 2015; Cilia et al. 2019; Kennedy et al. 2017; Klin et al. 2002a; Pelphrey et al. 2002; Sasson et al. 2008; Shic et al. 2013; Young et al. 2009). These studies have identified important differences in social attention among individuals with ASD, but have also produced a number of seemingly contradictory results. For example, though early studies of attention to faces suggested that attention to the eyes was typical and focus on the mouth was indicative of ASD (Jones et al. 2008; Klin et al. 2002b), numerous studies since have failed to find these differences (Chawarska et al. 2013; Elsabbagh et al. 2014; Shic et al. 2013) and suggest that attention to the mouth in early development is in fact related to better expressive language abilities in TD children and children with ASD (Elsabbagh et al. 2014; Habayeb et al. 2020; Tenenbaum et al. 2014a, b; Young et al. 2009). Given what we know about the relevance of context for gaze patterns, examination of time-course analysis rather than areas of interest could help resolve these seemingly disparate findings.

Use of Time-Course Data in Attention to Social Stimuli in ASD

To date, only a handful of eye-tracking studies in ASD have shifted from proportion of looking time to exploring time-locked scan paths for dynamic stimuli in relation to diagnostic status and level of impairment (Falck-Ytter et al. 2013a, b; Lönnqvist et al. 2017; Nakano et al. 2010; Wang et al. 2018). The studies that have made use of this granular information have approached the matter from a number of different angles, each with benefits and limitations.

Nakano et al. (2010) used multi-dimensional scaling (MDS) to quantify the typical scan path of children and adults with and without ASD. To do this, they calculated the absolute distance between every pair of gaze points across 104 participants at each time point in their stimuli of video clips from shows intended for young children. Using MDS, the resulting matrix of 103 dimensions was reduced to a two-dimensional plane and the median of the entire MDS distribution was established as the typical scan path. Distance from this typical path was greater for individuals with ASD than for TD controls. Among individuals with ASD, distance from the median path was correlated with symptoms of ASD as measured on the Japanese version of the autism spectrum quotient (Wakabayashi et al. 2006). This novel approach suggests that typical scan paths cluster together and the distance from that path is related to ASD symptomology. A limitation of this work is the difficulty in interpreting the MDS plane. There is no intuitive representation for the X or Y-axis in real space.

Improving on interpretability, Falck-Ytter et al. (2013a, b) identified a single reference point in their stimuli of two children interacting with a toy and calculated the distance from that reference point to each individual child's fixation over the course of their semi-naturalistic stimuli. Distance to the reference distinguished between children with ASD and controls and was associated with communication scores on the Vineland Adaptive Behavior Scales (Sparrow 2011) and verbal IQ on the Wechsler Preschool and Primary Scale of Intelligence (Wechsler 2002). Though this approach significantly improved upon the interpretability issue of Nakano et al. (2010)'s method, it requires that the areas of interest be at consistently different distances from the reference point, thus limiting the type of stimuli one can use with this approach.

Lönnqvist et al. (2017) expanded on these approaches by calculating the distance from each of a number of targets and identifying moments of group differences in distance from each of those targets across the length of their videos. The stimuli had been designed to assess for comprehension of pragmatically complex scenes in adolescents with ASD and the authors identified correlations between distance from specific targets and scores on the verbal tasks administered to assess comprehension.

In recent work, Wang et al. (2018) incorporated the use of high-cohesion time frames as a method for analyzing time course eye-tracking data in ASD. Using stimuli involving a woman engaged in dyadic bids,

sandwich-making, joint attention, and moving-toy conditions, the authors identified points in the video at which the 163 TD toddlers tended to focus on the same area. They then used the median fixation point among the TD group as a "typical" fixation for that high-cohesion time frame. Distance of the fixations of 112 toddlers with ASD and 36 with developmental delay were then calculated from the typical fixations during high-cohesion time frames to determine "typicality" of the participant's scan path. As was true of Nakano et al. and Flack-Ytter's approaches, distance from the typical fixations during high-cohesion time frames distinguished between groups and was related to ASD symptom severity as measured by the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2). While this approach has the advantage of interpretability of typical path and generalizability across stimuli, it too has its limitations. Specifically, this approach precludes analysis in contexts in which scan paths might shift back and forth between two points, as in the joint attention condition. Because a typical path in a joint attention bid shifts between the speaker's face and an object, joint attention bids were excluded from the analysis because no cohesion was observed during these segments. Given the relevance of joint attention for the development of communication and known impairments in joint attention among children with ASD (Charman 2003; Dawson et al. 2004; Mundy et al. 1990), this is a significant limitation.

Drawing on the strengths of each of these efforts to use time-course data in eye-tracking studies of ASD, in the current study, we took a different analytic approach. Using the same stimuli and clinical outcomes employed by Murias et al. (2018) to demonstrate that an area of interest analysis of eye gaze is associated with clinical measures of social communication, we explored parallel relations using time course analysis. We first calculated the mean fixation point of a group of TD participants frame-by-frame over the course of our 180 s video stimuli. We then measured the distance between the fixation of each participant with ASD and the mean TD fixation over the course of the videos. This approach had a number of advantages. First, mean fixation location is intuitively interpretable and allows for interpretable results even in situations when typical gaze might be expected in one of multiple locations (e.g., the face or object in a joint attention bid). Second, mean location of the typical gaze pattern inherently provides a measure of coherence because only when there is coherent typical fixation in a given location on the screen will the typical path veer from the center of the screen (which in the current stimuli fell just below the speaker's collarbone and had no meaningful content). In cases where there is coherent typical gaze in a particular region, the mean fixation shifts to that region and fixations that are far from that location are indicative of

aberrant scan paths. At times without a coherent fixation, the typical gaze is centralized, thereby limiting the potential distance of any other fixation.

In the current study, we used this mean distance from a typical scan path approach to better quantify the extent to which the gaze path of each individual child with ASD differed from the average path shown by the TD children and to determine whether levels of symptoms of ASD were related to distance from the average scan path of the TD children. We segmented the video stimuli based on what was occurring in the scene, isolating scenes with rich social and sensory content because we anticipated that scene content would influence gaze patterns. We hypothesized that greater ASD symptomology would be associated with greater distance from the typical scan path, and that associations with symptomatology would differ depending on scene content.

Methods

All caregivers/legal guardians of participants gave written, informed consent, and the study protocol was approved by the Duke University Health System Institutional Review Board. Methods were carried out in accordance with institutional, state, and federal guidelines and regulations.

Population

ASD Participants

Data were drawn from 176 participants with ASD who participated in a clinical trial. This was part of a single site, prospective, randomized, double-blind, parallel group study of a single intravenous autologous or allogeneic, unrelated cord blood (CB) infusion in children aged 2–7 years. The trial was conducted under IND #15,949. Only data from the baseline visit, which were collected before infusions, were used in this analysis. Clinical diagnosis of ASD was based on the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) (American Psychiatric Association 2013), and established by reliable clinical researchers using the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) (Lord et al. 2012) and the Autism Diagnostic Interview, Revised (ADI-R) (Le Couteur et al. 2003). Additional inclusion criteria included (1) stability on current medications for at least 2 months prior to the infusion, (2) participants and parents/guardians were English speaking, and (3) availability of autologous umbilical cord blood unit or ≥ 4/6 HLA-matched allogeneic unrelated umbilical cord blood unit from the Carolinas Cord Blood Bank. Exclusion criteria included (1) a history of prior cell therapy, (2) use of intravenous immunoglobulin (IVIG) or other anti-inflammatory medications (with the exception of

NSAIDs), (3) known genetic syndrome (e.g., Fragile X), presence of dysmorphic features, pathogenic mutation or copy number variation associated with ASD, and/or other significant medical and/or psychiatric comorbidity, (4) obvious physical dysmorphology, (5) an uncontrolled seizure disorder, (6) significantly impaired renal or liver function, (7) known active CNS infection, evidence of uncontrolled infection, and/or HIV positivity, (8) family unwilling or unable to commit to study-related assessments, and/or (9) clinically significant abnormalities in complete blood count.

Twenty-one subjects with ASD were excluded due to behavioral non-compliance with eye-tracking study procedures, poor calibration, and technical failures. Included participants were 155 children with ASD (122 males, 33 females) between 27 and 95 months of age ($M = 65.65$, SD 19.73). Because the TD sample described below included children who were younger than the participants with ASD, a reduced sample of 93 participants who fell within the age range of the TD comparison sample of 36 to 72 months (72 males, 21 females) ($M = 51.88$ months, SD 11.92) were included in a separate analysis. Furthermore, age was included as a covariate in analyses. Demographics for both the full and age-matched samples are shown in Table 1.

TD Participants

Forty-five children between the ages of 36 and 72 months who did not have a diagnosis or suspected diagnosis of ASD

Table 1 Demographic information for participants

	Full sample ASD (N=155)	Age-matched ASD (N=93)	TD (N=40)
Age (months)			
Mean (SD)	65.65 (19.73)*	51.88 (11.92)	49.95 (10.51)
Sex			
Female	33	21	17
Male	122	72	23
Ethnicity/race			
Minority	53	28	24
Non-minority	102	65	16
Nonverbal DQ			
Mean (SD)	76.47 (21.03)*	73.25 (19.99)*	118.47 (16.02)
ADOS-2 severity score			
Mean (SD)	7.99 (1.65)	7.90 (1.57)	

*Participants with autism spectrum disorder (ASD) in the full sample were significantly older than typically developing (TD) participants, $t(118) = 6.84$, $p < .01$. Ages were not significantly different between age-matched ASD participants and TD participants, $t(83) = .93$, $p = .18$. TD participants had significantly higher Nonverbal IQ than both the full ASD group, $t(105) = 14.34$, $p < .01$, and the age-matched sample, $t(91) = 14.39$, $p < .01$

were recruited from the community and the Duke Center for Autism and Brain Development research registry to participate in a study of preschool age children with and without autism. Children were eligible to be in the TD control group if they had scores on the Strengths and Difficulties Questionnaire (SDQ) that were within the normal range for all scales. The SDQ is a parent-report screening tool for measuring internalizing and externalizing difficulties in children (Goodman 1997, 2001; Goodman et al. 2000). Exclusion criteria for this group included having (1) a biological sibling or parent diagnosed with ASD or developmental delay (DD), (2) a genetic disorder (e.g., Fragile X), (3) vision or hearing problems, (4) a significant motor impairment (e.g., cerebral palsy), (5) chronic or acute medical illness, and (5) a seizure in the last year, a seizure disorder, or being on medication for seizures. Five TD participants were excluded due to non-compliance and calibration failures. Included participants were 40 children without ASD (23 males, 17 females) ($M = 49.95$ months, SD 10.51, range = 36 to 70 months).

Clinical Measures (ASD Participants Only)

Cognitive Assessments

Depending upon the age and verbal ability of the child, one of two measures was used to estimate IQ: the Mullen Scales of Early Learning (Mullen 1995) or the Differential Ability Scales, Second Edition (DAS-II) (Elliott 2007). Given the use of two measures, in lieu of standard scores, a nonverbal developmental quotient (NVDQ) was calculated for each participant as nonverbal age equivalent scores/chronological age.

Vineland Adaptive Behavior Scales-3

The Vineland Adaptive Behavior Scales-3 (VABS-3; Sparrow et al. 2016) is a well-standardized measure with strong reliability and validity which yields an overall composite score, as well as subscale standard scores in the following domains: Socialization, Communication, Daily Living Skills, and Motor Skills. The VABS-3 was administered as a caregiver interview with the primary caregiver. We included the Socialization and Communication subscales in our analyses due to the hypothesized relationship between visual attention and core symptoms of autism.

Aberrant Behavior Checklist-Community (ABC-C)

This parent-completed rating scale was used to measure aberrant behaviors associated with ASD (Aman et al. 1985). This measure yields a composite score as well as subscale scores in the following domains: Irritability, Lethargy/Social Withdrawal, Stereotypy, Hyperactivity, and Inappropriate

speech. Once again, the subscales we explored were those related to ASD symptomatology including Lethargy/Social Withdrawal and Stereotypy.

Autism Diagnostic Interview-Revised

The Autism Diagnostic Interview-Revised (ADI-R; Le Couteur et al. 2003) is a standardized, semi-structured, interview administered by trained clinicians to parents or caregivers of individuals referred for a possible autism spectrum disorder (ASD). The ADI-R includes 93 items in three domains of functioning: communication, reciprocal social interactions, and restricted, repetitive, and stereotyped patterns of behavior.

Pervasive Developmental Disorder Behavior Inventory

The Pervasive Developmental Disorder Behavior Inventory (PDD-BI; Cohen et al. 2003) is a caregiver questionnaire that is designed to measure symptoms within domains affected by ASD. The PDD-BI was standardized with a sample of caregivers and teachers of children with ASD from a range of racial, ethnic, and socioeconomic backgrounds. The PDD-BI was collected from each participant's primary caregiver.

Eye-Tracking Procedures

A Tobii TX300 (Tobii, Danderyd, Sweden) was used to measure corneal reflections and calculate eye-gaze direction using infrared (IR) light-emitting diodes and IR cameras. Data were collected with a sampling rate of 120 Hz in Tobii Studio Version 3.2.2 using standardized room lighting and movie volume for all participants. Participants sat in their parents' laps and watched a 3-min video of an actor engaged in child-directed speech while surrounded by four distracter toys (Chawarska et al. 2013). A research assistant stood next to participants and monitored the track status throughout the experiment; when subjects moved out of the trackable range they were re-positioned immediately. Data were analyzed offline. Trials in which at least one eye was tracked and the gaze fell on the media frame were considered valid trials and were included in the analysis. The media region was defined as the entire video frame and did not extend to the edges of the screen.

Stimuli

The video (Chawarska et al. 2012) displays a woman seated behind a table and flanked by a toy in each of the four corners of the screen. Four distinct conditions are included in the 3-min-long stimuli. At the start, the speaker greets the child using child-directed speech while looking directly into the camera. This *Dyadic Bid* condition is repeated throughout the stimuli an additional 10 times for a total duration of 69 s. In the *Sandwich* condition, the speaker directs her attention to the table while making a sandwich. This action is repeated once more for a total of two episodes and 63 s without direct gaze or speech. Though there is an element of joint attention at the start of the sandwich making segment, the significant movement of the speaker's hands and lack of gaze towards the child during the sandwich segment makes this less social than other portions of the video. In the four episodes of the *Joint Attention* condition (total of 30 s), the speaker looks into the camera, speaks briefly to the child, and then shifts attention to one of the four toys. She then maintains her gaze towards that toy for 4 s. In the *Moving Toys* condition, the speaker again looks into the camera and then shifts her gaze back to each of the four toys in succession (one toy per episode). In contrast with the joint attention condition, in the moving toys condition, while the speaker directs her gaze towards one toy, the toy on the opposite side of the screen is activated and begins to move and make noise. The speaker ignores this movement and maintains her gaze on the original toy on which she was fixating (total 27 s over four episodes). Rather than being presented as individual stimuli separated by attention getters, these four conditions are interleaved throughout the video without breaks to provide contextual cues more akin to naturalistic stimuli for shifting fixations.

Analytic Strategy

Statistical analyses were performed using RStudio version 3.5.1 (2018-07-02) (RStudio, Boston, MA, USA). Our first goal was to establish a typical scan path for individuals watching the 180 s stimuli. This was completed within the TD and ASD groups and plotted as mean fixations by group overlain onto the video stimuli. This video can be viewed in supplemental materials. A qualitative description of these scan paths is provided in results below.

Results

Descriptive Results of Mean Location

For descriptive purposes, dynamic mean fixations for both groups as well as standard deviations in the X and Y axes are shown in the video available in the supplemental materials. As the video illustrates, mean fixations in both groups followed generally predictable and similar patterns. On average, both TD and ASD participants focused on the speaker's mouth when she was speaking during dyadic bids and on her hands and the items she manipulated when she was making a sandwich. During joint attention segments, mean fixations fell midway between the speaker and the toys she referenced.

When the toys began to move, both groups shifted their attention to the moving toy. Very shortly after the onset of toy movement, gaze shifted between the speaker and the toy, resulting in mid-point fixations as in the joint attention conditions. Overall, these patterns held true for both groups.

Distance from Typical Scan Path Among ASD Participants

To determine the distance of each participant from the typical scan path, we calculated the mean of the typical scan path for each of the 21,600 frames collected over the 180 s video (120 frames/s) for the 40 TD participants and then determined the distance between that mean typical fixation for each of those 21,600 frames for each of the 155 ASD participants (and for the 40 TD participants for comparison). Mean fixation location was determined by calculating the mean X and mean Y values (coordinates provided by Tobii) at each frame across the TD participants. Distance was calculated as the Euclidian distance from the TD mean fixation for each participant's fixation for each frame. Mean fixations by group are shown in Fig. 1. The groups differed in mean distance from the typical scan path, $F(1,193)=36.13$, $p<.001$; TD: $M=181.27$, SD 31.64; ASD: $M=228.73$, SD 47.24. There was not a significant correlation between age and distance in the TD group ($r(38)=-.06$, $p=.73$). Age and NVDQ were both independently negatively correlated with mean distance in the ASD group (age: $r(155)=-.35$, $p<.001$; NVDQ: $r(155)=-.29$, $p<.001$). Mean distance from the typical scan path decreased as age and NVDQ increased.

Using the distance for each ASD participant at each time point, we then calculated the effect of context on distance by averaging each participant's distance within each of the four segment types: Dyadic Bid, Joint Attention, Sandwich, and Moving Toys (see Fig. 2). Analysis of variance revealed significant effects of segment type on distance from the typical scan path, $F(3,616)=47.17$, $p<.01$. Post-hoc comparisons revealed no significant differences in distance between the Joint Attention and Dyadic Bid conditions ($p=1.0$) but significant differences between Dyadic Bid and Sandwich segments ($p<.01$) and Dyadic Bid and Moving Toy segments ($p<.01$). Joint Attention segments were also significantly different from Sandwich ($p<.01$) and Moving Toy segments ($p=.01$). Finally, Sandwich and Moving Toy segments were significantly different from each other ($p<.01$). Given the convergence on moving objects (e.g., the speaker's hands), it was not surprising that the shortest distances were observed during the Sandwich condition. We were, however, surprised to find that the Moving Toy segments (which also included movement) had the greatest average distances between ASD fixations and the mean TD fixations. To explore the moment-by-moment changes in the video content that contributed to those results, we next explored frame-by-frame analyses of the distances between the TD mean and ASD fixations.

To do this, we plotted the mean distance across the 155 participants with ASD from the typical scan path (see Fig. 3). Figure 3 demonstrates a number of points at which the mean distance of ASD participants' fixations from the TD scan path diverges (peaks) as well as times when it converges (valleys). These brief instances of convergent and divergent fixations highlight the importance of time course data for understanding how context affects social attention.

Taking further advantage of this granular approach to explore how context affects the points at which scan paths converge and diverge in these groups, we identified frames

Fig. 1 Mean distance from typical scan path by group. Violin plots show the mean distance from the mean fixation for the typically developing (TD) participants. Distance was calculated as Euclidean distance from the mean X and mean Y locations for the 40 TD participants. Blue dots show the distance for each TD participant. Red dots show the mean distance for each participant with autism spectrum disorder (ASD). Violin plots show density. Because distance was calculated from the TD group data, mean TD distance was expected to be shorter than mean ASD distance, $F(1,193)=36.13$, $p<.001$. Px pixels
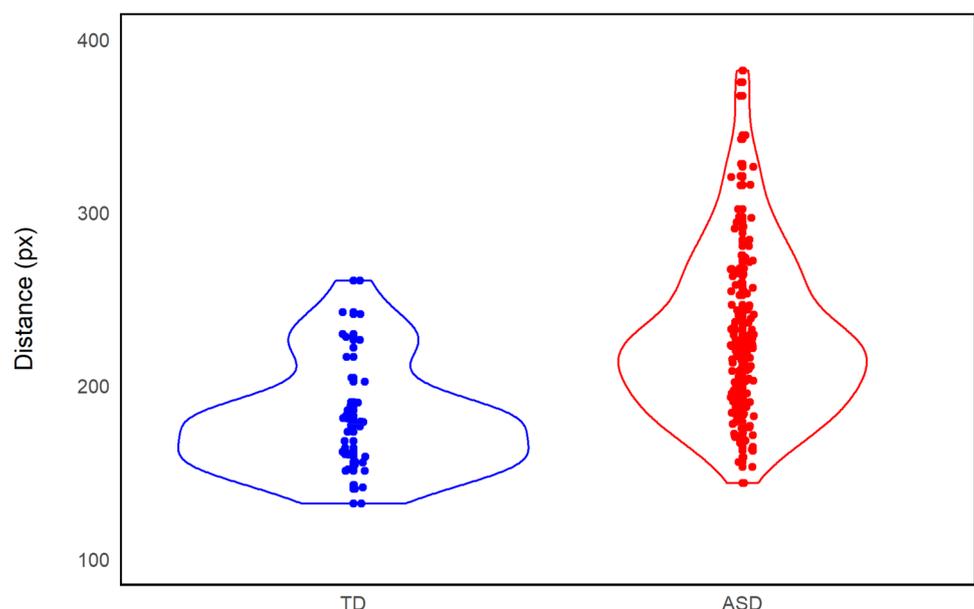
Fig. 2 Mean distance from typical scan path for participants with ASD by trial segment. Figure shows mean distance for participants with autism spectrum disorder (ASD) from the mean scan path of typically developing (TD) participants by video segment. Analysis of variance showed a significant effect of segment, $F(3,616) = 47.17$, $p < .01$. With the exception of the contrast between the Dyadic Bid and Joint Attention segments, all post-hoc pairwise comparisons were significant, $p \leq .01$. *Px* pixels
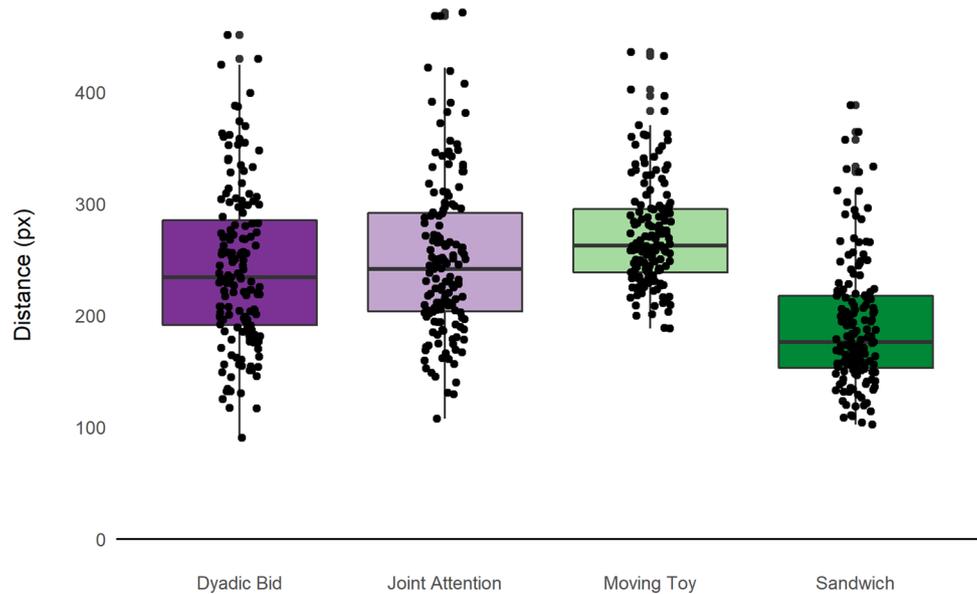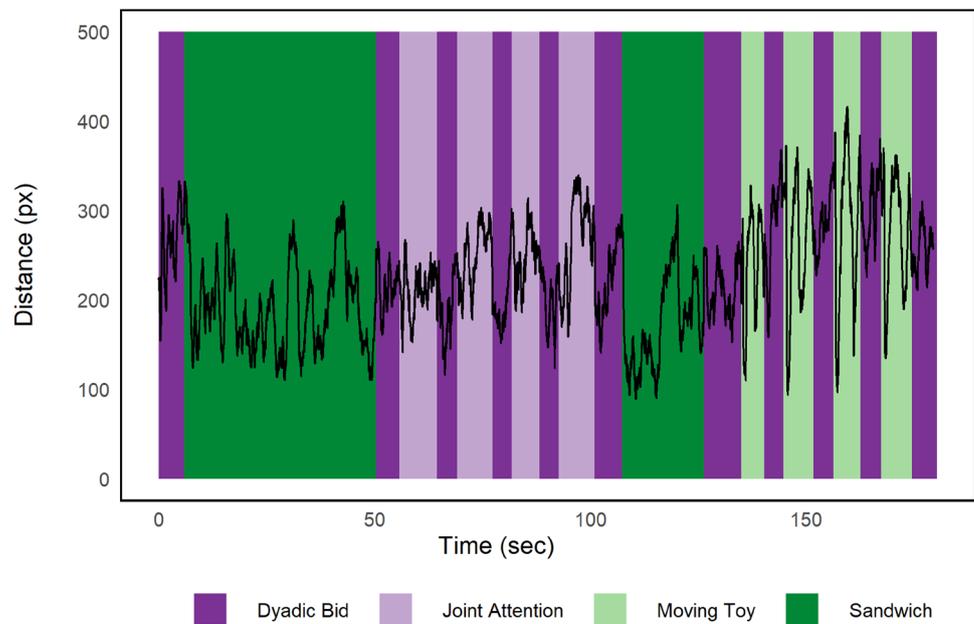

Fig. 3 Mean distance from typical scan path for participants with ASD over the course of the trial. Mean distance of participants with autism spectrum disorder (ASD) from the scan path for typically developing (TD) participants averaged across ASD participants over the 180 s video. Colors show trial segment. *Px* pixels

when the mean distance of ASD fixations from the TD scan path differed from the average mean distance for the ASD group by two or more standard deviations. This resulted in 714 frames (approximately 3% of the full stimulus). We next examined these frames for continuity of content (i.e., speaker was speaking, speaker shifted her gaze, speaker manipulated a specific object) and selected one single frame from each of these continuous events for analysis. In total, there were 28 separate events that resulted in mean distance greater than or less than two standard deviations from the mean distance of ASD fixations to the mean TD scan path. These included 14 valleys (ASD fixations were closest to the TD mean) and 14

peaks (ASD gaze was furthest from TD mean). These events ranged from 1 to 55 frames in length ($M = 20.85$, SD 19.37). Figure 4 illustrates the descriptions of each of these 28 peaks and valleys and includes analysis of the differences in the X and Y coordinates for the selected frame for each of the events. Differences were assessed separately with independent samples T-tests in the X and Y planes. As illustrated in Fig. 4, the shortest mean distances between ASD fixations and the mean TD fixations were nearly always in the context of salient movement absent social cues (e.g., onset of toy movement or speaker manipulating the sandwich). The longest mean distances between ASD fixations and the mean
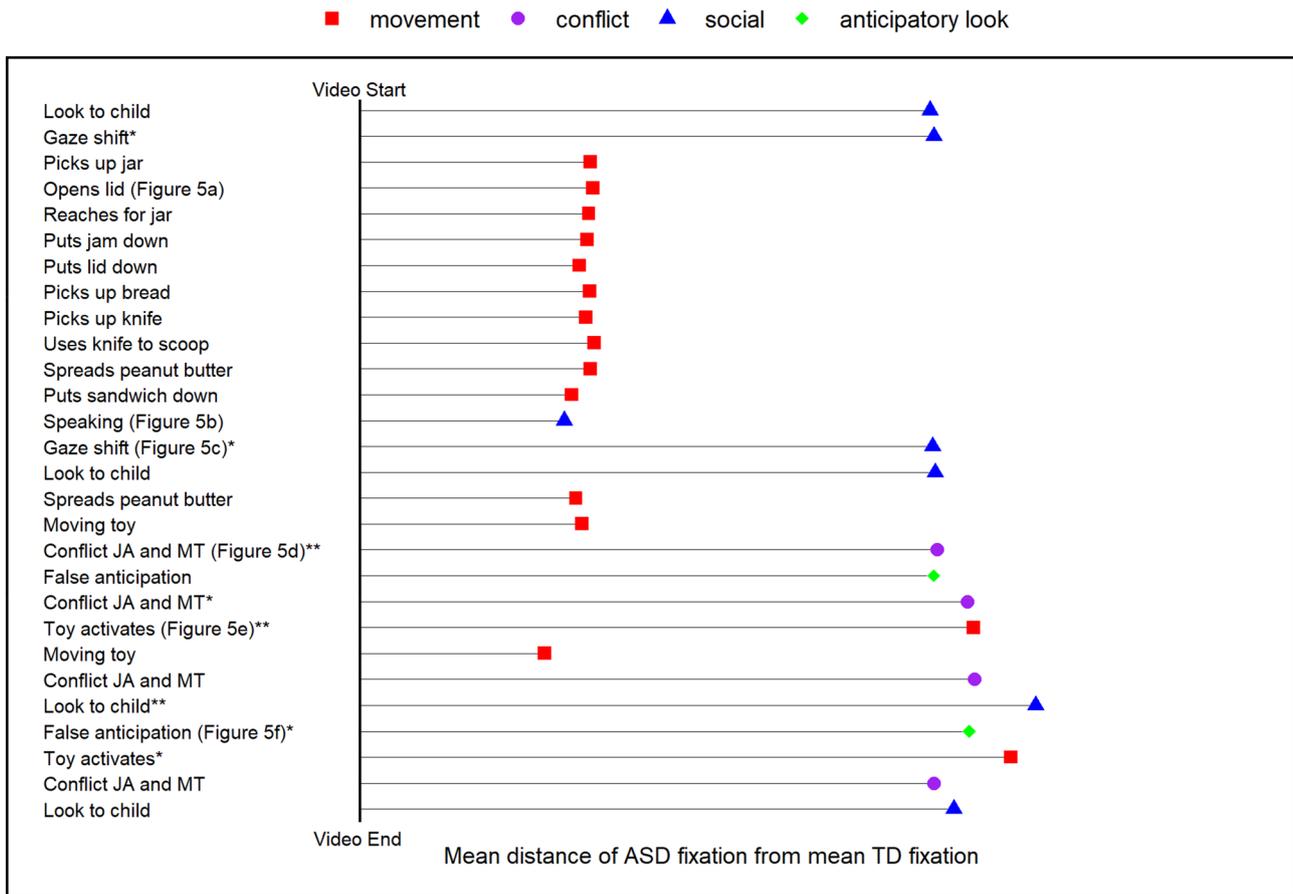
Fig. 4 Convergence and divergence by scene content. Shows the 28 instances when the mean distance of fixation for participants with autism spectrum disorder (ASD) from mean fixation of typically developing participants (TD) was greater than or less than two standard deviations (SD) from the mean of those distances. T-tests in X and Y dimensions were used to assess significance, *$p < .05$, **$p < .01$. *JA* joint attention, *MT* moving toys

TD fixations tended to be more social in nature (e.g., dyadic bids in which the speaker looked to the child or shifted gaze to an object), or were related to conflict between social and movement cues.

More specifically, of the 14 events in which the ASD fixations were closest to the TD path, 13 were salient movements (see example screen capture in Fig. 5a). These included overt actions on the part of the speaker (e.g., picking up a jar), or toy movement in the corners of the screen. Only one of these convergent events had strong social elements. This was one instance when the speaker was speaking, which also involves movement of the speaker's mouth (Fig. 5b).

The 14 events in which the ASD participants were furthest from the mean TD fixations included clusters of events that patterned in similar ways. Four of these instances with greatest distance from the mean TD fixations involved the speaker looking directly into the camera without speaking (e.g., at the start of the Dyadic Bid and Joint Attention conditions). During these moments, both groups were somewhat scattered in their fixations resulting in mean fixations

that were closer to the center of the screen but with TD gaze closer to the speaker's face and ASD fixations closer to the center of the screen. This suggests less of an attentional pull towards the face among the ASD participants at these moments of direct gaze. On two occasions when the speaker shifted her gaze to an object (Joint Attention bids), TD participants were more likely to look at the object than ASD participants resulting in a mean fixation that was closer to the target of the speaker's gaze (Fig. 5c).

Four of the divergent moments are best described as conflicting cues due to the overt conflict between the speaker's gaze shift to one object and movement of a different object on the opposite side of the screen. During these moments of conflict, TD participants were again more likely to be pulled to the target of the speaker's gaze, resulting in mean fixation closer to the speaker's head (midpoint between the two target toys) while ASD mean fixation reflected more looks to the moving toy (Fig. 5d). On two occasions in which ASD gaze was furthest from the mean TD gaze, one of the toys had just activated. In these instances, ASD participants
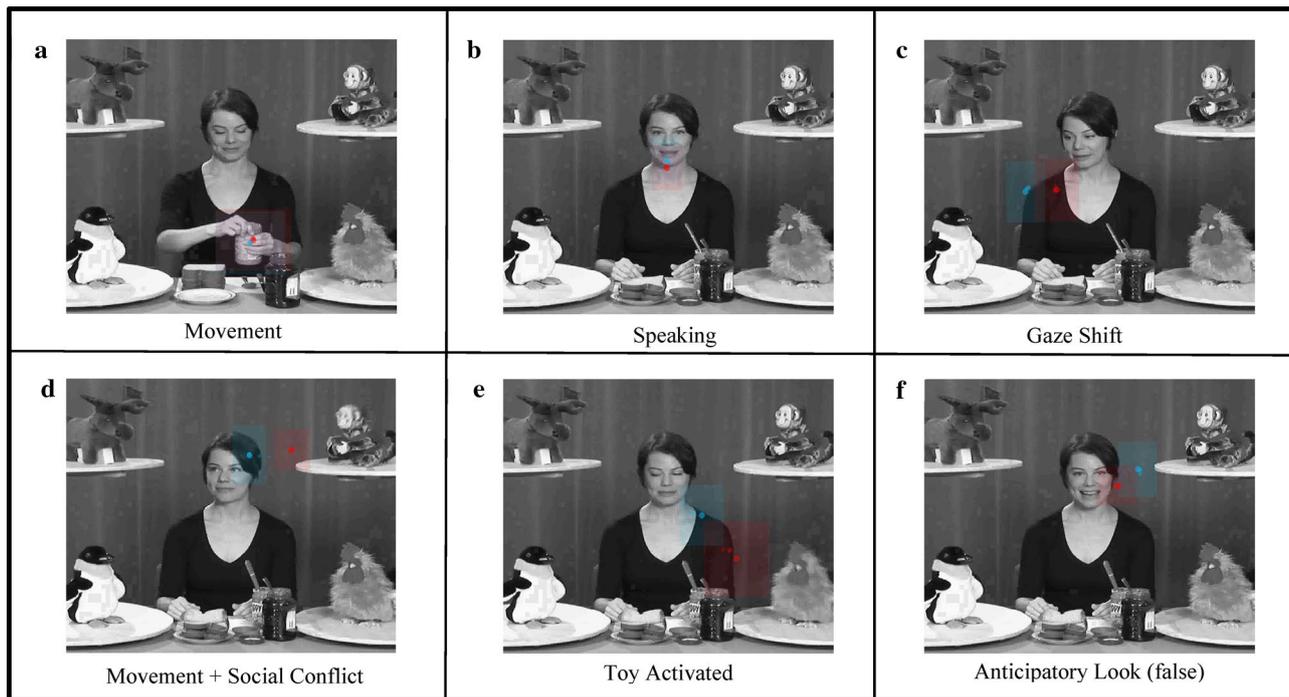
Fig. 5 Sample scene captures. Shortest (**a**, **b**) and longest distances (**c**–**f**) between mean fixation for typically developing (TD; blue) and autistic participants (ASD; red). Boxes show standard deviation. **a** Speaker opens jar. Both groups fixate her hands. **b** Speaker is speaking. Both groups fixate her mouth. **c** Speaker looks to penguin. Mean TD fixation is closer to penguin indicating more gaze following among TD participants. **d** Monkey activates but speaker looks at moose. Mean TD fixation is closer to face (midpoint between the conflicting cues). ASD fixation is closer to the moving monkey. **e** Chicken activates. ASD participants are faster to respond to the movement. **f** Monkey recently activated but has since stopped moving. Mean TD fixation continues to move towards the monkey as if anticipating further movement

were faster to respond to this activation, resulting in fixations among the ASD participants that were closer to those active objects (Fig. 5e).

Finally, on two separate occasions, after a toy had stopped moving, TD participants (but not ASD participants to the same degree) continued to shift their attention to those toys as if anticipating further movement in that area of the screen (Fig. 5f).

Correlations with Clinical Measures

To determine whether distance from a typical scan path is related to autism symptomology, we next examined correlations between the mean distance overall and within each of the four segment types (Dyadic, Joint Attention, Sandwich, and Moving Toys) for each of the participants. As described above, adaptive behavioral skills associated with the core features of autism (communication and socialization on the VABS-3) and autism symptomology were assessed with clinical measures including ADOS-2, ABC, PDDBI and ADI-R. Because this approach was exploratory in nature, we did not control for multiple comparisons and these results must therefore be interpreted with caution. As in Murias

et al. (2018), we did control for both age of the participant and nonverbal IQ (in this case, NVDQ as described above). Results of the pairwise partial correlations are shown in Table 2. All available data were included in these analyses.

With respect to adaptive behavior skills as measured by VABS-3, Communication scores were significantly negatively correlated with distance overall ($R^2 = -.19$, $p = .02$), distance during dyadic segments ($R^2 = -.16$, $p = .05$), and distance during moving toy segments ($R^2 = -.24$, $p = .002$) after controlling for age and NVDQ. Socialization scores from the VABS-3 were not correlated with distance after controlling for age and NVDQ. This suggests that ASD participants with greater distance in scan paths from the typical participants showed lower levels of communication skills. This relation was specific to the Dyadic and Moving Toy segments.

ASD symptomatology was assessed using ADOS-2, ABC, ADI-R and PDDBI. After controlling for age and NVDQ, mean distance during the Dyadic ($R^2 = .20$, $p = .01$) and Joint Attention segments ($R^2 = .18$, $p = .03$) was correlated with calibrated severity scores on the ADOS-2. This suggests that greater distance from the typical scan path during the more social segments of the

Table 2 Partial correlations controlling for age and non-verbal developmental quotient (NVDQ) for mean distance overall and mean distance within each segment type for the full sample of participants with autism spectrum disorder (ASD) ($n = 155$). $*p < .05$, $**p < .01$

	Overall distance	Dyadic	Joint attention	Sandwich	Moving toys
VABS-3 communication	**− .19***	**− .16***	− .10	− .11	**− .24****
VABS-3 socialization	.09	− .11	− .12	.01	− .05
ADOS-2 calibrated severity scores	.14	**.20***	**.18***	.01	− .08
PDD-BI expressive social communication composite	**− .23****	**− .22****	− .15	− .14	− .10
PDD-BI autism composite	.14	.07	.12	.14	.01
ABC lethargy/social withdrawal	**.33****	**.26****	**.34****	**.20***	**.19***
ABC stereotypy	**.18***	.11	**.22****	.13	.05
ADI-R reciprocal social interaction	**.17***	.11	**.18***	.13	.10
ADI-R verbal communication	− .08	− .10	.02	− .09	.03
ADI-R nonverbal communication	− .04	− .10	.13	.03	− .16
ADI-R restricted-repetitive behaviors	− .01	.00	− .11	.03	.03

Given lack of correction for multiple comparisons, all results should be interpreted with caution

stimuli specifically (Joint Attention and Dyadic Bids) was related to greater ASD symptomatology as measured by the ADOS-2.

PDD-BI expressive social communication was negatively correlated with distance from the typical scan path overall ($R^2 = − .23$, $p = .005$) and during the Dyadic Bids ($R^2 = − .22$, $p = .005$). This suggests that children who were further from the typical scan path had more symptoms associated with ASD on these communication scales.

On the ABC, the Lethargy/Social Withdrawal subscale was significantly correlated with overall distance ($R^2 = .33$, $p < .001$) and distance within each of the four segments [Dyadic ($R^2 = .26$, $p = .001$), Joint attention ($R^2 = .34$, $p < .001$), Sandwich ($R^2 = .20$, $p = .02$), and Moving Toys ($R^2 = .19$, $p = .02$)]. Stereotypy and Reciprocal Social Interaction scores from the ABC were also correlated with overall distance ($R^2 = .18$, $p = .03$) and distance during Joint Attention segments ($R^2 = .22$, $p = .007$) after controlling for age and NVDQ. The ABC Lethargy/ Social Withdrawal subscale includes the following survey items: listless, sluggish, seeks isolation, preoccupied, withdrawn, fixed facial expression(s), sits and watches others, resists physical contact, isolates self, sits/stands in one position, unresponsive to structured activity, difficult to reach/contact, prefers to be alone, communicates without words/gestures, inactive, responds negatively to affection, and shows few social reactions. This suggests that children with less typical scan paths were also more likely to display these behaviors. Finally, on the ADI-R, reciprocal social interaction was correlated with both overall distance ($R^2 = .17$, $p = .03$) and distance during Joint Attention segments ($R^2 = .18$, $p = .02$) after controlling for age and NVDQ.

Comparison with Clinical Measures for Age-Matched Participants

Because the children with ASD were significantly older than the TD comparison sample and distance was significantly correlated with age, we reran the correlations using a subset of the ASD sample that were age-matched to the TD sample. The following partial correlations between distance and clinical measures remained significant after controlling for NVDQ and age. Communication scores on the VABS-3 remained significantly negatively correlated with distance on the Moving Toys segments ($R^2 = − .23$, $p = .03$), suggesting that greater distance was associated with lower communication abilities. Calibrated severity scores on the ADOS-2 were again significantly positively correlated with distance from the typical scan path during the Dyadic segments ($R^2 = .24$, $p = .02$) suggesting that overall autism severity was associated with greater distances. Expressive social communication, as measured by the PDD-BI, was significantly negatively correlated with overall distance ($R^2 = − .21$, $p = .04$) suggesting that lower communication ability is associated with greater distance from the typical scan path. Finally, Lethargy/Social Withdrawal on the ABC also remained significantly positively correlated with distance overall ($R^2 = .30$, $p = .003$) and on the Dyadic ($R^2 = .23$, $p = .03$), Joint Attention ($R^2 = .30$, $p = .004$) and Sandwich segments ($R^2 = .23$, $p = .03$) (see Table 3).

Discussion

In the current study, we described the mean scan path of a group of 40 young typically-developing children while watching a video of a woman engaging in social and

Table 3 Partial correlations controlling for age and non-verbal developmental quotient (NVDQ) for mean distance overall and mean distance within each segment type for age-matched sample ($n = 93$). $*p < .05$, $**p < .01$

	Overall distance	Dyadic	Joint attention	Sandwich	Moving toys
VABS-3 communication	− .14	− .11	− .07	− .08	**− .23***
VABS-3 socialization	− .07	− .09	− .14	.00	− .03
ADOS-2 calibrated severity scores	.20[+]	**.24***	.20	.07	− .06
PDD-BI expressive social communication composite	**− .21***	− .19	− .14	− .14	− .10
PDD-BI autism composite	.06	.00	.03	.11	− .08
ABC lethargy/social withdrawal	**.30****	**.23***	**.30****	**.23***	.08
ABC stereotypy	.17	.12	.19	.11	.10
ADI-R reciprocal social interaction	.13	.10	.17	.11	.04
ADI-R verbal communication	− .16	− .12	− .03	− .20	.06
ADI-R nonverbal communication	− .06	− .14	.08	.04	− .18
ADI-R restricted-repetitive behaviors	− .07	− .08	− .17	.03	− .02

Given lack of correction for multiple comparisons, all results should be interpreted with caution

nonsocial behaviors. We then calculated the distance for each of 155 participants with ASD and for the TD participants from mean scan path of the typical group over the 180-s-long video. Overall, distance was lowest during the sandwich making segments of the video when most participants tended to cluster attention on the speaker's hands and the objects she was manipulating. Interestingly, though one might expect from watching the video (see supplemental materials) that segments involving moving toys would show the least distance given the convergence on the salient movement, the Moving Toy segments in fact had the greatest distance from the typical scan path for ASD participants.

Examination of the peaks and valleys of distance from the typical scan path among ASD participants revealed that periods of convergence (valleys) were nearly always associated with salient movement. Children with ASD, like their TD peers, fixated predominantly on the objects when the speaker manipulated them and on the moving toys. Peaks in the mean distance from the typical path revealed that unlike their TD peers, children with ASD were less likely to follow the speaker's gaze to a target object, particularly when a toy was moving in a different part of the screen. This is consistent with much evidence in the literature suggesting limited response to joint attention in ASD (e.g., Dawson et al. 2002, 2004; Sullivan et al. 2007). In contrast to limited looks to the target object in response to joint attention gaze shifts, ASD participants were faster to respond to the activation of the toy and slower to return to the speaker's face after it was activated. In general, the finding that ASD participants are drawn to salient movement and audiovisual synchrony is consistent with previous work (e.g., Klin et al. 2009).

Interestingly, ASD participants were also less likely to (falsely) anticipate movement of a previously activated toy than TD participants. This result provides some evidence

for recent suggestions that autism may be related to altered patterns of prediction (Sinha et al. 2014). Though in reality, the toy in question did not activate again, it is interesting that as a group, TD participants were more likely to anticipate that it would than their ASD peers.

To assess whether differences in scan paths were related to variations in clinical presentation for the children with ASD, mean distances from the typical scan path for each participant with ASD was correlated with clinical assessments of adaptive behavior and autism symptomology. We found that greater mean distance from the typical scan path among ASD participants related to several aspects of behavior, including lower communication abilities as measured by VABS-3, higher overall severity scores measured by ADOS-2, Lethargy/Social withdrawal, Stereotypies and Total Scores from the ABC, Reciprocal Social Interaction from ADI-R, and Expressive and Receptive Social composite scores from PDD-BI. These results lend support to previous findings showing relations between eye-tracking data and behavioral measures including autism symptoms as assessed by the Autism Spectrum Quotient (Nakano et al., 2010), social and communication abilities as assessed using VABS-3, BASC, and PDD-BI (Murias et al., 2018), verbal abilities as assessed by WPPSI (Falck-Ytter et al. 2013a, b), and social affect as measured by ADOS-2 (Wang et al. 2018).

Also of interest was the relation between mean distance from the typical scan path and communication scores from VABS-3, which was driven in part by mean distance during the moving toy segments of the video. This is likely because the moving toy conditions were in fact contrasts between joint attention bids and salient toy movement.

Our results add to a large body of literature suggesting that attention to social information including gaze following

and attention to faces is likely to be important for communication and social skills in individuals with ASD. When one is focused largely on the non-social movement in one's environment and missing the social cues and information available in the scene, it makes sense that these social cues would be less interesting, more difficult to interpret, and thus less informative than they are for individuals without ASD. This limited intake of socially relevant information may in turn lead to impairments in social communication skills (Arunachalam and Luyster 2015; Tenenbaum et al. 2017).

In addition to replicating previous findings of relations between eye-tracking and behavioral measures and contributing novel results for future exploration in this domain, the current study adds to the extant literature by providing a generalizable and intuitive approach to calculating atypical scan path to dynamic stimuli. This approach takes into account the variation in what constitutes a typical scan path over time. It also provides an inherent measure of coherence by using mean fixations which tend to converge in the center of the screen when there is a lack of coherent gaze among the typical participants. Finally, and importantly, this approach captures variability in scan paths even when there is not one single coherent point of gaze, as in joint attention bids.

There are a number of limitations to the approach used in this study. First, though it is our aim to incorporate time-course data at the frame level, in comparing distance to behavioral measures, we did collapse the distance measure over the course of the trial and segments. Second, the stimuli used in this study did not to contain much relevant information in the center of the screen, thus allowing for default means in a region of little interest and creating a convenient proxy for coherence in gaze patterns. This limits potential use of this approach to stimuli without areas of interest in the center of the screen. Third, although the children with ASD versus TD had overlapping age ranges, the mean age for the two groups differed with the ASD participants being roughly one year older. Thus, age was used as a covariate in our analyses. When we repeated the correlational analyses using age-matched groups, significant correlations between scan path distance from the typical path and behavior remained. Given the exploratory nature of these analyses and the multiple comparisons without correction, replication will be critical to confirm these findings.

As the field moves towards early screening for ASD using automated measures that can be assessed at a large scale, eye-tracking is likely to gain further appeal. The current approach provides a generalizable analytic method that lends itself well to big data and automated screening that may contribute in the future to the utility of these methods for identifying children at risk for ASD.

Compliance with Ethical Standards

Conflict of interest Dr. Dawson is on the Scientific Advisory Boards of Janssen Research and Development, Akili, Inc, LabCorp, Inc, Roche Pharmaceutical Company, and Tris Pharma, and is a consultant to Apple, Gerson Lehrman Group, Guidepoint, Inc, Axial Ventures, Teva Pharmaceutical, and is CEO of DASIO, LLC. Dr. Dawson has received book royalties from Guilford Press, Oxford University Press, Springer Nature Press. In addition, Dr. Dawson has the following patent applications: 1802952, 1802942, 15141391, and 16493754. Dawson and Carpenter helped develop technology for early screening for ASD. The technology has been licensed and Dawson, Carpenter, and Duke University have benefited financially. Dr. Howard reports personal fees from Roche.

References

Aman, M. G., Singh, N. N., Stewart, A. W., & Field, C. J. (1985). The aberrant behavior checklist: A behavior rating scale for the assessment of treatment effects. *American Journal Of Mental Deficiency, 89,* 485.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Arunachalam, S., & Luyster, R. J. (2015). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. *Autism Research, 9,* 810.

Charman, T. (2003). Why is joint attention a pivotal skill in autism? *Philosophical Transactions of The Royal Society of London Series B Biological Sciences, 358*(1430), 315–324. https://doi.org/10.1098/Rstb.2002.1199.

Chawarska, K., Macari, S., & Shic, F. (2012). Context modulates attention to social scenes in toddlers with autism. *Journal Of Child Psychology And Psychiatry, 53*(8), 903–913. https://doi.org/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*(3), 195–203. https://doi.org/10.1016/J.Biopsych.2012.11.022.

Chawarska, K., Ye, S., Shic, F., & Chen, L. (2015). Multilevel differences in spontaneous social attention in toddlers with autism spectrum disorder. *Child Development, 87,* 543.

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*(5), 620–628.

Chita-Tegmark, M. (2016). Social attention in asd: A review and meta-analysis of eye-tracking studies. *Research in Developmental Disabilities, 48,* 79–93.

Cilia, F., Aubry, A., Le Driant, B., Bourdin, B., & Vandromme, L. (2019). Visual exploration of dynamic or static joint attention bids in children with autism syndrome disorder. *Frontiers in Psychology, 10,* 2187.

Cohen, I. L., Schmidt-Lackner, S., Romanczyk, R., & Sudhalter, V. (2003). The pdd behavior inventory: A rating scale for assessing response to intervention in children with pervasive developmental disorder. *Journal of Autism and Developmental Disorders, 33*(1), 31–45.

Constantino, J. N., Kennon-Mcgill, S., Weichselbaum, C., Marrus, N., Haider, A., Glowinski, A. L., et al. (2017). Infant viewing of social

scenes is under genetic control and is atypical in autism. *Nature, 547*(7663), 340.

Dawson, G., Munson, J., Estes, A., Osterling, J., Mcpartland, J., Toth, K., et al. (2002). Neurocognitive function and joint attention ability in young children with autism spectrum disorder versus developmental delay. *Child Development, 73*(2), 345–358. https://doi.org/10.1111/1467-8624.00411.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology, 40*(2), 271–283. https://doi.org/10.1037/0012-1649.40.2.271.

Elliott, C. (2007). *Differential ability scales* (Vol. 22, pp. 128–132). San Antonio, TX: Harcourt Assessment.

Elsabbagh, M., Bedford, R., Senju, A., Charman, T., Pickles, A., Johnson, M. H., & Team B. (2014). What you see is what you get: Contextual modulation of face scanning in typical and atypical development. *Social Cognitive and Affective Neuroscience, 9*(4), 538–543.

Falck-Ytter, T., Bölte, S., & Gredebäck, G. (2013). Eye tracking in early autism research. *Journal of Neurodevelopmental Disorders, 5*(1), 28.

Falck-Ytter, T., Von Hofsten, C., Gillberg, C., & Fernell, E. (2013). Visualization and analysis of eye movement data from children with typical and atypical development. *Journal of Autism and Developmental Disorders, 43*(10), 2249–2258.

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581–586.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of The American Academy of Child & Adolescent Psychiatry, 40*(11), 1337–1345.

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the strengths and difficulties questionnaire (Sdq) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry, 177*(6), 534–539.

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.

Habayeb, S., Tsang, T., Saulnier, C., Klaiman, C., Jones, W., Klin, A., & Edwards, L. A. (2020). Visual traces of language acquisition in toddlers with autism spectrum disorder during the second year of life. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-020-04730-x.

Hanley, M., Mcphillips, M., Mulhern, G., & Riby, D. M. (2013). Spontaneous attention to faces in asperger syndrome using ecologically valid static stimuli. *Autism, 17*(6), 754–761.

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*(8), 946–954. https://doi.org/10.1001/Archpsyc.65.8.946.

Kennedy, D. P., D'onofrio, B. M., Quinn, P. D., Bölte, S., Lichtenstein, P., & Falck-Ytter, T. (2017). Genetic influence on eye movements to complex scenes at short timescales. *Current Biology, 27*(22), 3554–3560.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002a). Defining and quantifying the social phenotype in autism. *American Journal of Psychiatry, 159*(6), 895–908.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002b). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59*(9), 809–816. https://doi.org/10.1001/Archpsyc.59.9.809.

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

Le Couteur, A., Lord, C., & Rutter, M. (2003). *The autism diagnostic interview-revised (Adi-R)* (pp. 659–685). Los Angeles: Western Psychological Services.

Lönnqvist, L., Loukusa, S., Hurtig, T., Mäkinen, L., Siipo, A., Väyrynen, E., et al. (2017). How young adults with autism spectrum disorder watch and interpret pragmatically complex scenes. *The Quarterly Journal of Experimental Psychology, 70*(11), 2331–2346.

Lord, C., Rutter, M., Dilavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule–2nd edition (Ados-2)*. Los Angeles: Western Psychological Corporation.

Mastergeorge, A. M., Kahathuduwa, C., & Blume, J. (2020). Eye-tracking in infants and young children at risk for autism spectrum disorder: A systematic review of visual stimuli in experimental paradigms. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-020-04731-w.

Mullen, E. M. (1995). *Mullen scales of early learning*. Circle Pines, MN: American Guidance Service.

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*(1), 115–128.

Murias, M., Major, S., Davlantis, K., Franz, L., Harris, A., Rardin, B., et al. (2018). Validation of eye-tracking measures of social attention as a potential biomarker for autism clinical trials. *Autism Research, 11*(1), 166–174.

Nakano, T., Tanaka, K., Endo, Y., Yamane, Y., Yamamoto, T., Nakano, Y., et al. (2010). Atypical gaze patterns in children and adults with autism spectrum disorders dissociated from developmental changes in gaze behaviour. *Proceedings of The Royal Society B, 277*(1696), 2935–2943. https://doi.org/10.1098/Rspb.2010.0587.

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism & Developmental Disorders, 32*(4), 249–261.

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*(3), 238–248.

Sargezeh, B. A., Ayatollahi, A., & Daliri, M. R. (2019). Investigation of eye movement pattern parameters of individuals with different fluid intelligence. *Experimental Brain Research, 237*(1), 15–28.

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 viewing of complex social and nonsocial picture arrays. *Autism Research, 1*(1), 31–42. https://doi.org/10.1002/Aur.4.

Schomaker, J., Walper, D., Wittmann, B. C., & Einhäuser, W. (2017). Attention in natural scenes: Affective-motivational factors guide gaze independently of visual salience. *Vision Research, 133,* 161–175.

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.

Shic, F., Macari, S., & Chawarska, K. (2013). Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. *Biological Psychiatry, 75,* 231.

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., et al. (2014). Autism as a disorder of prediction. *Proceedings of The National Academy of Sciences, 111*(42), 15220–15225.

Sparrow, S. S. (2011). Vineland adaptive behavior scales. In J. Kreutzer, J. Deluca, & B. Caplan (Eds.), *Encyclopedia of clinical neuropsychology* (pp. 2618–2621). New York: Springer.

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland-3: Vineland adaptive behavior scales*. San Antonio: Pearson.

Speer, L. L., Cook, A. E., Mcmahon, W. M., & Clark, E. (2007). Face processing in children with autism: Effects of stimulus contents and type. *Autism, 11*(3), 265–277. https://doi.org/10.1177/1362361307076925.

Sullivan, M., Finelli, J., Marvin, A., Garrett-Mayer, E., Bauman, M., & Landa, R. (2007). Response to joint attention in toddlers at risk for autism spectrum disorder: A prospective study. *Journal of Autism and Developmental Disorders, 37*(1), 37–48.

Tenenbaum, E. J., Amso, D., Abar, B., & Sheinkopf, S. J. (2014). Attention and word learning in autistic, language delayed and typically developing children. *Frontiers in Psychology, 5,* 490.

Tenenbaum, E. J., Amso, D., Righi, G., & Sheinkopf, S. J. (2017). Attempting to "increase intake from the input": Attention and word learning in children with autism. *Journal of Autism and Developmental Disorders, 47*(6), 1791–1805.

Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2014). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language, 42,* 1173.

Wagner, J. B., Hirsch, S. B., Vogel-Farley, V. K., Redcay, E., & Nelson, C. A. (2013). Eye-tracking, autonomic, and electrophysiological correlates of emotional face processing in adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43*(1), 188–199.

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., & Tojo, Y. (2006). The autism-spectrum quotient (Aq) in Japan: A cross-cultural comparison. *Journal of Autism and Developmental Disorders, 36*(2), 263–270.

Wang, Q., Campbell, D. J., Macari, S. L., Chawarska, K., & Shic, F. (2018). Operationalizing atypical gaze in toddlers with autism spectrum disorders: A cohesion-based approach. *Molecular Autism, 9*(1), 25.

Wechsler, D. (2002). *Wechsler preschool and primary scale of intelligence (Wppsi-Iii)* (3rd ed.). San Antonio: Pearson.

Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6 months: Predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Developmental Science, 12*(5), 798–814.