Author Manuscript
Author Manuscript
Author Manuscript

Published in final edited form as:
Child Dev. 2016 ; 87(2): 543–557. doi:10.1111/cdev.12473.

Multilevel Differences in Spontaneous Social Attention in Toddlers With Autism Spectrum Disorder

Katarzyna Chawarska,
Yale University School of Medicine

Saier Ye,
Yale University

Frederick Shic, and
Yale University School of Medicine

Lisha Chen
General Electric Capital

Abstract

This study examined the latent structure of spontaneous social attention in 11- to 26-month-olds with autism spectrum disorder (ASD, $n = 90$) and typically developing ($n = 79$) controls. Application of the *joint and individual variance explained* decomposition technique revealed that attention was driven by a condition-independent tuning into the dynamic social scenes construct and context-specific constructs capturing selection of the most relevant social features for processing. Gaze behavior in ASD is characterized by a limited tuning into the social scenes and by a selection of atypical targets for processing. While the former may be due to early disruption of the reward circuitry leading to limited appreciation of the behavioral relevance of social information, the latter may represent secondary deficits reflecting limited knowledge about social partners.

For sighted observers, participation in the social world requires rapid deployment of attention in a task-relevant manner. A mother might make eye contact with her child, utter a remark about a cat on the windowsill, then walk across the room to prepare a snack. To gauge the mother's intentions and anticipate the outcomes of her actions, the child needs to fluidly adjust his attentional focus, for example, to look at the mother's face when she speaks, to follow the direction of her gaze, or to observe the mother's movements as she peels an apple, all while inhibiting attention to aspects of his environment that at a given moment are not most informative. Note that in such situations the key distinction may not be between attending to social versus nonsocial stimuli; instead, it may be selecting the particular aspect of the social scene that is most relevant to the task at hand. That is, a child looking at the mother's face may be advantageous when she speaks directly to him, but doing so when she demonstrates how to get from a whole apple to tasty morsels might limit his access to essential observational learning opportunities. Monitoring such dynamic,

Correspondence concerning this article should be addressed to Katarzyna Chawarska, Yale Child Study Center, 40 Temple Street, Suite 7D, New Haven, CT 06511. katarzyna.chawarska@yale.edu.

multimodal flow of events is complex and requires selecting task-relevant stimuli for processing in a rapid and flexible manner. This set of skills is essential for drawing inferences about other's intentions, mental states, and goals (Blakemore & Decety, 2001; Frischen, Bayliss, & Tipper, 2007; Tomasello, 2000), coordinating attention and actions with others in order to reach common goals (Sebanz, Bekkering, & Knoblich, 2006), and observational learning (Rizzolatti & Craighero, 2004).

Children with autism spectrum disorder (ASD), a neurodevelopmental disorder characterized by impairments in social functioning and communication (American Psychiatric Association, 2013), have marked difficulties monitoring the people around them. In the descriptions of the first reported 11 cases of autism, Kanner (1943) noted that children with autism exhibit marked differences in orienting to some of the most quintessential social stimuli around them, "Comings and goings, even of the mother, did not seem to register. Conversation going on in the room elicited no interest" (pp. 245–247). Subsequent observational and experimental work has indicated that social attention atypicalities (i.e., when selecting and encoding information about other people) represent one of the earliest symptoms of ASD (see Chawarska, Macari, Volkmar, Kim, & Shic, 2014, for a review). Considering the centrality in social learning of intact attention to others (Sebanz et al., 2006), such early manifestations and persistent atypicalities in endogenously driven attention to social targets are likely to have a negative impact on the development and function of neural systems involved in social cognition throughout the life span. Moreover, individual differences along the dimension of social attention may contribute to some of the observed heterogeneity of levels of functioning and outcomes observed in ASD (Campbell, Shic, Macari, & Chawarska, 2014; Elsabbagh et al., 2014; Norbury, 2014). Despite the prominence of atypical social attention in ASD, neither the mechanisms that give rise to such differences nor their roles in autistic psychopathology are well understood.

Considering the intricate link between gaze and attention, measurement of gaze allocation presents as one of the more promising paradigms for studying attentional processes in dynamic environments (Deubel & Schneider, 1996). In the past decade, studies utilizing eye movements have shown great promise for elucidating the processes associated with atypical social attention in ASD (see Falck-Ytter, Bölte, & Gredebäck, 2013; Guillon, Hadjikhani, Baduel, & Rogé, 2014; Senju & Johnson, 2009, for reviews). Specifically, there is a growing body of research focused on development of social attention in young children with ASD. These studies often involve static images of faces presented in a highly controlled manner in isolation from a broader context and largely concern the question of detection, discrimination, or recognition of faces. This line of research has generated important insights into the perceptual, attentional, and learning strategies associated with early stages of the disorder. Although toddlers with ASD detect faces among distractors as rapidly as the controls (Elsabbagh et al., 2013), unlike typically developing (TD) and developmentally delayed controls, they more quickly disengage their attention from dynamic faces (Chawarska, Klin, & Volkmar, 2003; Chawarska, Volkmar, & Klin, 2010). These findings suggest that although faces may trigger reflexive orienting in toddlers with ASD, they may not be able to hold their attention (Cohen, 1972) presumably due to limited depth of processing or engagement with these stimuli (Bloom & Mudd, 1991; Coin & Tiberghien, 1997). When examining a novel face, toddlers with ASD employ atypical face-scanning strategies and require more time to extract the invariant features necessary for identity recognition (Chawarska & Shic, 2009). Such studies reflect upon the atypical functioning of cortical networks involved in the structural analysis of faces in ASD (Haxby, Hoffman, & Gobbini, 2002). However, in real life, faces appear as dynamic stimuli embedded within perceptually and semantically complex contexts, and their processing relies on multisensory brain areas involved in the perception of biological motion, speech, and social cognition, including the posterior superior temporal sulcus (Allison, Puce, & McCarthy, 2000; Haxby et al., 2002) and dorsal medial prefrontal cortex (Calvert & Campbell, 2003; Wagner, Kelley, & Heatherton, 2011). Moreover, in experiments targeting elementary face-processing skills, such as capture or recognition, stimuli are typically presented in rapid succession, and attention to them is artificially supported by extraneous cues, such as their sudden onset on the screen or presentation of “attention getters” (e.g., central fixation stimuli prior to the target onset). Such manipulations are necessary when the intention is to obtain a large number of valid trials from individual participants and rely heavily on exogenous (i.e., reflexive, bottom-up) orienting; however, they do not allow for making inferences regarding endogenous (i.e., spontaneous, top-down) orienting to social stimuli. Therefore, the paradigms built on picture viewing, although helpful for examining specific aspects of social attention, may not be appropriate for explaining endogenously driven gaze behaviors in natural environments (Henderson, 2007; Tatler, 2014).

Several recent studies have examined spontaneous gaze responses to dynamic multimodal social stimuli in toddlers and preschoolers to identify processes that may compromise the development of social attention in ASD. Such studies typically employ a free-viewing paradigm, where a child is presented with video recordings of a variety of social scenes and provided no explicit instructions. When toddlers with ASD view people trying to engage their attention through eye contact and child-directed speech (Chawarska, Macari, & Shic, 2012), or through anticipatory games and overtures (Jones, Carr, & Klin, 2008), or when these children observe interacting adults (Nakano et al., 2010; von Hofsten, Uhlig, Adell, & Kochukhova, 2009), they tend to look less at faces than children with other developmental outcomes. Toddlers with ASD also are less likely to look at the objects attended to by others (Bedford et al., 2012), and they show limited attention to the goal-oriented activities of others (Shic, Bradshaw, Klin, Scassellati, & Chawarska, 2011). Importantly, these studies suggest that the differences in social attention observed in natural environments can also be observed and examined through analysis of gaze behaviors using free-viewing eye-tracking experiments. It is hoped that by parsing experimentally the contexts that give rise to the atypical gaze patterns in ASD, we will facilitate the discovery of their underlying cognitive and neural mechanisms.

One of the most common approaches to the analysis of eye-tracking data has been the region of interest (ROI) approach (Holmqvist et al., 2011). In this approach, regions to be analyzed are determined a priori based on underlying assumptions about the structure of the visual field under consideration, such as the distinctness of faces (Chawarska & Shic, 2009; Jones et al., 2008), activities (Shic et al., 2011), or areas of high perceptual salience (Freeth, Chapman, Ropar, & Mitchell, 2010; Shic, Chawarska, Lin, & Scassellati, 2008). Subsequently, looking times at the specified ROIs are compared between groups. The ROI analytic approach typically produces a set of highly negatively correlated variables reflecting the all-or-nothing nature of overt visual selection (e.g., if one's point of regard is directed at the face, it cannot simultaneously be deployed to the background). Mapping such trade-offs can illustrate individual differences in scene perception, as limited looking time at a person's face can result not only from enhanced monitoring of her hands engaged in a goal-oriented activity, but also from looking at irrelevant objects in the background. Moreover, another important and often underappreciated aspect of gaze behaviors is that they may arise from several underlying cognitive processes, some of which may be domain general and related to, that is, inattention or distractibility, whereas others might be domain specific and related to particular classes of stimuli (e.g., faces or eyes) or specific contexts (e.g., face-to-face interaction). These inherent properties of gaze behaviors call for analytic approaches that take advantage of the richness of the information carried by all of the available looking indices as well as their interactions.

In the present study, we examined the latent structure of gaze behaviors of toddlers with ASD and TD controls in response to a naturalistic stream of events that a child may encounter readily in real life. By examining toddlers with and without ASD in a single, combined sample, we aim to characterize the natural variability in gaze patterns of those affected by social disability and those with typical social development. This approach is consistent with the current dimensional conceptualization of traits associated with autism (Ronald & Hoekstra, 2011) and with National Institute of Mental Health research domain criteria recommendations (Insel et al., 2010). To this end, we presented a large group of 11- to 26-month-old children with ASD (n = 90) and TD controls (n = 79) with a 3-min free-viewing spontaneous social monitoring (SSM) task depicting a woman engaged in four activities: initiating bids for dyadic (face-to-face interaction) and triadic (joint attention) engagement, performing a goal-directed action (making a sandwich), or acting in the presence of highly salient nonsocial dynamic stimuli (mechanical toys). The task required the participants to adjust their scanning strategies depending on the content of the scene (e.g., to look at her face when she spoke, or to monitor her activity when she made a sandwich). Gaze behaviors in toddlers with ASD were evaluated at the time of their first diagnosis shortly after the period when behavioral symptoms first become apparent in ASD and prior to the initiation of treatment, therefore allowing us to access features that may more likely reflect primary characteristics of ASD in late infancy.

To evaluate the latent structure of attention to dynamic social scenes in affected and unaffected infants, we employed a novel data analytic approach: Joint and Individual Variation Explained (JIVE; Lock, Hoadley, Marron, & Nobel, 2013). JIVE represents an extension of the standard principal component analysis (PCA) designed for the integrated analysis of multiple data types. JIVE takes advantage of the hierarchical structure in experimentally manipulated data, allowing for the quantification of both the variance shared among experimental conditions and the variance specific to each condition. Although standard PCA techniques can be applied to all conditions combined to reveal a shared structure or can be applied to each condition separately to reveal the structure particular to each condition, it cannot perform both functions at once. This limitation hinders the discovery of the internal structure of complex experimental data that may be driven by the interplay of factors that may be common to all conditions and factors that may play a role only in specific social contexts.

Method

Procedure

Participants—Participants consisted of 169 toddlers ages 11–26 months (Visit 1; *M* = 20.5 months, *SD* = 3.77) who were either referred to a specialized university clinic for evaluation and diagnosed with ASD (*n* = 90) or who were TD controls (*n* = 79) recruited through advertisements. Seventy-six (84%) of those with ASD underwent a confirmatory diagnostic assessment at 36 months (Visit 2). The diagnostic assessment battery at Visits 1 and 2 targeted verbal and nonverbal skills using the Mullen Scales of Early Learning (Mullen, 1995), adaptive functioning skills using the Vineland Adaptive Behavior Scales (Vineland–II; Sparrow, Balla, & Cicchetti, 2005), and autism symptoms using the Autism Diagnostic Observation Schedule–Generic (ADOS–G; Lord et al., 2000). Two expert clinicians assigned a clinical best estimate (CBE) diagnosis of ASD based on the review of all standardized tests, as well as the child's medical and developmental histories. Considering the 90%–100% stability of the CBE ASD diagnosis in the 2nd year in clinic-referred samples (Chawarska, Klin, Paul, Macari, & Volkmar, 2009; Guthrie, Swineford, Nottke, & Wetherby, 2012; Lord et al., 2006), such as the one in this study, it is unlikely that diagnostic shifts (i.e., from ASD to non-ASD disorder) among the 14 infants who did not return for Visit 2 would occur with enough frequency to compromise the integrity of our analysis. The participants were recruited between 2008 and 2012 from the Northeastern part of the United States including Connecticut, Massachusetts, New York, and New Jersey. Seventy-five percent of the parents declared their race as Caucasian, and 12% declared a Hispanic ethnicity. The ASD group was 84% male as compared to 71% male in the TD group (p = .04). Children with ASD were seen on average at 21.4 (*SD* = 2.9) months compared to 19.3 (*SD* = 4.3) months in TD controls ($p < .001$). The ASD group had significantly lower verbal (*M* = 49, *SD* = 26, $p < .001$) and nonverbal (*M* = 80, *SD* = 17, $p < .001$) DQ scores compared to TD controls (*M* = 104, *SD* = 18 and *M* = 109, *SD* = 11, respectively); their mean ADOS–G Module 1 total algorithm score was 18.8 (*SD* = 5.0).

An additional 25 toddlers (8 TD and 17 ASD) were tested but contributed data to fewer than four conditions due to calibration errors (see the Data Reduction section) or insufficient data (< 20% of valid eye-tracking data) resulting from technical difficulties or inattention. These participants were excluded from the subsequent analysis, for one of the requirements of the JIVE approach is that all participants contribute data to all considered conditions. The excluded toddlers did not differ significantly from the retained sample in age (p = .39), Verbal Developmental Quotient (VDQ; p = .57), Nonverbal DQ (NVDQ; p = .91), or ADOS–G total algorithm score (p = .84, ASD only). Toddlers with a gestational age below 34 weeks, nonfebrile seizures, uncorrected vision or auditory impairments, or known genetic abnormalities were not included in this study. All parents provided informed consent in adherence to the Yale School of Medicine Human Investigation Committee requirements.

Stimuli—The stimulus consisted of a 3-min video of an actress filmed in a setting containing four toys and a table with ingredients for making sandwiches (see Figure 1a). The video comprised four conditions, with each condition presented over multiple episodes. In the *dyadic bid* condition, the actress engaged in child-directed speech while looking directly at the camera (11 episodes, total duration 69 s), resembling a bid for dyadic (face-to-face) engagement. The content of the actress's speech was related to the events presented in the video and included greeting the viewer ("How are you, baby?"), complimenting ("You look so cute there!"), or commenting on events occurring in the movie ("Let's make a sandwich!"). In the *joint attention* condition, the actress looked up briefly at the camera, making eye contact with the viewer, and then exclaimed, "uh-oh!" as she turned toward one of the toys and remained still for 4 s (four episodes, total duration 30 s), exhibiting a bid for joint attention. In the *sandwich* condition, the actress looked down at the table and made a sandwich; no direct gaze or speech cues were present (two episodes, total duration 63 s). Finally, in the *moving toys* condition, after the actress looked up at the camera, a toy began to move and make noises; this was immediately followed by the actress turning to look at the toy on the opposite side of the moving toy (four episodes, total duration 27 s). Children watched identical videos. To avoid extraneous attentional cues and disruptions in processing of the social scenes (Hirose, Tatler, & Kennedy, 2011), there were no breaks in the video, and the conditions were interleaved in a way suggestive of a natural flow of events. This type of display required the toddlers to adjust their gaze patterns depending on context, as they would in real life, without the benefit of extraneous cues directing their attention to the screen in general or to any of the specific elements of the scene.

Apparatus—Gaze trajectories were recorded at a sampling rate of 60 Hz using a SensoMotoric Instruments IView X™ RED eye-tracking system (SensoMotoric Instruments GmbH, Teltow, Germany). Eye-tracking data were processed using custom software written in Matlab (MathWorks, 2009). The software accommodated standard techniques for processing eye-tracking data, including blink detection, data calibration, recalibration, and ROI analysis (Duchowski, 2003; Shic, 2008).

Procedure—Toddlers were seated in a car seat, in a dark and soundproof room 75 cm in front of a 24-in. wide-screen LCD monitor. At this distance, the scene subtended 27° × 21° of visual angle, the actress's face 3.9° × 5.6°, her mouth 3.5° × 2.0°, the activity area 8.6° × 6.9°, and each of the toys approximately 5.8° × 6.4°. Each session began with a video to help the child get settled. A 5-point calibration procedure was then initiated with calibration points consisting of dynamic targets presented simultaneously with sound (e.g., a meowing, walking cartoon tiger). Subsequently, each participant was presented with the video described in the Stimuli section.

Analytic Strategy

Data Quality Assurance—Participants for whom algorithmic calibration uncertainty was less than approximately 1.2 visual degrees (Shic, 2008) were excluded from analysis. ROIs were dilated by 1.25 visual degrees in order to compensate for possible calibration error. Average calibration error was < 0.75°. Participants with any condition containing < 20% of valid data were also excluded.

Data Reduction—The visual scene was divided into ROIs (Figure 1b). Dependent variables were based on the proportions of time spent looking at each of the regions (i.e., the total amount of time the participant's point of regard was located within each region) and included (a) overall attention to the scene (%Scene), (b) proportion of attention directed toward the person (including %Person, %Face, %Eyes, %Mouth, %Hands/Activity Area), and (c) attention toward distractors (%Toys) and background (%Background). The proportion of total looking time (%Scene) was standardized by the total duration of the video display; the remaining variables were standardized by the total looking time at the scene (see Table 1 for average dwell time in each of the ROIs in the ASD and TD groups).

Statistical Analysis—To reveal the latent structure of the data that best explains the variation across all four conditions and the two groups, we employed the JIVE analytic approach (Lock et al., 2013). This approach takes advantage of the unique structure of the data related to the four distinct conditions and quantifies the variance shared by all the conditions as well as the variance specific to each condition. JIVE analysis was conducted on the combined ASD and TD groups ($N = 169$). Eight variables were considered (%Scene, %Person, %Face, %Hands/Activity, %Toys, %Eyes, %Mouth, and %Background), collected in four conditions: joint attention, dyadic bid, sandwich, and moving toys. The joint and individual variations were computed iteratively (Lock et al., 2013), yielding orthogonal PCs, or latent factors, that corresponded to variation specific to each condition (individual variation) and variation shared across conditions (joint variation). The number of PCs was determined in a stepwise manner. We first determined the number of joint PCs. Starting with one joint PC, we employed the criterion that the next joint PC needed to capture over 10% of variance in each of the conditions to be retained for the analysis. Next, the number of condition-specific, or individual, PCs was determined such that the sum of the variance explained by the model was at least 70% of the total variance in each condition. Subsequently, we computed PC scores by multiplying PC loadings by standardized gaze variables and compared the scores in ASD and TD groups using general linear models with group, age, and Age × Group factors. Effect sizes were computed using Cohen's *d*. Associations between PC scores and concurrent clinical features (VDQ and NVDQ, severity of autism symptoms) were evaluated using Pearson's correlation coefficient, partialed out for the effects of age with Bonferroni correction for multiple ($n = 15$) comparisons ($p < .0033$). The analysis was conducted using R (Team RDC, 2011) and in SAS Statistical Software (SAS Institute Inc., Cary, NC) packages.

Results

Latent Structure Analysis

Preliminary analyses revealed an almost identical latent structure for two of the conditions involving bids for social attention (dyadic bid and joint attention; see Figure 2, represented in yellow and black, respectively); thus, the two conditions were combined before the final JIVE analysis was performed. This resulted in the following conditions: social bid (joint attention and dyadic bid combined), sandwich, and moving toys. The results of JIVE analysis suggest that the structure of the gaze data could be decomposed into one joint PC reflecting overall attention to the complex dynamic video, regardless of its content; two individual PCs capturing variability in response to scenes containing social bids (social bid condition); an individual PC capturing variability in monitoring goal-oriented activity of another person (sandwich condition); and an individual PC capturing variability in gaze allocation when there is a direct competition between a socially and perceptually salient set of stimuli (moving toys condition; see Table 2 for variance partition). The five PCs accounted jointly for 75.5%, 71.4%, and 72.7% of variance in the social bid, sandwich, and moving toys conditions, respectively.

The joint PC was primarily associated with the amount of time each participant spent looking at, as opposed to away from, the entire scene, regardless of the social context (Scene On–Off PC; see Table 3, for PC loadings) and accounted for 35.2%, 38.1%, and 22.9% of variance in gaze patterns in response to the social bid, sandwich, and moving toys conditions, respectively. JIVE analysis revealed two PCs specific to the conditions involving social bids: Social Bid PC 1 was defined primarily by attention to the person's face as opposed to the toys (Face–Toys PC), and Social Bid PC 2 captured increased attention to the mouth and decreased attention to the eyes of the speaker (Mouth–Eyes PC). The Face–Toy PC accounted for 23.3% of variance and the Mouth–Eyes PC for 17.0% of variance in the combined social bid condition. The PC specific to the sandwich condition primarily captured attention to the activity area including hands and sandwich ingredients at the expense of monitoring the face of the sandwich maker (Activity–Face PC). The Activity–Face PC accounted for 33.3% of variance in the sandwich condition. Finally, a PC specific to the moving toys condition mainly captured attention to the toys relative to the person (Toys–Person PC) and accounted for 49.8% of variance in this condition.

Subsequently, we compared the groups on PC scores (Figure 3). Toddlers with ASD had significantly lower scores on the Scene On–Off PC, $F(1, 168) = 21.86$, $p < .001$ ($d = -.52$), but there was neither an effect of age ($p = .144$) nor a Group × Age interaction ($p = .233$). Toddlers with ASD also had lower Mouth–Eyes PC scores, $F(1, 168) = 11.04$, $p = .004$ ($d = -.48$). The effect of age was significant, $F(1, 132) = 2.12$, $p = .034$, and there was a trend toward a Group × Age interaction ($p = .069$). Further analysis indicated no correlation between Mouth–Eyes PC and age in the ASD group, $r(90) = .02$, $p = .875$, but a significant positive correlation in the TD group, $r(79) = .28$, $p = .011$. Toddlers with ASD also had lower scores on the Activity–Face PC, $F(1, 168) = 7.79$, $p = .012$ ($d = -.46$), but there was no effect of age ($p = .323$) and no Group × Age interaction ($p = .769$). There were no significant effects for the Face–Toys PC (neither group: $p = .833$, nor age: $p = .191$, nor Group × Age interaction: $p = .218$) or for the Toys–Person PC (neither group: $p = .403$, nor age: $p = .967$, nor Group × Age interaction: $p = .905$). Thus, the attentional differences observed in toddlers with ASD include a more general difficulty attending to a complex scene, as well as difficulties in selecting the most relevant, in a given context, element of the scene.

Associations Between Scanning Patterns and Clinical Features

Subsequently, we evaluated whether performance on the eye-tracking task indexed by the five PCs was concurrently associated with individual clinical characteristics. Verbal and nonverbal DQ scores were available for the entire sample of ASD and TD toddlers, whereas ADOS–G total scores were available only for the ASD group, for this measure was not designed to quantify social functioning in unaffected children. In the ASD group, those with better attention to the scene in general (Scene On–Off PC) had a lower autism symptom severity score as indexed by the ADOS–G (Table 4). In the combined ASD and TD sample, greater attention to the scene was also associated positively with both VDQ and NVDQ. Greater attention to the speaker's mouth (Mouth–Eyes PC) and to the toys (Toys–Person PC) was associated with a higher VDQ, and better attention to the goal-oriented activities of others (Activity–Face PC) was associated with higher NVDQ.

Discussion

To the best of our best knowledge, this is the first study that examines directly the latent structure of gaze behaviors in response to dynamic social scenes in TD toddlers and toddlers with ASD. The study indicates that their gaze behaviors are driven by two sets of orthogonal factors: (a) general orienting and sustained attention to dynamic social scenes and (b) selection for processing the most informative features of the social scene. Compared to TD controls, toddlers with ASD spent less time tuning into the scenes in general and were less likely to select for processing the most context-relevant targets among an array of other social targets. Characteristics of gaze behaviors were associated in a meaningful manner with clinical features in both ASD and TD toddlers. Taken together, this work suggests that differences in endogenously driven attention to social targets in toddlers with ASD are expressed along multiple axes and are related to clinical characteristics that are highly relevant to long-term outcomes in ASD. These findings are discussed in turn.

The application of a novel approach to latent structure analysis revealed that approximately one third of the variance in gaze behaviors across conditions was accounted for by orienting to and sustaining attention to dynamic scenes. It is not clear if this PC is specific to scenes containing people or extends to other dynamic, complex scenes, reflecting, therefore, a domain-general function supporting a range of social and nonsocial cognitive processes reliant on the salience network of the brain (see Barrett & Satpute, 2013, for a review). Nevertheless, by isolating this construct, we have identified a major contributor to spontaneous allocation of attention in the 2nd year of life both in typically and atypically developing children. Moreover, a large amount of variance in gaze behaviors was accounted for by condition-specific latent components, which were independent from the joint PC; that is, they captured the underlying structure of gaze behaviors once the toddlers looked at the scene. In response to a person trying to engage the viewer using eye contact and speech cues, prioritizing the person's face over toys and her mouth over eyes accounted for 40% of variance in gaze behaviors. When a person performed a goal-directed action and refrained from bids for shared engagement, approximately 33% of variance in gaze behaviors was accounted for by the selection of the activity area as opposed to her face, an area typically of great social relevance but less informative in this particular context. Finally, in a condition in which salient nonsocial events took place, allocation of attention toward the toys and away from the person captured almost 50% of the variance in gaze behavior. These findings suggest that a parent's attempt to engage a baby in a playful face-to-face interaction may posit a challenge to the child on several levels: (a) tuning into the most salient/relevant area of the sensory field (e.g., parent acting within a broader natural environment), (b) selecting for processing the parent's face while inhibiting attention to irrelevant objects in the visual field, and (c) distributing attention between the parent's eyes and mouth in a way that maximizes the extraction of emotional and linguistic content. Similarly, learning by observing the parent's actions requires not only orienting to the parent and the surrounding scene, but also directing attention to the parent's activity in a way that will facilitate the extraction of the parent's goals and intentions along with invariant features of the action. Importantly, a disruption at any of these facets of attention regulation may perturb the child's engagement in and learning from the most quintessential of social experiences.

Toddlers with ASD spent less time tuning into the social scenes altogether. In this group, decreased attention to the scenes was associated with greater autism severity symptoms and lower verbal and nonverbal skills. Thus, the toddlers with decreased spontaneous attention to sources of social information were less likely to avail themselves of learning opportunities, which, in turn, may hinder their skill acquisition across domains. This interpretation is consistent with results of a prospective study that suggested that 2-year-olds with autism who exhibit the most limited attention to social scenes exhibit more severe symptoms of ASD and are lower functioning at the age of 3–4 years than toddlers with autism who show more robust spontaneous attention to such scenes (Campbell et al., 2014). This effect does not appear to be specific to ASD, as in a sample of TD infants, the overall attention to the social scene at 12 months uniquely predicted later developmental outcomes at 18 and 24 months (Tenenbaum, Sobel, Sheinkopf, Malle, & Morgan, 2014).

Results of our study also indicate that once the toddlers with ASD tuned into the scene, they spent less time monitoring the speaker's mouth compared to TD controls. Although an initial report suggested that toddlers with ASD may display enhanced attention to the speaker's mouth compared to children with other developmental outcomes (Jones et al., 2008), more recent work suggests that these differences are not consistently observed in ASD (see Guillon et al., 2014; Senju & Johnson, 2009, for reviews). Instead, consistent with our findings, a number of studies have reported prospective links between preferential orienting to the speaker's mouth in infancy and later language outcomes in younger siblings of children with ASD (Elsabbagh et al., 2014; Young, Merin, Rogers, & Ozonoff, 2009), TD infants (Tenenbaum et al., 2014), as well as clinic-referred toddlers with ASD (Campbell et al., 2014). The significance of this finding is linked with extensive evidence suggesting that the deployment of attention to the mouth region of an interactive partner provides direct access to the tightly coupled and redundant patterns of visual and auditory speech information (Chandrasekaran, Trubanova, Stillittano, Caplier, & Ghazanfar, 2009), and therefore facilitates speech perception and acquisition in young children (Frank, Vul, & Saxe, 2012; Lewkowicz & Hansen-Tift, 2012; Tenenbaum, Shah, Sobel, Malle, & Morgan, 2013). Consistent with other work (Frank et al., 2012), TD toddlers in our study showed an increase in attention to the mouth versus eyes over the course of the 2nd year. This was, however, not the case in the ASD group, suggesting persisting limited utilization of audiovisual cues in speech perception throughout the 2nd year of life.

Finally, similar to work by Shic et al. (2011), in our study toddlers with ASD spent less time monitoring the activity of the actress, and poorer performance in this condition was associated with lower levels of nonverbal skills. Observing the actions of others plays a role not only in the encoding of visual properties of the goal-directed movement, but also in understanding others' goals or intentions without any other form of symbolic communication (Umilta et al., 2001). Moreover, such observations can elicit complementary (joint) actions in the observer and play a key role in observational learning (Rizzolatti & Craighero, 2004). Taken together, the results suggest that the atypical social attention observed in toddlers with ASD is driven by several distinct underlying processes. Early variability in attentional responses to social scenes may contribute to the observed heterogeneity in phenotypic presentation in preschoolers with ASD with regard to severity of social impairment and levels of verbal and cognitive functioning.

After the overall attention to the scene was accounted for, there were no differences between groups with regard to the distribution of attention between dynamic mechanical toys and the person in the moving toys condition, suggesting absence of prepotent and indiscriminate interest in objects as opposed to people in toddlers with ASD. Moreover, there were no differences between groups in the distribution of attention between face and toys in the social bid condition. This finding was inconsistent with several studies utilizing similar sets of dynamic stimuli. However, these studies typically view proportion of gaze to faces as an independent outcome measure and do not consider the variance that might be accounted for by the overall ability to tune into the social scenes. In addition, the studies do not take into account what might be the most parsimonious trade-offs in looking patterns given attentional constrains; that is, they consider the face in insolation from the rest of the scene rather than the face in relation to the most salient distractor such as toys in the background. Thanks to the novel approach to identifying the latent structure of spontaneous orienting, our study demonstrates what the gaze patterns of young, affected toddlers are after the very elementary orienting toward complex stimuli is taken under consideration. Thus, it is plausible that limited attention to faces reflects “collateral damage” related to limited general tuning into the social scenes. Support for this hypothesis comes from studies that, instead of relying on endogenous orienting to faces, employ a variety of strategies to enhance reflexive or exogenous attention to these stimuli by, for instance, employing “attention getters” in attention cueing (Chawarska et al., 2003; Chawarska et al., 2010), or visual paired comparison (Chawarska & Shic, 2009; Chawarska & Volkmar, 2007; de Klerk, Gliga, Charman, & Johnson, 2014) paradigms. In such contexts, toddlers with ASD exhibit no generalized atypicalities in baseline attention to faces. Thus, the limited attention to faces within complex dynamic scenes observed in toddlers with ASD may be related to the general limited significance of such scenes rather than a face-specific deficit.

The attentional atypicalities reported in our study were observed shortly after the onset of behavioral symptoms of ASD and, therefore, had not yet been altered by, for example, the remedial effects of early intervention or extensive cumulative effects of abnormal social interaction patterns. Naturally, a question arises concerning whether emergence of the atypical gaze patterns simply co-occurs with the onset of behavioral symptoms that typically happen around or after 12 months of age (Ozonoff et al., 2010; Paul, Fuerst, Ramsay, Chawarska, & Klin, 2011), or whether these patterns may be present during the prodromal stage of the disorder. Recent work on infant siblings of children with ASD who, due to familial factors, are also at risk for developing ASD, suggests that differences in endogenous social attention can be observed in experimental settings in infancy. Specifically, 6-month-old infants who later developed ASD exhibited particularly limited attention to social scenes when tested in two free-viewing eye-tracking paradigms: the SSM task (Chawarska, Macari, & Shic, 2013) and a dynamic face scanning task (Shic, Macari, & Chawarska, 2014). At 6 months, attention to faces was also decreased but no differences in distribution of attention between eyes and mouth were noted. Taken together, this work suggests that attention to complex social scenes is perturbed before behavioral symptoms of autism begin to emerge, and this perturbation is still observed at the time when children become symptomatic in the 2nd year of life. Thus, atypical facets of endogenous regulation of attention to dynamic social targets are likely to constitute some of the primary manifestations of ASD and inform the search for candidate markers for ASD in infancy and early childhood.

Considering that the control of eye movements in response to dynamic social scenes relies on a complex interplay between top-down and bottom-up processes (see Tatler, 2014; Tatler et al., 2013, for reviews), it is not clear what specifically accounts for the observed group differences. Similar to models of gaze behaviors in natural environments, models that could account for the dynamics of gaze behaviors in free-viewing tasks in typical individuals are still in development. Nonetheless, current evidence suggests the key roles of *behavioral relevance* (i.e., costs and benefits of gaze behaviors in acquisition of goal-relevant information) and *prior knowledge* (i.e., learned models of the environment) as forces directing gaze in complex environments in service of foraging for information relevant to survival (Tatler, Hayhoe, Land, & Ballard, 2011). These two factors drive the mechanisms controlling what we should attend to based on where we will gain information for fulfilling behavioral goals. In this context, the information acquired during a fixation can be thought of as a secondary reward, mediating learning of gaze patterns by virtue of their significance for adaptation. This principle is exemplified by studies illustrating that monkeys are willing to give up food (i.e., “pay per view”) in order to obtain visual information about members of their social group (e.g., their availability for mating; Deaner, Khera, & Platt, 2005; Shepherd, Deaner, & Platt, 2006). Sensitivity to reward is manifest throughout the saccadic eye movement circuitry and the neurons involved in saccadic targeting respond in a graded manner to both the magnitude of the expected reward and the probability of a reward prior to execution of the saccade (Hikosaka, Nakamura, & Nakahara, 2006). Sensitivity to both of these factors is critical for learning and consequently linking fixation patterns to task demands.

We propose that the observed limited attention to scenes containing conspecifics in a complex experiment was driven by diminished appreciation for the potential relevance or reward value that these scenes hold for gathering essential information for adaptation and survival. This premise is consistent with a hypothesis linking social impairments in ASD with a disruption in the reward system (Chevallier, Kohls, Troiani, Drodkin, & Schults, 2012; Kohls et al., 2012). Importantly, once the toddlers in the present study looked at the social scene, they had difficulty selecting for processing of the most *task-relevant* social features. That is, they did not appear to “know” that in certain contexts, mouth is more informative than eyes, or that monitoring goal-oriented action is more informative than looking at the face of the person performing the action. Although it is possible that these are highly domain-specific effects, we argue instead that they may represent secondary effects of limited spontaneous monitoring of conspecifics. Thus, while the diminished monitoring of conspecifics may be primary in ASD, limited selection of the most relevant social targets may reflect the secondary effects of atypical experiences with interactive partners and, therefore, limited knowledge about them. Support for this hypothesis comes from studies suggesting that although there is a continuity in social attention impairment from prodromal to early syndromal stages of the disorder, there is also an evolution of such impairment from a limited attunement to people in general (Chawarska et al., 2013) to more specific impairments in attention toward their social and communicative bids (Chawarska et al., 2012). This pattern is consistent with the interactive specialization model, which suggests that in the postnatal period many of the brain regions are poorly specialized, but in the subsequent months undergo fine-tuning to more specific classes of stimuli in an experience-dependent manner (Greenough, Black, & Wallace, 2002). As such, improving strategies for information gathering about conspecifics may represent one of the novel intervention targets for ASD in early development.

Limitations

Future work will need to consider the analysis of temporal patterns of social attention in response to dynamic social cues as potentially informative in the further decomposition of processes underlying atypical social attention in autism. Considering that the 2nd year of life is a period of rapid developmental transformations in attention and cognition, it remains an empirical question whether the identified latent structure generalizes to other developmental epochs. Although earlier studies suggest that deficits in attention to social scenes are not present in toddlers with global developmental delays, investigations into the specificity of the observed patterns of deficits in relation to such disorders, (e.g., attention deficit disorder or specific language impairment) are warranted. Pending technological advances, the natural extension of our work would be to experimentally evaluate the regulation of attention in fully naturalistic, interactive, immersive contexts.

Conclusions

Employment of a novel approach to latent structure analysis revealed that foraging for social information appears to be compromised in toddlers with ASD by limited tuning into social scenes in general and, specifically, by limited appreciation for what social features are most relevant, given broader contexts. We hypothesize that while the former atypicalities may be due to a core disruption of reward circuitry, the latter may stem from limited knowledge about social partners, secondary to the disruption of the reward system manifesting in early infancy. These findings reinforce the need for further prospective longitudinal investigation into developmental dynamics underlying the development of social and nonsocial attention in ASD.

Acknowledgments

The study was supported by the National Institute of Child Health and Development, PO1 HD003008, and the National Institutes of Mental Health P50 MH081756, Project 2 (PI: Katarzyna Chawarska) and R01 MH087554 (PI: Katarzyna Chawarska). We thank Lauren DiNicola for help on preparation of the manuscript and express our appreciation to the families and their children who participated in the study. The authors declare no conflict of interest.

References

Allison T, Puce A, McCarthy G. Social perception from visual cues: Role of the STS region. Trends in Cognitive Sciences. 2000; 4:267–278. DOI: 10.10161/s1364-6613(00)01501-1 [PubMed: 10859571]

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5. Washington, DC: Author; 2013.

Barrett LF, Satpute AB. Large-scale brain networks in affective and social neuroscience: Towards an integrative functional architecture of the brain. Current Opinion in Neurobiology. 2013; 23:361–372. DOI: 10.1016/j.conb.2012.12.012 [PubMed: 23352202]

Bedford R, Elsabbagh M, Gliga T, Pickles A, Senju A, Charman T, Johnson MH. Precursors to social and communication difficulties in infants at-risk for autism: Gaze following and attentional engagement. Journal of Autism and Developmental Disorders. 2012; 42:2208–2218. DOI: 10.1007/s10803-012-1450-y [PubMed: 22278030]

Blakemore SJ, Decety J. From the perception of action to the understanding of intention. Nature Reviews Neuroscience. 2001; 2:561–567. DOI: 10.1038/35086023 [PubMed: 11483999]

Bloom LC, Mudd SA. Depth of processing approach to face recognition: A test of two theories. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1991; 17:556–565. DOI: 10.1037/0278-7393.17.3.556

Calvert G, Campbell R. Reading speech from still and moving faces: The neural substrates of visible speech. Journal of Cognitive Neuroscience. 2003; 15:57–70. DOI: 10.1162/089892903321107828 [PubMed: 12590843]

Campbell DJ, Shic F, Macari S, Chawarska K. Gaze response to dyadic bids at 2 years related to outcomes at 3 years in autism spectrum disorders: A subtyping analysis. Journal of Autism and Developmental Disorders. 2014; 44:431–442. DOI: 10.1007/s10803-013-1885-9 [PubMed: 23877749]

Chandrasekaran C, Trubanova A, Stillittano S, Caplier A, Ghazanfar AA. The natural statistics of audiovisual speech. PLoS Computational Biology. 2009; 5:e1000436.doi: 10.1371/journal.pcbi.1000436 [PubMed: 19609344]

Chawarska K, Klin A, Paul R, Macari S, Volkmar F. A prospective study of toddlers with ASD: Short-term diagnostic and cognitive outcomes. Journal of Child Psychology and Psychiatry. 2009; 50:1235–1245. DOI: 10.1111/j.1469-7610.2009.02101.x [PubMed: 19594835]

Chawarska K, Klin A, Volkmar F. Automatic attention cueing through eye movement in 2-year-old children with autism. Child Development. 2003; 74:1108–1122. DOI: 10.1111/1467-8624.00595 [PubMed: 12938707]

Chawarska K, Macari S, Shic F. Context modulates attention to social scenes in toddlers with autism. Journal of Child Psychology and Psychiatry. 2012; 53:903–913. DOI: 10.1111/j.1469-7610.2012.02538.x [PubMed: 22428993]

Chawarska K, Macari S, Shic F. Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. Biological Psychiatry. 2013; 74:195–203. DOI: 10.1016/j.biop-sych.2012.11.022 [PubMed: 23313640]

Chawarska, K., Macari, S., Volkmar, F., Kim, SH., Shic, F. ASD in infants and toddlers. In: Volkmar, FR.Paul, R.Rogers, SJ., Pelphrey, KA., editors. Handbook of autism and pervasive developmental disorders. Hoboken, NJ: Willey; 2014.

Chawarska K, Shic F. 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. 2009; 39:1663–1672. DOI: 10.1007/s10803-009-0803-7 [PubMed: 19590943]

Chawarska K, Volkmar F. Impairments in monkey and human face recognition in 2-year-old toddlers with autism spectrum disorder and developmental delay. Developmental Science. 2007; 10:266–279. DOI: 10.1111/j.1467-7687.2006.00543.x [PubMed: 17286849]

Chawarska K, Volkmar F, Klin A. Limited attentional bias for faces in toddlers with autism spectrum disorders [Research Support, N.I.H., Extramural]. Archives of General Psychiatry. 2010; 67:178–185. DOI: 10.1001/archgenpsychiatry.2009.194 [PubMed: 20124117]

Chevallier C, Kohls G, Troiani V, Drodkin E, Schults RT. The social motivation theory of autism. Trends in Cognitive Sciences. 2012; 16:231–239. DOI: 10.1016/j.tics.2012.02.007 [PubMed: 22425667]

Cohen LB. Attention-getting and attention-holding processes of infant visual preferences. Child Development. 1972; 4:869–879. DOI: 10.2307/1127638

Coin C, Tiberghien G. Encoding activity and face recognition. Memory. 1997; 5:545–568. DOI: 10.1080/741949479 [PubMed: 9415321]

Deaner RO, Khera AV, Platt ML. Monkeys pay per view: Adaptive valuation of social images by rhesus macaques. Current Biology. 2005; 15:543–548. DOI: 10.1016/j.cub.2005.01.044 [PubMed: 15797023]

de Klerk CCJM, Gliga T, Charman T, Johnson MH. BASIS Team. Face engagement during infancy predicts later face recognition ability in younger siblings of children with autism. Developmental Science. 2014; 17:596–611. DOI: 10.1111/desc.12141 [PubMed: 24314028]

Deubel H, Schneider WX. Saccade target selection and object recognition: Evidence for a common attentional mechanism. Vision Research. 1996; 36:1827–1837. DOI: 10.1016/0042-6989(95)00294-4 [PubMed: 8759451]

Duchowski, AT. Eye tracking methodology: Theory and practice. New York, NY: Springer; 2003.

Elsabbagh M, Bedford R, Senju A, Charman T, Pickles A, Johnson MH. What you see is what you get: Contextual modulation of face scanning in typical and atypical development. Social Cognitive and Affective Neuroscience. 2014; 9:538–543. DOI: 10.1093/scan/nst012 [PubMed: 23386743]

Elsabbagh M, Gliga T, Pickles A, Hudry K, Charman T, Johnson MH, Team B. The development of face orienting mechanisms in infants at-risk for autism. Behavioural Brain Research. 2013; 251:147–154. DOI: 10.1016/j.bbr.2012.07.030 [PubMed: 22846849]

Falck-Ytter T, Bölte S, Gredebäck G. Eye tracking in early autism research. Journal of Neurodevelopmental Disorders. 2013; 5doi: 10.1186/1866-1955-5-28

Frank MC, Vul E, Saxe R. Measuring the development of social Attention using free-viewing. Infancy. 2012; 17:355–375. DOI: 10.1111/j.1532-7078.2011.00086.x

Freeth M, Chapman P, Ropar D, Mitchell P. Do gaze cues in complex scenes capture and direct the attention of high functioning adolescents with ASD? Evidence from eye-tracking. Journal of Autism and Developmental Disorders. 2010; 40:534–547. DOI: 10.1007/s10803-009-0893-2 [PubMed: 19904597]

Frischen A, Bayliss AP, Tipper SP. Gaze cueing of attention: Visual attention, social cognition, and individual differences. Psychological Bulletin. 2007; 133:694–724. DOI: 10.1037/0033-2909.133.4.694 [PubMed: 17592962]

Greenough, WT., Black, JE., Wallace, CS. Experience and brain development. In: Johnson, M.Munakata, Y., Gilmore, RO., editors. Brain development and cognition: A reader. 2. Malden, MA: Blackwell; 2002. p. 186-216.

Guillon Q, Hadjikhani N, Baduel S, Rogé B. Visual social attention in autism spectrum disorder: Insights from eye tracking studies. Neuroscience & Biobehavioral Reviews. 2014; 42:279–297. DOI: 10.1016/j.neu-biorev.2014.03.013 [PubMed: 24694721]

Guthrie W, Swineford LB, Nottke C, Wetherby AM. Early diagnosis of autism spectrum disorder: Stability and change in clinical diagnosis and symptom presentation. Journal of Child Psychology and Psychiatry. 2012; 54:582–590. DOI: 10.1111/jcpp.12008

Haxby JV, Hoffman EA, Gobbini M. Human neural systems for face recognition and social communication. Biological Psychiatry. 2002; 51:59–67. DOI: 10.1016/s0006-3223(01)01330-0 [PubMed: 11801231]

Henderson JM. Regarding scenes. Current Directions in Psychological Science. 2007; 16:219–222. DOI: 10.1111/j.1467-8721.2007.00507.x

Hikosaka O, Nakamura K, Nakahara H. Basal ganglia orient eyes to reward. Journal of Neurophysiology. 2006; 95:567–584. DOI: 10.1152/jn.00458.2005 [PubMed: 16424448]

Hirose, Y., Tatler, B., Kennedy, A. Eye movement control in movies: The effect of visual and narrative continuity across a cut. Paper presented at the Perception; 2011.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., Van de Weijer, J. Eye tracking: A comprehensive guide to methods and measures. New York, NY: Oxford University Press; 2011.

Insel TR, Cuthbert BN, Garvey MA, Heinssen RK, Pine DS, Quinn KJ, … Wang PS. Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. American Journal of Psychiatry. 2010; 167:748–751. DOI: 10.1176/appi.ajp.2010.09091379 [PubMed: 20595427]

Jones W, Carr K, Klin A. 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. 2008; 65:946–954. DOI: 10.1001/archpsyc.65.8.946 [PubMed: 18678799]

Kanner L. Autistic disturbances of affective contact. Nervous Child. 1943; 2:217–250.

Kohls G, Schulte-Rüther M, Nehrkorn B, Müller K, Fink GR, Kamp-Becker I, … Konrad K. Reward system dysfunction in autism spectrum disorders. Social Cognitive and Affective Neuroscience. 2012; doi: 10.1093/scan/nss033

Lewkowicz DJ, Hansen-Tift AM. Infants deploy selective attention to the mouth of a talking face when learning speech. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109:1431–1436. DOI: 10.1073/pnas.1114783109 [PubMed: 22307596]

Lock EF, Hoadley KA, Marron J, Nobel AB. Joint and Individual Variation Explained (JIVE) for integrated analysis of multiple data types. The Annals of Applied Statistics. 2013; 7:523–542. DOI: 10.1214/12-AOAS597 [PubMed: 23745156]

Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism from 2 to 9 years of age. Archives of General Psychiatry. 2006; 63:694–701. DOI: 10.1001/archpsyc.63.6.694 [PubMed: 16754843]

Lord C, Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC, … Rutter M. The Autism Diagnostic Observation Schedule–Generic: A standard measure of social and communication deficits associated with the spectrum of autism. Journal of Autism and Developmental Disorders. 2000; 30:205–223. DOI: 10.1023/A:1005592401947 [PubMed: 11055457]

MathWorks. MATLAB 7.8 (Release R2009a). 2009. Retrieved from http://www.mathworks.com

Mullen, E. Mullen Scales of Early Learning. Circle Pines, MN: American Guidance Serivce; 1995.

Nakano T, Tanaka K, Endo Y, Yamane Y, Yamamoto T, Nakano Y, … Kitazawa S. Atypical gaze patterns in children and adults with autism spectrum disorders dissociated from developmental changes in gaze behaviour. Proceedings of the Royal Society B: Biological Sciences. 2010; 277:2935–2943. DOI: 10.1098/rspb.2010.0587 [PubMed: 20484237]

Norbury CF. Sources of variation in developmental language disorders: Evidence from eye-tracking studies of sentence production. Philosophical Transactions of the Royal Society B: Biological Sciences. 2014; 369:20120393.doi: 10.1098/rstb.2012.0393

Ozonoff S, Iosif AM, Baguio F, Cook IC, Hill MM, Hutman T, … Sigman M. A prospective study of the emergence of early behavioral signs of autism. Journal of the American Academy of Child & Adolescent Psychiatry. 2010; 49:256–266. DOI: 10.1016/j.jaac.2009.11.009 [PubMed: 20410715]

Paul R, Fuerst Y, Ramsay G, Chawarska K, Klin A. Out of the mouths of babes: Vocal production in infant siblings of children with ASD. Journal of Child Psychology and Psychiatry. 2011; 52:588–598. DOI: 10.1111/j.1469-7610.2010.02332.x [PubMed: 21039489]

Petersen SE, Posner MI. The attention system of the human brain: 20 years after. Annual Review of Neuroscience. 2012; 35:73–89. DOI: 10.1146/annurev-neuro-062111-150525

Rizzolatti G, Craighero L. The mirror-neuron system. Annual Review of Neuroscience. 2004; 27:169–192. DOI: 10.1146/annurev.neuro.27.070203.114230

Ronald A, Hoekstra RA. Autism spectrum disorders and autistic traits: A decade of new twin studies. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2011; 156:255–274. DOI: 10.1002/ajmg.b.31159

Sebanz N, Bekkering H, Knoblich G. Joint action: Bodies and minds moving together. Trends in Cognitive Sciences. 2006; 10:71–76. doi:10.1016.j.tics.2005.12.009.

Senju A, Johnson MH. Atypical eye contact in autism: Models, mechanisms and development. Neuroscience and Biobehavioral Reviews. 2009; 33:1204–1214. DOI: 10.1016/j.neubiorev.2009.06.001 [PubMed: 19538990]

Shepherd SV, Deaner RO, Platt ML. Social status gates social attention in monkeys. Current Biology. 2006; 16:R119–R120. DOI: 10.1016/j.cub.2006.02.013 [PubMed: 16488858]

Shic, F. Computational methods for eye-tracking analysis: Applications to autism. New Haven, CT: Yale University; 2008.

Shic F, Bradshaw J, Klin A, Scassellati B, Chawarska K. Limited activity monitoring in toddlers with autism spectrum disorder. Brain Research. 2011; 1380:246–254. DOI: 10.1016/j.brainres.2010.11.074 [PubMed: 21129365]

Shic, F., Chawarska, K., Lin, D., Scassellati, B. The computational modeling of perceptual biases of children with ASD in naturalistic settings. Paper presented at the International Meeting for Autism Research; London, UK. 2008.

Shic F, Macari S, Chawarska K. Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. Biological Psychiatry. 2014; 75:231–237. DOI: 10.1016/j.biopsych.2013.07.009 [PubMed: 23954107]

Sparrow, SS., Balla, DA., Cicchetti, DV. Vineland Adaptive Behavior Scales. 2. Circle Pines, MN: American Guidance Service; 2005.

Tatler, BW. Eye movements from laboratory to life. In: Horseley, M.Elliot, M.Knight, BA., Reilly, R., editors. Current trends in eye tracking research. New York, NY: Springer; 2014. p. 17-35.

Tatler BW, Hayhoe MM, Land MF, Ballard DH. Eye guidance in natural vision: Reinterpreting salience. Journal of Vision. 2011; 11(5):5.doi: 10.1167/11.5.5

Tatler BW, Hirose Y, Finnegan SK, Pievilainen R, Kirtley C, Kennedy A. Priorities for selection and representation in natural tasks. Philosophical Transactions of the Royal Society B: Biological Sciences. 2013; 368:20130066.doi: 10.1098/rstb.2013.0066

Team, R. D. C. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2011.

Tenenbaum EJ, Shah RJ, Sobel DM, Malle BF, Morgan JL. Increased focus on the mouth among infants in the first year of life: A longitudinal eye-tracking study. Infancy. 2013; 18:534–553. DOI: 10.1111/j.1532-7078.2012.00135.x [PubMed: 23869196]

Tenenbaum EJ, Sobel DM, Sheinkopf SJ, Malle BF, Morgan JL. Attention to the mouth and gaze following in infancy predict language development. Journal of Child Language. 2014; :1–18. DOI: 10.1017/S0305000914000725 [PubMed: 25023492]

Tomasello M. Culture and cognitive development. Current Directions in Psychological Science. 2000; 9:37–40. DOI: 10.1111/1467-8721.00056

Umilta MA, Kohler E, Gallese V, Fogassi L, Fadiga L, Keysers C, Rizzolatti G. I know what you are doing: A neurophysiological study. Neuron. 2001; 31:155–165. DOI: 10.1016/S0896-6273(01)00337-3 [PubMed: 11498058]

von Hofsten C, Uhlig H, Adell M, Kochukhova O. How children with autism look at events. Research in Autism Spectrum Disorders. 2009; 3:556–569. DOI: 10.1016/j.rasd.2008.12.003

Wagner DD, Kelley WM, Heatherton TF. Individual differences in the spontaneous recruitment of brain regions supporting mental state understanding when viewing natural social scenes. Cerebral Cortex. 2011; 21:2788–2796. DOI: 10.1093/cercor/bhr074 [PubMed: 21527789]

Young GS, Merin N, Rogers SJ, Ozonoff S. 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. 2009; 12:798–814. DOI: 10.1111/j.1467-7687.2009.00833.x [PubMed: 19702771]

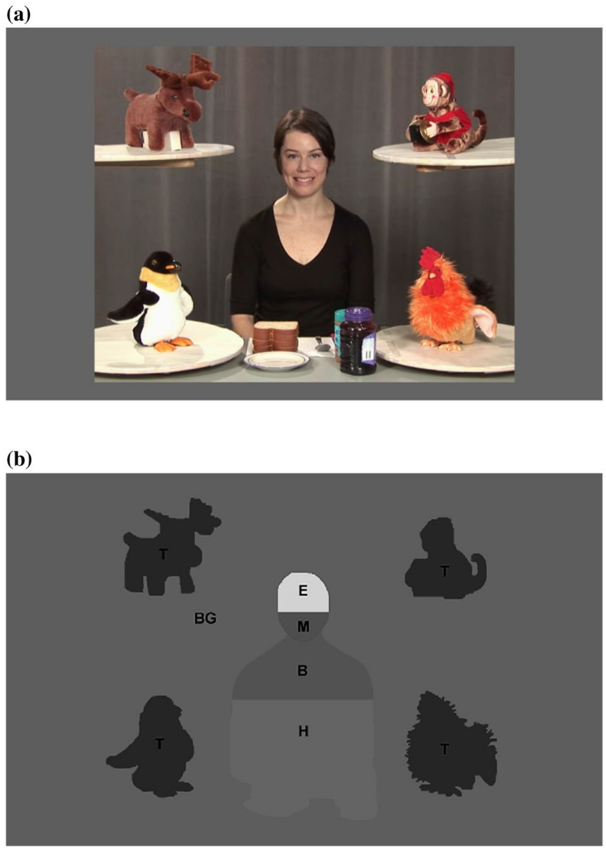

Figure 1.

(a) Frame from video stimulus used with (b) region of interest in the analysis. Regions were eyes (E), mouth (M), face (Eyes + Mouth), person (Face + Body + Hand/Activity), hand/activity area (H), toys (T), and background (BG).

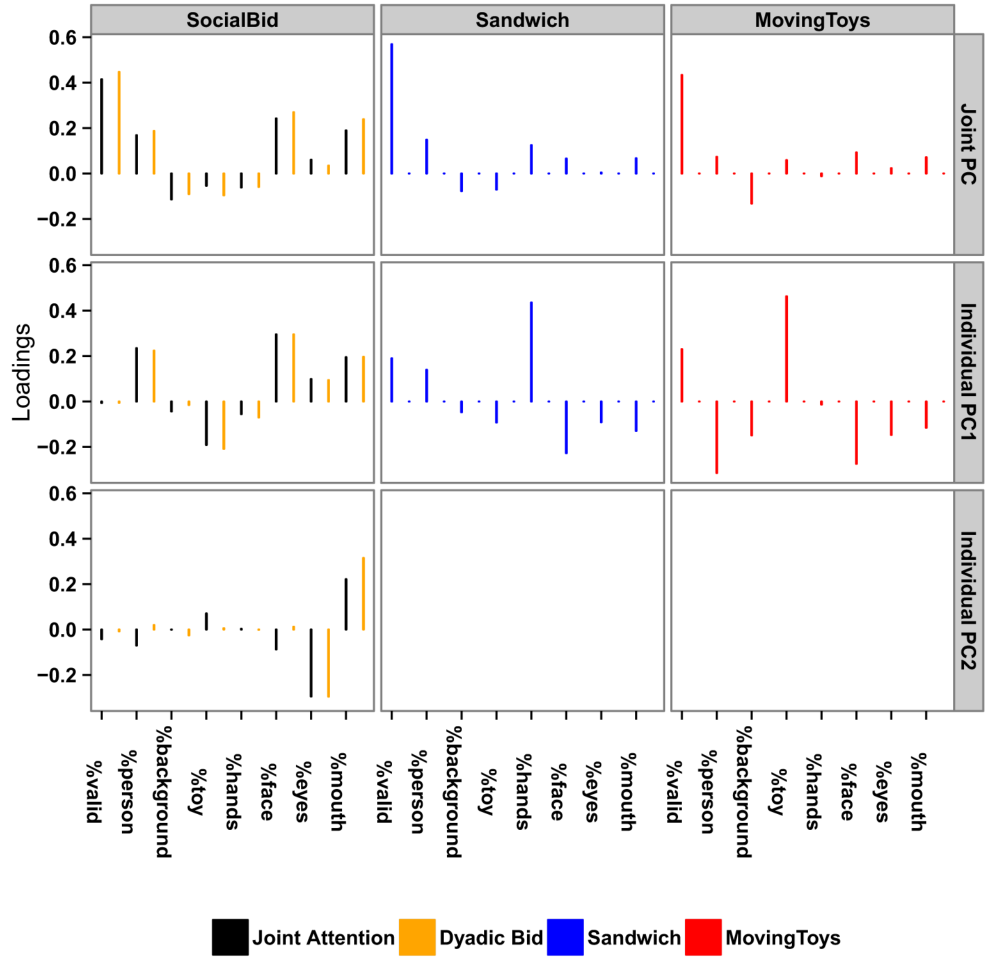

Figure 2.

Joint and individual variance explained latent structure analysis: Loadings for joint principal component (PC): Scene On-Off, two individual PCs for social bid condition: Face-Toys and Mouth-Eyes, one individual PC for sandwich condition: Activity-Face and one individual PC for moving toys condition: Toys-Person.

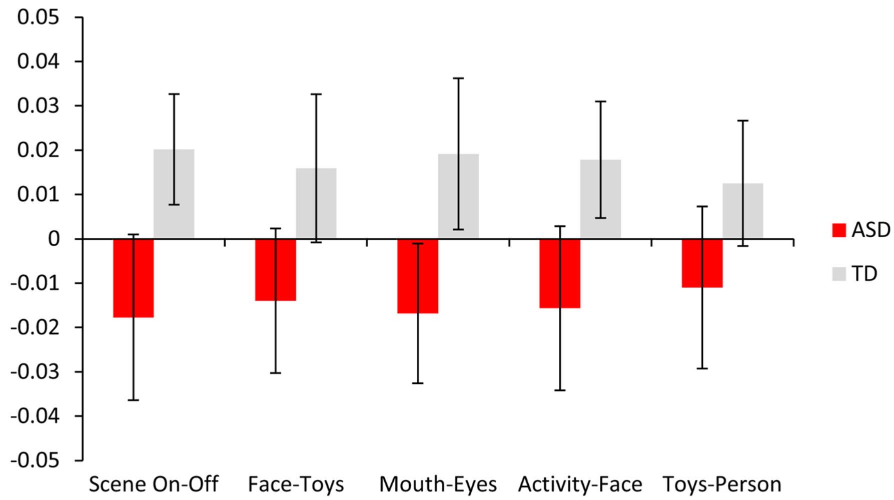

Figure 3.
Mean (± 2 *SE*) principal component scores in autism spectrum disorder (ASD) and typically developing (TD) groups.

Table 1

Ave"rage Percentage of Total Dwell Time in Specific Region of Interest (ROI) in Autism Spectrum Disorder (ASD) and Typically Developing (TD) Groups in Response to Dyadic Bid (DB), Joint Attention (JA), Sandwich (SA), and Moving Toys (MT) Conditions

		Condition			
		DB	**JA**	**SA**	**MT**
ROI	**Group**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**
%Scene	ASD	73 (21)	68 (28)	75 (23)	78 (21)
	TD	84 (15)	84 (17)	84 (15)	85 (17)
%Person	ASD	68 (15)	68 (21)	87 (10)	23 (13)
	TD	76 (11)	72 (6)	90 (8)	19 (10)
%Face	ASD	54 (19)	54 (25)	16 (11)	18 (13)
	TD	67 (14)	63 (17)	13 (7)	15 (9)
%Eyes	ASD	19 (14)	22 (17)	5 (5)	8 (9)
	TD	16 (14)	20 (17)	4 (5)	6 (6)
%Mouth	ASD	35 (17)	32 (22)	10 (9)	10 (8)
	TD	51 (16)	43 (16)	9 (5)	10 (7)
%Hands/Activity	ASD	10 (08)	10 (13)	64 (16)	3 (4)
	TD	7 (5)	6 (6)	72 (11)	2 (3)
%Toys	ASD	26 (13)	23 (15)	9 (7)	69 (18)
	TD	20 (10)	24 (14)	8 (8)	75 (14)
%Background	ASD	6 (6)	9 (14)	4 (5)	8 (10)
	TD	4 (4)	4 (6)	2 (2)	5 (10)

Note. DB and JA conditions were later combined into a single social bid condition in the latent structure analysis.

Table 2

Proportions of Variance Explained by the Joint and Individual Variance Principal Components (PCs) in the Social Bid, Sandwich, and Moving Toys Conditions

			Proportion of variance explained		
PC type	**Condition**	**PC**	**Social bid (%)**	**Sandwich (%)**	**Moving toys (%)**
Joint	All	Scene On–Off	35.2	38.1	22.9
Condition specific	Social bid	Face–Toys	23.3	—	—
		Mouth–Eyes	17.0	—	—
	Sandwich	Activity–Face	—	33.3	—
	Moving toys	Toys–Person	—	—	49.8
Total explained variance			75.5	71.4	72.7

Table 3

Joint and Individual Variance Explained Latent Structure Analysis: Principal Component (PC) Loadings in Social Bid, Sandwich, and Moving Toys Conditions

Region of interest	Joint PC Scene On–Off	Condition-specific PCs			
		Face–Toys	Mouth–Eyes	Activity–Face	Toys–Person
Social Bid					
%Scene	0.43	−0.01	−0.03	—	—
%Person	0.18	0.23	−0.03	—	—
%Background	−0.10	−0.03	−0.02	—	—
%Toys	−0.08	−0.20	0.04	—	—
%Hands/Activity	−0.06	−0.07	0.00	—	—
%Face	0.26	0.30	−0.04	—	—
%Eyes	0.05	0.10	−0.30	—	—
%Mouth	0.22	0.20	0.27	—	—
Sandwich					
%Scene	0.57	—	—	0.19	—
%Person	0.15	—	—	0.14	—
%Background	−0.08	—	—	−0.05	—
%Toys	−0.07	—	—	−0.09	—
%Hands/Activity	0.12	—	—	0.44	—
%Face	0.07	—	—	−0.23	—
%Eyes	0.00	—	—	−0.09	—
%Mouth	0.07	—	—	−0.13	—
Moving Toys					
%Scene	0.43	—	—	—	0.23
%Person	0.07	—	—	—	−0.31
%Background	−0.13	—	—	—	−0.15
%Toys	0.06	—	—	—	0.46
%Hands/Activity	−0.01	—	—	—	−0.01
%Face	0.09	—	—	—	−0.27
%Eyes	0.02	—	—	—	−0.15

Region of interest	Joint PC	Condition-specific PCs			
	Scene On–Off	Face–Toys	Mouth–Eyes	Activity–Face	Toys–Person
%Mouth	0.07	—	—	—	−0.12

Table 4

Partial (Age) Pearson's r Correlation Coefficient Analysis Between Principal Component Scores and Measure of Verbal and Nonverbal Functioning in TD and ASD Groups Combined As Well As Severity of Impairment in ASD Group

Measure	Scene On–Off	Face–Toys	Mouth–Eyes	Activity–Face	Toys–Person
ADOS–G Total[a]	**−0.31**	−0.01	−0.07	0.08	−0.16
	< 0.003	*ns*	*ns*	*ns*	*ns*
MSEL VDQ	**0.37**	0.11	**0.24**	0.22	**0.25**
	< 0.0001	*ns*	**< 0.003**	< 0.006	**< 0.002**
MSEL NVDQ	**0.38**	0.05	0.17	0.26	0.22
	< 0.001	*ns*	0.025	**< 0.001**	< 0.006

Note. Bonferroni correction for multiple comparisons was employed ($p < .0033$); (in bold). Correlations that were statistically significant but did not survive the Bonferroni correction are reported but not interpreted. ASD = autism spectrum disorder; TD = typically developing; ADOS–G = Autism Diagnostic Observation Schedule–Generic; MESL = Mullen Early Scales of Learning; VDQ = Verbal Developmental Quotient; NVDQ = Nonverbal DQ.

[a]ASD group only.