Published in final edited form as:

J Autism Dev Disord. 2014 March ; 44(3): 664–673. doi:10.1007/s10803-013-1925-5.

Bottom-Up Attention Orienting in Young Children with Autism

Dima Amso,

Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, 190 Thayer St., Box 1821, Providence, RI 02912, USA

Sara Haas,

Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, 190 Thayer St., Box 1821, Providence, RI 02912, USA

Elena Tenenbaum,

Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, 190 Thayer St., Box 1821, Providence, RI 02912, USA

Julie Markant, and

Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, 190 Thayer St., Box 1821, Providence, RI 02912, USA; Center for the Study of Children at Risk, Warren Alpert Medical School, Brown University, Providence, RI, USA

Stephen J. Sheinkopf

Center for the Study of Children at Risk, Warren Alpert Medical School, Brown University, Providence, RI, USA

Abstract

We examined the impact of simultaneous bottom-up visual influences and meaningful social stimuli on attention orienting in young children with autism spectrum disorders (ASDs). Relative to typically-developing age and sex matched participants, children with ASDs were more influenced by bottom-up visual scene information regardless of whether social stimuli and bottom-up scene properties were congruent or competing. This initial reliance on bottom-up strategies correlated with severity of social impairment as well as receptive language impairments. These data provide support for the idea that there is enhanced reliance on bottom-up attention strategies in ASDs, and that this may have a negative impact on social and language development.

Keywords

Bottom-up attention; Saliency; Visual attention; Autism; Eye tracking; Social attention

Introduction

The earliest description of children with autism spectrum disorders (ASDs) documented an atypical response to social stimuli and interactions (Kanner 1943). Since then, myriad

empirical papers verified this social disruption in this population (e.g., Dalton et al. 2005; Klin et al. 2002; Osterling and Dawson 1994; Sigman et al. 1986). However, differences in visual and attention abilities in ASDs, relative to typically developing (TD) children, have also been documented. The critical issue, addressed in this work, is whether social and visual influences have differential weight in guiding attention orienting in TD relative to young children diagnosed with ASDs.

Visual attention supports the selection of salient stimuli, and the suppression of interfering stimuli, from a cluttered visual world (Desimone and Duncan 1995). As such, it serves as an information filter, determining what is subsequently processed for perception and memory beginning early in postnatal life (e.g., Craik et al. 1996; Markant and Amso 2013). Visual processing plays an important role in bottom-up attention orienting. Bottom-up attentional search is driven by visually ‘salient’ events in the world (Itti and Koch 2001; Koch and Ullman 1985; Niebur and Koch 1996; Treisman and Gelade 1980; Wolfe 1994). Within this framework, perceptual information is processed in parallel over the entire visual field, giving rise to multiple visual feature maps (e.g., color, motion, intensity, orientation). These maps quantify the amount of a visual feature at each location in the field. Locations that differ from their surround are tagged. Individual feature maps are then combined into a single topographically organized ‘saliency map’ and attention orienting is biased to the ‘winner-take-all’ location with the highest saliency value. After the selection is made, an inhibitory process suppresses the attended location through an inhibition of return (IOR) mechanism, and the next salient location in the series is attended. Notably, this saliency framework is most relevant to bottom up attention in the initial processing of an image, where the first several fixations are directed to high saliency regions in sequence (Itti and Koch 2001; Koch and Ullman 1985; Niebur and Koch 1996; Parkhurst et al. 2002). Typically, top-down information, like meaningful stimuli or prior knowledge, serve to then guide attention orienting (Desimone and Duncan 1995; Itti and Koch 2001). The current work is concerned with precisely the interaction between bottom-up visual, and top-down social, influences on visual exploration in children with ASDs as compared to TD children.

Relative to TD participants, individuals with ASDs have been shown to attend to irrelevant aspects of the visual environment and over-perform on visual attention search tasks (Joseph et al. 2009; Kaldy et al. 2011; Mottron et al. 2006; Wang et al. 2007). O’Riordan and Plaisted (2001) provided an early suggestion that the ASD advantage in search derives from aberrant bottom-up visual processing of stimulus attributes. The idea is supported by vision research showing poorer chromatic discrimination (e.g., Franklin et al. 2008), enhanced orientation detection (e.g., Bertone et al. 2005), and lower dynamic contrast sensitivity thresholds (e.g., Bertone et al. 2005) in ASDs compared to TD groups. Other work has shown abnormal gamma activity over visual regions in adolescents with ASDs (Brown et al. 2005). The *Enhanced Perceptual Functioning* model (e.g., Mottron et al. 2006) suggests that these differences may stem from over-functioning of perceptual systems in autism. Joseph et al. (2009) used dynamic and static visual search tasks and showed that the visual search advantage in children with ASDs, relative to TD children, derived from differences in enhanced discrimination of stimulus features. Keehn et al. (2008) used event-related functional magnetic resonance imaging (fMRI) in similar static visual search task. They

found increased occipital and frontoparietal activation in ASDs relative to TD groups. The increased activation of occipital cortex is consistent with the idea that superior visual search abilities may result from enhancements in low level visual processing. Frontoparietal activation differences indicate increases in top-down modulation of visual attention in the ASDs group.

Freeth et al. (2011) examined bottom-up attention orienting in static natural scenes using the saliency framework in adults with ASDs. Participants freely viewed color photographs, all of which had human faces in them. They found very few differences between groups when the head areas were not salient, as determined by a saliency algorithm (Saliency Toolbox, Walther and Koch 2006). When heads were visually salient, typical adults arrived at the face on average one fixation earlier than adults with ASDs. However, there is a lifetime of development and visual experience between young children and adults and it is not clear that patterns of performance from one age group can be applied to another. Therefore, we revisited these questions using a similar paradigm in a group of children, recently diagnosed with ASDs, as compared to age and sex matched TD children.

Specifically, this work examined whether there are differences in bottom-up attention orienting in minimally verbal early diagnosed children with ASDs, relative to TD children, and whether this is modulated by the presence of socially meaningful information. In order to examine the possibility that children with ASDs use exclusively bottom-up attention orienting mechanisms for visual exploration, attention orienting data from both the initial viewing interval and for the duration of trials were examined. We relied on the standard saliency algorithm (Saliency Toolbox, Walther and Koch 2006) to generate two analytic strategies. The first is the prevailing approach, wherein the algorithm was used to determine the ‘winner-take-all’ most salient areas in each scene, some of which will be social and others nonsocial in content. This addressed whether children with ASDs and TD children differ in their attentional selection of this bottom-up salient information in general, and as a function of whether the salient region is social.

The second analytic strategy used the saliency algorithm in a novel way. It examined the predictive strength of each color, orientation, intensity, and saliency maps on free viewing in group-level analyses. As noted, visual processing has been shown to be atypical in ASDs. Using this analysis, the study addressed whether children diagnosed with ASDs and TD children weight individual visual features differently to drive scanning in natural scenes. Bottom-up and social attention processing necessarily occur in parallel at all times. Therefore, we additionally asked whether bottom-up orienting differed when bottom-up and social attention were congruent (faces are also the most salient parts of the display) and when social information was in competition with bottom-up information. Finally, the visual attention system develops early and visual exploration strategies have been shown to impact what information is selected for subsequent perception and learning as early as 3 months of age (Amso and Johnson 2006). Therefore, observed visual exploration strategies in children with ASDs were correlated with performance on clinically obtained measures of language and social interaction.

Methods

Participants

See Table 1 for details on group characteristics. Fifteen predominantly non-verbal to minimally verbal children aged 2–5 years with a diagnosis of an autism spectrum disorder (Autistic Disorder: $N = 14$, PDD-NOS: $N = 1$) were recruited from early intervention centers, medical and mental health clinics and by referral from local pediatricians. Fifteen TD age and sex-matched participants aged 2–5 years were included in the sample for comparison. TD children were recruited from the community via advertisements and birth records. TD children were screened prior to enrollment to ensure that they were eligible to be in the study. Children were excluded from the study if they had been born early (<36 weeks), had low birth weight (<5 lbs), or had serious birth complications or a history of psychiatric or neurological problems.

Diagnostic Procedures

Diagnoses were confirmed with Module 1 ($N = 14$) or Module 2 ($N = 1$) of the Autism Diagnostic Observation Schedule (ADOS, Lord et al. 2000, see Table 1). Participants with ASDs were tested on this paradigm as part of a larger study battery. As such, they underwent extensive cognitive and language testing, supervised by a licensed clinical psychologist. Cognitive IQ testing involved either the Bayley Scales of Infant Development, 3rd Edition (Bayley 2005), or the Wechsler Preschool and Primary Scale of Intelligence-III (5 children, Wechsler 2002), depending on the age and verbal ability of the child. All children with an ASD were tested using the Preschool Language Scales (PLS)—4th Edition (Zimmerman et al. 2002). If cognitive or language testing had been completed by a licensed professional within the last 6 months, the child's existing scores were used. In one case, this resulted in scores obtained using the 5th Edition of the PLS.

Materials

Eye Tracking Apparatus—Children sat about 70 cm from a 22" (55.9 cm) monitor. Eye movements were recorded using a remote eye tracker (SMI SensoMotoric Instruments RED system). In addition, a standard digital video camera (Canon ZR960) was placed above the computer screen to record children's head movements. All calibration and task stimuli were presented using the Experiment Center software from SMI. Before beginning the task, point-of-gaze (POG) was calibrated by presenting an attractive, looming stimulus in the upper left and lower right corners of the screen. The same calibration stimulus was then presented in the four corners of the screen in order to validate the accuracy of the calibration.

Free Viewing Task—Participants were eye tracked as they freely scanned sixteen color photographs ($1,680 \times 1,050$ pixels), 12 experimental and 4 foil images, depicting both indoor and outdoor scenes (see Fig. 1 for examples). All images contained a person. All images were presented for 5 s. Perfect control over natural scenes is not possible. Nonetheless, in selecting the 12 experimental images for this study, we attempted to include images with similar layout and complexity, that were indoor or outdoor, that had similar size of face, and whether eyes were looking forward or to the side etc. Four foil images were included in an attempt to prevent participants from becoming clued into searching for

prominent faces in the 12 experimental images. These foils also contained people, but these were not prominent or central to the scene. The order of image presentation was randomized. A central fixation target was used to return participants' POG to the center of the screen between images.

Data Preprocessing

Proportion Attended Salient Areas of Interest (AOIs)—This study examined possible differences across groups in bottom-up attention orienting to visually salient social and nonsocial stimuli. As such, analyses were conducted for the dependent variables for the first second of each trial, accounting for the first several fixations of viewing per image (Freeth et al. 2011), as compared to the total trial duration.

AOIs that encompass the human 'head' were defined as the face and hair regions combined. However, we refer to it as 'face' in the remainder of the document to keep with convention. In one set of 6 images (labeled Congruent), the face was identified as the 'winner-take-all' most visually salient location (Face Primary AOI, Figs. 1a, 2a). In the remaining 6 images (labeled Incongruent), a non-face AOI was the 'winner-take-all' most visually salient location or Non-Face Primary AOI (Figs. 1b, 2b). Secondary AOIs were also defined for comparison. Secondary AOIs were regions identified by the Saliency Toolbox to be visually prominent. The Saliency Toolbox generated the 'winner-take-all' most salient AOI and the next 4 visually salient locations per image. These are illustrated in Fig. 2b, for example, as the AOIs with the cooler colors. Both Face and Non-Face Secondary AOIs were drawn from this group and were roughly matched based on area size. The percent of image occupied by Primary AOIs was $M = 4\%$, $SD = 3\%$ of images in the Congruent condition and $M = 3\%$, $SD = 1\%$ in the Incongruent condition. The size of Secondary AOIs encompassed $M = 3\%$, $SD = 1\%$ in the Congruent condition, and $M = 3\%$, $SD = 2\%$ in the Incongruent condition. The proportion of images that each child looked at these Primary and Secondary AOIs was calculated for the following: Face Primary AOI, Non-Face Primary AOI, Face Secondary AOI, and Non-Face Secondary AOIs. This was calculated as the proportion of trials in which the participant fixated the AOI type at least once during the 1 s interval and for the full trial duration (calculated as number of trials attended/total possible trials). A fixation was defined as 100 ms per 100 pixel dispersion.

Beta Coefficients—First, a down-sampled 16×25 region matrix grid for each image was calculated (Fig. 3a). Next, duration of looking data for each matrix grid location (MGL) in each image per subject (Fig. 3c) was extracted, effectively producing a gaze distribution map per subject per image. We then used Saliency Toolbox to extract map values for color, intensity, orientation, and their linear combinations for each location in the 16×25 matrix grid for each image (Fig. 3b). These conspicuity or 'feature' maps delineate MGLs where there is a difference in value from one square to its nearest neighbor along these visual feature dimensions. We then determined the fit of the looking data to both the individual feature maps and the saliency map. Specifically, multiple regressions were used to generate beta coefficients per subject and image for each map. The dependent variable was the child's duration of looking per MGL in the 16×25 matrix. The predictors were color, orientation, intensity, their individual linear combinations, and the overall saliency maps per MGL, as

extracted from the Saliency Toolbox. These resulting coefficients per participant and image were then used in ASD versus TD group level analyses. Beta values were derived 4 times per participant for all images separately, first including the MGLs encompassing the face and separately again deleting the face MGLs prior to running the individual regressions. This procedure was repeated for the first 1 s of the display and again for the total 5 s duration of each trial. This resulted in beta coefficients for color, intensity, orientation, and overall saliency maps for (1) face present and 1 s duration (2) face absent and 1 s duration (3) face present and total duration, and (4) face absent and total duration.

Results

Oculomotor Metrics

Preliminary analyses examined possible differences in oculomotor disengagement that might bear on free viewing performance among ASD and TD groups. Previous work has shown that ASD groups have difficulty with disengagement of fixations (e.g., Elsabbagh et al. 2009; Elison et al. 2013). Hence, we asked whether ASD and TD Groups differed in the number and duration of fixations during free viewing. Fixations are defined as 100 ms per 100 pixel dispersion. Multivariate analysis of variance examined the dependent variables of mean fixation count and mean fixation duration calculated per individual for the total duration of each trial (5 s) and separately for the first 1 s of each trial. We found statistically marginal effects (Bonferroni alpha corrected to $p = .01, .05/4$) of Group for mean fixation count in the first second of each trial, $F(1,27) = 5.79, p < .05, \eta_p^2 = .18$ (ASD $M = 2.67$ fixations, $SD = .44$, TD $M = 3.03$ fixations, $SD = .37$). In order to be statistically conservative, we model the mean fixation count metric in subsequent analyses in order to examine and account for its effects on bottom-up guided attention orienting across groups.

Proportion Attended AOIs

We examined how bottom-up processing influences attention orienting across groups in the first few fixations after display presentation (1 s) and for the total trial duration (5 s). This was done to consider the possibility that children may orient based on bottom-up visual input throughout the trial. We compared proportion of attended AOIs as a function of AOI Type (two levels: Face and Non-Face) $\times$ Salience Level (two levels: Primary and Secondary AOIs) $\times$ Time (two levels: 1 s and Total Duration) across Groups (two levels: ASD and TD) with Age and Average Fixation Count for the 1 s interval treated as continuous variables in the model. Only statistically reliable results pertaining to the Group variable are reported. The analysis resulted in an interaction between Salience Level, AOI Type, Time, and Group, $F(1,26) = 8.01, p < .01, \eta_p^2 = .24$. Follow-up analyses considered AOI Type and Salience Level for both Groups in the two time intervals separately (all Bonferroni alpha levels set to $p = .01, .05/4$). In the 1 s interval, children with ASDs showed reliable main effects of AOI Type, $F(1,14) = 85.83, p = .000, \eta_p^2 = .86$ and Salience, $F(1,14) = 16.08, p = .001, \eta_p^2 = .54$. Face AOIs were attended on proportionally more trials ($M = 63\%$) than were Non-Face stimuli ($M = 7\%$). With respect to visual salience, Primary AOIs ($M = 41\%$) were also attended more than Secondary AOIs (29 %) regardless of AOI Type. TD children showed only a similar main effect of AOI Type, $F(1,14) = 82.12, p = .000, \eta_p^2 = .85$, with Faces ($M =$

62 %) attended on proportionally more trials than Non-Face AOIs ($M = 4\%$). Figure 4 illustrates these results. This is our first indication that children with ASDs use bottom-up attention to a greater degree than TD children. The same analysis for the full 5 s trial duration yielded only a main effect of AOI Type for both ASD, $F(1,14) = 116.81, p = .000, \eta_p^2 = .89$ and TD Groups, $F(1,14) = 221.76, p = .000, \eta_p^2 = .94$.

Beta Coefficients

In order to use a more precise metric to ask whether there are bottom-up differences in attention orienting between groups, we compared ASD and TD children's beta coefficients for the saliency map, as well as for color, orientation and intensity individual feature maps. Recall that these beta coefficients represent the explanatory power of each feature, and the combined saliency map value, on the distribution of looking. For example, a high color beta coefficient means that a child used color to guide attention orienting, while a low intensity coefficients suggests they did not use intensity. The values naturally vary with image statistics. For example, color might be dominant in one image and orientation in another. As such, this analysis is only valuable for understanding relative *group differences* in free viewing. We report here only statistics relevant to the Group variable and its interactions.

We generated these beta values twice, once with the matrix grid locations or MGLs encompassing the face, and again deleting the face MGLs from the matrix prior to running the individual regressions. This was intended to answer the question: How was the remainder of the image scanned in the presence of a social stimulus? We were additionally interested in whether different patterns of gaze distribution emerged as a function of whether faces were congruent or competing with bottom-up image salience statistics. Those images where the face is salient Primary AOI are labeled Congruent (Figs. 1a, 6 images), as bottom-up and socially meaningful information are the same in this case. Incongruent images (Figs. 1b, 6 images) are images where there is a conflict between bottom-up information, a Non-Face Salient Primary AOI, and the competing presence of the face.

Our initial analysis was an omnibus ANCOVA with four within and one between- subjects factors. Specifically, we included the within subjects factors of Time (two levels: 1 s and Total Duration), Image Type (two levels: Congruent and Incongruent), Matrix Type (two levels: Face Present or Face Absent), and Coefficient Maps (four levels: saliency, color, orientation, intensity), and the between subjects factor of Group (two levels: ASD and TD) with Age and Average Total Fixation Counts modeled as continuous variables. We found a five way interaction of Time, Image Type, Matrix, Maps, and Group, $F(3,78) = 3.11, p < .05, \eta_p^2 = .11$. Notably, there were no Maps by Group interaction, $F(3,78) = .63, p = ns$, and no main effect of Group, $F(1,26) = .08, p = ns$, in this model. This indicates that there were no differences in the explanatory power of color, orientation, intensity, and overall saliency map metrics across Groups on bottom-up attention that were independent of image content.

We followed up on the 5-way interaction, separating analyses into Face Present and Face Absent Matrix types (all Bonferroni corrected alpha levels set to $p = .025, .05/2$). The ANCOVAs had three within and one between- subjects factors. We included the within subjects factors of Time (two levels: 1 s and Total Duration), Image Type (two levels:

Congruent and Incongruent), and Coefficient Maps (four levels: saliency, color, orientation, intensity), and the between subjects factor of Group (two levels: ASD and TD) with Age and Average Total Fixation Counts as covariates. The analysis of data with Face MGLs Present resulted in no main effects and interactions relevant to Group. Figure 5 illustrates distribution of looking across both Congruent and Incongruent conditions. The groups used image properties similarly and had no difference in bottom-up distribution of attention orienting relevant to image visual statistics. Both groups had higher beta coefficients for the Congruent condition and lower beta coefficients for the Incongruent condition. Overall proportion of total looking at the face (over 5 s) in the Incongruent condition was $M = 25\%$, $SD = .03$ for the children with ASDs, and $M = 27\%$, $SD = .03$ for TD children. Therefore, low beta coefficients in the Incongruent condition were driven by looking at the social stimulus. These data, however, do not provide clear insight into how the remainder of the scene is being processed in the presence of a social stimulus.

How is bottom-up attention orienting distributed in the presence of meaningful social stimuli? In the Face Absent analysis, we found a Maps, Time, and Group interaction, $F(3, 78) = 4.75$, $p < .005$, $\eta_p^2 = .15$. We conducted simple effects tests separately on each feature Coefficient Map by Time and Group (Bonferroni corrected alpha set to $p = .01$, $.05/4$). The Time by Group effect was specific to the Saliency Coefficient Map, $F(1, 26) = 7.74$, $p = .01$, $\eta_p^2 = .23$. Figure 6 shows that, only in children with ASDs, the use of bottom-up saliency information to guide attention orienting was higher for the first second than for the total duration of trials, $t(14) = 2.19$, $p < .05$. One-sample t-tests (Bonferroni corrected alpha set to $p = .01$, $.05/4$) verified that children with ASDs had beta coefficients statistically higher than 0 for the 1 s interval, $t(14) = 3.18$, $p < .01$, but not for the Total Duration, $t(14) = .65$, $p = ns$. TD children did not have saliency beta coefficient values statistically higher than 0 for either the 1 s interval, $t(14) = .24$, $p = ns$, or the Total Duration, $t(14) = 1.55$, $p = ns$.

Taken together with the Face Present data, TD children are heavily influenced by the presence of the face for attention orienting and show little evidence of bottom-up reliance unless faces also happen to be the ‘winner-take-all’ most salient AOI. Children with ASDs are also heavily influenced by the presence of the face, but also attend to the remaining bottom-up visual information in the scene that is nonsocial in nature in the first few fixations after image presentation.

We reasoned that this reliance on bottom-up driven orienting in the presence of social stimuli might bear negatively on how social information is gathered from the environment (Amso and Johnson 2006; Schlesinger et al. 2007). Therefore, we generated a difference score for the saliency map beta coefficient for Face Present minus Face Absent values for the first second of viewing. A high positive difference score suggests that beta coefficients were driven by participants looking predominately at faces and not elsewhere. A zero value suggests they distributed looking evenly across the face MGLs and elsewhere in the matrix. A negative score suggests that individuals were looking mostly off of the face. We focus on the Congruent Image Type only, because the face region was not high in beta values in the Incongruent condition, and therefore looking at that region would not result in increases in beta coefficients, rendering the subtraction difficult to interpret. We asked if these scores

bore any relationship to data collected from the PLS language inventories or ADOS scores from the revised ADOS algorithm (Gotham et al. 2007). Partial correlations, controlling for Age and Average Fixation Counts in the 1 s interval, revealed a negative relationship between ADOS Social Affect scores and the Congruent saliency beta coefficient difference scores, $r(11) = -.62, p < .05$, and a trending positive relationship between this metric and the PLS Receptive, $r(11) = .53, p = .06$, language scores. As the difference score approached zero, meaning that participants were looking more off of the salient face Primary AOI and at other bottom-up salient image Secondary AOIs, the Social Affect score increased, indicating severity on this metric. Conversely, as the difference score indicated sustained attention to the face, more consistent with the TD pattern, receptive language values increased (Fig. 7).

Discussion

There is an attempt to shift diagnosis and treatment of developmental disorders to their earliest possible manifestations, with the goals of correcting atypical developmental trajectories and improving long-term developmental outcomes. As such, a growing literature has considered visual and attentional development in individuals with ASDs, reasoning that these mechanisms emerge early in postnatal development and function primarily to gather information for subsequent perception and learning. There is strong evidence in the literature that the oculomotor component of visual attention is atypical in young children with ASDs and infants at-risk for the disorder (Elison et al. 2013; Elsabbagh et al. 2009). However, there is less data about the precise bottom-up computations that determine *what* is looked at in free viewing in young children with ASDs, rather than *how* the gaze shift is executed. These data from minimally verbal young children diagnosed with ASDs shed light on this issue.

Overall, our data suggest that, relative to TD children, children with ASDs rely more on bottom-up information for attention orienting. First, Children with ASDs had a higher proportion of selecting visually salient image regions than did TD children, regardless of social content (Fig. 3). Second, while we found no evidence that attending to faces differed across groups, children with ASDs additionally continued to scan the scene consistent with secondary visual feature salience in the first few fixations after image presentation, as revealed by the beta coefficient analyses (Fig. 6). This suggests that whatever biased TD children to sustain attention to faces in the initial viewing period did not operate in children with ASDs. It is not clear whether the pattern in children with ASDs was driven by an inability to inhibit the bottom-up influence when a social stimulus was encountered, or whether the meaningfulness of the social stimulus differed across groups.

This interesting pattern of results may have important implications for information gathering and learning in children with ASDs. Most of the visual attention differences reported in the literature are subtle, and these data are no different. It seems there is a reliance on bottom-up input for orienting in this group in the first few fixations of scene viewing. We suggest that social interactions are complex and often subtle, and natural scenes can change on a moment-to-moment basis as a result of the motion of the stimulus or eye and body movements of the viewer. Therefore, the 5 s opportunity to freely scan a static image afforded in this experiment is not often present in the natural world. Shifts away from the

social stimulus during the initial viewing period may mean fewer opportunities to learn about it in ASDs, as natural scenes quickly change and new bottom-up salient and attention-attracting information becomes available. Consistent with this idea are the correlations between this effect and ADOS Social Affect score and the PLS Receptive language measures (Fig. 7). Data showed a negative relationship between ADOS Social Affect score and the saliency map beta difference scores in the initial 1 s interval for congruent images. As the difference score value indicated that children with ASDs were distributing looking evenly to face and other salient image regions, the ADOS Social Affect score increased indicating severity on this metric. In addition, as the difference score indicated sustained attention to the face, Receptive PLS language values increased. Although correlational, these results are consistent with research that has demonstrated the importance of attending to social information for learning about language (Brooks and Meltzoff 2005, 2008; Carpenter et al. 1998; Morales et al. 2000; Mundy et al. 2007).

It is unlikely that the observed patterns of bottom-up dominance in ASDs was driven by oculomotor differences between groups. The only oculomotor difference we found across our groups involved a higher average fixation count in TD children. This is consistent with previous literature showing longer oculomotor disengagement times in children with ASDs (Elsabbagh et al. 2009; Elison et al. 2013). However, if children with ASDs had difficulty disengaging from a salient face in the Congruent condition, for example, they should have had little bottom-up driven exploration of the remainder of the image, rather than the pattern observed of greater bottom-up driven orienting than TD children (Fig. 6).

We highlight possible limitations of this work. First, we are limited by the lack of naturalistic motion in our displays. The addition of motion in scenes will be very important in future work. A second limitation is the lack of ADOS and PLS measures in TD children. We cannot at this point speak to whether the observed correlations with language measures are part of a spectrum of typical language development or whether they relate specifically to patterns in ASDs. Future work will examine these and other important issues related to the interaction of visual and social attention across development.

Finally, we focused only on images that had a social stimulus, as our primary question was focused on how bottom-up and socially meaningful information interacts to guide attention orienting. As such, we over weighted the need to get as much eye tracking data as was feasible from this minimally verbal group using images that allowed us to examine our questions. As a consequence, we cannot speak to whether the same ASD reliance on bottom-up influences would be apparent in the absence of a social stimulus. Regardless, natural scenes tend to involve social stimuli, and attention to these stimuli is critical to learning about the world. Data on the developmental course of influences on visual attention would have implications for both an understanding of the emergence of autism over time and potentially on improved treatment targets for very early intervention. Future work will examine how early this reliance on bottom-up attention is observable in infants at risk for ASDs and whether this can serve as a biomarker for the disorder.

Acknowledgments

Special gratitude to the Center for the Study of Children at Risk at Brown University and the children and families who participated in this research. Also thanks to the James S. Mc-Donnell Award in Human Cognition (DA) and Autism Speaks Translational Postdoctoral Fellowship (ET).

References

- Amso D, Johnson SP. Learning by selection: Visual search and object perception in young infants. *Developmental Psychology*. 2006; 42(6):1236–1245. [PubMed: 17087555]
- Bayley, N. Bayley scales of infant development. 3rd ed. PsychCorp; San Antonio: 2005.
- Bertone A, Mottron L, Jelenic P, Faubert J. Enhanced and diminished visuospatial information processing in autism depends on stimulus complexity. *Brain*. 2005; 128:2430–2441. [PubMed: 15958508]
- Brooks R, Meltzoff AN. The development of gaze following and its relation to language. *Developmental Science*. 2005; 8(6):535–543. [PubMed: 16246245]
- Brooks R, Meltzoff AN. Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language*. 2008; 35(1):207–220. [PubMed: 18300435]
- Brown C, Gruber T, Ripon G, Brock J, Boucher J. Gamma abnormalities during perception of illusory figures in autism. *Cortex*. 2005; 41(2005):364–376. [PubMed: 15871601]
- Carpenter M, Nagell K, Tomasello M. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*. 1998; 63(4):i–vi. 1–143. [PubMed: 9835078]
- Craik FIM, Govoni R, Naveh-Benjamin M, Anderson ND. The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*. 1996; 125:159–180. [PubMed: 8683192]
- Dalton KM, Nacewicz BM, Johnstone T, Schaefer HS, Gernsbacher MA, Goldsmith HH, et al. Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience*. 2005; 8(4):519–526.
- Desimone R, Duncan J. Neural mechanisms of selective visual attention. *Annual Reviews of Neuroscience*. 1995; 18:193–222.
- Elison JT, Paterson SJ, Wolff JJ, Reznick JS, Sasson NJ, Gu H, et al. White matter microstructure and atypical visual orienting in 7-month-olds at risk for autism. *American Journal of Psychiatry*. 2013; 170(8):899–908. [PubMed: 23511344]
- Elsabbagh M, Volein A, Holmboe K, Tucker L, Csibra G, Baron-Cohen S, et al. Visual orienting in the early broader autism phenotype: Disengagement and facilitation. *Journal of Child Psychological Psychiatry*. 2009; 50(5):637–642.
- Franklin A, Sowden P, Burley R, Notman L, Alder E. Color perception in children with autism. *Journal of Autism and Developmental Disorders*. 2008; 38:1837–1847. [PubMed: 18449634]
- Freeth M, Foulsham T, Chapman P. The influence of visual saliency on fixation patterns in individuals with Autism Spectrum Disorders. *Neuropsychologia*. 2011; 49(1):156–160. [PubMed: 21093466]
- Gotham K, Risi S, Pickles A, Lord C. The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*. 2007; 37(4): 613–627. [PubMed: 17180459]
- Itti L, Koch C. Computational modeling of visual attention. *Nature Reviews Neuroscience*. 2001; 2:194–203.
- Joseph R, Keehn B, Connolly C, Wolfe JM, Horowitz TS. Why is visual search superior in autism spectrum disorder? *Developmental Science*. 2009; 12:1083–1096. [PubMed: 19840062]
- Kaldy Z, Kraper C, Carter AS, Blaser E. Toddlers with autism spectrum disorder are more successful at visual search than typically developing toddlers. *Developmental Science*. 2011; 14(5):980–988. [PubMed: 21884314]
- Kanner L. Autistic disturbances of affective contact. *Nervous Child*. 1943; 2:217–250.

- Keehn B, Brenner L, Palmer E, Lincoln AJ, Müller R. Functional brain organization for visual search in ASD. *Journal of the International Neuropsychiatric Society*. 2008; 14:990–1003.
- Klin A, Jones W, Schultz R, Volkmar F, Cohen D. Defining and quantifying the social phenotype in autism. *American Journal of Psychiatry*. 2002; 159(6):895–908. [PubMed: 12042174]
- Koch C, Ullman S. Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*. 1985; 4:219–227. [PubMed: 3836989]
- Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, et al. 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(3):205–223. [PubMed: 11055457]
- Markant J, Amso D. Selective memories: Infants' encoding is enhanced in selection via suppression. *Developmental Science*. 2013 doi:10.1111/desc.12084.
- Morales M, Mundy P, Delgado CEF, Yale M, Messinger D, Neal R, et al. Responding to joint attention across the 6-through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology*. 2000; 21(3):283–298.
- Mottron L, Dawson M, Soulières I, Hubert B, Burack J. Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*. 2006; 36:27–43. [PubMed: 16453071]
- Mundy P, Block J, Delgado C, Pomares Y, Van Hecke AV, Parlade MV. Individual differences and the development of joint attention in infancy. *Child Development*. 2007; 78(3):938–954. [PubMed: 17517014]
- Niebur, E.; Koch, C. Control of selective visual attention: Modeling the 'where' pathway. In: Touretzky, D.; Mozer, M.; Hasselmo, M., editors. *Neural information processing systems*. MIT Press; Cambridge, MA: 1996. p. 802-808.
- O'Riordan M, Plaisted K. Enhanced discrimination in autism. *Journal of Experimental Psychology*. 2001; 54A:961–979.
- Osterling J, Dawson G. Early recognition of children with autism: A study of first birthday home videotapes. *Journal of Autism and Developmental Disorders*. 1994; 24(3):247–257. [PubMed: 8050980]
- Parkhurst D, Law K, Niebur E. Modeling the role of salience in the allocation of overt visual selective attention. *Vision Research*. 2002; 42:107–123. [PubMed: 11804636]
- Schlesinger M, Amso D, Johnson SP. Simulating the role of visual selective attention during the development of perceptual completion. *Developmental Science*. 2007; 15(6):739–752. [PubMed: 23106728]
- Sigman M, Mundy P, Sherman T, Ungerer J. Social interactions of autistic, mentally retarded and normal children and their caregivers. *Journal of Child Psychology and Psychiatry*. 1986; 27(5): 647–655. [PubMed: 3771681]
- Treisman A, Gelade G. A feature-integration theory of attention. *Cognitive Psychology*. 1980; 12(1): 97–136. [PubMed: 7351125]
- Walther D, Koch C. Modeling attention to salient proto-objects. *Neural Networks*. 2006; 19(9):1395–1407. [PubMed: 17098563]
- Wang J, Clementz BA, Keil A. The neural correlates of feature-based selective attention when viewing spatially and temporally overlapping images. *Neuropsychologia*. 2007; 45:1393–1399. [PubMed: 17161441]
- Wechsler, D. Wechsler preschool and primary scale of intelligence, third edition (WPPSI-III). Pearson; San Antonio: 2002.
- Wolfe JM. Guided search 2.0. A revised model of visual search. *Psychology Bulletin and Review*. 1994; 1(2):202–238.
- Zimmerman, IL.; Steiner, VG.; Pond, RE. Preschool language scale, fourth edition (PLS-4). Pearson; San Antonio: 2002.

Fig. 1.

Depicts sample images shown to participants. **a** Congruent Condition—Faces are ‘winner-take-all’ most salient (Face Primary AOI). **b** Incongruent Condition—Faces are not ‘winner-take-all’ most salient (Face Secondary AOI)

Fig. 2.

a Example of Congruent condition and associated heat map. Intensity bar on the left indicates saliency values. Hot colors represent higher saliency values. The face region is the ‘winner-take-all’ salient area in the Congruent condition. **b** Example of Incongruent condition and associated heat map. The ‘winner-take-all’ salient region is not the face in this condition. All heat maps derived from MATLAB SaliencyToolbox (Walther and Koch 2006)

Fig. 3.

a Input image with superimposed 16×25 matrix grid. **b** Individual feature maps. Hotter colors represent locations that differ from their neighbors along each intensity, orientation, or color dimension. Values were extracted from the Saliency Toolbox for each location in the 16×25 matrix grid per feature map. These were used as predictors to derive beta coefficients for each individual subject's data, illustrated in **(c)**

Fig. 4.
Depicts proportion of salient primary AOIs versus secondary salient AOIs attended in children with ASDs relative to TD children

Fig. 5.

Beta coefficients for the total trial durations in the face present analysis by group for both congruent and incongruent conditions

Fig. 6.
Saliency map beta coefficient data with faces absent from the matrix prior to individual subject regressions

Fig. 7.

Correlations between the difference score of saliency map values in the congruent condition (faces present–faces absent) and **a** ADOS social affect scores and **b** PLS receptive language scores in children with ASDs. High positive values along the x -axis indicate a pattern more consistent with TD children

Table 1

Age ranges, sex, and ADOS test scores (ASD only)

	ASD (<i>N</i> = 15)	TD (<i>N</i> = 15)
Age	2–5 years, 3.53 years (1.06)	2–5 years 3.73 years (1.03)
Sex	9 Male, 6 female	9 Male, 6 female
ADOS		
Social affect	18.47 (2.77) ^a	–
Repetitive	3.47 (1.99)	–
Score	21.80 (3.91)	–
Severity	8.47 (1.23)	–
PLS		
Expressive	18.87 (10.77)	
Receptive	17.33 (12.60)	
IQ	68.8 (12)	

^aMean (SD)