Please note:
This is the accepted manuscript version of the article. The final version is published in Research in Autism Spectrum Disorders, with DOI:
http://dx.doi.org/10.1016/j.rasd.2015.06.007

Gaze performance during face-to-face communication: a live eye tracking study of typical children and children with autism

Author: Terje Falck-Ytter1,2*

*corresponding author: Center of Neurodevelopmental Disorders (KIND, Dept. of Women’s & Children’s Health, Karolinska Institutet, Gävlegatan 22B, SE-11330 Stockholm, Tel.: +46 (0)70 4581475 Fax.: +46 (0) 517 79696. Email adress: terje.falck-ytter@ki.se

Affiliations:

1 Center of Neurodevelopmental Disorders at Karolinska Institutet (KIND), Pediatric Neuropsychiatry Unit, Department of Women’s & Children’s Health, Karolinska Institutet, Stockholm, Sweden
2 Uppsala Child & Babylab, Department of Psychology, Uppsala University, Uppsala, Sweden
Abstract: Autism spectrum disorder (ASD) is characterized by socio-communicative impairments, and limited attention to other people's faces is thought to be an important underlying mechanism. Here, non-invasive eye-tracking technology was used to quantify the amount of time spent looking at another person’s face during face-to-face communication in children with ASD (n=13, age 6 years) and age and IQ-matched neurotypical children (n=27, 6 years). We found that in one context of high ecological relevance – listening to an adult telling a children’s story – children with ASD showed a markedly reduced tendency to look at the adult’s face. In interactions between typical children and the adult, the amount of gaze to the other’s face aligned between the two individuals. No such relation was found when the ASD group interacted with the adult. Despite these differences in the storytelling context, we also observed that social looking atypicalities did not generalize to another and more structured context, implying that social looking cannot not be considered fundamentally disrupted in children with ASD.

Keywords: Eye tracking; attention; social interaction

1. Introduction

When we act, our eyes are free to move and facilitate perceptual access to the most useful visual information given the intention of the individual and the current situational constraints (Land & Furneaux, 1997; Rosander & von Hofsten, 2011). In social contexts, looking at other people’s faces provides important information about their intentions and emotions, and can facilitate encoding of speech (McGurk & Macdonald, 1976; Senju & Cibra, 2008; Smith, Cottrell, Gosselin, & Schyns, 2005; Vatikiotis-Bateson, Eigsti, Yano, & Munhall, 1998). Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder marked by severe socio-communicative deficits, and influential models of ASD hold that reduced or altered attention to social stimuli is a key component of the condition (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012; Dalton, et al., 2005; Klin, Jones, Schultz, Volkmar, & Cohen, 2002). Behavioral observation studies of children with ASD support this view (Ozonoff, et al., 2010), and reduced looking time to social stimuli is considered to be a cardinal sign of autism in gold standard diagnostic tests (Lord, et al., 2000). Nevertheless, because the eyes tend to
move several times per second, there are limits to what behavioral observation of children can reveal about the nature of gaze behavior in ASD in everyday life.

Eye tracking allows assessment of the gaze distribution with high spatial and temporal resolution, and can be used to quantify gaze performance in complex naturalistic environments. Although many previous eye-tracking studies have provided support for the view that atypicalities in social looking are indeed present in ASD (e.g. Falck-Ytter, Rehnberg, & Bölte, 2013; Klin, et al., 2002; Shic, Macari, & Chawarska, 2014), the validity of these studies are often compromised by the use of video or pictures as stimuli. Several studies point to the importance of distinguishing between live and non-live presentation. For example, studying typically developing adults, Hietanen, Leppanen, Peltola, Linna-Aho, & Ruuhiala (2008) showed that ERP and galvanic skin responses were strongly modulated by other people’s gaze direction in a live context, but not when the stimuli were presented as videos. Live presentation entails using multimodal three-dimensional objects as stimuli. Moreover, when the stimulus is a person, it entails bi-directional contingent responding. Both these aspects are key components of social interaction in everyday life.

Only a handful live eye-tracking studies have been published so far involving children with ASD (Falck-Ytter, Carlström, & Johansson, 2015; Hanley, et al., 2014; Nadig, Lee, Singh, Bosshart, & Ozonoff, 2010; Noris, Nadel, Barker, Hadjikhani, & Billard, 2012). One reason why so few studies of this type have been conducted may be that eye-tracking experiments involving live person-to-person interaction can be more methodologically challenging and time consuming than conventional (screen-based) eye tracking (Gredebäck, Fikke, & Melinder, 2010; Risko, Laidlaw, Freeth, Foulsham, & Kingstone, 2012).

Of the live eye tracking studies of ASD published so far, all except one (Falck-Ytter, et al., 2015) confirmed the prevailing view of reduced looking time to social stimuli in children with ASD. In contrast, Falck-Ytter et al. (2015) reported not only that children with ASD and typical controls looked equally long to an interlocutor’s face, but also that both groups modulated their gaze adaptively (and similarly) as a function of task demands (e.g. looking more at the lower part of the face when listening; looking higher up when answering, etc). An aspect of the Falck-Ytter et al. study that differed from all the other reports was that the eye tracking was conducted during a ‘maximum performance’ cognitive testing session. Specifically, the authors were measuring gaze performance during a short term memory (digit span) task, which is included in common intelligence tests, and in which the child is required to
perform at his or her very best. Thus, one potential explanation for the discrepancy across studies is that the atypicalities in social looking in ASD are found for some contexts but not others. However, because the aforementioned live eye tracking studies differed in a number of aspects in addition to the nature of the tasks involved, this conclusion remains premature. Another explanation could be that the sample studied in Falck-Ytter et al (2015) was non-representative, and that these particular children would display normal looking patterns in all contexts.

Here, we analyzed gaze data from a group of children that both were part of the experiment reported in Falck-Ytter et al. (2015) as well as in a less structured task where were no task instructions were given and where no response was required from the child; listening to the adult telling a brief children’s story. In addition to providing a contrast to the cognitive testing context, this task resembles many teaching situations the children encounter in school, and thus is of high ecological relevance.

All children participating in the current study contributed to the data reported previously (Falck-Ytter, et al., 2015). In addition, to increase sample size, data from five additional children (three with ASD) were included in the current study. Based on the previous studies of gaze performance in social ‘average performance’ tasks (Hanley, et al., 2014; Nadig, et al., 2010; Noris, et al., 2012), we predicted reduced gazing to the adult’s face in the ASD group relative to the TD group in the current setting.

2. Materials and methods

2.1 Participants. Thirteen high-functioning individuals with ASD (1 girl and 12 boys, age range 4.9 to 10.4 years) and 27 typically developing (TD) children (8 girls and 19 boys, age range 5.1 to 10.2 years) were studied (final samples, after exclusion; Table 1). Autistic children were recruited from habilitation centers in the greater Stockholm region and had a clinical community diagnosis of ASD (Autistic Disorder, Asperger Syndrome, or Pervasive Developmental Disorder–Not Otherwise Specified) according to DSM-IV (AmericanPsychiatricAssociation, 1994). We selectively recruited children judged by their clinical psychologist at the habilitation center to have an IQ above 70 (subsequently confirmed by formal testing). A this particular habilitation center, the use of the ADOS (Lord, et al., 2000) or the ADI-R (Lord, Rutter, & Lecouteur, 1994) is standard when diagnosing ASD, and in our sample the medical records specified the use of these instruments in ten out of 13 cases. The TD group was recruited from birth records from the same geographical area as the ASD group. None of the TD children had relatives (including second degree) with ASD. None of the children in this study had
major medical conditions (e.g., epilepsy). In all children, we assessed intelligence using the Wechsler Intelligence Scales (Wechsler, 1967/2002, 2003), autistic symptomatology using the Social Responsiveness Scale (Constantino, 2005; Constantino & Gruber, 2009), and socio-economic status using an in-house form (Table 1). Recruited families were predominantly Caucasian.

One participant with ASD was excluded because of technical failure and two (one with ASD) because of too little eye-tracking data (cut-off for exclusion: data from less than 30% of the duration of the story). Parents provided written consent, and the study was approved by the Local Ethics Committee in Stockholm and conducted in accordance with the standards specified in the 1964 Declaration of Helsinki. As compensation for their participation, the families received a gift voucher (value ~25 euro).

2.2 Procedure and stimuli. The storytelling task was administered during the same eye tracking session as the cognitive testing session reported on earlier (Falck-Ytter, et al., 2015). The whole session lasted approximately ten minutes. Please note that the procedure and other methodological aspects of this study have already been described elsewhere (Falck-Ytter, et al., 2015) and are repeated here with minor modifications for clarity only.

After familiarization, the participants were brought to the study room with their caregivers. The room included as few distractions as possible (e.g., white curtains covered the walls and technical equipment, except for the eye tracker itself). To produce a typical classroom environment, two colorful illustrations were hung on a wall behind the experimenter, separated from the head by ~5 visual degrees. This setting was chosen to increase the ecological validity of the study, keeping both social and non-social stimuli within the field of view of the child and the area covered by the eye tracker (Figure 1A).

The experimenter and the child were seated in chairs, facing each other, with a distance of approximately 1.5 meters. The experimenters were young adult males, one with blond hair and a short beard and the other with dark brown hair and no beard. The experimenters tested an equal proportion of children from the two groups. The eye tracker (Tobii TX300; 60Hz; Tobii Technology, Stockholm, Sweden) was placed on a table approximately 70 cm in front of and 30 degrees below the participant’s line of vision. A camera provided video recordings of the scene (Figure 1A). This video was used to define events and areas of interest (AOIs) for subsequent analysis of eye movements. A portable
calibration surface with five holes (calibration points) was built to calibrate the eye tracker, using the procedure required by the eye-tracker software (Tobii Studio, Tobii Technology, Stockholm, Sweden). During calibration, a thin object was moved sequentially from hole to hole, accompanied by verbal instruction to the child to look at it. The calibration surface was placed where the face of the experimenter would be located during the subsequent experiment. The calibration procedure was repeated until a good calibration was obtained, which was checked online via a monitor placed behind the child.

2.2.1 The storytelling phase. Once the calibration procedure was completed, the experimenter told the story. The story was 10 sentences long (147 words), lasted approximately one minute, and was designed to be appropriate for the participants’ age, easy to understand, and interesting (the experimenter told about an incident when the experimenter’s, angry little dog bit a police horse). The story was told face-to-face with the children, there was no book or other material involved. The cognitive testing started immediately after the storytelling ended.

2.2.2 Rating of experimenter behavior during storytelling. An independent rater, blind to group membership, examined the experimenters’ behavior in all videos. In all sessions, the experimenters’ attitude/expression towards the child was coded as ‘mild positive affect/expression with occasional smiles’. The amount of direct gaze used by the experimenter varied spontaneously between children; these data are reported in the Results section.

2.2.3 The cognitive testing session. The cognitive testing session started directly after the storytelling. For full details regarding this part, please see Falck-Ytter et al. (2015). The task was a modified version of the digit span task taken from Wechsler’s Intelligence Scale for Children (WISC) (Wechsler, 2003). This particular sub-scale measures primarily short time memory. The task is to remember a series of digits given orally by the experimenter and report the digits back directly afterwards, in the same order. Before the testing phase, the experimenter determined each child’s level of performance on the task individually by gradually increasing the difficulty of the task. Then, in the testing phase, the child was given series of digits at their maximum level. On half of the trials the experimenter gazed at the child (direct gaze condition), on half he was looking down to the test protocol (averted gaze condition). Thus, on average, the experimenter looked at the child for 50% of the time during this session.
2.3 Data reduction. We used the Tobii Studio software (Tobii, Stockholm, Sweden) to code events on the basis of the audio and video recording obtained during the experiment and to export text files with the event codes along with eye-tracking data (expressed as pixel coordinates of the scene video).

In-house computer programs written in MATLAB (MathWorks Inc., Natick, MA, USA) were used to 1) define a dynamic AOI covering the face of the experimenter (defined throughout the scene video recording, individually for each participant), as well as two static AOIs covering the two pictures (objects AOI); 2) calculate the duration of looking in the face AOI relative to the time spent looking at non-AOI areas; 3) calculate the horizontal gaze position for data falling within the face AOI (normalized to AOI height); 4) calculate the vertical gaze position for data falling within the face AOI (normalized to AOI width); and 5) produce summary statistics for the gaze data.

We avoided defining small AOIs covering specific face parts because small AOIs would not produce accurate results in a live eye-tracker setting when the model occasionally moves away from the calibrated stimulus surface (e.g., towards or away from the child) (Falck-Ytter, et al., 2015). Therefore, we chose to use the vertical coordinate of the gaze within the face AOI as a proxy for eye versus mouth looking (Falck-Ytter, von Hofsten, Gillberg, & Fernell, 2013). The face AOI was an ellipse, covering the whole face of the experimenter (vertical/horizontal visual angle ~5/7 degrees). The center of the face AOI was defined as the tip of the nose of the experimenter. Given that the dependent measures used were based on to looking time over extended periods of time, raw data was used in the analyses (i.e., no fixation filter was applied).

2.3.1 Comparison between the storytelling and the cognitive testing session. As noted in the introduction, Falck-Ytter et al. (2015) found no group difference in terms of duration of looking to the face during a cognitive testing session when comparing ASD and TD children. To test if looking time to the face was differently modulated across the two contexts in the two groups, we calculated a difference score defined as the difference between the relative duration of looking time in the face AOI in the two situations (cognitive test versus storytelling contexts). We compared this measure between groups using a non-parametric test rather than a regular ANOVA due to the data not fulfilling criteria for parametric testing (non-equal variances; as reported in Results). To match the storytelling and the cognitive testing phases as much as possible, we only included data from the cognitive testing session when the experimenter was speaking, i.e. when he was administering the digit series. Thus, in both the storytelling context and the cognitive testing context, gaze was measured while the adult was talking.
2.4 Statistical analysis. Unless otherwise specified, data fulfilled criteria for parametric testing (normality; homogeneity of variance) and were tested using ANOVAs/ANCOVAs, t-tests, or Pearson correlation. The $n$ values for individual tests correspond to the global $N$ (Table 1) unless otherwise stated. Alpha level was .05 (two-tailed) throughout.

3. Results

The time it took to tell the story was significantly shorter in the TD group than in the ASD group (TD: $M$=55.95 s, $SD$=6.66 s; ASD: $M$=61.82 s, $SD$=9.65 s; $t_{38}$=2.252, $P$=.03; independent samples t-test). Also, the total looking duration anywhere in the scene (Figure 1A) relative to the duration of the story was shorter in the ASD group than in the TD group (TD: $M$=83.38%, $SD$=16.66%; ASD: $M$=66.81%, $SD$=22.09%; $t_{38}$=2.647, $P$=.012; independent samples t-test). Therefore, both of these variables were used as covariates in the main analyses below.

[insert Table 1 about here]

The duration of looking in the face Area of Interest (AOI) relative to total looking duration anywhere in the scene was significantly shorter in the ASD group than the TD group (TD: $M$=77.72%, $SD$=12.89%; ASD: $M$=50.38%, $SD$=24.54%; $P$<.001; independent samples Mann–Whitney U Test). These data included two potential outliers in the ASD group with low scores on this measure (causing non-homogeneity of variances and necessitating a non-parametric test). Repeating the analysis without these two participants did not change the observed pattern ($t_{36}$=3.925, $P$<.001). Using this reduced sample (to fulfill criteria for parametric testing), we also confirmed that the difference remained, controlling for the duration of the story ($F_{1,35}$=11.054, $P$=.002; ANCOVA), as well as controlling for the percent of looking time anywhere in the scene ($F_{1,35}$=9.019, $P$=.005; ANCOVA). To control for the gender imbalance in the samples, we conducted the comparison for boys only. Again, the ASD group scored significantly lower than the TD group ($P$<.001, independent samples Mann–Whitney U Test). How much the children looked at the face AOI did not differ between the two experimenters ($t_{38}$ =
Taken together, these results confirmed that children with ASD looked less at the adult’s face in the storytelling context.

The duration of looking in the objects AOIs relative to total looking duration anywhere in the scene was significantly longer in the ASD group than the TD group (TD: \(M=10.96\%, SD=9.13\%\); ASD: \(M=26.56\%, SD=15.77\%; P=.001\); independent samples Mann–Whitney U Test). Similarly, the duration of looking outside the face and objects AOIs relative to total looking duration anywhere in the scene was significantly longer in the ASD group than the TD group (TD: \(M=11.32\%, SD=9.79\%\); ASD: \(M=23.07\%, SD=17.44\%; P=.004\); independent samples Mann–Whitney U Test). To test whether the two groups prioritized differently between the pictures and the remaining background, we calculated the proportion of looking time in the objects AOI relative to both the objects AOIs and the non-AOI area. This measure did not differentiate between the groups (TD: \(M=48.57\%, SD=24.34\%\); ASD: \(M=54.02\%, SD=20.59\%, P = .549\), Independent samples Mann-Whitney U Test), suggesting that when the two group did no look at the face, they spent an equal proportion of looking at the two pictures relative to the non-AOI area.

Next, we evaluated whether group differences in looking duration in the face AOI changed over the course of the story. Because the duration of the story always exceeded 45 s, we compared performance during the first 45 s between groups, divided into three 15-s intervals (Figure 1B). Although descriptively the groups diverged more as a function of presentation time, time-by-group interaction effect was not significant. The groups differed already in the first 15 second period (\(P = .007\), Independent samples Mann-Whitney U Test). There was no difference between the two groups in terms of their within-person variability across the three time-bins (\(t_{38} = 1.192, P = .241\); independent samples t-test).

During cognitive testing reported earlier (Falck-Ytter, et al., 2015) we manipulated the direction of the experimenter’s gaze direction as a within subjects experimental variable. We found that both groups adjusted their amount of looking back at the experimenter’s face as a function of this manipulation. Although we did not manipulate gaze direction during the storytelling, the amount of direct gaze used by the experimenter varied between children (spontaneously). In the ASD group, the
experimenter looked at the child 79.20% of the time on average ($SD = 20.60\%$), and in the TD group, the experimenter looked at the child 78.5% of the time on average ($SD = 15.62\%$). Neither the variances ($F = 2.467, P = .125$) nor the means ($t_{38} = 0.122, P = .904$) differed between groups. However, in the TD group, the duration of the experimenter’s direct gaze at the child when telling the story was positively related to the children’s looking time at the experimenter’s face ($r = .391, P = .044$; Pearson correlation). A non-significant association in the other direction was found in the ASD group ($r = -.008, P = .980$; Pearson correlation).

In the cognitive testing context (see Methods), the ASD group looked at the face 68.53% on average ($SD = 23.76\%$), while the TD group looked at the face 77.57% on average ($SD = 14.36\%$). A non-parametric test comparing the groups in terms of their tendency to differentiate between the two contexts (percent looking at the face AOI in storytelling context minus percent looking at the same AOI in the cognitive testing session) confirmed that looking time in the face AOI was differently modulated by context in the two groups ($P = .019$, independent samples Mann–Whitney U Test). Only the ASD group changed their amount of looking to the face as a function of context ($P = .033$; one sample Wilcoxon signed-rank test; most looking time to face during cognitive testing).

As noted in the Methods section, during the cognitive testing, the experimenter looked at the child 50% of the time, which is lower than the experimenter’s average use of direct gaze in the storytelling context (~79%, see above). Falck-Ytter et al. (2015) found that during cognitive testing, both groups increased their looking time to the face when the experimenter gazed directly at them, compared to when he was looking away. Thus, it is highly unlikely that the relatively low amount of looking at the face AOI in the ASD group during the storytelling is due to the fact that the experimenter tended to look at directly at the child during this session (assuming that the change in context does not reverse the direction of the effect of direct gaze). To show this formally, we subtracted the amount of looking time in the face AOI in the storytelling context from the amount of looking time in the face AOI in the direct gaze condition of the cognitive testing session (i.e. excluding all averted gaze trials, resulting in 100% direct gaze), and compared this difference score between groups. As expected, an independent samples Mann-Whitney U test showed that the difference between the groups remained (in the same direction; $P = .013$).

Both groups tended to look low in the face on average, and the vertical position of the gaze (expressed as percentage of the height of the AOI, where 0 is bottom) did not differ between groups
(TD: $M=38.55\%$, $SD=11.20\%$; ASD: $M=40.73\%$, $SD=9.37\%$; $t_{38}=.607$, $P=.548$; independent samples t-test). Both groups tended to look close to the vertical midline of the face on average, and the horizontal gaze position (expressed as percentage of the width of the AOI; 0 is left) did not differ between the groups (TD: $M=47.79\%$, $SD=7.92\%$; ASD: $M=49.00\%$, $SD=12.81\%$ ($t_{38}=.368$, $P=.715$; independent samples t-test). Thus, we found no indication that the groups looked differently within the face AOI (Figure 2).

4. Discussion

We found that when listening to a brief children’s story, children with ASD looked significantly less at the speaker’s face compared to age and IQ matched neurotypical children. This finding suggests that during face-to-face (receptive) communication – i.e. a situation relevant for school situations – children with ASD may miss important information that is available to other children. In addition to the risk of reduced encoding of the verbal material because of attention capture by task-irrelevant stimuli, reduced attention to the face will reduce the observer’s capacity to process non-verbal visual signals that naturally accompany speech, such as emotional expression and ostensive cues.

In the typically developing group, spontaneous use of direct gaze by the experimenter was positively associated with the child’s tendency to look at the experimenter’s face. Thus, in these pairs of typically developing individuals, attention to the other’s face seemed to spontaneously align between the two. In the ASD group, a non-significant correlation in the other direction was observed. These findings are reminiscent of a recent study that showed that the tendency to look at as well as to imitate a model (shown on a computer screen) was influenced by the model’s gaze direction in typical children, but not in children with ASD (Vivanti & Dissanayake, 2014). In contrast, the study by Falck-Ytter et al. (2015), which involved the same experimenters, the same experimental setting, and largely the same children as the current study found strong modulation of gaze to the experimenter’s face in both groups as a function of experimental manipulation of the experimenter’s gaze direction. This shows that the current lack of correlation cannot be due to an inability to regulate one’s own looking based on gaze cues in the ASD group. Nevertheless, it is not unlikely that the lack of spontaneous alignment of looking in contexts resembling the current storytelling situation limit the quality of the interaction, which may have downstream consequences for learning and development.
Context modulated gaze performance differently in the two groups. Specifically, when we compared the amount of time spent looking at the face during storytelling and cognitive testing, we found that the change in context significantly reduced the magnitude of the group difference. Typical children looked at the face for a high percentage of time in both contexts. The ASD group looked less at the face during storytelling than during cognitive testing (note that we only included data from the encoding phase of the cognitive testing session; thus, both during storytelling and cognitive testing, the adult was speaking and intending to communicate with the child). That reduced looking time to other people’s faces was found in one context but not another (Falck-Ytter, et al., 2015) suggests that visual attention in social contexts cannot be considered fundamentally disrupted in ASD (Schultz, 2005; Senju & Johnson, 2009; Shah, Gaule, Bird, & Cook, 2013). It also argues against the possibly that the difference is due to general problems with shifting attention in a flexible way (Elison, et al., 2013; Elsabbagh, et al., 2013), as such impairments would be expected to influence performance in all contexts. Finally, because there were no differences in gaze patterns between children with ASD and typical children during a cognitively demanding (verbal) testing session (Falck-Ytter, et al., 2015), the low level of gaze to the face in the storytelling context reported here is unlikely to reflect an adaptive strategy to improve verbal encoding. If that was the case, one would expect more pronounced group differences when the cognitive load increases, but this was not observed. Rather, because the data presented here and in (Falck-Ytter, et al., 2015) suggest that the level of gaze to others’ faces can be normalized in children with ASD through changes in context, the results point to a potential role of motivational factors (Chevallier, et al., 2012; Kylliainen, et al., 2012; Scott-Van Zeeland, Dapretto, Ghahremani, Poldrack, & Bookheimer, 2010). It is notable that the behavioral difference in ASD was obtained without any use of non-social rewards.

On average, it took slightly longer to tell the story to the ASD group than the TD group. One possibility is that the experimenters spontaneously detected that the visual attention of the children with ASD was low or atypical, and that they adjusted their behavior accordingly. In live experiments focused on ecologically validity, it is not always possible or even meaningful to ‘control for’ all confounds, as controlling for such factors would reduce the ecological validity of the task. As noted in the introduction, the ecological validity of the current task entails at least two important components, 1) that the child is sitting face-to-face with a real three-dimensional person and 2) that both parties are influenced by the other’s behavior. Importantly, the results remained when controlling for the duration
of the task. Moreover, as can be seen in Figure 1B, the group difference was present already from the onset of the storytelling (first 15 seconds), showing that the process that differentiates between the two groups has a rapid onset.

That the two groups did not differ in terms of where in the face they looked accords with the results presented in Falck-Ytter et al. (2015). However, it is notable that the two measures used to study how the children distributed their gaze within the face are not direct measures of eye vs mouth looking. This is a limitation of the current study. Another limitation is the fact that the story was always told before the cognitive testing session, hence introducing potential order effects. However, given previous evidence from interactions of various lengths (Hanley, et al., 2014; Nadig, et al., 2010; Noris, et al., 2012) as well as the fact that the proportion of looking time to the face in ASD did not change throughout the duration of the story in the current study it seems unlikely that performance in ASD normalized simply as a function of time during the course of the eye tracking session. Indeed, if anything, across the first 45 seconds of the story, the groups seemed to diverge rather than converge (Figure 1B). It is important to emphasize that the current report is based on a small sample of individuals with ASD, and that further study is required to verify the stability of the effect, particularly in light of the heterogeneity of the disorder. Yet another limitation is the reliance on community diagnoses. Finally, further studies are required in order to understand exactly what aspects of the two contexts (e.g. performance expectation, amount of prompting, predictability) that differentially affected the two groups.

In summary, when interacting with an adult who tells them a short story, children with ASD have markedly reduced tendency to look at adult’s face. Moreover, while typical children and the experimenter spontaneously aligned their level of looking to the other’s face, no such association was found in the ASD group. Both these findings stand in sharp contrast to a previous related study (Falck-Ytter et al., 2015), where we observed highly similar patterns of gaze to faces in the two groups, adaptive adjustments of gaze as a function of task changes in both groups, and, finally, that both groups looked more at the experimenter when he looked at them than when he looked away. Together, the two
studies suggest that although ‘social looking’ was clearly atypical in the storytelling context, attention to social stimuli does not seem to be fundamentally disrupted in children with ASD.

Acknowledgements

I thank all children and parents for their participation, as well as Christoffer Carlström and Martin Johansson for help during data collection. This research was supported by The Bank of Sweden Tercentenary Foundation [P12-0270:1], the Swedish Research Council in partnership with FAS, FORMAS, and VINNOVA [Cross-disciplinary research programme concerning children’s and young people’s mental health; grant number 259 - 2012-24], the ESF COST Action BM1004 ‘Enhancing the Scientific Study of Early Autism’ (ESSEA), Niclas Öberg Foundation, EU MSCA Initial Training Network (BRAINVIEW) and Sällskapet Barnavård.


## Tables

### Table 1

*Study group characteristics, final samples (M / SD)*

<table>
<thead>
<tr>
<th>Measure</th>
<th>TD</th>
<th>ASD</th>
<th>Pairwise comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=27</td>
<td>n=13</td>
<td></td>
<td>(p value(^1))</td>
</tr>
<tr>
<td>Age (years)</td>
<td>6.8/1.2</td>
<td>6.9/1.5</td>
<td>.983(^2)</td>
</tr>
<tr>
<td>WPPSI-III/WISC-IV Total</td>
<td>101/13</td>
<td>100/19</td>
<td>.793</td>
</tr>
<tr>
<td>WPPSI-III/WISC-IV Verbal</td>
<td>103/11</td>
<td>98/20</td>
<td>.391(^2)</td>
</tr>
<tr>
<td>WPPSI-III/ WISC-IV Non-Verbal</td>
<td>106/15</td>
<td>106/19</td>
<td>.893</td>
</tr>
<tr>
<td>SES (a.u.)</td>
<td>47/6</td>
<td>42/11</td>
<td>.102(^2)</td>
</tr>
<tr>
<td>SRS-Total</td>
<td>42/4</td>
<td>79/15</td>
<td>&lt;0.001(^2)</td>
</tr>
</tbody>
</table>

\(^1\)Independent samples t-test; \(^2\) equal variances not assumed
**Figure captions**

**Figure 1.** Stimuli and results. **A)** The scene with AOIs superimposed. **B)** Children with ASD looked less at the face AOI when listening to the story than children with typical development (TD). Here data from the first 45 seconds are displayed (thin lines represent individual data, thick lines the median for each group; TD = black; ASD = red), divided into three 15-second bins. Already in the first bin, the groups differed significantly. As noted in the main text, the group difference remained even when excluding the two children in the ASD group who looked least at the face AOI.

**Figure 2.** When looking at the face AOI, the two groups did not differ in terms of neither the vertical nor the horizontal coordinate of their gaze. Error bars show standard deviation in the vertical dimension, for
simplicity error bars for the horizontal dimension are omitted (included in main text). It should be noted that these data show the average, and not the typical landing point of gaze.