Inferring mental states from dynamic faces in adolescents with ASD: insights from eye-tracking

Dr Elisa Back

Kingston University London

Short running title: Inferring mental states in ASD and eye-tracking

Conflict of interest statement: The author has no conflict of interest to declare

Acknowledgements: I would like to acknowledge Dr Stephanie Lietz and Ms Hayley Hunt for their assistance with data collection and Dr Jo Van Herwegen and Ms Wendy Ross for their feedback on drafts of the manuscript.

Correspondence concerning this article should be addressed to:
Dr Elisa Back
Department of Psychology
School of Law, Social and Behavioural Sciences
Faculty of Business and Social Sciences
Kingston University London
Penrhyn Road
Kingston upon Thames
Surrey
KT1 2EE
England
Tel: (0044) (0)208 4172831
Electronic mail: e.back@kingston.ac.uk

Abstract

There is mixed evidence concerning whether individuals with Autism Spectrum Disorder (ASD) can infer mental states from the eyes. This study aims to elucidate whether they use less efficient strategies. Sixteen adolescents with ASD (11-16 year olds) were compared to a chronological age- and IQ- matched sample of 16 typically developing (TD) adolescents. Eight mental states were presented as full dynamic faces and in conditions altering the presence of expressive dynamic information from the eyes and mouth. Bayes factors revealed that adolescents with ASD had similar accuracy, response times (less conclusive), and fixations to TD adolescents. Findings imply that adolescents with ASD spontaneously fixate on the eyes and not all individuals with ASD have difficulties inferring mental states from faces.

Keywords: Autism, mental states, facial expressions, ASD, fixations

Inferring mental states from dynamic faces in adolescents with ASD: insights from eye-tracking

The ability to infer mental states from facial expressions is key to social communication. It has been claimed that people with autism have difficulty interpreting mental states from faces and especially from the eyes (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen, Wheelwright, & Jolliffe, 1997). However, these findings have not always been replicated (Ponnet, Roeyers, Buysee, De Clercq, & Van Der Heyden, 2004; Roeyers, Buysse, Ponnet, & Pichal, 2001), particularly with more natural dynamic faces (Back, Ropar, & Mitchell, 2007).

Back et al. (2007) found that children and adolescents with Autism Spectrum Disorder (ASD) were as accurate as typically developing (TD) individuals when inferring mental states (such as “relieved” or “worried”) from the eyes. Mental states were presented as full dynamic faces or edited to “freeze” the eye or mouth regions of the face in a neutral position while the rest of the face remained dynamic and expressive. Participants with ASD performed worse in the eyes-frozen and mouth-frozen conditions than with full dynamic faces, indicating that they benefitted from cues in those regions. An overall difference between groups was found in Experiment 1 where TD children were more accurate than those with ASD at inferring mental states from faces, however children with ASD performed significantly above chance indicating that they did not have a complete deficit. In sum, individuals with ASD were more accurate at extracting mental state information from faces and specifically the eyes than previous studies suggested (e.g., Baron-Cohen et al., 2001).

One possible reason for the mixed findings of previous studies is that accuracy is not the most sensitive measure for detecting processing differences in people with ASD and investigating response times is more appropriate; in everyday life we need to be able to interpret facial expressions quickly otherwise they may be no longer relevant to conversation. Studies investigating response times in relation to the recognition of basic emotions have found no differences between ASD and TD children (e.g., Fink, de Rosnay, Wierda, Koot, &

Begeer, 2014). However, the speed of responding to more subtle dynamic mental states has yet to be investigated. Therefore, the first aim of the study is to investigate the speed of attributing mental states to facial expressions. The second aim of the study is to investigate the role face processing strategies may play in this and whether initial attraction to the face or subsequent face processing strategies differ in ASD.

Visual attention to faces has been found to correlate with theory of mind abilities and it is suggested that because the eyes convey mental state information (e.g., Baron-Cohen et al., 2001), eye gaze is an essential part of mind reading. Therefore, monitoring participants' fixations allows additional investigation into whether attention to particular facial regions occurs spontaneously or voluntarily. Findings are mixed as to whether individuals with ASD have a deficit in spontaneously attributing mental states or whether they use compensatory strategies (see Senju, 2013). People with ASD might be slow to judge mental states (albeit relatively accurate) because they may not fixate on the eyes (Klin et al., 2002) and instead attend to this region via parafoveal processing, which may require more time than direct fixation of the eyes. Findings concerning gaze behaviours to faces and eyes are mixed. Some studies have found that individuals with ASD do not have reduced attention to the eyes (see Guillon, Hadjikhani, Baduel, & Rogé, 2014) and Elsabbagh et al. (2013) also demonstrated typical initial orienting towards faces in infants that go on to develop ASD. Overall though, research has shown reduced social attention to faces (see Chita-Tegmark, 2016a) and a meta-analysis has demonstrated that overall attention allocation to social information is atypical in autism (see Chita-Tegmark, 2016b).

An alternative explanation for atypical face processing is suggested by Freeth, Chapman, Ropar, and Mitchell (2010). They found that individuals with ASD do not differ from control participants in their overall looking time to faces when presented with pictures of social scenes but, rather, were significantly slower to first fixate the face. Moreover, Guimard-Brunault et al. (2013) found that spontaneous visual attention to a static face presented on a computer screen was affected by autism severity. Nevertheless, a study by Elsabbagh et al., (2013) demonstrated that, at least early in development, infants with ASD do not lack an initial attraction to the face, however it is not known if this is the case in adolescence when inferring mental states. Therefore, it will be fruitful to investigate whether it is the initial attentional attraction to a face or the subsequent processing that differs in ASD. Distinguishing between these possibilities will give valuable information about what underlying mechanisms might explain atypical responses and subsequently help explain why those with ASD may be less effective in their social interactions.

The current study investigated whether the strategies used to infer mental states differ between ASD and typically developing (TD) adolescents. Response times and fixations (as well as the traditional measure of accuracy) were utilized as the combination of these more sensitive measures allows investigation of *how* mental states are inferred from facial expressions. Moreover, to examine potential differences in face processing strategies, the eyes and mouth frozen conditions from Back et al. (2007) were included to investigate how important expressive dynamic information from these regions is for processing mental states. The study focused on pre-adolescence and adolescence as these periods are important for the processing of social information and developing expertise with faces. Mechanisms related to mentalising and more broadly social cognition are still developing late in adolescence (Blakemore, 2012). Therefore, research investigating adolescents is imperative to gain a full understanding of how mental states are inferred from faces. Specifically, it was hypothesised that adolescents with ASD will be as accurate as TD adolescents at inferring mental states but they may be slower. Furthermore, it was predicted that fixations may differ between groups; those with ASD may look at the eyes for less time than TD participants but they will still be able to attribute mental states correctly. In the frozen conditions (where there is no expressive information from either the eyes or the mouth), it is expected that fixations will be focussed on the more informative area (e.g., the mouth when the eyes are frozen and the eyes when the mouth is frozen) but that adolescents with ASD may be slower at switching their attention to the more informative region due to executive function difficulties. Previous research has found that individuals with ASD have difficulties with shifting their attention due to a lack of cognitive flexibility (Hill, 2004). Therefore, this study aimed to address whether this may also be applicable to switching attention to more informative parts of the face when inferring mental states whilst using naturalistic dynamic stimuli.

Method

Participants

Sixteen adolescents with Autism Spectrum Disorder (ASD) were individually matched to 16 TD adolescents on chronological age (CA), gender, and full scale IQ (FSIQ) using the Wechsler Abbreviated Scales of Intelligence (WASI-2; Wechsler, 2011). Participants were aged between 11 to 16 years old and there were 30 males and two females. Independent samples *t*-tests revealed no significant differences between groups on CA, $t(30) = -.302$, $p > .05$ and FSIQ, $t(30) = -1.327$, $p > .05$. All participants in the ASD group had received a diagnosis by a clinician for ASD in accordance with the DSM-5 criteria (American Psychiatric Association, 2013) but not for any other developmental condition (e.g., attention deficit hyperactivity disorder). The ASD group's diagnosis was reconfirmed using the Autism Diagnosis Observation Schedule (ADOS-2; Lord, Rutter, & DiLavore, 2012) and ten adolescents reached cut-off for autism and six for autism spectrum. The Childhood Autism Spectrum Test (CAST; Scott, Baron-Cohen, Bolton, & Brayne, 2002) was also completed by parents of both ASD and TD adolescents. This confirmed the presence of a significant amount of autism features in the ASD group (above the cut-off 15) but this was not the case in the TD group. The mean CAST score was 19.5 (SD= 4.1, range= 15 to 27) in the ASD group compared to 2.6 (SD= 2.1, range= 1 to 6) in the TD group. Participants were recruited from schools in Greater London and Surrey. Parents completed a demographic background questionnaire. Although there was variability in parental education and family income, participants were from a middle socioeconomic status background. Participants were from different ethnic backgrounds (90% White and 10% Asian). All had English as their first language and all participants had normal or corrected to normal vision. Table 1 displays participants' details.

[Table 1 goes here]

Design and stimuli

Faces were randomised in five different orders and each participant experienced one of these orders (approximately six participants viewed each order). On each trial, a fixation cross appeared for one second, followed by a video clip of a face (approximately five seconds) depicting a mental state, then a word appeared that was either a correct or incorrect term to describe the facial expression. Participants were encouraged to respond as quickly as possible using a button box to judge whether the mental state word accurately described what the person was thinking/feeling. Participants' fixations were recorded throughout. Eight mental states (that originated from Back et al., 2007, see Figure 1) were presented; deciding, disapproving, don't trust, not interested, not sure, relieved, surprised and worried. These had been extensively validated in previous studies as being the correct mental state label for each facial expression and respective validated foils (incorrect answers) were also used (Back et al., 2007; Back & Jordan, 2014). Each mental state was presented four times (twice when the word correctly corresponded and twice when the same incorrect word corresponded to the face) in each of the three different display types (full face dynamic, eyes frozen and mouth frozen). In the frozen conditions, the facial area remained static and neutral while the rest of the face was expressive and dynamic (see supplementary materials for examples of the frozen stimuli). It took participants approximately 20 minutes to complete the 96 trials and they were given a short break half way through.

[Figure 1 goes here]

Procedure

The study was approved by the Faculty ethics committee at Kingston University and informed consent was given by parents of participants prior to their inclusion in the study. Participants took part in the following sessions; the ADOS (ASD participants only), the WASI-2, and the experimental task with eye-tracking. Participants were seated in front of a 17-inch monitor at a viewing distance of 60 cm. Eye movements were recorded using a T120 Tobii eye-tracker. The eye-tracker was calibrated with 9 dots using Tobii Studio software. Standardised instructions appeared on the screen along with two practice trials.

Accuracy, response times, and fixations were recorded. All fixations less than 100ms were removed from the data analysis as it would be unlikely information could be extracted from such short fixations (Manor & Gordon, 2003). Outliers were removed that were 2SD above the mean for both response times and eye movements. The eyes and the mouth were the two areas of interest (AOI) and the following eye-tracking measures were used: 1) time to first fixation (the amount of time it takes to look at an AOI from the onset of the face), this is a measure of spontaneous looking which is indicative of early processing, 2) first fixation duration (how long the first fixation lasted for) and this is also an early processing measure, and 3) total fixation duration (sum of the duration for all fixations), which is a voluntary and late processing measure.

Mixed design ANOVA's were carried out and Bayesian ANOVA's were subsequently undertaken using JASP (www.jasp-stats.org). The reported study was limited in power to fully accept the null hypothesis (that there are no differences) using traditional statistical testing, therefore Bayes factors were calculated that allows us to draw inferences about the probability of the data under the null hypothesis (relative to the alternative).

Results

Accuracy analysis

Accuracy scores were analysed using a two-way ANOVA on Group (ASD, TD) x Condition (full face, eyes frozen, mouth frozen). A significant main effect of Condition was obtained, $F(2, 60) = 13.059$, $p < .001$, $\eta_p^2 = .303$ but there was no main effect of Group, $F(1, 30) = .348$, $p = .560$, $\eta_p^2 = .011$ (BF01= 3417, more likely to support null, extreme support) and no interaction between Group and Condition, $F(2, 60) = .197$, $p = .822$, $\eta_p^2 = .007$ (BF01= 12.683, more likely to support null, strong support). Pairwise comparisons with Bonferroni's adjustment revealed that accuracy scores were higher for the full dynamic face than the eyes frozen ($p = .006$) and the mouth frozen ($p < .001$). Moreover, accuracy scores were higher for the eyes frozen than the mouth frozen condition ($p = .035$). Mean accuracy scores can be seen in Table 2 and Figure 2a.

Response times analysis

Response times were analysed for correct answers only and outliers were removed that were more than 2SD above the mean (above 5295.09ms). This resulted in 122 outliers (out of 2026 data points) being removed (6% of data). There were no extremely fast responses (all above 500ms) so none were removed. A two-way ANOVA on Group (ASD, TD) x Condition (full face, eyes frozen, mouth frozen) was carried out and a significant main effect of Condition was found, $F(2, 60) = 5.466$, $p = .007$, $\eta_p^2 = .154$. There was no main effect of Group, $F(1, 30) = 1.760$, $p = .195$, $\eta_p^2 = .055$ (BF01= 5.854 in support of null, moderate support) and no significant interaction between Group and Condition, $F(2, 60) = 1.957$, $p = .150$, $\eta_p^2 = .061$ (BF01= 2.121, anecdotal support). Pairwise comparisons with Bonferroni's adjustment for the effect of Condition revealed faster response times for the full dynamic face compared to the eyes frozen ($p = .021$) and the mouth frozen ($p = .003$). There was no difference in response times between eyes frozen and mouth frozen conditions ($p = .713$). Mean response times can be found in Table 2 and Figure 2b.

[Table 2 goes here]

Fixation analyses

For time to first fixation, seven outliers (above 15.17s) were removed due to technical difficulties (out of 192 data points). A three-way ANOVA was carried out on Group (ASD, TD) x Condition (full face, eyes frozen, mouth frozen) x AOI (eyes, mouth) for time to first fixation. This revealed no significant main effect of Group, $F(1, 30) = .098$, $p= .757$, $\eta_p^2 = .003$ (BF01= 481352 more likely to support the null, extreme support). There was a significant main effect of Condition, $F(2, 60) = 6.659$, $p= .002$, $\eta_p^2 = .182$. Pairwise comparisons with Bonferroni's adjustment revealed that participants were significantly faster to fixate either AOI in the frozen eyes than frozen mouth condition ($p = .011$). There was a significant main effect of AOI, where participants were faster to fixate on the eyes ($M = 1.81$) than the mouth ($M = 3.88$), $F(1, 30) = 11.449$, $p = .002$, $\eta_p^2 = .276$. However, there was no significant interaction between Group and Condition, $F(2, 60)= .914$, $p = .407$, $\eta_p^2 = .030$ (BF01= 1.371e+6 in support of null, extreme support) and there was no significant interaction between Group and AOI, $F(1, 30)= .448$, $p = .508$, $\eta_p^2 = .015$ (BF01=1461.127, in support of null, extreme support). There was a significant interaction between Condition and AOI, $F(2, 60) = 3.589$, $p = .034$, $\eta_p^2 = .107$. To break-down this interaction, paired samples *t*-tests were carried out showing that in the full dynamic face condition, participants were faster to fixate on the eyes ($M = 2.19$) than the mouth ($M = 4.19$), $t(31) = 2.361$, $p = .025$, $d = 0.42$. In the eyes frozen condition, there was no significant difference in the time to first fixate on the eyes or the mouth, $t(31) = 1.479$, $p > .1$. Finally in the mouth frozen condition, participants were also faster to first fixate on the eyes ($M = 1.30$) than the mouth ($M = 4.58$), $t(31) = 3.699$, $p = .001$, $d = 0.70$. There were no other significant effects or interactions.

With respect to first fixation duration, four outliers (above 1.70s) were removed due to technical difficulties (out of 192 data points) and a three-way ANOVA was carried out on Group (ASD, TD) x Condition (full face, eyes frozen, mouth frozen) x AOI (eyes, mouth). A significant interaction was found between AOI and Group, $F(1, 30) = 5.154$, $p = .031$, $\eta_p^2 = .147$. Independent samples *t*-tests revealed that the mouth was fixated on for longer by the TD group than the ASD group, $t(30) = 2.109$, $p = .043$, $d = 0.75$ whereas there was no group difference for the eyes, $t(30) = .168$, $p = .868$ (BF01= 3.325 in support of the null, moderate support). There was also no significant main effect of Group, $F(1, 30) = 2.313$, $p = .139$, $\eta_p^2 = .072$ (BF01= 1.897 in support of the null, anecdotal support). There were no other significant effects or interactions.

For total fixation duration (no outliers were removed), a proportional analysis was carried out. A three-way ANOVA was conducted on Group (ASD, TD) x Condition (full face, eyes frozen, mouth frozen) x AOI (eyes, mouth). A main effect of Condition was found, $F(2, 60) = 46.095$, $p < .001$, $\eta_p^2 = .606$ and Bonferroni pairwise comparisons revealed that participants spent longer looking at either AOI in the full dynamic face ($M = .734$) than in the eyes ($M = .168$, $p < .001$) and mouth ($M = .189$, $p < .001$) frozen conditions. Moreover, participants looked longer at either AOI when the mouth was frozen compared to the eyes frozen condition (*p*= .016). Furthermore, a significant main effect of AOI was obtained, *F*(1, 30)= 10.437, *p*= .003, η_p^2 = .258, where participant spent longer looking at the mouth (*M* =.487) than the eyes (*M* = .240). There was also a significant interaction between Condition and AOI, *F*(2, 60)= 50.801, *p* < .001, η_p^2 = .629. Paired samples *t*-tests revealed that in the mouth frozen condition, participants spent longer looking at the eyes than the mouth, *t*(31)= 7.778, *p*< .001, *d* = 1.59. In the full face condition, participants spent longer looking at the mouth than the eyes, *t*(31)= 5.633, *p*< .001, *d* = 1.11. Whereas there was no significant difference between the eyes and the mouth in the eyes frozen condition, *t*(31)= 1.050, *p*= .302 (BF01= 3.197, moderate support). There was no significant main effect of Group, *F*(1, 30)= .232, *p* = .633, η_p^2 = .008 (BF01= 4.952 in support of the null, moderate support). There were no significant interactions between Group and Condition, *F*(2, 60)= .342, *p* = .711, η_p^2 = .011 (BF01= 5.000e +16, extreme support for the null) and between Group and AOI, *F*(1, 30)= .291, *p* = .594, η_p^2 = .010 (BF01= 5.415e +23, extreme support for the null). There were no further interactions for total fixation duration. Mean scores for all fixation measures can be found in Table 2 and Figure 2.

[Figure 2 goes here]

Discussion

This study investigated whether the strategies used to infer mental states differ between ASD and TD adolescents using finely grained measures of response time and fixation to the eyes and mouth region. Accuracy scores showed that adolescents with ASD successfully attributed mental states to faces and they had similar response times as typically developing adolescents in making these inferences. Fixation patterns were also similar to TD adolescents when inferring mental states. Three sources of evidence; accuracy, fixations and, to a more limited extent, response times converge to support the view that individuals with ASD in the current study were able to efficiently infer mental states from facial expressions. Fixation data revealed that they used similar strategies to those without ASD to infer mental states. There was no evidence that adolescents with ASD have difficulty spontaneously fixating on the eyes when inferring mental states as there was no difference between groups in the time to first fixate on the eyes. This is contrary to previous research suggesting that individuals with ASD do not spontaneously attend to the eyes (e.g., Pelphrey et al., 2002). This may be because dynamic stimuli of mental states were used instead of static stimuli of basic emotions. This implies that using more real-world dynamic faces and more complex expressions facilitates spontaneous fixation on the eyes.

There was no evidence to suggest that adolescents with ASD were slower than TD adolescents at switching their attention to more informative regions as the interaction between group, condition and AOI for time to first fixation was not significant. Instead, there was a trend for adolescents with ASD being faster at switching their attention from the eyes (1.90) to the mouth (2.09) in the eyes frozen condition compared to TD adolescents (respective means for eyes was 1.99 and mouth 3.62). This could perhaps be related to their superior local processing skills (Happé, 1999) or enhanced perceptual functioning (Mottron, Dawson, Soulières, Hubert, & Burack, 2006) that override executive function difficulties in switching their attention from one region of the face to another.

Indeed, the only group difference to emerge across all the measures was that adolescents with ASD had a shorter first fixation duration to the mouth than TD adolescents. Overall looking times to the eyes and the mouth were similar to TD adolescents as seen by later measures such as total fixation duration. Findings may at first seem to be contrary to Freeth et al.'s (2010) study. However, the current study used dynamic faces rather than pictures of more complex social scenes and this could explain why in the current study no difference was found between groups regarding time to first fixation to the eyes. Moreover, findings corroborated that overall time spent looking at the eyes were similar for those with and without ASD.

It is known that individuals with ASD can perform well on structured tasks so the clear task to carry out may explain the lack of group differences with respect to processing the eyes. Additionally, this was a group of just 16 individuals with ASD, so the generalizability is limited; it can only be concluded that not all individuals with ASD have a deficit in inferring mental states. Nevertheless, the Bayesian analyses confirmed that there was more support for the null relative to the alternative hypothesis when comparing the groups across a range of measures. Further research is required especially regarding speed of responding to mental states. This was not as conclusive and it appears that response time is a sensitive measure to potentially reveal differences between groups in the way they infer mental states from facial expressions.

Nonetheless, the current study supports the use of more naturalistic methods, such as dynamic full faces over just presenting the static eye region (as is traditional in research such as Baron-Cohen et al., 2001) to enable a clearer picture of relevant abilities to emerge as in daily life we rarely see just the static eye region. Findings showed that adolescents with ASD can infer mental states using a paradigm that involves verifying whether a word correctly or incorrectly described what the person was thinking or feeling. This has implications for interventions with regards to how scaffolding (providing potential labels) could improve their ability to infer mental states and subsequently their social relationships. Findings appear to differ to Experiment 1 of Back et al.'s (2007) study, where they found that those with ASD did not infer mental states to the same extent as control participants to faces. However, Back et al., (2007) used a four-way forced choice procedure whereas in this study participants just had to verify whether or not a word was a correct or incorrect label for the facial expression.

This could be viewed as an easier task and therefore potentially explains why participants had higher accuracy scores in the current study.

Previous studies investigating the attribution of mental states to facial expressions have found mixed results, which could be due to methodology and participant characteristics, such as different measures (only accuracy has been used in previous studies), the selection of words used in forced-choice procedures, participants' IQ, matching procedures, the severity of autism, and age of participants. Future research should consider these variables and explore the development of inferring mental states from faces across different age groups as this could be particularly informative when delivering interventions at the appropriate developmental stage. Facial expressions were presented for approximately five seconds (including the onset and offset of the expression) so there is a need to see if differences occur when facial expressions are more fleeting. The processing of biological motion has been found to be impaired in ASD (Gepner, Lainé & Tardiff, 2005) and that presenting slowed down facial expression information can be beneficial for individuals with ASD (Gepner, Deruelle, & Grynfeltt, 2001). However, future research using more real-life paradigms such as briefly presented mental states will provide further insight into the social communication difficulties of individuals with ASD and importantly begin to bridge the gap between research and practice.

To conclude, findings from the current study support previous research that suggests that not all individuals with ASD have a deficit in inferring mental states from faces (Back et al., 2007, Ponnet et al., 2004; Roeyers et al., 2001) but goes a step further and shows that there are similarities in the way a face is processed. This study raises awareness that some individuals with ASD can efficiently interpret mental states from facial expressions during social interactions.

References

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

Back, E. & Jordan, T. R. (2014). Revealing variations in perception of mental states from dynamic facial expressions: a cautionary note. *PloS one, 9*, e84395. ISSN (online). doi: 10.1371/journal.pone.0084395

Back, E., Ropar, D., & Mitchell, P. (2007). Do the eyes have it? Inferring mental states from animated faces in autism. *Child Development, 78,* 397-411. doi: 10.1111/j.1467 8624.2007.01005.x

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 4,* 241-251. doi: 10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition, 4,* 311-331. doi: 10.1080/713756761

Blakemore, S.J. (2012). Development of the social brain in adolescence. *Journal of the Royal Society of Medicine, 105* (3), 111-116. doi: 10.1258/jrsm.2011.110221

Chita-Tegmark, M. (2016a). Social attention in ASD: A review and meta-analysis of eye tracking studies. *Research in Developmental Disabilities, 48,* 79-93. doi: 10.1016/j.ridd.2015.10.011

Chita-Tegmark, M. (2016b). Attention allocation in ASD: a review and meta-analysis of eye tracking studies. *Review Journal of Autism and Developmental Disorders, 3*, 209-223. doi: 10.1007/s40489-016-0077-x

Elsabbagh, M., Bedford, R., Senju, A., Charman, T., Pickles, A., & Johnson, M.H. (2013). What you see is what you get: contextual modulation of face scanning in typical and atypical development. *Social Cognitive and Affective Neuroscience*, *9,* 538–543. doi: 10.1093/scan/nst012

Fink, E., de Rosnay, M., Wierda, M. Koot, H. M., & Begeer, S. (2014). Brief report: accuracy and response time for the recognition of facial emotions in a large sample of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44,* 2363-2368. doi: 10.1007/s10803-014-2084-z.

Freeth, M., Chapman, P., Ropar, D. & Mitchell, P. (2010). 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*, *40*, 534-547. doi: 10.1007/s10803-009-0893-2

Gepner, B., Deruelle, C., & Grynfeltt, S (2001). Motion and emotion: a novel approach to the study of face processing by young autistic children. *Journal of Autism and Developmental Disorders, 31* (1), 37-45. doi: 10.1023/A1005609629218

Gepner, B., Lainé, F. & Tardif, C. (2005). *E-Motion mis-sight* and other temporal processing disorders in autism. *Current Psychology of Cognition/Cahiers de Psychologie Cognitive, 23*, 104-121.

Guillon, Q., Hadjikhani, N., Baduel, S., & Rogé, B. (2014). Visual social attention in autism spectrum disorder: insights from eye tracking studies. *Neuroscience & Biobehavioural Reviews, 42,* 279-297. doi: 10.1016/j.neubiorev.2014.03.013

Guimard-Brunault, M., Hernandez, N., Roché, L., Roux, S., Barthélémy, C., Martineau, J., & Bonnet-Brilhault, F. (2013). Back to basic: do children with autism spontaneously look at screen displaying a face or an object? *Autism Research and Treatment*. doi: 10.1155/2013/835247

Happé, F. (1999). Autism: Cognitive deficit or style? *Trends in Cognitive Sciences, 3,* 216 222. Doi: 10.1016/S1364-6613(99)01318-2

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

Lord, C., Rutter, M., & DiLavore, P. C. (2012). *Autism Diagnostic Observation Schedule* (ADOS 2). 2nd edition. Los Angeles, CA: Western Psychological Services.

Manor, B. R., & Gordon, E. (2003). Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. *Journal of Neuroscience Methods*, *128,* 85-93. doi: 10.1016/S0165-0270(03)00151-1

Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced Perceptual Functioning in Autism: An Update, and Eight Principles of Autistic Perception. *Journal of Autism and Developmental Disorders, 36* (1), 27-43. doi: 10.1007/s10803-005-0040-7

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders, 32*, 249–261. doi: 10.1023/A:1016374617369

Ponnet, K. S., Roeyers, H., Buysee, A., De Clercq, A., & Van Der Heyden, E. (2004). Advanced mind-reading in adults with Asperger syndrome. *Autism, 8,* 249-266. doi: 10.1177/1362361304045214

Roeyers, H., Buysse, A., Ponnet, K., & Pichal, B. (2001). Advancing advanced mind-reading tests: Empathic accuracy in adults with a pervasive developmental disorder. *Journal of Child Psychology and Psychiatry, 42,* 271-278. doi: 10.1017/S0021963001006680

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002). The CAST (Childhood Asperger Syndrome Test): Preliminary Development of a UK Screen for Mainstream

Primary-School Age Children. *Autism, 6*, 9-31. doi: 10.1177/1362361302006001003

Senju, A. (2013). Atypical development of spontaneous social cognition in autism spectrum disorders. *Brain & Development*, *35*, 96-101. doi: 10.1016/j.braindev.2012.08.002

Wechsler, D. (2011). Wechsler Abbreviated Scale of Intelligence–Second Edition (WASI II). San Antonio, TX: NCS Pearson

Table Titles

Table 1: Participant characteristics

Table 2: Means (and standard deviations) for each group across conditions: Accuracy, response times and fixations (to each AOI: Eyes and mouth)

Figure Titles

Figure 1: Stimuli

Figure 2: Results for accuracy, response times and fixations

Table 1: Participant characteristics

	ASD participants	Typically developing participants
Age (years; months)		
Mean	14:4	14;5
SD	1.33	1.58
Range	10;9-16;5	11;3-16;9
Full-Scale IQ		
Mean	100.13	106.06
SD	17.66	11.55
Range	70-140	87-125

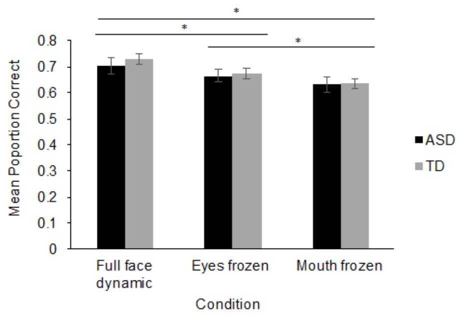
Table 2: Means (and standard deviations) for each group across conditions: Accuracy, response times and fixations (to each AOI: Eyes and mouth)

Group	Condition	Accuracy (proportions)	RT (ms)	Time to first fixation (s)		First fixation duration (s)		Total fixation duration (proportions)	
				Eyes	Mouth	Eyes	Mouth	Eyes	Mouth
ASD	Full face	0.70 (0.13)	1914 (468)	2.97 (3.72)	4.94 (3.79)	0.39 (.22)	0.31 (.17)	0.21 (0.14)	0.92 (0.35)
	Eyes frozen	0.66 (0.09)	2131 (526)	1.90 (2.30)	2.09 (2.69)	0.37 (.32)	0.38 (.38)	0.18 (0.13)	0.15 (0.09)
	Mouth frozen	0.62 (0.12)	2077 (545)	1.46 (1.47)	3.92 (4.27)	0.36 (.24)	0.34 (.28)	0.31 (0.18)	0.07 (0.04)

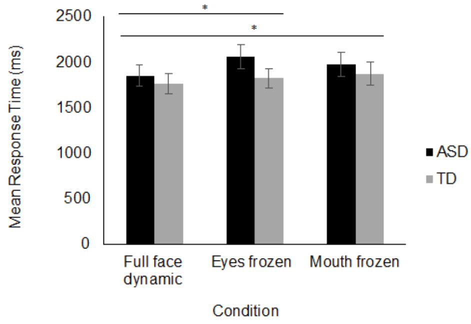
Figure 1

Eyes and mouth AOI

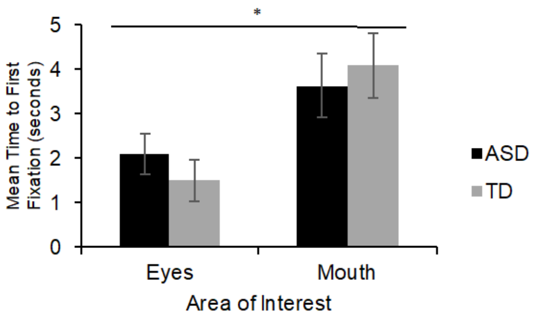
Figure 2: Results for accuracy, response times and fixations



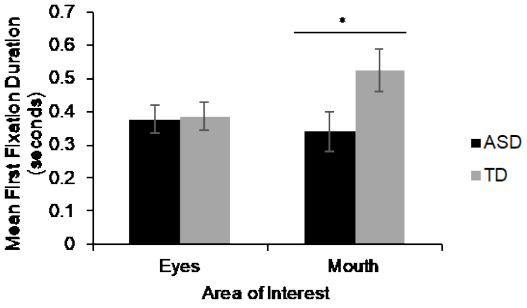
a) Accuracy across conditions and groups.
* denotes significant mean difference between Full face dynamic and Eyes frozen; between Full face dynamic and Mouth frozen; between Eyes frozen and Mouth frozen conditions.



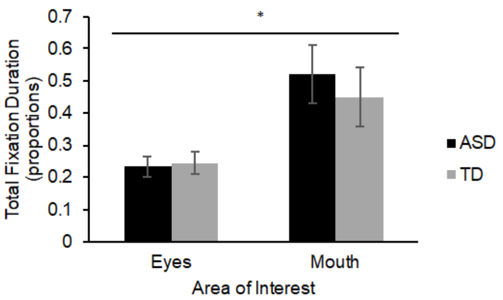
b) Response times across conditions and groups.
* denotes significant mean difference between Full face dynamic and Eyes frozen; between Full face dynamic and Mouth frozen condition.



c) Time to first fixation of eyes and mouth for each group
* denotes significant mean difference between the Eyes and Mouth.



d) First fixation duration to eyes and mouth for each group.
* denotes significant mean difference between ASD and TD for the Mouth.



e) Total fixation duration to eyes and mouth for each group.
* denotes significant mean difference between the Eyes and Mouth.