Social scene perception in autism spectrum disorder: An eye-tracking and pupillometric study

Morgan Frost-Karlsson, Martyna Alexandra Galazka, Christopher Gillberg, Carina Gillberg, Carmela Miniscalco, Eva Billstedt, Nouchine Hadjikhani & Jakob Åsberg Johnels


To link to this article: https://doi.org/10.1080/13803395.2019.1646214
Social scene perception in autism spectrum disorder: An eye-tracking and pupillometric study

Morgan Frost-Karlsson, Martyna Alexandra Galazka, Christopher Gillberg, Carina Gillberg, Carmela Miniscalco, Eva Billstedt, Nouchine Hadjikhani, and Jakob Åsberg Johnels

ABSTRACT

Typically, developing humans innately place subjective value on social information and orient attention to it. This can be shown through tracking of gaze patterns and pupil size, the latter of which taps into an individual’s cognitive engagement and affective arousal. People with Autism Spectrum Disorder (ASD) present with atypical social, communicative and behavioral patterns, but underlying substrates of these behavioral differences remain unclear. Moreover, due to high comorbidity with other neurodevelopmental disorders, it is often difficult to distinguish which differences are distinctive to ASD. In this study, a group of 35 adolescents and young adults with neurodevelopmental disorders were tested to investigate the processing of social and non-social scenes in individuals who meet the diagnostic criteria for autism and those who do not. Eye tracking and pupillometry measures were collected while participants observed images of tightly controlled natural scenes with or without a human being. Contrary to individuals without autism diagnosis, participants with autism did not show greater pupillary response to images with a human. Participants with autism were slower to fixate on social elements in the social scenes, and this latency metric correlated with clinical measures of poor social functioning. The results confirm the clinical relevance of eye-tracking and pupillometric indices in the field of ASD. We discuss the clinical implications of the results and propose that analysis of changes in visual attention and physiological level to social stimuli might be an integral part of a neurodevelopmental assessment.

Introduction

Humans are a social species. We live and thrive in large social groups and exhibit collective behavior, which provides a number of evolutionary advantages. The evolution of the human brain is very closely linked with the increasing complexity of social interactions and the cognitive abilities they require (Emery, 2000). In primates, the primary cues for these processes rely on visual input, which has led to an increasingly elaborate visual social processing system including more than 30 brain regions in humans (Adolphs, 2009; Emery, 2000).

Eye tracking studies have shown that typically developing (TD) humans have an innate tendency to attend to the central facial region of other humans regardless of scene complexity or salience of other, non-social objects (Birmingham, Bischof, & Kingstone, 2009; Constantino et al., 2017). The attentional bias for faces and other social stimuli is present from infancy (Amso, Haas, & Markant, 2014; Chita-Tegmark, 2016; Constantino et al., 2017; Fitzgerald, 1968; Kelly, Duarte, Meary, Bindemann, & Pascalis, 2019), possibly as early as the third trimester of fetal life (Reid et al., 2017).

In addition to eye tracking, social processing is observable through pupillometry due to its control by sympathetic arousal systems and resulting physiological association with cognitive engagement and affective arousal (Bradley, Miccoli, Escrig, & Lang, 2008; Laeng, Sirois, & Gredebäck, 2012). One central mechanism of this association is the locus coeruleus (LC) – norepinephrine (NE) system. The LC, a noradrenergic “hub” used in the brain as a pacemaker for attentional processes, also controls the muscles in the iris (Laeng et al., 2012), and changes in pupil diameter have been shown to directly correlate with activation changes in LC neurons that signify attentional shifts (Laeng et al., 2012; Rajkowski, Majczynski, Clayton & Aston-Jones, 2004). This psychophysiological signature...
is also presumed to reveal social processing development and differences, as social perception is a product of attentional shifts due to the prioritized attention placed on social stimuli (Bast, Poustka, & Freitag, 2018). Indeed, it has been shown that infants as young as one month show differential (larger) pupil dilation in response to social vs. non-social stimuli (Fitzgerald, 1968). Moreover, pupillometry has been argued to be an especially useful tool for studying non- or low-verbal populations such as infants and individuals with developmental disorders affecting social skills and processes (Graur & Siegle, 2013). Notably, Graur and Siegle (2013) suggest that pupillometry has moved from being a basic research tool toward a clinical use in the domain of psychiatry, and that this development might be instrumental in linking neuroscience and psychiatric clinical practice more closely together.

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by a number of social-communication deficits with restricted and repetitive behaviors and interests (American Psychiatric Association, 2013). Among the most prominent social deficits in ASD is a reduced ability to use social cues from facial expressions and body language for learning and social interactions. Although results from different studies have been mixed, atypical facial gaze patterns and eye contact are among the most discernible indications of social impairments in ASD, described already in the first descriptions of the disorder (Kanner, 1944), and included in standardized diagnostic criteria including the DSM-V (American Psychiatric Association, 2013). With single faces as stimuli, atypical social gaze in ASD patients has been observed as early as in toddlerhood and in high-risk infants even prior to diagnosis (cf., Hutt & Ounsted, 1966; Senju & Johnson, 2009). Decreased social attention in ASD has also been shown in studies using videos of social scenes versus geometric shapes (Crawford et al., 2016), images of scenes containing people (Constantin et al., 2017; Riby & Hancock, 2008), and films of social interactions (Klin, Jones, Schultz, Volkmar, & Cohen, 2002). As mentioned, however, previous results are also quite mixed, as many similarities in looking times at faces for ASD and TD individuals were found in several other studies (Chawarska, Macari, & Shic, 2012; Freeth, Chapman, Ropar, & Mitchell, 2010). Notably, Freeth et al. (2010) found that in complex scenes containing one person, ASD individuals were slower to fixate to faces than TD controls, but once fixated, both groups spent comparable time looking at faces. This highlights the potential importance of an analytical distinction between latency of orienting versus time of exploration in studies of visual social perception in ASD.

Most eye-tracking equipment also records pupil size data, though these data have quite rarely been used until relatively recently in ASD research. Studies using pupillometry have suggested reduced cognitive engagement and/or affective arousal in ASD during social processing, as indicated by decreased pupil diameter in ASD participants compared to controls in response to faces (Anderson, Colombo, & Shaddy, 2006; Martineau et al., 2011) as well as in response to dynamic social videos (Muller, Baumeister, Dziobek, Banaschewski, & Poustka, 2016). By contrast, Wagner, Luyster, Tager-Flusberg and Nelson (2016) found that infants at high-risk for ASD showed increased pupil dilation in response to emotionally expressive faces compared to low-risk controls, suggesting heightened cognitive/affective arousal in ASD. The differences in results reflect a larger contested question in autism research, namely, whether the atypical social processing in ASD is due to a relative indifference or due to an overarousal to socio-affective information such as eye contact (e.g., Hadjikhani et al., 2017).

Complicating the issue further, it has become well established in the past decade that ASD is highly comorbid and shares symptoms, genes, and environmental risk factors with other neurodevelopmental disorders to the extent that according to Gillberg (2010), all can be said to fall under the umbrella term “Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Evaluation – ESSENCE”. For this reason, it is often difficult – yet highly relevant from both research and clinical perspectives – to determine psychophysiological mechanisms associated with social processing differences in ASD and in other commonly comorbid disorders.

The aim of the current study was to characterize the gaze patterns and the psychophysiological underpinnings of social processing in ASD by comparing findings with those of other neurodevelopmental/ESSENCE disorders. We examined the visual scanning pattern and pupil dilation in response to tightly controlled social and non-social images of a group of adolescents with either a diagnosis of ASD or another ESSENCE diagnosis (e.g., ADHD, dyslexia). Considering the existing literature highlighting the extent of social impairments in ASD, it was expected that the ASD group would take longer to fixate to, and possibly spend less time fixating on social elements of the scenes. In addressing the scarcity and lack of clarity in prior research, pupil size changes were also investigated in order to evaluate cognitive engagement/affective arousal in response to social and non-social visual stimuli. Finally, correlational investigation was carried out to examine the extent that a participant’s social symptom severity scores, according to the clinical assessment, associate with their scanning and pupil size characteristics.
Methods

Participants

The targeted group consisted of 37 adolescents who had been referred to the Child Neuropsychiatric Clinic (CNC) in Gothenburg, Sweden between year 2005 and 2009. These individuals had received childhood diagnosis of ASD and/or another ESSENCE disorder(s) at the time of initial assessment. All diagnoses were based on criteria from the DSM-IV TR (in the DSM IV TR, the Autism Spectrum Disorder was coined Pervasive Developmental Disorder, with several subcategories such as autistic disorder, Asperger syndrome, etc). A follow-up evaluation was carried out at the time of this study in which participants were re-evaluated for a variety of factors including diagnosis, IQ, social functioning, cognitive functioning, well-being, and social perception (eye-tracking and pupillometry).

One participant was excluded from the study due to experimenter error during the registration leaving the final sample size for this study at N = 36 (9 females, 27 males).

The age range for the participants was 12.9–20.0 years (M = 16.45, SD = 1.65). The IQ range for this group, based on either the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV) (Wechsler, 2003a) or the Wechsler Adult Intelligence Scale (WAIS-IV) (Wechsler, 2003b), was variable, 65–120 (M = 90.6, SD = 12.5). IQ score was unreported for one subject.

The subjects were categorized according to two main diagnostic subgroups: ASD (n= 23) and non-ASD (n = 13). The non-ASD group displayed subtler, yet still a variety of, often comorbid neurodevelopmental diagnoses. Specifically, all 13 individuals had a diagnosis of ADHD. Among these, two also received a specific reading disability/dyslexia diagnosis and two had borderline intellectual functioning.

No difference was found between the two groups regarding age (p = .361) or IQ (p = 1.0).

Standard clinical metrics/assessment

The Five to Fifteen (FTF/5–15) questionnaire (Kadesjö et al., 2004) was given at the follow up. The FTF is a parental questionnaire developed to assess development and behavioral difficulties in children and adolescents consisting of the following eight domains: motor function, attention/executive function, language, memory, learning, social competence, internalizing and externalizing behavior problems. For the purposes of this study, the social competence scores were used in data analysis. Items are scored from 0 (does not apply), 1 (applies sometimes or to some extent), or 2 (definitely applies); and thus, the higher the number of a patient’s score, the lower the functioning of the child within the given domain. Scores are unreported for three individuals in the non-ASD group. In this study sample, FTF social competence scores in the ASD group ranged from 0.19 to 1.37 (M = 0.83; SD = 0.34) and in the non-ASD group scores were 0–0.63 (M = 0.17; SD = 0.21). The difference was significant (p < .001) between groups, which is to be expected since a key diagnostic factor in ASD is poor social functioning.

The Global Assessment of Functioning (GAF) (Yamauchi, Ono, & Ikegami, 2001) is a clinical assessment of a patient’s overall level of functioning, on a scale ranging from 1 to 100 (ranging from extremely low functioning/risk for one’s own life to superior functioning). Physical and environmental limitations are not assessed; rather it is a scale of a patient’s social, communication, psychological, and occupational functioning. In the ASD group, the GAF scores ranged from 45 to 65 (M= 54.83; SD = 6.01). In the non-ASD group, GAF scores ranged from 45 to 75 (M = 61.92; SD = 9.67). There was a between-group difference (p < .05) for GAF scores in this sample, with the ASD group overall having lower functioning that the non-ASD group.

The Vineland Adaptive Behavioral Scale (VABS) (Sparrow, Cicchetti, & Balla, 2005) is a ranking of adaptive functioning given by a patient’s caregiver and covers four general domains: Daily Living, Communication, and Socialization. This study utilized the patients’ socialization scores on the VABS scale. Scores are unreported for six individuals (two ASD and four non-ASD). In the ASD group, the VABS social scores ranged from 56 to 107 (M = 80.48; SD = 15.52). The non-ASD group scores ranged from 92 to 113 (M = 106.00; SD = 8.35). There was a between-group difference on this scale as well (p < .001), again expected due to the overall low social functioning that is a defining characteristic of ASD.

Apparatus

Gaze and pupil information were measured with a Tobii T120 (Tobii Technology Inc., Stockholm, Sweden), which records near-infrared reflections of both eyes at 60 Hz as the subject watches an integrated 17-in. monitor. Subjects were seated 60 cm away from the monitor. A 9-point calibration procedure was used before the experiment, in which a small animated object is shown to the participant and collects data at nine points on the screen, repeating until data are collected for all points to ensure proper calibration.
Stimuli

We developed and used a total of 22 images, divided into 11 pairs. Each pair consisted of the same scene with (social) and without (non-social) a person present in the image. Stimuli were tightly controlled: other than presence or absence of a person, the scenes within each pair were identical (see figure 1). Subjects were divided into two groups, version A and version B, and shown a series of 11 images. Version A contained six social images and five non-social images, and version B contained five social images and six non-social images. No subject was shown two images within the same pair to avoid conflicting gaze patterns due to image recognition.

Procedure

After arriving at the Gillberg Neuropsychiatry Centre, subjects were seated approximately 60 cm from a monitor in a dimly lit, sound-proof testing room which had no natural light source. All participants completed and passed the standard 9-point calibration procedure in which a small animated object is presented at nine different positions on the screen. Gaze position relative to the position of the animated object is calculated and the resulting error vectors are presented schematically. Unfortunately, the version of Tobii studio used in our study does not allow formal calculations of calibration quality. Instead, calibration data were inspected visually where large error vectors signaled need for recalibration.

Following calibration, subjects were instructed to look freely at the photographs presented on the monitor. Stimuli were presented for 12 s each, a time course not uncommon for scene scanning studies that entail memorization (e.g., Birmingham et al., 2009; Hayes & Henderson, 2018) but relatively long for free-viewing studies of scene perception as well as pupillometric studies (e.g., Martineau et al., 2011); for this reason, in our study we focus our pupil data analyses on the first 5 s. Moreover, as pupillary responses have rapid onset (even before conscious awareness in some cases, Laeng et al., 2012) and return to baseline over time (Bast, Poustka, & Freitag, 2018), analyzing the onset rather than a longer period of time is more telling for the purposes of our study.

Data analysis

Pupil data were exported from Tobii studio and reduced using TimeStudio extension for MATLAB, an open-access analysis tool designed specifically for analyzing time series data (Nyström, Falck-Ytter, & Gredebäck, 2016). Data were inspected for various artifacts such as blinks and head movements, and rejected if these artifacts were present for more than 50% of the trace duration within the analyzed time frame (0–5 s). As recommended (Nyström et al., 2016), gaps with five or fewer consecutive poor data points were interpolated, and moving average was filtered with a five-sample window. Pupil size was averaged for each subject across both eyes. To eliminate individual differences within subjects and trials, the data were adjusted and normalized according to a baseline calculation over the initial first second (0–1 s) after stimulus onset to account for the pupil’s luminance response (e.g., Mathot, 2018). Mean pupil size for each subject was then calculated over social and non-social images over the period of 1–5 s in order to observe the response onset rather than the entirety of the relatively long time course for examining pupil data. Due to lack of sufficient quality according to the description above, pupil data were missing for both the social and the non-social condition in one individual, and in the social condition for an additional one. Note that fixation data for these individuals were included. These data were imported into SPSS (version 24) for analysis. A related samples Wilcoxon

Figure 1. Examples of the non-social (left) and social (right) images used in the study.
A Wilcoxon signed-rank test was run to determine if there was a within-group difference in each group for pupil size in response to social versus non-social scenes. In the ASD group (social pupil $M = 0.030 \text{ mm, } SD = 0.156$; non-social pupil $M = -0.007, SD = 0.119$) the test showed no significant difference between scene types ($p = 0.394$). By contrast, in the non-ASD group, a difference was found (social pupil $M = 0.083, SD = 0.145$; non-social pupil $M = -0.042, SD = 0.089, p = 0.019$), such that there was larger mean pupil size for social than the non-social scenes. Figure 2

**Discussion**

This study aimed to investigate the gaze patterns and pupillary response in response to social and non-social scene images in a clinical group of adolescents with neurodevelopmental disorders, divided into subgroups depending on presence or absence of an ASD diagnosis. Based on findings from previous research, it was expected that the ASD group would attend less to the social (face) information in the scenes. Arousal-related reactions to social stimuli were difficult to predict from previous research that have provided evidence in favor of both reduced social motivation/reactivity and in favor of theories of over-sensitivity to social information in ASD. The present study also made comparisons between clinical and empirical measures to determine whether participants’ symptom severity correlated with their response to social stimuli. With regard to methodology, a considerable strength of our study in comparison to previous work of scene processing is the tight control of the characteristics of the scenes, where the presence or absence of a person constitutes the only difference between the two scene conditions.

For the ASD group, there was no significant difference in the physiological reaction measured through pupil size.
between images that contained a social component versus those that did not, indicating a relative reduction in attention and cognitive/affective arousal and load when viewing these types of stimuli. At the same time, pupil size differences were evident in the non-ASD clinical group, suggesting that the lack of a pupillary response seems to be indicative of autism in particular. The ASD group was also slower at fixating on the social information within the images compared with the non-ASD group. In line with most current research the slowed fixation to social information in individuals with ASD in particular, especially young children, serves as another distinguishing factor between those with an autism diagnosis and other neurodevelopmental/ESSENCE disorders (for a meta-analysis, see Guillon, Hadjikhani, Baduel, & Rode, 2014).

Finally, for all participants (regardless of diagnosis) correlations were found between various social clinical measures (specifically VABS social and FTF social competence) and a participant’s latency to social AOIs. This has important clinical implications, mainly in reinforcing the clinical relevance and external validity of these assessments.

Interestingly, though there was a possible trend, no significant difference was found between groups for the total amount of time spent fixating to social information. This contradicts results from some (Guillon et al., 2014), but certainly not all earlier studies (e.g., Anderson et al., 2006). Methodological differences may potentially explain this variability among studies: we used static images, whereas some previous studies have used multimodal and/or dynamic social stimuli in their studies. Similar to us, Anderson et al. (2006) used static social and non-social images and did not see decreased looking time to social images in participants with ASD. Baranek et al. (2013), however, used a combination of social and non-social tactile, auditory, and visual stimuli and found decreased response to social stimuli in all categories. A meta-analysis of eye tracking studies by Guillon et al. (2014) revealed that in infants and young children, ASD patients showed decreased looking time to social information in general when the stimuli were dynamic and/or interactive, and showed no clear difference from controls when the stimuli were static images.
An additional methodological consideration in our study was the use of complex social scenes where faces occupied only 3–5% of the total image many existing studies have used only faces as stimuli in their experiments. Freeth et al. (2010) found that in response to static pictures of complex scenes, ASD adolescents spent a similar amount of time fixating to faces as adolescents without ASD. Fletcher-Watson, Leekam, Benson, Frank, and Findlay (2008) also found that ASD spent a normative amount of time looking at faces within social scenes, and when person-present scenes were presented alongside person-absent scenes, ASD adults showed preference for the person-present scenes. There are several other studies with scenes, however, which show that ASD patients attend less to social information (Guillon et al., 2014), so this may be an incomplete explanation. Even so, the fact that we observed correlations between various clinical measures and social orienting (latency) but not with total fixation time further reinforces the conclusion that it is important to distinguish between these two types of eye-tracking metrics in autism research on social perception, and that clear associations with social impairments might only be apparent for latency measures. One also should keep in mind the fact that some individuals with ASD are being actively trained to look at faces throughout their education, and though this practice may be less common in Sweden, it may have impacted the fixation time findings, causing them to perform better with age.

Two potential caveats of this study have to do with combining pupillometric analysis with unconstrained scanning of natural scenes. First, luminosity might be an issue. If the two subgroups (ASD and no-ASD) do not all the time attend to the same places on the screen, differences in luminosity on these diverging areas of interest would potentially influence pupil dilation over and above the role of social arousal. We attempted to control for this by calculating the luminosity on most-viewed areas of each social (p = .692) and non-social image (p = .879) in both groups and found that did not differ. Also, important to highlight is that the key comparison in our study was between the social and the non-social scenes, and from this perspective, it is hard to imagine a study with greater control of lower-level features of the scenes. Even so, differences in luminosity may still have introduced noise in the data set.

The second potential concern of this study is the pupil foreshortening error, i.e., the error in pupil diameter measurement having to do with gaze location (Brisson et al., 2013; Gagl, Hawelka, & Hutzler, 2011; Hayes & Petrov, 2016). Specifically, the fact that the screen and the camera are flat while the eye rotates means that if the gaze angle changes so might the precision in the pupil measurement. However, we do not think this error severely skews the results in our study since unconstrained scanning of natural scenes – while potentially problematic from the perspective of luminosity – do not introduce any systematic errors arising from PFE provided that certain areas of interests are not placed at only one part of the scene. That being said, some error related to pupil foreshortening have likely added further noise in the dataset. These potential drawbacks are not unique to this particular study, but characteristic of most research in the field.

A parallel to further application of pupillometry and gaze tracking in clinical research and clinical settings is a great need to understand more the nature, and ways of overcoming, such technical confounds.

In addition to examining pupillary reactions to natural scenes, the present study differed in that typically developing controls were not used. Instead, the comparison group was composed of individuals with other neurodevelopmental/ESSENCE diagnoses. This was done because ASD is highly comorbid with other developmental disorders, and these disorders often display similar symptoms. In order to outline subtle, more exact differences and provide a clear picture of social processing in ASD that is relevant for clinical work, it is necessary to compare an ASD group with individuals who have similar challenges but not ASD. For diagnostic and research purposes, it is thus valuable to utilize a neurodevelopmental/ESSENCE control group to identify ASD specific markers and differences. That being said, the non-ASD subgroup in the current study was rather small and in future research larger groups and an additional “TD” comparison group would be preferable.

In the introduction, we discussed a contested and highly fundamental issue in current autism research, namely whether the atypical social behavior of individuals with ASD reflects an indifference versus overarousal/oversensitivity to socio-affective information. The efforts to unravel that issue has led researchers in different directions, and we have previously reported several kinds of data pointing towards oversensitivity or over-arousal being a central response in ASD during processing of emotionally expressive faces with direct gaze (e.g., Hadjikhani et al., 2018; Lassalle et al., 2017), a profile that is particularly clear when the participants are asked to focus their attention at the eye region. Indeed, these stimuli may be perceived as threatening, as corroborated by increased amygdala activation (Lassalle et al., 2017). By contrast, the current study, which utilized information from spontaneous eyetracking and pupil dilations, aligns more with those studies.
that suggest a lack of interest and a lack of prioritized attention to social information in ASD (e.g., Chevalier, Kohls, Troiani, Brodkin, & Schultz, 2012; Moriuchifl et al., 2016). We speculate that part of the reason why there are so many conflicting results related to this topic could be vastly different stimuli used across studies, the components of which can lead to different engagement and arousal levels in ASD participants. First, visuospatial constraints of the scene might make a difference, such as the complexity of the scene and/or the scale and centrality of the social information (e.g., the face) in the scene. Second, the socioaffective intensity of the scenes also varies from one study to another with regard to i) the emotionality of the socioaffective information, and ii) the gaze direction of the faces in the scenes. We posit that people with ASD react more strongly compared to typically developing individuals to social information in some types of scenes (e.g., of emotionally expressive faces with direct gaze), and that they may react less strongly toward other social stimuli, including the ones used in the current experiment (e.g., in bigger scenes that include a lot of background and people without direct gaze), as a consequence of reduced development of the social brain that results from gaze avoidance and less time spent learning from facial expressions (Hadjikhani et al., 2017). Also, the heterogeneity of autism might contribute to the inconsistency of previous findings in research on arousal levels during social processing. Future research using eye-tracking and pupillometric data should systematically explore these ideas in order to shed further light on the social impairment in autism, including the cascading effects of social processing differences in autism in response to different kinds of social stimuli.

In sum, our study revealed subtle differences in the way adolescents with ASD visually attend and physiologically react to social scenes. Our findings have both clinical and experimental implications since these features were observed to align with both standard clinical measures of social functioning and to distinguish ASD from a clinical comparison group without ASD.

Acknowledgments

We would like to express our gratitude to the participants in this study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Stena Foundation.

References


