scientific reports

Check for updates

OPEN

Social attention in the wild - fixations to the eyes and autistic traits during a naturalistic interaction in a healthy sample

Raimund Buehler[1]✉, **Ulrich Ansorge**[2,3,4] & **Giorgia Silani**[1]✉

Attention to social stimuli is a key component of social behavior and facilitates the development of fundamental social skills. Studies investigating social attention in neurotypical or neurodiverse populations have often relied on screen-based experiments using static images or videos, which lack the sensory richness and reciprocity present in real-life social interactions. This can possibly be attributed to the challenges one encounters when creating naturalistic experiments, such as dealing with dynamically moving areas of interest (AOIs), which require either time-intensive manual coding or restraining of participants. Here, we present findings from an experimental paradigm using unrestrained mobile eye-tracking and a face detection algorithm (MTCNN) to measure fixation rates during a semi-structured, face-to-face interview. Data from $N = 62$ healthy adult participants was analyzed for gaze behavior and related to participants' autistic traits. We observed a significant negative correlation between fixation rates on the eye region averaged over the entire interaction and scores on the autism spectrum quotient (AQ) ($r = -0.14$), indicating participants with high autistic traits fixated less frequently on the eye region. We also compared different types of interview questions (open vs. closed) to explore whether the reduction in fixation rates was more pronounced for specific time intervals during the interview. Lastly, we discuss both possibilities for extensions as well as limitations of the presented paradigm that could serve as inspiration for future research.

At almost every moment of our life, we are surrounded by an overwhelming amount of information. Selectively attending to only the most relevant stimuli is vital for adaptive functioning and well-being. Social stimuli, such as faces, voices, or bodily movements, carried a particular importance for human behavior during evolution[1] and are often preferentially attended to by typically developing individuals[2]. This bias towards social stimuli, present from early infancy onwards, facilitates social interaction and the acquisition of social skills across development[3,4].

In autism spectrum disorder (ASD), an overall reduction in attention to social stimuli has been observed[2], with the largest effect sizes found for the eye region[5]. This effect is present across children[6], adolescents[7], and adults[8]. Excess gaze to the mouth has sometimes been reported in children[9] and adults[7], however this failed to replicate in a number of later studies[10]. Presumed aberrations in visual social attention seem to generalize to the broader autism phenotype, operationalized by autistic trait questionnaires such as the autism spectrum quotient (AQ)[11–15], the social skills subscale of the AQ[16], the Social Responsiveness Scale (SRS)[17], or when examining parents of ASD individuals[18].

Notably, to investigate mechanisms of social attention, researchers have predominantly relied on screen-based experiments[19]. In these settings, static or dynamic stimuli are displayed, such as pictures of faces or videos of social scenes[12,19,20]. This has resulted in a wealth of knowledge about specific mechanisms of social attention such as gaze cueing[21], emotion recognition[22,23] and inferring mental states from the eye region of others[24]. Additionally, a bias for the eyes has been observed across a range of stimuli, such as pictures of faces[25], complex social scenes[26], and dynamic videos of social interactions[27].

However, several researchers have questioned the generalizability of findings generated in such settings to real-life social situations[11,28], pointing out important differences in results gained from screen-based and naturalistic experiments[12,16,27]. Importantly, self-reported difficulties and clinical assessments of eye contact

[1]Department of Clinical and Health Psychology, Faculty of Psychology, University of Vienna, Liebiggasse 5, 1010 Vienna, Austria. [2]Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria. [3]Vienna Cognitive Science Hub, University of Vienna, Vienna, Austria. [4]Research Platform Mediatized Lifeworlds, University of Vienna, Vienna, Austria. ✉email: raimund.buehler@univie.ac.at; giorgia.silani@univie.ac.at

occur during real, face-to-face social interactions[29,30]. This has led to calls for more experiments to investigate social attention "in the wild"[28,31,32], involving the actual presence of another person[12,16], with the possibility to study the temporal and reciprocal nature of social interactions.

One major challenge in creating naturalistic experiments is dynamically moving areas of interest (AOIs), which require time-consuming manual coding procedures[11,33] or restraining of participants. Some researchers have therefore opted to reduce the complexity of the interaction, for example by creating a video-call setup, to ensure static AOIs across the duration of the experiment[11,34]. Others have implemented computer vision algorithms to automate the detection of faces and facial landmarks[33,35–38]. Early studies have used comparatively simple face detection algorithms[38], however more sophisticated approaches involving convolutional neural networks (CNNs) trained for face detection are emerging[35], with promising results when compared to manual coding procedures[33,36].

Still, the implementation of computer vision algorithms remains challenging and most studies in naturalistic settings continue to use manual coding of AOIs[33]. The goal of this study was therefore to investigate the feasibility of using face detection algorithms in naturalistic experiments and to explore the effects of autistic traits on gaze behavior using unrestrained mobile eye tracking and computer vision assisted coding of AOIs. We also aimed to explore characteristics of gaze behavior over the timecourse of the interaction, specifically in response to different types of questions asked during a semi-structured interview[34,39].

Materials and methods

Participants

All participants gave written informed consent before participation. Experimental procedures were carried out in accordance with the Declaration of Helsinki and experimental protocols were approved by the ethics committee of the Medical University of Vienna (EK-1393/2017). Informed consent for the publication of identifiable images in an online open-access format was obtained from depicted members of the research team (Fig. 1).

The total recruited sample consisted of $N = 74$ participants. 13 participants were psychology students, who received course credit for participation. 61 participants were not studying psychology and were recruited via another channel. These participants received 10€ for compensation. The sample size was based on an a priori power analyses using G*Power, assuming a medium effect size of $r = 0.3$ and a power of $\beta = 0.8$. This resulted in

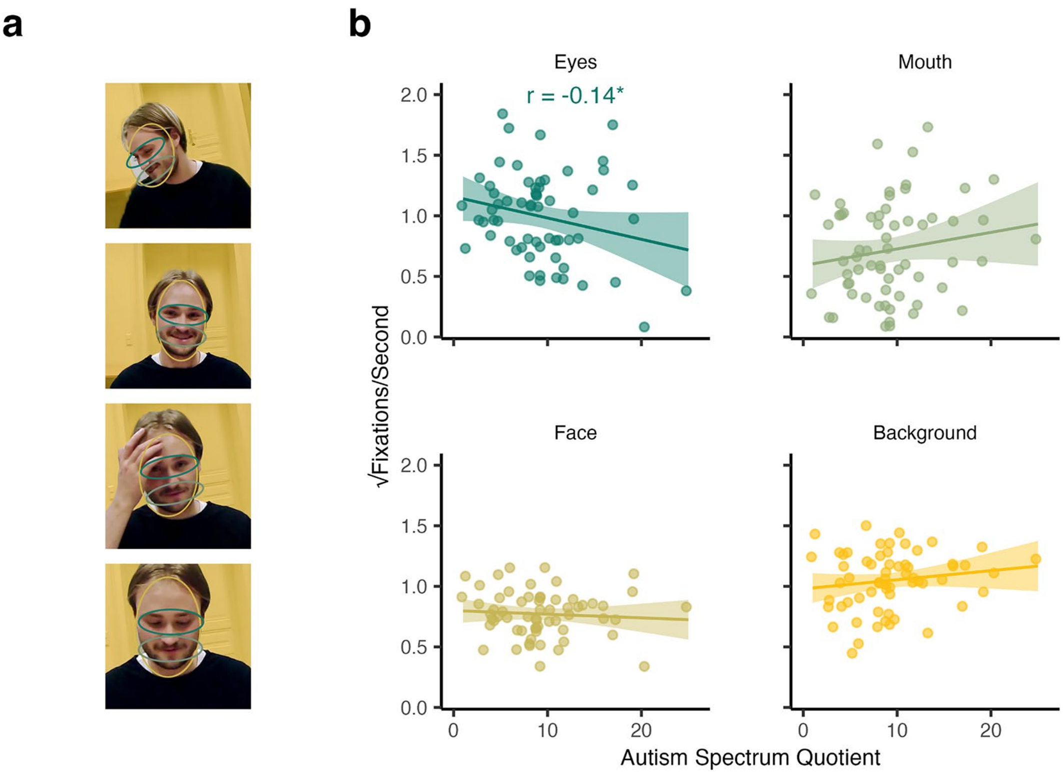

Fig. 1. **a** Exemplary output frames of the ellipses fitted to the face and facial regions of the experimenter using *cv2* and the *facenet-pytorch* package. The face and facial landmarks are accurately detected in a range of conditions, for example when looking to the side, in frontal view, when the face is partially occluded or when looking down. **b** Scatterplots for fixation rates on each AOI, averaged over the entire interview. Every data point represents a participant, slightly jittered for improved visibility. Regression lines are depicted as coloured lines and confidence intervals in shaded colors. Fixation rates were square-root transformed to account for non-normality of residuals in linear models. Summary statistics for untransformed fixation rates can be found in the supplementary material.

an estimated sample size of 64 participants, however we decided to include slightly more participants, to account for potential missing data.

Inclusion criteria comprised an age of 18–65 years, fluency in English and absence of psychiatric or neurological disorders including drug and alcohol addiction or abuse. Participants were screened for these criteria via single items on a self-report questionnaire. To control for potential differences in gaze behavior based on the sex of the interaction partner, only heterosexual participants were included, and all participants were interviewed by an experimenter of the same sex.

Eye-tracking data for 5 participants was missing due to recording errors and therefore excluded from analysis. 5 participants were tested but did not fulfill inclusion criteria and were excluded from analysis (3 due to non-heterosexual orientation, 1 due to a history of alcohol abuse and 1 due to regular drug use). 2 participants were excluded due to missing data on the prescreening questionnaire. The final sample entering analysis consisted of $N = 62$ participants (33 females, 29 males), 19–63 years of age ($M = 26.39$, $SD\text{t} = 6.84$) (see Table S1).

Procedure

After completing the prescreening questionnaire, participants were invited to a lab appointment. First, participants gave informed consent and were instructed about the experimental procedure. Next, the eye-tracking headset was put on and calibration was performed using the single-marker calibration choreography. Instructions and calibration lasted approximately 10 min, after which the eye tracking recording was started with a button press. Timepoints for interview questions were created by pressing buttons with the respective question label. An additional computer-based task was administered during the course of the experiment related to another research question. The order of tasks was counterbalanced to avoid sequence effects. Thus, the beginning of the interview was not constant across participants, but either 10 or 40 min after the start of the experiment. After the interview was completed, the recording was terminated with a button press. Average interview duration over all participants was $M = 7.03$ min, $SD = 4.16$ min.

When both tasks had been completed, participants were given the AQ[40] in its shortened version, translated to German and abbreviated to 33 items with most discriminatory power[41]. The AQ is a self-report measure designed to screen for autistic traits in the general population. In this study, it was used as a proxy measure of the broader autism phenotype. Mean of AQ was $M = 9.34$, $SD = 4.82$ with a range of 1–25 and 4 participants scoring above the threshold of 17 points[41] (see Table S1). Additionally, participants were asked to rate how sympathetic the experimenter was perceived (attractiveness). Attractiveness and gender were used as covariates in the main analysis, in order to control for the confounding effect of these variables.

Eye-tracking interview

A semi-structured interview using a set of 12 questions was conducted during the eye tracking recording. Participant and experimenter were sitting at a table facing each other at a distance of 1 m. The participant was wearing an eye-tracking headset described below in more detail. Due to pandemic restrictions, a fully opaque plastic screen was placed between participant and experimenter.

Difficulties in eye contact in ASD and the broader autism phenotype are reported to occur during everyday social interactions[30,42]. Likewise, diagnostic assessment of eye contact is largely based on short interactions between clinician and patient[29]. Thus, similar to previous studies[12,34,43], we created a set of questions meant to resemble everyday conversation such as: “Did you sleep well last night?” or “Did you see a movie last week?”. As indicated by findings of reduced eye contact during social interactions[30,42], we expected to observe a reduction of fixations to the eyes for individuals high in autistic traits.

However, we were also interested in whether this effect would be more or less pronounced for certain types of questions. To investigate this, we included both open and closed questions. The interview started out with a number of closed questions (Question 1–6, e.g., “Did you sleep well last night?”), to which participants could answer with a simple “yes” or “no” and continued with a set of open questions, which required a more elaborate answer and possibly more social engagement (Question 7–12, e.g., “What was the most interesting thing that happened to you in the last week?”). We hypothesized that open questions would exhibit a stronger relationship between autistic traits and fixations to the eyes compared to closed questions.

Question markers were created using the annotation feature of the pupil labs capture software by pressing a set of keys corresponding to the respective question. The list of interview questions can be found in the supplementary material and on this paper’s OSF repository (https://osf.io/y9t5q/).

Eye-tracking equipment

Eye-tracking data was collected using the Pupil Labs Core eye-tracking platform (consisting of a wearable eye-tracking headset and open-source recording and analysis software)[44]. The manufacturer reports the headset to achieve 0.6° visual angle (VA) accuracy under ideal conditions (defined as the angular offset between estimated fixation and corresponding fixation target location). The headset consists of two pupil cameras for gaze estimation and one world camera on the forehead of the participant, recording the participants’ field of view.

Eye movements were recorded with 120 Hz sampling frequency. Pupil locations were estimated using the “dark pupil” detection method, which first identifies dark parts and contrasting edges of a converted-to-grayscale image to which an ellipse is fitted[44,45]. We used the 3D pupil detection method, based on a 3D model of the eyes, due to its ability to compensate for movement of the head and slippage of the headset. To further improve data quality, we instructed participants to avoid large movements of the head. To calibrate the headset, we used the single marker calibration choreography with a printed physical marker. The marker was held up in front of the experimenter’s face during calibration to ensure appropriate accuracy for this depth level. Mean accuracy over all recordings was $M = 1.70°$ VA, $SD = 0.59°$ VA. At a distance of 1 m, this corresponds to an area of 2.96 cm, roughly the diameter of a large coin (e.g., a 2€ coin).

As the experimenter moved during the interview, his face was not a constant size in the participant's visual field. We used the following formula to calculate the degrees of visual angle of the face in the participant's field of view:

$$\alpha = 2 \cdot \arctan\left(\frac{\text{Object size}}{2 \cdot \text{Distance}}\right) \tag{1}$$

Assuming the face to be approximately 20 cm in height and 10 cm in width, the face took up $\alpha = 11.42°$ VA vertically and $\alpha = 5.72°$ VA horizontally at 1 m distance. In the world camera (1280 × 720 resolution), the face took up approximately 150 pixels vertically and 90 pixels horizontally.

Data preprocessing and face detection

Fixations were extracted from gaze data using the built-in fixation detector plugin of the pupil labs software. Maximum dispersion was set to 1.5° with a minimum duration of 80 ms and a maximum duration of 220 ms. Blinks (short drops of confidence in pupil detection) were removed using the built-in blink detector plugin. Furthermore, all detected fixations below a threshold of 0.8 confidence were filtered during a separate preprocessing step.

We analyzed fixation locations within the participant's field of view frame by frame using openCV (*cv2* package) and the *facenet-pytorch* package in python[46]. This package uses a multi-task convolutional neural network (MTCNN) to detect faces in a given image and takes inspiration from David Sandberg's TensorFlow implementation of the FaceNet face recognition model[47,48]. It also includes face detection models pretrained on the VGGFace2 and CASIA-Webface datasets[49,50]. The code to analyze recordings has been published as a github-repository and is linked to this papers OSF-page (https://osf.io/y9t5q/*)* or can be accessed on GitHub directly (https://github.com/raimund-buehler/face_detection_small*)*.

Each frame of the world video was first analyzed by the face detection algorithm, yielding the location of the experimenter's face as well as the position of facial landmarks, such as eyes and corners of the mouth. Ellipses were automatically fitted to the face, eyes, and mouth (see Fig. 1). The ellipse for the face was fitted using the center, width and height values of the bounding box of the face returned by the face detection algorithm. For eyes and mouth, the midpoint between left and right eye or left and right corner of the mouth was computed and used as center for the ellipse. The width value of the facial bounding box was used to fit the ellipse while the height value was scaled down by a factor of 0.3. For eyes and mouth, ellipses were also angled according to the angle formed by the left and right eye or corners of the mouth. Thus, ellipses for eyes and mouth automatically adjusted to the position and size of the experimenter's face in the participant's visual field.

Lastly, if a fixation for a frame was present, it was tested whether it fell onto one of the ellipses (with all ellipses being mutually exclusive). Fixations for each region were then counted and averaged over different time intervals (e.g., total interview duration or question duration), leading to a measure of average fixation rate in Hz (fixations / second) on the respective region during that time interval. This measure served as the dependent variable for further analysis.

Statistical analysis

The main analyses for this study were preregistered on OSF (https://osf.io/tgcqm*)*. We fit a linear model with fixation rate as dependent variable, an interaction of AOI and AQ as independent variables and gender and perceived attractiveness of the experimenter (Attr.) as covariates. AQ and AOI were defined as interactive effects in order to investigate the effect of AQ at each level of AOI, for instance specifically for the eye or mouth region. Due to non-normality of residuals (see supplementary results), we used a square root transformation for the dependent variable, as described before for similar eye-tracking data[34]. This transformation was also preregistered.

A number of additional exploratory analyses were conducted, which were not part of the preregistration: First, we investigated the fixation rates on eyes and mouth over the timecourse of the interaction by creating longitudinal plots for each participant (see Fig 2). Second, we looked at differences in fixation rates to the eyes when averaged within questions. Lastly, we compared sensitivity to autistic traits for open vs. closed questions by fitting a linear mixed model with AQ and question type (open vs. closed) as fixed effects, including an interaction term between AQ and question type as well as random intercepts by subject.

For fitting linear models, we used the base *lm()* command in R. For mixed models, the *lmer()* command of the *lmerTest*-package[51] was used. The *anova()* command was used to obtain *F*-tests for main effects and interactions. Significant interactions were further investigated in a simple slope analysis using the *emmeans*-package[52]. Plots were created using *ggplot*[53] and the *interactions*-package[54].

Model assumptions were investigated using plots of residuals versus fitted values, conditional boxplots within levels of AOI, and qqPlots (see supplementary results). Importantly, for a model without transformation of the dependent variable, histograms and qqPlots indicated a significant deviation of the residuals from a normal distribution, confirming the necessity of a square-root transformation for robust results.

We have also created a synthetic version of our dataset using the *synthpop*-package in R. This dataset is a scrambled and anonymized version of the original dataset, which preserves the most important summary aspects[55]. The dataset is freely available on the OSF repository (https://osf.io/y9t5q/*)* along with a summary of the R code used for plotting, analysis, and model diagnostics. Although the synthetic dataset does not fully capture the results presented here, it permits researchers to reproduce the most important aspects of our analysis.

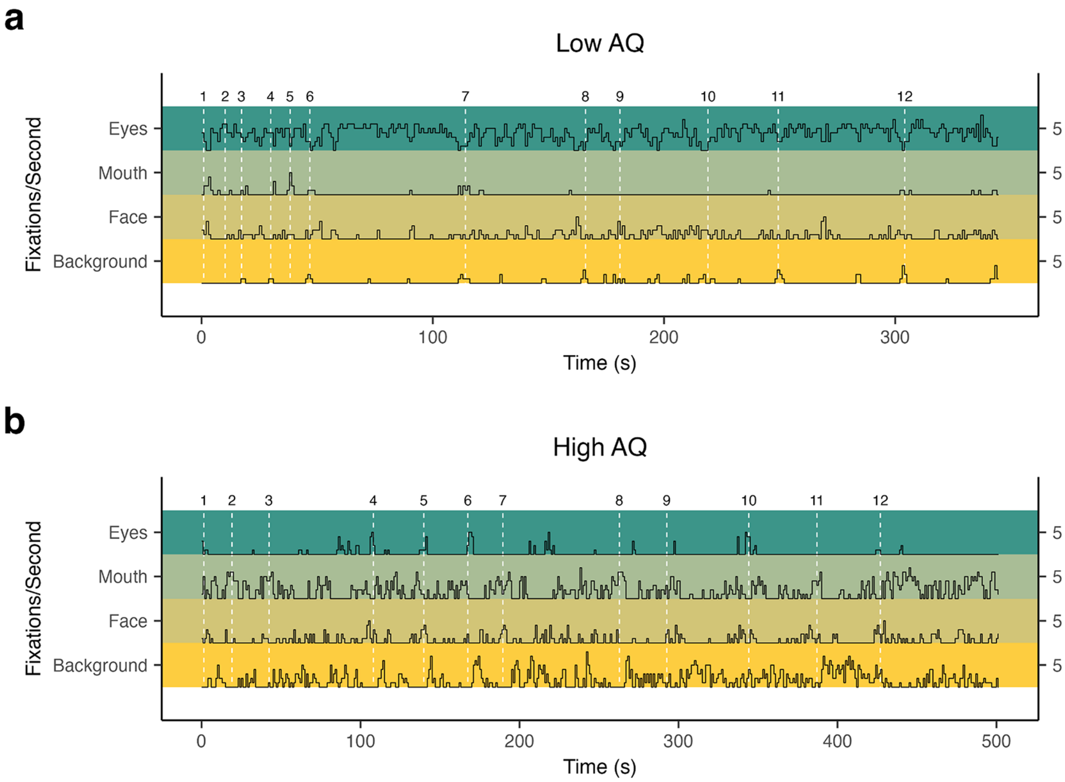

Fig. 2. - Longitudinal plots of fixation rates (binned over 1s intervals) for two exemplary participants with low AQ (5 points) and high AQ (17 points). Question onsets are vertical, dashed lines. The high AQ participant is fixating more on the mouth and background region over the entire interview, whereas the low AQ participant is fixating mostly on the eyes.

Results

Main analysis

Fitting the model described above yielded the following results: the main effect of AOI was significant, $F(3, 238) = 17.37$, $p < 0.001$, suggesting overall differences in the mean amount spent fixating on the different AOIs. Averaged over the entire interview, participants mostly fixated on the eyes ($M = 1.15$ fix/s, $SD = 0.74$) and the background ($M = 1.14$ fix/s, $SD = 0.47$) and less on the mouth ($M = 0.67$ fix/s, $SD = 0.65$) or the rest of the face ($M = 0.63$ fix/s, $SD = 0.29$) (see Table S2). Post-hoc contrasts revealed that significantly more fixations fell on the eye and background region, when compared to the mouth and the rest of face (all $p < 0.001$). However, there were no significant differences between eyes and background or between the mouth and rest of face (all $p > 0.05$). AQ was not significant as a main effect ($F(1, 238) = 0.001$, $p = 0.97$), but the interaction of AOI × AQ was significant, $F(3, 238) = 2.81$, $p = 0.04$. This suggests the effect of AQ to be dependent on AOI.

Post-hoc analyses of the interaction revealed it to be driven by a negative effect of AQ for fixations to the eye region (See Fig. 1). In a simple slope analysis, this effect was significantly different from zero ($t(238) = -2.17$, $p = 0.031$). Converting the t-value of the simple slope analysis to a partial correlation coefficient resulted in an effect size of $r = -0.14$. Submitting this effect size to a post-hoc power analysis revealed an actual power of $\beta = 0.58$.

The simple slope analysis also showed that the effect of AQ did not differ significantly from zero for any other AOI (all $p > 0.05$). The main effects of gender ($F(1, 238) = 1.70$, $p = 0.19$) and attractiveness ($F(1, 238) = 0.06$, $p = 0.81$) were not significant. Comprehensive results of these analyses in table format can be found in the supplementary material (S3 and S4).

Exploratory analyses

We conducted several additional exploratory analyses to investigate the timecourse of gaze behavior during the interview. First, we created longitudinal plots for each participant, depicting dynamic fixation rate changes over time and during question intervals. Exemplary plots for two participants with low and high AQ-Score are included in this paper (see Fig. 2). As apparent, the high AQ participant fixated mostly on the mouth and less on the eyes over the entire interaction. Conversely, the low AQ participant fixated mostly on the eyes and less on the mouth.

Additionally, we were interested in whether the association of autistic traits and fixation rates to the eyes was a constant effect over the entire interview, or whether it was particularly pronounced for certain parts of the interaction and/or for specific questions (open vs. closed questions).

First, to explore overall differences between questions, we computed the average fixation rate to the eyes for each question and fit a linear mixed model with question number as a main effect and random intercepts by subject. As can be seen in Fig. 3, most fixations to the eyes were made during question 4 ("Have you been out of the country in the past week?") and least for question 9 ("What was the most interesting thing that happened to you in the past week?"). Question 4 (a closed question), differed significantly from question 9, 10 and 11 (open questions, all $p < 0.0001$), with less fixations to the eyes made for open questions.

Lastly, we were also interested in whether open vs. closed questions were differently sensitive to autistic traits. Figure 4 depicts the average fixation rate to the eyes by AQ Score for each question separately. We fit a linear mixed model with question type (open vs. closed) and AQ as main effects, as well as an interaction of question type by AQ and random intercepts by subject (Table S5). As in the model above, the main effect of question type was significant, $F(1, 653.67) = 6.41$, $p = 0.01$, suggesting differences in the average fixation rate on the eyes between open and closed questions with less fixations to the eyes made for open questions. AQ was not a significant main effect in this model, $F(1, 60.62) = 2.12$, $p = 0.15$. Furthermore, the interaction between AQ and question type was not significant $F(1, 658.22) = 0.51$, $p = 0.48$, indicating that open or closed questions were not differentially sensitive to the effect of autistic traits. Thus, while an effect of autistic traits was present when averaging fixation rates over the entire interaction, this was not related to any specific event during the interview, but rather acted as a general effect in this experiment.

Discussion

In light of the need for experiments investigating social attention "in the wild"[28,31,32], the goal of this study was to explore the use of face detection algorithms in a naturalistic experiment, to investigate the effects of autistic traits, and to examine the timecourse of gaze behavior over a face-to-face interaction. Using the MTCNN face detection algorithm, we observed a significant negative relationship between AQ and the average fixation rate to the eye region. The effect of autistic traits was not significant for any of the other AOIs. Thus, contrary to some screen-based experiments[7], we did not observe a significant increase in fixations to the mouth for high AQ participants. Some open questions elicited less fixations to the eyes compared to closed questions, potentially because they required a more elaborate response and participants would fixate more on the background while deliberating their reply. However, open questions were not more sensitive towards autistic traits compared to closed questions, as indicated by a non-significant interaction of AQ and question type. Thus, in this study, the reduction of fixations to the eyes was a constant effect over the entire interaction and was not related to the onset of a particular type of question. Generally, the reduction of fixations to the eyes for high AQ participants observed in our study is in line with findings gained in screen-based settings[2,5,10,56] and studies in naturalistic environments, using non-mobile eye tracking or different approaches to AOI coding[11,14,37]. However, our results are in contrast to some studies, which found effects in screen-based, but not in live interaction conditions, for ASD[43,57] or autistic traits[14]. This can likely be attributed to differences in task design and experimental procedure, but more research is needed to determine the precise circumstances in which effects of ASD or autistic traits can be observed.

ASD and other disorders of social behavior are inherently marked by deficits in social interactions in the real world[29,30,42]. It is therefore important to study these behaviors in situations in which they actually occur,

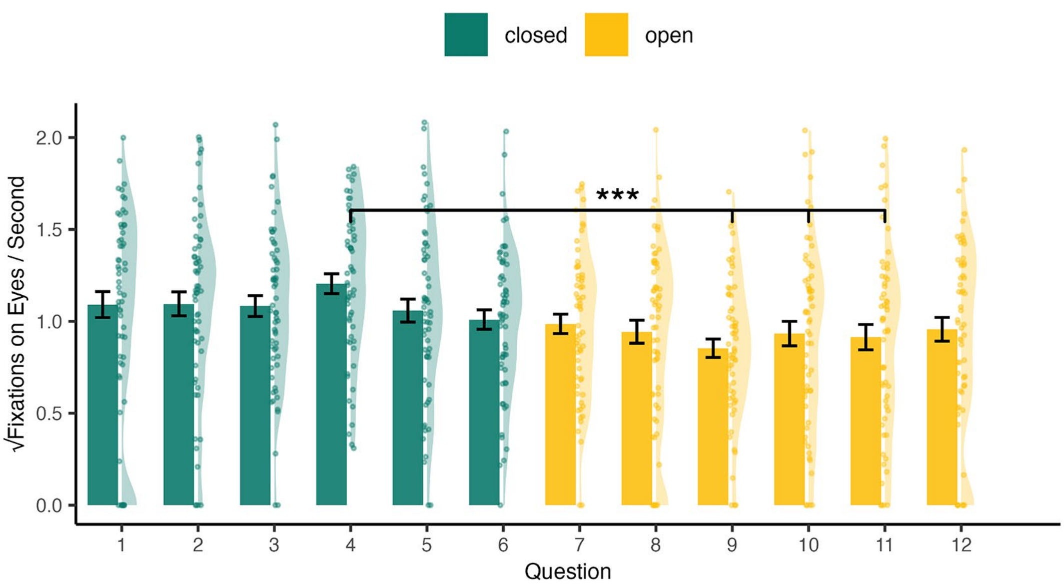

Fig. 3. Average fixation rates within questions for fixations to eyes only. Questions are color coded for open and closed questions. Means and standard errors are depicted in black, significant differences are marked with stars ($p < 0.0001$). Fixation rates were square-root transformed to account for non-normality of residuals in linear models.

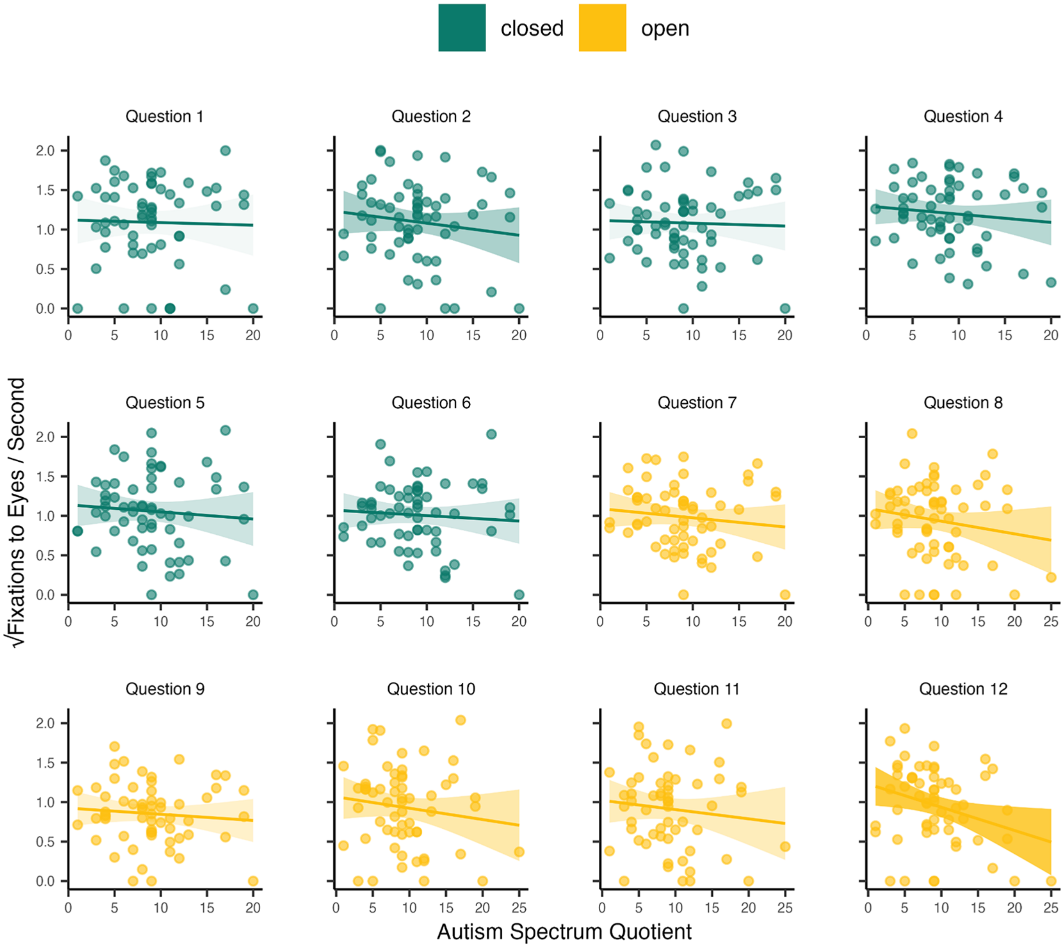

Fig. 4. Fixation rates to the eyes by AQ Score for each question. Questions are color coded for open and closed questions. . The strength of the relationship is mapped to the transparency of confidence intervals. Darker confidence intervals represent stronger associations, lighter confidence intervals weaker associations. Fixation rates were square-root transformed to account for non-normality of residuals in linear models.

involving the presence of another person and including the complexity and richness of real social interactions. Furthermore, one of the major advantages of naturalistic over screen-based experiments is the possibility to study temporal and reciprocal aspects of social attention in everyday interactions. Here, we investigated the timecourse of gaze behavior by plotting fixation rates on AOIs over time and by modelling fixation rates within interview questions. In our study, fixation rates for participants with high AQ were not particularly affected by specific question types but rather reduced over the entire interaction. However, longitudinal analyses of gaze behavior could be expanded to capture interactive aspects more fully, for example in response to episodes of speech or the display of emotional expressions by the experimenter.

Emotion recognition algorithms would be a relatively straightforward addition to the face detection algorithm and would allow to automatically classify time intervals when emotional expressions are displayed (smiles, frowns etc.). Episodes of speech could be decoded from an audio signal of the recording[58], although speaker detection is still a challenging aspect. Temporal profiles of gaze behavior, for instance via growth-curve modeling[59], could provide a fruitful statistical approach for longitudinal analysis of gaze behavior. Other studies have also used generalized additive mixed models (GAMMs) to study the temporal distribution of fixations on facial regions[60]. Growth-curve modeling or GAMMs could be used to model the timecourse of gaze behavior in the first seconds after the onset of emotional expressions or speech and investigate whether gaze behavior in these critical time periods differs for ASD or high AQ individuals.

Taken together, our results demonstrate how the utilization of mobile eye-tracking in combination with face detection algorithms offer a promising venue for research in social attention. In light of differences between findings gained in screen-based settings and naturalistic interactions[16], with some results not generalizing to naturalistic interactions[43,57], developing paradigms that involve realistic social interactions is necessary to generate findings that extend beyond lab settings[61]. Adopting technological advances in face detection algorithms can aid in removing barriers towards such an approach.

Some limitations of the experiment also need to be taken into account. First, our sample did not include participants with an ASD diagnosis, but consisted of a sample of healthy individuals. We used scores on the AQ to capture the broader autism phenotype in the general population. We believe that we have achieved an appropriate range in AQ scores with some participants scoring above the threshold indicative of ASD. Still, results may differ compared to individuals with an ASD diagnosis. Additionally, the study was not at ideal power level ($\beta = 0.58$), as some participants due to exclusion criteria, missing data on questionnaires, or missing eye tracking data. However, we believe power and sample size were adequate and comparable to other studies in the field.

Furthermore, as participants tend to move their head during conversation (e.g., when nodding) slippage of the headset is to be expected and deteriorates the accuracy of gaze estimation. We have tried to tackle this problem by using the 3D-estimation method, which is able to compensate for movement and headset slippage. We also instructed participants not to move their head too much. Instances of movement should primarily influence the confidence of gaze estimation and relatedly the confidence of fixation detections. Fixations below a certain confidence level were removed during preprocessing. Thus, any potential impact of movement would likely manifest as a general reduction in fixation rates across all AOIs. Given that our findings show an effect of autistic traits specifically for the eye region but not for other AOIs, it is unlikely that movement confounded this result. Furthermore, no significant main effect of AQ was found on overall fixation rates. Nevertheless, the effect of head movement on data quality is important and future studies should address this point more exhaustively.

Third, and relatedly, the size of ellipses fitted to the face, eyes, and mouth was based on the height and width of the facial bounding box returned by the face detection algorithm. Furthermore, while ellipses adjust to changes in the experimenter's head pose, they did not adjust in size. AOI size affects the classification of fixations on AOIs: We used comparatively large AOIs, which compensate for false negative classifications in situations of low accuracy, but at the cost of increasing false positives[62]. Choosing smaller AOIs may have resulted in less fixations classified as falling on the eyes and mouth and more fixations falling on the rest of the face and background, which potentially could have influenced results.

Lastly, the presence of an eye-tracking headset and explicit behavioral instructions may impact the realism of the social situation and, therefore, influence gaze behavior[43,63]. The tradeoff between the invasiveness of the method and the realism of the experimental setup has been described in previous research[64]. While our setup allows for natural interaction, the eye-tracking headset remains noticeable to the participant. More recent approaches, such as appearance-based gaze estimation using OpenFace, which estimate gaze direction based on facial features and head-poses in video recordings, offer promising less invasive alternatives[65]. However, they currently lack the precision of dedicated eye-tracking hardware. Future developments may improve the accuracy of these approaches, offering more ecologically valid methods without compromising data quality.

Despite these limitations, we are confident that the paradigm presented here is sensitive towards interindividual differences in autistic traits and captures important aspects of social attention in face-to-face interactions. Future research may investigate the timecourse of gaze behavior more comprehensively in a longitudinal way and uncover characteristics of gaze in response to certain events during the interaction, such as the display of emotions. We encourage researchers to use or adapt our face detection code or experimental design to achieve these goals.

Data availability

A synthetic version of the dataset as well as a summary of the code used for analysis can be found under the following link (Scientific_Reports/R Analysis): https://osf.io/y9t5q/files/osfstorage While this dataset does not fully capture the results presented here, it permits researchers to reproduce the most important aspects of our analysis.

Received: 5 March 2024; Accepted: 6 November 2024
Published online: 03 December 2024

References

1. Emery, N. J. The eyes have it: the neuroethology, function and evolution of social gaze. *Neurosci. Biobehav. Rev.* **24**, 581–604. https://doi.org/10.1016/S0149-7634(00)00025-7 (2000).
2. Chita-Tegmark, M. Social attention in ASD: A review and meta-analysis of eye-tracking studies. *Res. Dev. Disabil.* **48**, 79–93. https://doi.org/10.1016/j.ridd.2015.10.011 (2016).
3. Dawson, G., Webb, S. J. & McPartland, J. Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. *Dev. Neuropsychol.* **27**, 403–424. https://doi.org/10.1207/s15326942dn2703_6 (2005).
4. Schultz, R. T. Developmental deficits in social perception in autism: The role of the amygdala and fusiform face area. *Int. J. Dev. Neurosci.* **23**, 125–141. https://doi.org/10.1016/j.ijdevneu.2004.12.012 (2005).
5. Frazier, T. W. et al. A Meta-analysis of gaze differences to social and nonsocial information between individuals with and without autism. *J. Am. Acad. Child Adolesc. Psychiatry* **56**, 546–555. https://doi.org/10.1016/j.jaac.2017.05.005 (2017).
6. Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G. & Jones, W. Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature* **459**, 257–261. https://doi.org/10.1038/nature07868 (2009).
7. Klin, A., Jones, W., Schultz, R., Volkmar, F. & Cohen, D. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch. Gen. Psychiatry* **59**, 809–816. https://doi.org/10.1001/archpsyc.59.9.809 (2002).
8. Pelphrey, K. A. et al. Visual scanning of faces in autism. *J. Autism Dev. Disord.* **32**, 249–261. https://doi.org/10.1023/A:1016374617369 (2002).
9. Jones, W., Carr, K. & Klin, A. Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Arch. Gen. Psychiatry* **65**, 946–954. https://doi.org/10.1001/archpsyc.65.8.946 (2008).
10. Guillon, Q., Hadjikhani, N., Baduel, S. & Rogé, B. Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neurosci. Biobehav. Rev.* **42**, 279–297. https://doi.org/10.1016/j.neubiorev.2014.03.013 (2014).
11. Hessels, R. S., Holleman, G. A., Cornelissen, T. H. W., Hooge, I. T. C. & Kemner, C. Eye contact takes two—autistic and social anxiety traits predict gaze behavior in dyadic interaction. *J. Exp. Psychopathol.* https://doi.org/10.5127/jep.062917 (2018).

12. Freeth, M., Foulsham, T. & Kingstone, A. What affects social attention? Social presence, eye contact and autistic traits. *PLoS ONE* **8**, e53286. https://doi.org/10.1371/journal.pone.0053286 (2013).
13. Nummenmaa, L., Engell, A. D., Von dem Hagen, E., Henson, R. N. A. & Calder, A. J. Autism spectrum traits predict the neural response to eye gaze in typical individuals. *NeuroImage* **59**, 3356–3363. https://doi.org/10.1016/j.neuroimage.2011.10.075 (2012).
14. Vabalas, A. & Freeth, M. Brief report: Patterns of eye movements in face to face conversation are associated with autistic traits: Evidence from a student sample. *J. Autism Dev. Disord.* **46**, 305–314. https://doi.org/10.1007/s10803-015-2546-y (2016).
15. Wegner-Clemens, K., Rennig, J. & Beauchamp, M. S. A relationship between Autism-Spectrum Quotient and face viewing behavior in 98 participants. *PLoS ONE* **15**, e0230866. https://doi.org/10.1371/journal.pone.0230866 (2020).
16. Laidlaw, K. E. W., Foulsham, T., Kuhn, G. & Kingstone, A. Potential social interactions are important to social attention. *Proc. Natl. Acad. Sci.* **108**, 5548–5553. https://doi.org/10.1073/pnas.1017022108 (2011).
17. Kleberg, J. L. et al. Autistic traits and symptoms of social anxiety are differentially related to attention to others' eyes in social anxiety disorder. *J. Autism Dev. Disord.* **47**, 3814–3821. https://doi.org/10.1007/s10803-016-2978-z (2017).
18. Nayar, K., Shic, F., Winston, M. & Losh, M. A constellation of eye-tracking measures reveals social attention differences in ASD and the broad autism phenotype. *Mol. Autism* **13**, 18. https://doi.org/10.1186/s13229-022-00490-w (2022).
19. Risko, E. F., Laidlaw, K. E. W., Freeth, M., Foulsham, T. & Kingstone, A. Social attention with real versus reel stimuli: Toward an empirical approach to concerns about ecological validity. *Front. Hum. Neurosci.* **6**, 143. https://doi.org/10.3389/fnhum.2012.00143 (2012).
20. Chevallier, C. et al. Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Res.* **8**, 620–628. https://doi.org/10.1002/aur.1479 (2015).
21. Frischen, A., Bayliss, A. P. & Tipper, S. P. Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychol. Bull.* **133**, 694–724. https://doi.org/10.1037/0033-2909.133.4.694 (2007).
22. Hernandez, N. et al. Exploration of core features of a human face by healthy and autistic adults analyzed by visual scanning. *Neuropsychologia* **47**, 1004–1012. https://doi.org/10.1016/j.neuropsychologia.2008.10.023 (2009).
23. Leppänen, J. M. & Nelson, C. A. Tuning the developing brain to social signals of emotions. *Nat. Rev. Neurosci.* **10**, 37–47. https://doi.org/10.1038/nrn2554 (2009).
24. Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. & Plumb, I. The, "reading the mind in the eyes" test revised version: a study with normal adults, and adults with asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry* **42**, S0021963001006643. https://doi.org/10.1017/S0021963001006643 (2001).
25. Itier, R. J., Villate, C. & Ryan, J. D. Eyes always attract attention but gaze orienting is task-dependent: Evidence from eye movement monitoring. *Neuropsychologia* **45**, 1019–1028. https://doi.org/10.1016/j.neuropsychologia.2006.09.004 (2007).
26. Birmingham, E., Bischof, W. F. & Kingstone, A. Social attention and real-world scenes: the roles of action, competition and social content. *Q. J. Exp. Psychol.* **61**, 986–998. https://doi.org/10.1080/17470210701410375 (2008).
27. Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J. & Kingstone, A. Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition* **117**, 319–331. https://doi.org/10.1016/j.cognition.2010.09.003 (2010).
28. Risko, E. F., Richardson, D. C. & Kingstone, A. Breaking the fourth wall of cognitive science: Real-world social attention and the dual function of gaze. *Curr. Dir. Psychol. Sci.* **25**, 70–74. https://doi.org/10.1177/0963721415617806 (2016).
29. MacDonald, K. Patient-clinician eye contact: Social neuroscience and art of clinical engagement. *Postgrad. Med.* **121**, 136–144. https://doi.org/10.3810/pgm.2009.07.2039 (2009).
30. Trevisan, D. A., Roberts, N., Lin, C. & Birmingham, E. How do adults and teens with self-declared Autism Spectrum Disorder experience eye contact? A qualitative analysis of first-hand accounts. *PLoS ONE* **12**, e0188446. https://doi.org/10.1371/journal.pone.0188446 (2017).
31. Schilbach, L. et al. Toward a second-person neuroscience. *Behav. Brain Sci.* **36**, 393–414. https://doi.org/10.1017/S0140525X12000660 (2013).
32. Schilbach, L. Eye to eye, face to face and brain to brain: novel approaches to study the behavioral dynamics and neural mechanisms of social interactions. *Curr. Opin. Behav. Sci.* **3**, 130–135. https://doi.org/10.1016/j.cobeha.2015.03.006 (2015).
33. Jongerius, C. et al. Eye-tracking glasses in face-to-face interactions: Manual versus automated assessment of areas-of-interest. *Behav. Res. Methods* **53**, 2037–2048. https://doi.org/10.3758/s13428-021-01544-2 (2021).
34. Auyeung, B. et al. Oxytocin increases eye contact during a real-time, naturalistic social interaction in males with and without autism. *Transl. Psychiatry* **5**, e507. https://doi.org/10.1038/tp.2014.146 (2015).
35. Chong, E. et al. Detecting gaze towards eyes in natural social interactions and its use in child assessment. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **1**, 1–20. https://doi.org/10.1145/3131902 (2017).
36. de Belen, R. A. J., Bednarz, T., Sowmya, A. & Del Favero, D. Computer vision in autism spectrum disorder research: A systematic review of published studies from 2009 to 2019. *Transl. Psychiatry* https://doi.org/10.1038/s41398-020-01015-w (2020).
37. Falck-Ytter, T. Gaze performance during face-to-face communication: A live eye tracking study of typical children and children with autism. *Res. Autism Spectr. Disord.* **17**, 78–85. https://doi.org/10.1016/j.rasd.2015.06.007 (2015).
38. Noris, B., Nadel, J., Barker, M., Hadjikhani, N. & Billard, A. Investigating gaze of children with ASD in naturalistic settings (2012).
39. Hanley, M. et al. The use of eye-tracking to explore social difficulties in cognitively able students with autism spectrum disorder: A pilot investigation. *Autism* **19**, 868–873. https://doi.org/10.1177/1362361315580767 (2015).
40. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, Malesand females, scientists and mathematicians. *J. Autism Dev. Disord.* **31**, 5–17 (2001).
41. Freitag, C. M. et al. Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ)—die Kurzversion AQ-k. *Z. Für Klin. Psychol. Psychother.* **36**, 280–289. https://doi.org/10.1026/1616-3443.36.4.280 (2007).
42. Senju, A. & Johnson, M. H. Atypical eye contact in autism: Models, mechanisms and development. *Neurosci. Biobehav. Rev.* **33**, 1204–1214. https://doi.org/10.1016/j.neubiorev.2009.06.001 (2009).
43. Cañigueral, R., Ward, J. A. & de Hamilton, A. F. C. Effects of being watched on eye gaze and facial displays of typical and autistic individuals during conversation. *Autism* **25**, 210–226. https://doi.org/10.1177/1362361320951691 (2021).
44. Kassner, M., Patera, W. & Bulling, A. Pupil. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* 1151–1160 (ACM, New York, NY, USA, 2014). https://doi.org/10.1145/2638728.2641695.
45. Świrski, L., Bulling, A. & Dodgson, N. Robust real-time pupil tracking in highly off-axis images. In *Proceedings of the Symposium on Eye Tracking Research and Applications* 173–176 (ACM, New York, NY, USA, 2012). https://doi.org/10.1145/2168556.2168585.
46. Tim Esler. facenet-pytorch. *GitHub Repository* https://github.com/timesler/facenet-pytorch (2020).
47. David Sanberg. facenet. https://github.com/davidsandberg/facenet (2016).
48. Schroff, F., Kalenichenko, D. & Philbin, J. FaceNet: A unified embedding for face recognition and clustering. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 815–823 (IEEE, 2015). https://doi.org/10.1109/CVPR.2015.7298682.
49. Cao, Q., Shen, L., Xie, W., Parkhi, O. M. & Zisserman, A. VGGFace2: A dataset for recognising faces across pose and age (2017).
50. Yi, D., Lei, Z., Liao, S. & Li, S. Z. Learning Face Representation from Scratch. (2014).
51. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* https://doi.org/10.18637/jss.v082.i13 (2017).
52. Lenth, R., Singmann, H., Love, J., Buerkner, P. & Herve, M. Package 'emmeans'. (2019).
53. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
54. Long, J. A. interactions: Comprehensive, user-friendly toolkit for probing interactions. (2019).

55. Nowok, B., Raab, G. M. & Dibben, C. Synthpop: Bespoke creation of synthetic data in R. *J. Stat. Softw.* https://doi.org/10.18637/jss.v074.i11 (2016).
56. Chen, F. S. & Yoon, J. M. D. Brief report: Broader autism phenotype predicts spontaneous reciprocity of direct gaze. *J. Autism Dev. Disord.* **41**, 1131–1134. https://doi.org/10.1007/s10803-010-1136-2 (2011).
57. Grossman, R. B., Zane, E., Mertens, J. & Mitchell, T. Facetime versus screentime: Gaze patterns to live and video social stimuli in adolescents with ASD. *Sci. Rep.* **9**, 12643 (2019).
58. Rogers, S. L., Speelman, C. P., Guidetti, O. & Longmuir, M. Using dual eye tracking to uncover personal gaze patterns during social interaction. *Sci. Rep.* **8**, 1–9. https://doi.org/10.1038/s41598-018-22726-7 (2018).
59. Del Bianco, T. et al. Temporal profiles of social attention are different across development in autistic and neurotypical people. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **6**, 813–824. https://doi.org/10.1016/j.bpsc.2020.09.004 (2021).
60. Maes, P., Stercq, F. & Kissine, M. Brief report: Temporal distribution of visual attention between the eyes and mouth in young autistic children. *Res. Autism Spectr. Disord.* **110**, 102292. https://doi.org/10.1016/j.rasd.2023.102292 (2024).
61. Leong, V. & Schilbach, L. The promise of two-person neuroscience for developmental psychiatry: Using interaction-based sociometrics to identify disorders of social interaction. *Br. J. Psychiatry* **215**, 636–638. https://doi.org/10.1192/bjp.2019.73 (2019).
62. Vehlen, A., Standard, W. & Domes, G. How to choose the size of facial areas of interest in interactive eye tracking. *PLoS ONE* **17**, e0263594. https://doi.org/10.1371/journal.pone.0263594 (2022).
63. Cañigueral, R., Hamilton, A. & Ward, J. A. Don't Look at Me, I'm Wearing an Eyetracker! In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* 994–998 (ACM, Singapore Singapore, 2018). https://doi.org/10.1145/3267305.3274123.
64. Valtakari, N. V. et al. Eye tracking in human interaction: Possibilities and limitations. *Behav. Res. Methods* **53**, 1592–1608. https://doi.org/10.3758/s13428-020-01517-x (2021).
65. Zhang, X., Sugano, Y., Fritz, M. & Bulling, A. Appearance-based gaze estimation in the wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 4511–4520 (2015).

Acknowledgements
We would like to thank the CoBeNe doctoral school and the University of Vienna for providing funds for this project, Daniel Gugerell for providing the eye-tracking equipment, our master students who assisted in data collection (Libor Potočar, Nina Sireč, and Serap Özlü), the CogSciHub-SPP team (David Cserjan), and all the participants who took part in the study.

Author contributions
R.B.: Conceptualization, Software, Supervision of Data Collection, Analysis, Visualization, Writing - Original draft. G.S.: Conceptualization, Supervision, Writing - Reviewing and Editing. U.A.: Conceptualization, Resources (Eye Tracking), Writing - Reviewing and Editing.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-79210-8.

Correspondence and requests for materials should be addressed to R.B. or G.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.