



Published in final edited form as:

Autism. 2019 May ; 23(4): 846–857. doi:10.1177/1362361318788071.

Perceptions of Self and Other: Social judgments and gaze patterns to videos of adolescents with and without ASD

Ruth B. Grossman^{1,2}, Julia Mertens¹, and Emily Zane¹

¹Emerson College, Dept. of Communication Sciences and Disorders

²University of Massachusetts Medical School Shriver Center

Abstract

Neurotypical (NT) adults often form negative first impressions of individuals with Autism Spectrum Disorder (ASD) and are less interested in engaging with them socially. In contrast, individuals with ASD actively seek out the company of others who share their diagnosis. It is not clear, however, whether individuals with ASD form more positive first impressions of ASD peers when diagnosis is not explicitly shared. We asked adolescents with and without ASD to watch brief video clips of adolescents with and without ASD and answer questions about their impressions of the individuals in the videos. Questions related to participants' perceptions of the social skills of the individuals in the video, as well as their own willingness to interact with that person. We also measured gaze patterns to the faces, eyes, and mouths of adolescents in the video stimuli. Both participant groups spent less time gazing at videos of ASD adolescents. Regardless of diagnostic group, all participants provided more negative judgments of ASD than NT adolescents in the videos. These data indicate that, without being explicitly informed of a shared diagnosis, adolescents with ASD form negative first impressions of ASD adolescents that are similar to, or lower, than those formed by NT peers.

Individuals with Autism Spectrum Disorder¹ (ASD) have significant difficulties with social relatedness and often experience high rates of loneliness, depression, and even suicidal ideation (Berns, 2016). Even those on the autism spectrum who have preserved cognitive and language skills are frequently subjected to bullying and are often unemployed, under-employed, or unable to complete secondary education (Schroeder, Cappadocia, & Weiss, 2014, Maïano et al. 2016, Taylor & Seltzer, 2011). In recent years, there has been an increased awareness that these interactional difficulties are not founded exclusively on the social communication difficulties of individuals with ASD. Rather, there is often a breakdown in reciprocity between social partners – those with and those without ASD - who do not share the same understanding or interpretation of social rules and behavior (Milton, 2012).

Corresponding Author: ruth_grossman@emerson.edu.

Conflict of Interest Statement: None of the authors have any conflicts to declare

¹In the clinical literature, person-first language is preferred. However, many individuals with ASD prefer the term autistic adult/child. To reflect this dichotomy we use the two terminologies interchangeably

A likely starting point for this breakdown has been proposed in recent research showing that neurotypical (NT) adults and children form negative judgments of individuals with ASD within mere seconds of exposure (Sasson et al. 2017, Faso, Sasson, & Pinkham, 2015; Grossman, 2015). Even individuals with ASD who have typical language and cognitive abilities are subject to such judgments. In fact, autistic individuals who have preserved cognitive and language skills often report *higher* levels of stigmatization, possibly because they “look” typical, but aren’t (Shtayermman, 2009). For instance, Stagg and others (2014) showed NT children deciding that children with ASD were undesirable social partners after a single exposure (Stagg, Slavny, Hand, Cardoso, & Smith, 2014). This type of negative first impression could lead to lasting social exclusion of individuals with ASD (Lobst et al., 2009; Swaim & Morgan, 2001).

In recent years, a growing number of online and in-person communities have been established where autistic individuals can interact (Komeda, 2015). Members of these communities often report a sense of shared identity and shared traits, sometimes expressed as coming from the “same planet” (Sinclair, 2010). This suggests that individuals with ASD may prefer the company of one another to the company of NT individuals. However, these communities are created with the explicit knowledge that all members share the same diagnostic status. Recent evidence shows that knowledge of an ASD diagnosis leads to more positive perceptions by others (Sasson & Morrison, 2017). Without this explicit knowledge, it is unclear whether autistic individuals form more favorable first impressions of each other than those formed by NT peers. Without being explicitly informed of a shared diagnosis, do social signals of unfamiliar individuals with ASD reduce the possibility of interactions with other autistic peers, similar to the effect seen for impressions formed by NT peers?

Although there are many factors that can influence first impressions, such as body posture and gestures (Ambady & Skowronski, 2008), we focus our investigation on how individuals with ASD process social information from two highly salient sources: facial and vocal expressions. The evidence on face processing is mixed, with some studies reporting significant deficits, but others describing preserved skills (for review see Dawson, Webb, & McPartland, 2005 or Jemel, Mottron, & Dawson, 2006; also Walsh, Creighton, & Rutherford, 2016). Eyetracking studies have also yielded contradictory results, with some studies indicating reduced gaze to faces (e.g. Klin et al., 2002, Nakano et al., 2010, Pelphrey et al., 2002; Tanaka and Sung, 2016) and others reporting gaze patterns to faces that are analogous to those of NT peers (Fletcher-Watson, Leekam, Benson, Frank, and Findlay, 2009; Fletcher-Watson, Findlay, Leekam, and Benson, 2008; McPartland, Webb, Keehn, and Dawson, 2011).

There are similarly mixed findings in the literature delineating how individuals with ASD process social information from vocal cues (see McCann and Peppé, 2003 for review). Some studies show clear deficits in processing social and linguistic information from prosodic patterns (Diehl, Bennetto, Watson, Gunlogson, & McDonough, 2008; Golan, Baron-Cohen, Hill, & Rutherford, 2007; Peppé, McCann, Gibbon, O’Hare, & Rutherford, 2007). Others, however, indicate preserved abilities to perceive emotional and/or lexical signals from vocal signals (Grossman, Bemis, Skwerer, & Tager-Flusberg, 2010; Hubbard, Faso, Assmann, & Sasson, 2017), although this ability deteriorates when emotional expressivity is less intense

(Grossman and Tager-Flusberg, 2012). Recent data have shown that at least some of the variability in findings across vocal and facial perception may be related to the wide range in stimuli used across all studies, especially as they relate to their ecological validity (Chevallier et al., 2015). Overall, the existing literature does not indicate a blanket deficit in processing of facial and vocal social signals in ASD.

Most of this work focuses on whether autistic individuals can determine identity from faces, or read emotion from facial and vocal cues. Only a few studies have attempted to examine the way individuals with ASD process the complex social signals of first impression formation, and they have focused on the visual domain. Autistic adults with preserved cognitive and language skills reported impressions of a virtual “job applicant” that were similar to those of their NT peers (Kuzmanovic, Schilbach, Lehnhardt, Bente, and Vogeley 2011). Similarly, autistic children aged 5–13 looking at silent video clips of athletes during break times, were as successful as their NT peers at distinguishing between athletes who were winning and those who were losing (Furley & Schweizer, 2014; Ryan, Furley, & Mulhall, 2016). Kuzmanovic et al. (2011) showed that adults with ASD formed similar impressions of silent virtual characters as their NT peers, although the impressions of autistic participants were more susceptible to incongruent written information. These studies demonstrate that individuals with ASD use subtle visual signals to form first impressions from brief exposures to social information, although they leave open the question of how auditory information might affect this process. Since NT individuals use both visual and auditory information to form negative first impressions of autistic individuals (Sasson* et al. 2017), and because natural social interactions typically contain simultaneous visual and auditory information, this is an important question to pursue.

The formation of social relations is crucial during adolescence and first impressions can have a significant impact on social engagement and inclusion. We therefore wanted to better understand the formation of first impressions within and across these diagnostic groups. We asked participants to provide explicit impressions of social stimuli and used eyetracking to measure implicit gaze patterns to faces during this task. We specifically chose to include both visual (facial expressions, body position, etc.) and auditory (language, prosody) information in this study, so that our stimuli mirrored the multi-modal social cues people are exposed to during day-to-day interactions. We also increased ecological validity by using videos of aged-matched peers, rather than computer-generated or manipulated stimuli, and by asking questions relevant to daily lives of adolescents. This study poses two fundamental questions: 1.) Do adolescents with ASD form negative first impressions of age-matched peers with ASD when no diagnosis information is explicitly provided? 2.) Do adolescents with and without ASD visually explore the faces of autistic vs. NT peers differently? We expected this task to be more difficult for adolescents with ASD than tasks in previous studies for multiple reasons: 1.) Stimuli are realistic videos, rather than virtual characters, which often produce more salient social behaviors. 2.) The task presented here depends on interpretation of subtle social cues, rather than canonical emotional expressions. 3.) We provide no explicit diagnosis information about the adolescents in the videos. We therefore expect ASD participants not to differentiate their judgments or gaze patterns for peers with and without ASD. In contrast, we hypothesize that NT participants will report more negative social judgments of ASD than NT peers and will also gaze at them relatively less.

Method

Stimuli

We used videos of adolescents with and without ASD. During recording, we asked participants to retell an adventure story to the camera in a way that would be engaging to an imagined audience of young children. From the resulting corpus of story-retelling videos, we extracted 2–4-second-long video clips. We cut clips to include an entire phrase or sentence from beginning to end. All emotions targeted in the stories (happiness, fear, anger, and positive surprise) were represented in the clips. We rated all clips for video and audio quality and for the child's verbal production (no mispronunciations, grammatical errors, etc.) on a scale from 0 (lowest) to 3 (highest), and we excluded all clips rated below 2. We randomly selected several angry, fearful, happy, and surprise videos of NT stimulus producers and matched them with clips of the same sentence/emotion from ASD producers, resulting in six angry, four fearful, eight happy, and six surprise video clips for a total of 24 clips in the stimulus set. The uneven numbers of emotions represented in the videos was due to fewer good quality clips being available to match across diagnoses for some emotions.

Each clip shows the adolescents from approximately mid-chest upward, including shoulders, upper arms, and head. Adolescents in the videos were reading notes to assist their retelling of the stories from cue cards that were placed directly below the camera lens, ensuring that their gaze is directed toward the camera, albeit slightly below. The video background is a white wall flanked by two empty book-cases. Adolescents in the videos are also wearing 32 small (4mm in diameter) reflective motion-capture markers on their faces. We used those markers to track the movement of their facial features for a separate study. During the instructions for the current study, participants were informed about these markers and asked to ignore them. Since the markers are identical in all videos, their impact on perception and gaze patterns should be consistent across stimuli and not lead to group effects. The final stimulus set of 24 clips includes 12 videos of males with ASD and 12 videos of NT males. Seven adolescents in each diagnostic group are shown in the video clips, with no more than two videos of each person. We did not include videos of the same person saying the same phrase. Impression ratings of the same person remain stable over multiple exposures (Sasson* et al. 2017), so we did not expect two exposures to the same adolescent to affect the resulting data. For the 14 adolescents shown in the final set of stimuli videos, there were no between-group differences in non-verbal IQ (Leiter-R; Roid & Miller, 1997, ASD $M=105.1$, NT $M=114.6$, $F(1,13) = 1.84$, $p = .2$), receptive vocabulary (PPVT-4; Dunn & Dunn, 2007, ASD $M=119.1$, NT $M=131.7$, $F(1,13) = 1.94$, $p = .19$), or age (ASD $M=142.1$, NT $M=145$, $F(1,13) = .03$, $p = .86$). We created two pseudorandomized stimulus sequences that were counterbalanced across participants.

Participants

The Institutional Review Board of Emerson College approved this study and we obtained written informed consent from each participant. We recruited adolescents with and without ASD who completed the Core Language Subtests of the Clinical Evaluation of Language Fundamentals, 5th Edition (CELF-5; Semel, Wiig, and Secord, 2013) and the Kaufman Brief Intelligence Test, 2nd Edition (K-BIT-2; Kaufman & Kaufman, 2004). ASD diagnosis was

confirmed via the Autism Diagnostic Observation Schedule, 2nd edition (ADOS-2, Rutter et al. 2012) by administrators who achieved research reliability with a certified trainer. Participants ranged in age from 10:6 to 17:10 and the two groups (ASD $N = 22$, mean age 13:11, NT $N = 30$, mean age 13:7) were not significantly different in age, ($F(1, 51) = .18, p = .67$), IQ ($F(1, 51) = .64, p = .43$), language ability ($F(1, 51) = .32, p = .57$), or gender ($\chi^2 = .95, p = .33$, Table 1).

Procedure

Participants were led into a quiet room and seated at a comfortable viewing distance and angle from a 24" computer screen. All videos completely filled the screen. We explained to participants that they would see video clips of adolescents retelling snippets from a story to clarify that the individuals in the videos did not choose the text they were saying, and to clarify that they were not speaking spontaneously. During the experimental task, all videos completely filled the screen. We asked participants to provide first and honest impressions of the person in each clip by answering five questions using a non-graduated, continuous slider bar. The anchor points of the slider were "not likely" and "very likely" and the marker was at mid-point at the start of each question. We asked two questions about the rater's willingness to engage with the person in the video: "How likely is it that you would sit at lunch with this person?" and "How likely is it that you would start a conversation with this person?" We also asked three questions regarding the rater's assumptions about the person in the video: "How likely is it that this person gets along well with others?", "How likely is it that this person is socially awkward?", and "How likely is it that this person spends a lot of time alone?". We displayed task directions on the computer screen before each video clip and verbally verified that participants understood the instructions. The five questions appeared individually on the screen after each video and participants self-paced their responses and presentation of subsequent stimuli.

Behavioral data

We calculated the average slider response of each diagnostic group for each question using a range of -250 ("not likely") to 250 ("very likely"). For three questions, high positive slider numbers indicate higher social skills (i.e. kids in the videos are more likely to get along with others, more likely to have peers want to sit with them at lunch, or more likely to have peers want to start a conversation with them). For two questions higher slider numbers indicate lower social skills (i.e. kids in the videos are more likely to spend time alone and more likely to be socially awkward). We therefore multiplied the slider responses to the latter two questions by -1 to normalize the polarity of responses across all five questions. All results are based on these normalized response scores. For ease of visualization, we also reversed graph labels for these two questions ("less social awkwardness" and "less social isolation") to clarify that higher scores on all questions indicate more positive evaluations.

Eyetracking

We performed a dynamic five-point (one central point and four corners) calibration of the SensoMotoric Instruments (SMI™) RED eyetracker, aiming for < 1 degree of deviation in either axis. When gaze to the first fixation point was captured, the calibration automatically proceeded to the next fixation point. We analyzed gaze data for presentation of all videos,

but not for the time participants responded to the questions. We defined three Areas of Interest (AOI) for each video clip: face, eyes, and mouth, using SMI software to draw AOIs for each adolescent in the stimulus video for the entire duration of each clip. We ensured that the face AOI adhered only to the face and not the hairline or other features in the video. The eye AOI was created as a rectangle, encompassing the eyebrows and eyes. The mouth AOI was drawn as an oval to capture the entire mouth. We made frame-by-frame location and shape adjustments for all AOIs in all stimulus videos to account for whole head and feature movements (e.g. mouth opening) during speech production. The face AOI did not exclude the eye and mouth region to capture gaze to the entire face in a single AOI.

Eyetracking analysis focused on two gaze values: 1.) Percent fixations for each AOI. Fixations were defined as gaze lasting a minimum of 60ms within a maximum dispersion area of 30 pixels. This variable is expressed as a percentage of fixations to each AOI relative to the entire screen. We include this measure because it is the most commonly used metric for determining meaningful gaze patterns in this population. 2.) Fixation frequency, expressed as the number of fixations per second within a given AOI. We include this measure to account for the fact that increased fixation can be constituted of frequent short or infrequent longer fixations. We visually inspected all gaze data to check for flickering, unstable, or intermittent gaze tracking and determined that data with a tracking ratio of less than 60% were not reliable enough to be included. The remaining eyetracking dataset includes 19 ASD and 29 NT participants; the two groups do not differ on language, age, IQ scores, and gender distribution (Table 2).

Results

Behavioral data

We conducted a 2 (diagnostic *group*, ASD or NT) by 2 (stimulus *type*, ASD or NT) by 5 (question) repeated measures ANOVA to determine overall patterns in responses to all stimuli and questions. Sphericity assumption was not met, so we are reporting results with Greenhouse-Geisser correction. Results show a main effect for question type ($F(1,50) = 32.55, p < .001, \text{partial } \eta^2 = .39$), with both participant groups reporting higher ratings for NT stimulus videos. We also see a main effect for question ($F(2.32,116.18) = 7.1, p = .001, \text{partial } \eta^2 = .12$). There is no main effect for participant diagnosis ($F(1,50) = 2.11, p = .15, \text{partial } \eta^2 = .04$), but a significant question-by-diagnosis interaction ($F(2.32,116.18) = 3.78, p = .02, \text{partial } \eta^2 = .07$). To verify that participants' responses did not change based on seeing some of the adolescents in the videos a second time, we conducted the same analysis using data from only the first video clip of each stimulus producer. Results of ratings based on this stimulus set show the same pattern. There is a main effect for question type ($F(1,50) = 21.77, p < .001, \text{partial } \eta^2 = .3$) and a main effect for question ($F(2.39,119.57) = 6.45, p = .001, \text{partial } \eta^2 = .12$). There is no main effect for participant diagnosis ($F(1,50) = 2.47, p = .12, \text{partial } \eta^2 = .05$), but a significant question-by-diagnosis interaction ($F(2.39,119.57) = 3.57, p = .02, \text{partial } \eta^2 = .07$). All subsequent analyses are therefore based on the full dataset, to maximize available power.

To better understand the patterns of main effects and interactions, we divided questions into two categories: 1.) Questions about participants' perception of others ("person gets along

with others,” “person is socially awkward,” and “person spends a lot of time alone”); and 2.) Questions about participants’ willingness to interact with others (“I would sit at lunch with that person,” “I would start a conversation with that person”).

Questions about others. A 2 (diagnostic *group*, ASD or NT) by 2 (stimulus *type*, ASD or NT) by 3 (question) repeated measures ANOVA reveals a main effect for stimulus type ($F(1,50) = 34.41, p < .001$, partial $\eta^2 = .41$), with both participant groups reporting higher ratings for NT stimulus videos. We also see a main effect for question ($F(1.74,100) = 6.32, p = .004$, partial $\eta^2 = .11$), with the lowest ratings for both stimulus types given for the question of social awkwardness. There is a main effect for participant diagnosis ($F(1,50) = 4.79, p = .03$, partial $\eta^2 = .09$), with the NT group providing higher ratings overall. We also find a significant question-by-diagnosis interaction ($F(1.74,100) = 4.58, p = .017$, partial $\eta^2 = .08$).

To follow up on this interaction, we conducted a one-way ANOVA with participant diagnosis as the between-group variable. It reveals that ASD participants judge adolescents in both stimulus types significantly more negatively than do NT participants on the questions of social awkwardness (ASD stimuli $F(1,51) = 8.2, p = .006$, NT stimuli $F(1,51) = 4.9, p = .032$) and spending time alone (ASD stimuli $F(1,51) = 6.3, p = .016$, NT stimuli $F(1,51) = 4.2, p = .045$). There is no significant difference between groups on judgments of whether adolescents in the videos get along well with others (ASD stimuli $F(1,51) = .36, p = .55$, NT stimuli $F(1,51) = .12, p = .73$, Fig. 1).

In both participant groups, ratings of social awkwardness and social isolation (spends time alone) are significantly positively correlated for both stimulus types (all correlations data in Table 3). Interestingly, the NT participants provide ratings for *gets along with others* that are highly correlated with ratings of *less social awkwardness* and *less social isolation* for both stimulus types (*ASD gets along with ASD socially awkward* $r = .75, N = 30, p < .0001$, *ASD gets along with ASD social isolation* $r = .73, N = 30, p < .0001$, *NT gets along with NT socially awkward* $r = .63, N = 30, p < .0001$, *NT gets along with NT social isolation* $r = .72, N = 30, p < .0001$). In contrast, ratings from ASD participants on these questions are not correlated (*ASD gets along with ASD less socially awkward* $r = .26, N = 22, p = .25$, *ASD gets along with ASD less social isolation* $r = .25, N = 22, p = .27$, *NT gets along with NT less socially awkward* $r = .29, N = 22, p = .19$, *NT gets along with NT less social isolation* $r = .13, N = 22, p = .57$), see Figure 2.

Questions about willingness for own engagement. A 2 (diagnostic *group*, ASD or NT) by 2 (stimulus *type*, ASD or NT) by 2 (questions about self-engagement) repeated measures ANOVA reveals a main effect for stimulus type ($F(1,50) = 23.55, p < .001$, partial $\eta^2 = .32$), with all participants providing higher ratings of videos showing NT adolescents. There is no main effect for question or diagnosis and no question by type, or type by diagnosis interaction.

Eyetracking data

To determine overall visual attention to the task for the two groups, we investigated percent fixation to the parts of the screen that do not contain social information (i.e. everything but

the face) using a two-tailed, independent groups t-test. Results show that participants with ASD ($M = 13.06$, $SD = 10.22$) gaze less at the non-social screen than NT participants ($M = 21.31$, $SD = 14.08$, $t(46) = 2.2$, $p = .033$). We therefore calculated fixations to the social AOIs (eyes, mouth, face) as a proportion of each participant's gaze to the overall screen. Since the mouth and eye AOIs overlap with the face AOI, we conducted the analysis for the face AOI separately.

Percent fixation. For the face AOI, we used a 2 (participant *group*) by 2 (stimulus *type*) Repeated Measures ANOVA for percent fixation and fixation frequency. Results show a main effect for stimulus type ($F(1,46) = 30.31$, $p < .0001$) with both groups gazing less at the faces of ASD adolescents. There is no main effect for participant diagnosis ($F(1,46) = 1.21$, $p = .28$) and no stimulus type by diagnosis interaction ($F(1,46) = 1.69$, $p = .2$).

We also conducted a 2 (participant *group*) by 2 (stimulus *type*) by 2 (AOI: mouth, eyes) Repeated Measures ANOVA for percent fixation and fixation frequency. The sphericity assumption is not met, so we present results with Greenhouse-Geisser correction. Results show a main effect for stimulus type ($F(1,46) = 10.28$, $p = .002$, partial $\eta^2 = .18$) with less gaze to ASD stimuli and a main effect for AOI ($F(1,46) = 22.71$, $p < .0001$), with both groups fixating more on the mouth than the eye AOI. There is no effect for diagnostic group ($F(1,46) = 1.98$, $p = .17$, partial $\eta^2 = .04$). To follow-up on the main effect for stimulus type, we conducted paired t-tests with a Bonferroni adjusted alpha-level of .017. Results show higher fixation percentage to the eyes ($t(48) = 3.3$, $p = .002$), face ($t(48) = 5.86$, $p < .0001$), and mouth ($t(48) = 3.86$, $p < .0001$) of NT vs. ASD videos (Figure 3).

Fixation frequency. The 2 (group) by 2 (stimulus type) Repeated Measures ANOVA for the face AOI shows a main effect for stimulus type ($F(1,46) = 28.39$, $p < .0001$, partial $\eta^2 = .38$), but no main effect for diagnosis ($F(1,46) = 2.49$, $p = .12$, partial $\eta^2 = .05$).

The 2 (group) by 2 (stimulus type) by 2 (AOI: eyes, mouth) repeated measures ANOVA reveals a main effect for stimulus type ($F(1,46) = 9.02$, $p = .004$, partial $\eta^2 = .16$) and a main effect for AOI ($F(1,46) = 18.72$, $p < .0001$, partial $\eta^2 = .29$), with both groups having higher fixation frequency to the mouth than the eyes. There is also a marginal main effect for diagnosis ($F(1,46) = 3.92$, $p = .05$, partial $\eta^2 = .078$), with the NT cohort having significantly higher fixation frequency to both AOIs than ASD participants. Post-hoc paired t-tests to investigate stimulus type (α set at .017) show significantly greater fixation frequency in both participant cohorts to the face ($t(48) = 5.63$, $p < .0001$) and mouth ($t(48) = 3.6$, $p = .001$), but not the eyes ($t(48) = 2.19$, $p = .03$) of NT vs. ASD videos.

Discussion

We predicted that NT participants would rate NT peers in videos more highly than they rated ASD peers and would look at the faces of NT peers more than at the faces of ASD peers. We also predicted that ASD participants would not differentiate between peers with and without ASD in either judgment formation or gaze patterns. The data confirm our hypotheses about the NT, but not the ASD cohort. As predicted, NT participants rated NT peers in videos more favorably than they rated peers with ASD across all five questions and also looked

more at the faces of NT than ASD peers. Counter to our predictions, however, ASD participants also showed preferential looking to NT faces and rated NT peers more favorably. On questions related to their assessment of others' social skills, particularly ASD participants even rated adolescents with ASD more harshly than NT participants did.

While unexpected, these results are partially supported by recent data showing that ASD and NT children had similar accuracy rates for identifying emotional expressions of ASD and NT adults (Brewer et al., 2016). Similarly, Hubbard et al. (2017) found that both NT and ASD adults judged ASD speakers' emotional speech to be less natural, despite more clearly expressing the emotion. In another study, a small sample of adults with ASD were asked to judge whether virtual, animated "job applicants" exhibited dominant, neutral, or submissive attitudes (Schwartz, Dratsch, Vogetley, & Bente, 2014). Similar to our results, participants with ASD were as capable as their NT peers at providing judgments based on social cues from the stimuli. However, in contrast to our findings, adults with ASD judged the characters more positively than did NT adults. The difference between our findings and the ones reported by Schwartz et al. may depend on many factors, including the fact that our stimuli consisted of videos of real adolescents with and without ASD, while theirs presented virtual characters whose expressions are often simplified, making them potentially more salient to individuals with ASD (Rosset et al., 2008). Task demands in each study also focused on different social attributes in different social contexts. Still, both sets of results confirm an important finding that adolescents and adults with ASD are able to extract subtle social cues from exposures as brief as a few seconds.

These data highlight the saliency of social-signal differences found in facial and vocal expressions of ASD adolescents. There are several studies showing that expressive prosody of individuals with ASD is marked by unusual patterns in pitch, dynamic range, and rhythm (e.g. Hubbard et al. 2017, Diehl & Paul, 2013; Grossman, Bemis, Skwerer, & Tager-Flusberg, 2010; Grossman & Tager-Flusberg, 2012; Peppé et al., 2007 for review). Performing acoustic analyses of these prosodic metrics allows for objective interpretation of autistic vocal signals that could be foundational to the fact that this cohort is perceived as "awkward" (Bone, Black, Ramakrishna, Grossman, & Narayanan, 2015). Objective analyses of facial expressions is more difficult to perform, since it often depends on a time-intensive process whereby human coders who have undergone extensive training identify subtle facial movement patterns (Ekman & Friesen, 1978; Sato & Yoshikawa, 2007). This makes it impractical to perform large-scale, objective analyses of facial movements in ASD. One possible alternative is to use facial motion capture to quantify facial feature movement during dynamic expressions. The stimuli used in this study contain videos of several individuals repeating different sentences expressing a variety of emotions. The resulting variability in the facial movements across stimuli combined with the small number of samples per participant do not allow for a meaningful motion-capture analysis of the facial expressions produced in this stimulus set. However, we have successfully used facial motion capture to investigate a larger dataset of mimicked facial expressions of ASD and NT adolescents from the same cohort, including the adolescents who produced the stimulus videos. Our analyses of those data show higher levels of asynchrony between movements of facial regions (Metallinou, Grossman, & Narayanan, 2013) and reduced complexity of dynamic facial feature movements in adolescents with ASD (Guha, Yang, Grossman, &

Narayanan, 2016). Although these findings do not directly correspond to the stimuli used in this study, they do provide some insight into underlying differences between the facial movement patterns of the autistic vs. NT adolescents in the videos. These differences may have contributed to participants' higher ratings of and increased gaze to videos of NT adolescents. Further investigation into facial feature movements of individuals with ASD is needed to better understand underlying causes of social deficits that seem salient from brief exposures.

The finding that both participant groups perceive subtle social cues of ASD expressions is particularly interesting when viewed in combination with our eyetracking data, which show that both participant groups looked less at ASD than NT adolescents. This finding was consistent across all three gaze measures, showing that adolescents with ASD do not differ from their NT peers in the frequency or amount of fixations to these stimuli. The social motivation hypothesis of ASD (Chevallier, Kohls, Troiani, Brodtkin, & Schultz, 2012; Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 2005; Dawson, 2008) suggests that the social impairments of individuals with ASD may be caused by an early lack of interest in and visual attention to social stimuli, including faces. This lack of interest is proposed to lead to a reduced expertise in producing and processing social cues. Our data do not support this claim, since adolescents with ASD in our study did not show decreased looking to faces overall nor decreased expertise at decoding social cues from these stimuli.

When looking at correlations between ratings data, the NT cohort shows a significant correlation between responses to the question of whether a person gets along well with others and whether they spend time alone. Interestingly, the ASD cohort doesn't show this correlation, indicating that ASD adolescents do not assume that social ability (getting along well with others) entails social interaction (not spending their time alone). Based on these limited data, it is possible to speculate, that from the perspective of an autistic person, the desire to spend a lot of time with other people may not depend on an ability to get along with others and vice versa. These data align with findings showing that adults with ASD often express a desire for *better quality* interactions with friends and family, but not *more frequent* interactions (van Asselts-Goverts, Embregts, Hendriks, Wegman, & Teunisse, 2015). Despite reporting higher incidents of loneliness than adult NT peers (Sasson* et al., 2017), ASD individuals may not see frequent interactions with friends as an important or necessary aspect of social integration.

Our results show that both participant groups fixated more and more frequently to the mouth region of the face, rather than the eyes. This result contrasts with previous literature suggesting that individuals with ASD show pervasive preference for mouth-directed gaze (Falck-Ytter, Bölte, & Gredebäck, 2013) that is different from the eye-directed gaze of NT individuals (Tanaka & Sung, 2015). However, recent reviews of eyetracking literature suggest that gaze patterns of autistic individuals to faces may not reflect a rigid diagnosis-based (ASD vs. NT) difference, but instead are strongly dependent on the nature of the stimuli and the specifics of the task demands (Chita-tegmark, 2016; Falck-Ytter & von Hofsten, 2011; Guillon, Hadjikhani, Baduel, & Rogé, 2014). In the case of our study, it is likely that both participant groups looked at the mouths more because the speech movements attracted visual attention. Additionally, both participant groups may have gazed less at the eyes because

individuals in the stimulus videos were not looking directly into the camera. As described in the methods, adolescents in the videos were reading cue cards, which were held directly below the recording camera. This resulted in stimulus producers gazing slightly below the camera. This may have reduced the pressures of direct eye gaze on participants with ASD in our study and made the gaze patterns of the two cohorts more comparable. Even if videos had shown adolescents looking directly at participants, it is possible that gaze patterns would still be similar across groups simply because pre-recorded videos do not provide the same social pressures as live interactions. It is therefore important to push the field further into developing live-viewing eyetracking paradigms for older children and adults in interactive social contexts (Guillon et al., 2014).

Limitations

These data are based on a relatively small sample of adolescents and cannot be generalized without replication. It is also important to consider that the videos we asked participants to rate were based on elicited retellings of stories with adolescents in the videos, rather than live interactions. Stimulus producers in the videos were also wearing motion capture markers. Given that all these factors apply to videos of children with and without ASD, they should not have caused differences in how videos of either group of stimulus producers was perceived by participants. Nevertheless, future studies should focus on videos captured during natural conversations, rather than elicited narratives, and present individuals without reflective markers on their faces.

Conclusion

Although it is somewhat encouraging that ASD adolescents in our study were able to “read” the subtle social cues that may mark their own expressions as more socially awkward and less socially capable, the fact that individuals with ASD are not only perceived negatively by NT peers, but also by peers who share their diagnosis could have significant repercussions for the success of their social interactions. Importantly, participants’ negative social perceptions and unwillingness to engage are not merely present in explicit responses to judgment questions, but are even reflected implicitly, in relatively reduced face-directed gaze to adolescents with ASD by both participant groups. Given the oft-reported desire of individuals with ASD to interact with and support each other, as well as the need to facilitate better integration between ASD and NT individuals, these data highlight the importance of creating intentional spaces where information about diagnosis is shared so that such interactions can be fostered and encouraged. Otherwise, if left to chance encounters in school lunchrooms, both autistic and NT adolescents may perpetuate the social rejection of this group.

Acknowledgments:

This work was supported by NIH-NIDCD 1R01DC012774 and NIH-NIDCD R21DC010867. We thank Dr. Rhiannon Luyster for her very helpful insights about these data and Kayla Neumeyer for introducing us to the writings of autistic adults. We are also grateful to the children and families who dedicated their time to participate in our research.

References

- Ambady N, & Skowronski JJ (2008). *First impressions*. Guilford Press.
- Berns AJ (2016). Perceptions and experiences of friendship and loneliness in adolescent males with high cognitive ability and autism spectrum disorder. The University of Iowa.
- Bone D, Black MP, Ramakrishna A, Grossman R, & Narayanan S (2015). Acoustic-prosodic correlates of “awkward” prosody in story retellings from adolescents with autism. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* (Vol. 2015–Janua).
- Brewer R, Biotti F, Catmur C, Press C, Happe F, Cook R, & Bird G (n.d.). Can neurotypical individuals read autistic facial expressions? Atypical production of emotional facial expressions in Autism Spectrum Disorders.
- Chevallier C, Kohls G, Troiani V, Brodtkin ES, & Schultz RT (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, 16(4), 231–9. [PubMed: 22425667]
- Chevallier C, Parish-Morris J, McVey A, Rump KM, Sasson NJ, Herrington JD, & Schultz RT (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Research*, 8(5), 620–628. [PubMed: 26069030]
- Chita-tegmark M (2016). Attention Allocation in ASD: a Review and Meta-analysis of Eye-Tracking Studies. *Review Journal of Autism and Developmental Disorders*, 48, 209–223.
- Dawson G (2008). Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Development and Psychopathology*, 20(3), 775–803. [PubMed: 18606031]
- Dawson G, Meltzoff AN, Osterling J, Rinaldi J, & Brown E (2005). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, 28, 479–485.
- Dawson G, Webb SJ, & McPartland J (2005). Understanding the Nature of Face Processing Impairment in Autism: Insights From Behavioral and Electrophysiological Studies. *Developmental Neuropsychology*, 27(3), 403–424. [PubMed: 15843104]
- Diehl JJ, Bennetto L, Watson D, Gunlogson C, & McDonough J (2008). Resolving ambiguity: A psycholinguistic approach to understanding prosody processing in high-functioning autism. *Brain and Language*, 106(2), 144–152. [PubMed: 18502497]
- Diehl JJ, & Paul R (2013). Acoustic and perceptual measurements of prosody production on the PEPS-C by children with autism spectrum disorders. *Applied Psycholinguistics*, 34(1), 135–161.
- Dunn DM, & Dunn LM (2007). *Peabody picture vocabulary test: Manual*. Pearson.
- Ekman P, & Friesen WV (1978). *The Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA: Consulting Psychologists Press.
- Falck-Ytter T, Bölte S, & Gredebäck G (2013). Eye tracking in early autism research. *Journal of Neurodevelopmental Disorders*, 5(1), 28. [PubMed: 24069955]
- Falck-Ytter T, & von Hofsten C (2011). How special is social looking in ASD. A review *Progress in Brain Research* (1st ed., Vol. 189). Elsevier B.V.
- Faso D, Sasson N, & Pinkham A (2015). Evaluating Posed and Evoked Facial Expressions of Emotion from Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 45(1), 75–89. [PubMed: 25037584]
- Fletcher-Watson S, Findlay JM, Leekam SR, & Benson V (2008). Rapid Detection of Person Information in a Naturalistic Scene. *Perception*, 37(4), 571–583. [PubMed: 18546664]
- Fletcher-Watson S, Leekam SR, Benson V, Frank MC, & Findlay JM (2009). Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia*, 47(1), 248–257. [PubMed: 18706434]
- Furley P, & Schweizer G (2014). The Expression of Victory and Loss: Estimating Who’s Leading or Trailing from Nonverbal Cues in Sports. *Journal of Nonverbal Behavior*, 38(1), 13–29.
- Golan O, Baron-Cohen S, Hill J, & Rutherford M (2007). The “Reading the Mind in the Voice” Test-Revised: A Study of Complex Emotion Recognition in Adults with and Without Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*, 37(6), 1096–1106. [PubMed: 17072749]

- Grossman RB (2015). Judgments of social awkwardness from brief exposure to children with and without high-functioning autism. *Autism: The International Journal of Research and Practice*, 19(5), 580–7. [PubMed: 24923894]
- Grossman RB, Bemis RH, Skwerer DP, & Tager-Flusberg H (2010). Lexical and affective prosody in children with high-functioning Autism. *Journal of Speech, Language, and Hearing Research*, 53(3), 778–793.
- Grossman RB, & Tager-Flusberg H (2012). Quality matters! Differences between expressive and receptive non-verbal communication skills in adolescents with ASD. *Research in Autism Spectrum Disorders*, 6(3), 1150–1155. [PubMed: 22773928]
- Grossman RB, & Tager-Flusberg H (2012). “Who said that?” Matching of low- and high-intensity emotional prosody to facial expressions by adolescents with ASD. *Journal of Autism & Developmental Disorders*, 42(12), 2546–2557. [PubMed: 22450703]
- Guha T, Yang Z, Grossman RB, & Narayanan SS (2016). A Computational Study of Expressive Facial Dynamics in Children with Autism. *IEEE Transactions on Affective Computing*, 1–1.
- Guillon Q, Hadjikhani N, Baduel S, & Rogé B (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience and Biobehavioral Reviews*, 42, 279–297. [PubMed: 24694721]
- Hubbard DJ, Faso DJ, Assmann PF, & Sasson NJ (2017). Production and perception of emotional prosody by adults with autism spectrum disorder. *Autism Research*, 10(12), 1991–2001. [PubMed: 28815940]
- Iobst E, Nabors L, Rosenzweig K, Srivorakiat L, Champlin R, Campbell J, & Segall M (2009). Adults’ perceptions of a child with autism. *Research in Autism Spectrum Disorders*, 3(2), 401–408.
- Jemel B, Mottron L, & Dawson M (2006). Impaired Face Processing in Autism: Fact or Artifact? *Journal of Autism and Developmental Disorders*, 36(1), 91–106. [PubMed: 16477517]
- Kaufman A, & Kaufman N (2004). *Manual for the Kaufman Brief Intelligence Test Second Edition*. Circle Pines, MN: American Guidance Service.
- 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 in General Psychiatry*, 59(9), 809–816.
- Komeda H (2015). Similarity hypothesis: understanding of others with autism spectrum disorders by individuals with autism spectrum disorders. *Frontiers in Human Neuroscience*, 9(March), 1–7. [PubMed: 25653611]
- Kuzmanovic B, Schilbach L, Lehnhardt FG, Bente G, & Vogeley K (2011). A matter of words: Impact of verbal and nonverbal information on impression formation in high-functioning autism. *Research in Autism Spectrum Disorders*, 5(1), 604–613.
- McCann J, & Peppé S (2003). Prosody in autism spectrum disorders: a critical review. *International Journal of Language & Communication Disorders*, 38(4), 325–350. [PubMed: 14578051]
- McPartland JC, Webb SJ, Keehn B, & Dawson G (2011). Patterns of visual attention to faces and objects in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 41(2), 148–157. [PubMed: 20499148]
- Metallinou A, Grossman RB, & Narayanan S (2013). Quantifying atypicality in affective facial expressions of children with autism spectrum disorders. In *Proceedings - IEEE International Conference on Multimedia and Expo*.
- Milton DEM (2012). On the ontological status of autism: the “double empathy problem.” *Disability & Society*, 27(6), 883–887.
- Nakano T, Tanaka K, Endo Y, Yamane Y, Yamamoto T, Nakano Y, ... Kitazawa S (2010). Atypical gaze patterns in children and adults with autism spectrum disorders dissociated from developmental changes in gaze behaviour. *Proceedings. Biological Sciences / The Royal Society*, 277(1696), 2935–2943.
- Pelphrey KA, Sasson NJ, Reznick JS, Paul G, Goldman BD, & Piven J (2002). Visual scanning of faces in autism. *Journal of Autism & Developmental Disorders*, 32(4), 249–261. [PubMed: 12199131]

- Peppé S, McCann J, Gibbon F, O'Hare A, & Rutherford M (2007). Receptive and expressive prosodic ability in children with high-functioning autism. *Journal of Speech Language & Hearing Research*, 50(4), 1015–1028.
- Roid GH, & Miller LJ (1997). *Leiter International Performance Scale - Revised*. Wood Dale, IL: Stoelting Co.
- Rosset D, Rondan C, Da Fonseca D, Santos A, Assouline B, & Deruelle C (2008). Typical Emotion Processing for Cartoon but not for Real Faces in Children with Autistic Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 38(5), 919–925. [PubMed: 17952583]
- Ryan C, Furley P, & Mulhall K (2016). Judgments of Nonverbal Behaviour by Children with High-Functioning Autism Spectrum Disorder: Can they Detect Signs of Winning and Losing from Brief Video Clips? *Journal of Autism and Developmental Disorders*, 46(9), 2916–2923. [PubMed: 27306454]
- Sasson* NJ, Faso DJ, Nugent J, Lovell S, Kennedy* DP, & Grossman* RB (2017). Neurotypical Peers are Less Willing to Interact with Those with Autism based on Thin Slice Judgments. *Scientific Reports*, (February), 1–10. [PubMed: 28127051]
- Sasson NJ, & Morrison KE (2017). First impressions of adults with autism improve with diagnostic disclosure and increased autism knowledge of peers. *Autism*, 136236131772952. 10.1177/1362361317729526
- Sato W, & Yoshikawa S (2007). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition*, 104(1), 1–18. [PubMed: 16780824]
- Semel E, Wiig EH, & Secord WA (2013). *Clinical evaluation of language fundamentals-5*. Bloomington, MN: NCS Pearson, Inc.
- Shtayermman O (2009). An Exploratory Study of the Stigma Associated With a Diagnosis of Asperger's Syndrome: The Mental Health Impact on the Adolescents and Young Adults Diagnosed With a Disability With a Social Nature. *Journal of Human Behavior in the Social Environment*, 19(3), 298–313.
- Sinclair J (2010). Being Autistic Together. *Disability Studies Quarterly*, 30(1).
- Stagg SD, Slavny R, Hand C, Cardoso A, & Smith P (2014). Does facial expressivity count? How typically developing children respond initially to children with autism. *Autism*, 18(6), 704–711. [PubMed: 24121180]
- Swaim KF, & Morgan SB (2001). Children's Attitudes and Behavioral Intentions Toward a Peer with Autistic Behaviors: Does a Brief Educational Intervention Have an Effect? *Journal of Autism and Developmental Disorders*, 31(2), 195–205. [PubMed: 11450818]
- Tanaka JW, & Sung A (2015). Impaired Face Processing in Autism: Fact or Artifact? *Journal of Autism & Developmental Disorders*, 46(5), 1538–1552.
- Tanaka JW, & Sung A (2016). The "Eye Avoidance" Hypothesis of Autism Face Processing. *Journal of Autism and Developmental Disorders*, 46(5), 1538–1552. [PubMed: 24150885]
- Taylor JL, & Seltzer MM (2011). Employment and post-secondary educational activities for young adults with autism spectrum disorders during the transition to adulthood. *Journal of Autism and Developmental Disorders*, 41(5), 566–574. [PubMed: 20640591]
- van Asselts-Goverts AE, Embregts PJCM, Hendriks AHC, Wegman KM, & Teunisse JP (2015). Do Social Networks Differ? Comparison of the Social Networks of People with Intellectual Disabilities, People with Autism Spectrum Disorders and Other People Living in the Community. *Journal of Autism and Developmental Disorders*, 45, 1191–1203. [PubMed: 25326258]
- Walsh JA, Creighton SE, & Rutherford MD (2016). Emotion Perception or Social Cognitive Complexity: What Drives Face Processing Deficits in Autism Spectrum Disorder? *Journal of Autism and Developmental Disorders*, 46(2), 615–623. [PubMed: 26439480]

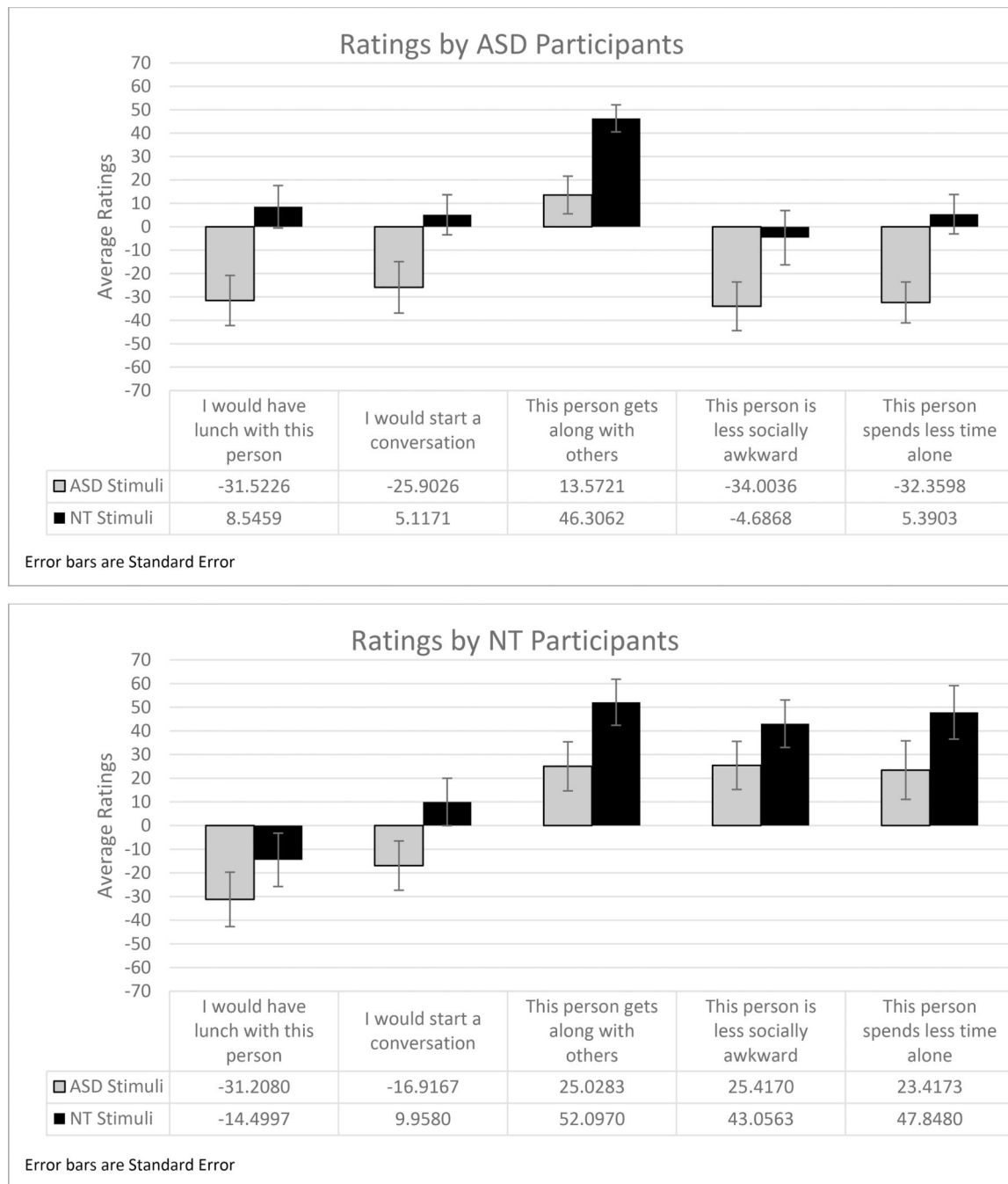


Figure 1. Average ratings for all questions

Note: these graphs present average ratings data, ranging from -32 to 65, in contrast to the full range of raw ratings data based on the +/- 250 range available to participants.

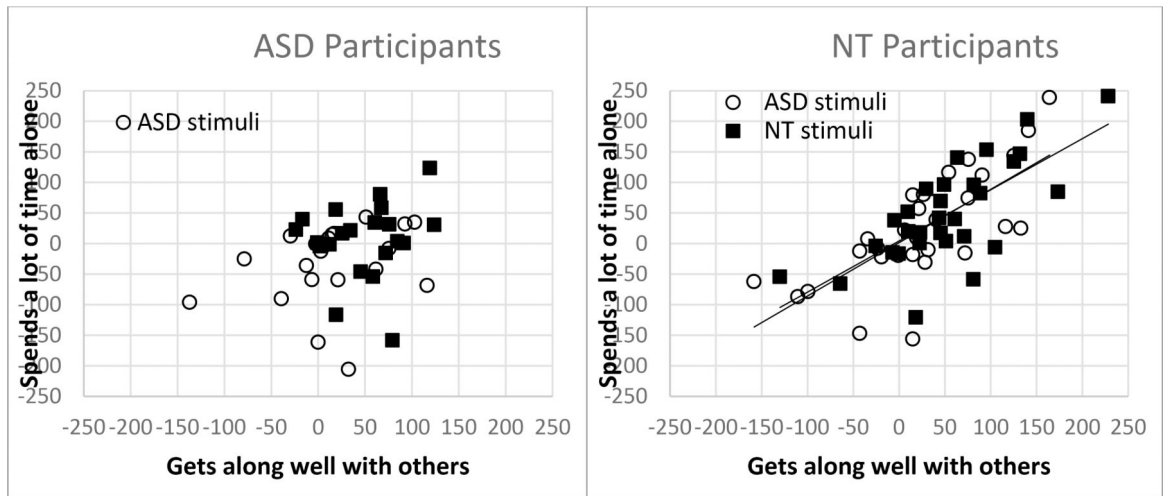


Figure 2.
Correlations of ratings across questions

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

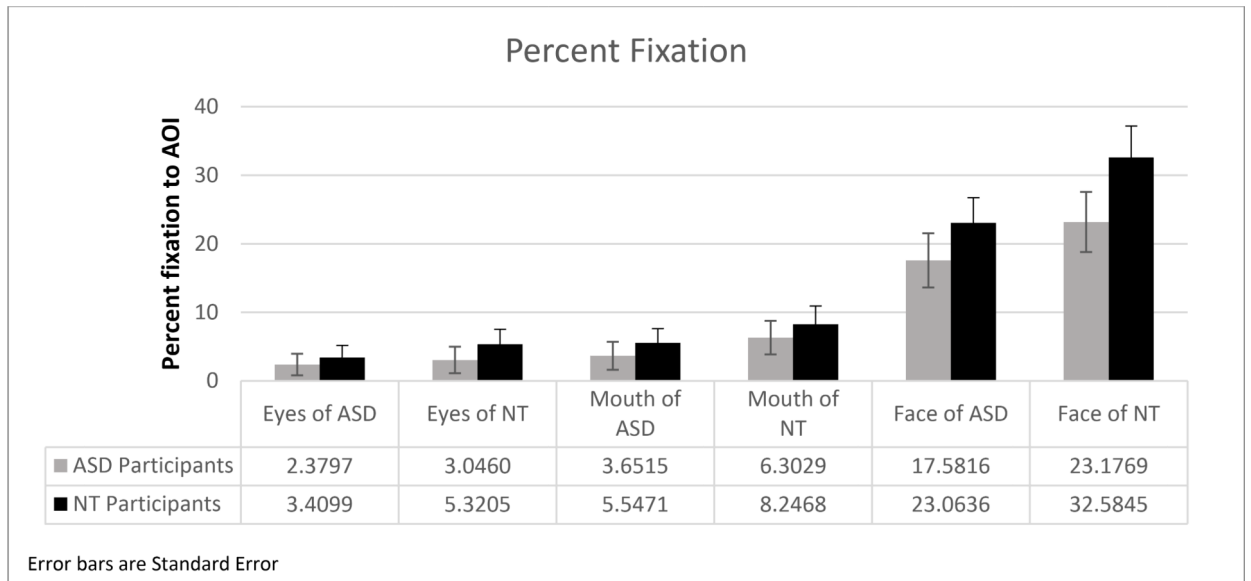


Figure 3.
Gaze data for all stimulus types

Table 1

Descriptive Characteristics of Participant groups

	ASD (n=22) <i>M</i> [95% CI]	NT (n=30) <i>M</i> [95% CI]	Significance
Age	13:7 [12:8,14:7] Range: 10:5 – 17:10	13:4 [12:8, 14:2] Range: 10:6 – 16:10	$F(1,51) = .18, p = .67$
Sex	18 male, 4 female	21 male, 9 female	$\chi^2(1, 52) = .95, p = .33$
IQ (K-BIT-2)	114.55 [105.59,123.5]	110.77 [105.63,115.91]	$F(1,51) = .64, p = .43$
Language (CELF-V)	110.68 [102.88,118.49]	113.3 [107.52,119.08]	$F(1,51) = .32, p = .57$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Descriptive Characteristics of Participants included in Eyetracking Analysis

	ASD (n=19) <i>M(SD)</i>	NT (n=29) <i>M(SD)</i>	Significance
Age	13:11 [12:10,14:10] Range: 10:8 – 17:10	13:7 [12:10,14:2] Range: 10:6 – 16:10	$F(1,46) = .26, p = .61$
Sex	15 male, 4 female	21 male, 8 female	$\chi^2(1, 48) = .26, p = .61$
IQ (K-BIT-2)	116.63 [107.01,126.25]	110.17 [104.99,115.35]	$F(1,46) = 1.78, p = .19$
Language (CELF-V)	113.05 [104.51,121.59]	112.76 [106.88,118.64]	$F(1,46) = .004, p = .95$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

ASD: autism spectrum disorder; NT: neurotypical, Bold font = ASD participants, normal font = NT participants

* significant at the .05 level

** significant at the .01 level

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript