



## How Young Adults with Autism Spectrum Disorder Watch and Interpret Pragmatically Complex Scenes

Linda Lönnqvist, Soile Loukusa, Tuula Hurtig, Leena Mäkinen, Antti Siipo, Eero Väyrynen, Pertti Palo, Seppo Laukka, Laura Mämmelä, Marja-Leena Mattila & Hanna Ebeling

**To cite this article:** Linda Lönnqvist, Soile Loukusa, Tuula Hurtig, Leena Mäkinen, Antti Siipo, Eero Väyrynen, Pertti Palo, Seppo Laukka, Laura Mämmelä, Marja-Leena Mattila & Hanna Ebeling (2016): How Young Adults with Autism Spectrum Disorder Watch and Interpret Pragmatically Complex Scenes, *The Quarterly Journal of Experimental Psychology*, DOI: [10.1080/17470218.2016.1233988](https://doi.org/10.1080/17470218.2016.1233988)

**To link to this article:** <http://dx.doi.org/10.1080/17470218.2016.1233988>

 [View supplementary material](#) 

 Accepted author version posted online: 12 Sep 2016.  
Published online: 12 Sep 2016.

 [Submit your article to this journal](#) 

 Article views: 11

 [View related articles](#) 

 [View Crossmark data](#) 

**Publisher:** Taylor & Francis & The Experimental Psychology Society

**Journal:** *The Quarterly Journal of Experimental Psychology*

**DOI:** 10.1080/17470218.2016.1233988

How Young Adults with Autism Spectrum Disorder Watch and Interpret Pragmatically  
Complex Scenes

**Running head:** COMPLEX SCENE PERCEPTION IN ASD

**Authors and affiliations**

Lönnqvist, Linda<sup>1</sup> | Loukusa, Soile<sup>1</sup> | Hurtig, Tuula<sup>2,3,4</sup> | Mäkinen, Leena<sup>1</sup> | Siipo, Antti<sup>5</sup> |  
Väyrynen, Eero<sup>6</sup> | Palo, Pertti<sup>7</sup> | Laukka, Seppo<sup>5</sup> | Mämmelä, Laura<sup>3,8</sup> | Mattila, Marja-Leena<sup>2,</sup>  
<sup>3</sup> | & Ebeling, Hanna<sup>2,3</sup>

1 = Logopedics, Faculty of Humanities, and Child Language Research Center, PO box 1000,  
90014, University of Oulu, Finland

2 = PEDEGO Research Unit, Child Psychiatry, PO box 5000, 90014, University of Oulu,  
Finland

3 = Clinic of Child Psychiatry, PO box 26, 90029, Oulu University Hospital, Finland

4 = Neuroscience Research Unit, Psychiatry, PO box 5000, 90014, University of Oulu,  
Finland

5 = Learning Research Laboratory, Research unit of Psychology, Faculty of Education, PO  
box 2000, 90014, University of Oulu, Finland

6 = BME Research Group, Department of Computer Science and Engineering, PO box 4500,  
90014 University of Oulu, Finland

7 = CASL Research Center, EH21 6UU, Queen Margaret University, Edinburgh, Scotland,  
UK

8 = Department of Psychology, Faculty of Social Sciences, PO box 35, 40014 University of  
Jyväskylä, Finland

### **Corresponding author**

**Name:** Linda Lönnqvist

**Address:** Logopedics, Faculty of Humanities, PO box 1000, 90014, University of Oulu,  
Finland

**Email:** linda.lonnqvist@icloud.com

**Phone number:** +358 400 729 674

### **Acknowledgements**

This research was supported by the Academy of Finland, Helsinki, Finland, under grant  
276578; the Alma and K. A. Snellman Foundation, Oulu, Finland; the Child Psychiatry

Research Foundation, Finland; the Emil Aaltonen Foundation, Tampere, Finland; Northern Ostrobothnia Hospital District, Finland; Rinnekoti Research Foundation / the Finnish Brain Foundation, Espoo / Helsinki, Finland; the Sigrid Jusélius Foundation, Helsinki, Finland. The authors declare no conflict of interest

### Abstract

The aim of the current study was to investigate subtle characteristics of social perception and interpretation in high-functioning individuals with autism spectrum disorders (ASDs), and to study the relation between watching and interpreting. As a novelty, we used an approach that combined moment-by-moment eye tracking and verbal assessment. Sixteen young adults with ASD and 16 neurotypical control participants watched a video depicting a complex communication situation while their eye movements were tracked. The participants also completed a verbal task with questions related to the pragmatic content of the video. We compared verbal task scores and eye movements between groups, and assessed correlations between task performance and eye movements. Individuals with ASD had more difficulty than the controls in interpreting the video and during two short moments there were significant group differences in eye movements. Additionally, we found significant correlations between verbal task scores and moment-level eye movement in the ASD group, but not among the controls. We concluded that participants with ASD had slight difficulties in understanding the pragmatic content of the video stimulus and attending to social cues, and that the connection between pragmatic understanding and eye movements was more pronounced for participants with ASD than for neurotypical participants.

**Keywords:** autism spectrum disorder, complex communication, eye tracking, moment-to-moment analysis, pragmatics

Autism spectrum disorders (ASDs) are neurodevelopmental conditions characterised by impairments in social and communicative skills, and the presence of repetitive behaviours or restricted interests (American Psychiatric Association, 2013; World Health Organization, 1993). It has been suggested that individuals with ASD commonly compensate for difficulties in spontaneous social information processing by using conscious reasoning and explicit knowledge about social behaviour (Baez et al., 2012; Ullman & Pullman, 2015). These compensatory mechanisms may mask social difficulties, especially when social skills are assessed in research environments, which are more structured than real-world situations, and therefore facilitates the use of nonspontaneous reasoning. This is true in particular for high-functioning individuals with ASD, that is, individuals who have ASD but no impairment in general cognitive functions (Loukusa & Moilanen, 2009; Tobin, Drager, & Richardson, 2014). Even subtle social difficulties may have negative long-term impact on a person's well-being, and it is therefore important to recognise the difficulties high-functioning individuals with ASD experience, and develop ways of supporting these individuals (Hanley et al., 2015; Tobin et al., 2014). Eye tracking methods provide the possibility to directly assess how high-functioning individuals with ASD watch social interactions, and may help bridge commonly occurring gaps between evaluation outcomes and real-world manifestations of the disorder (Guillon, Hadjikhani, Baduel, & Rogé, 2014; Hanley et al., 2015). It has been suggested that subtle but clinically relevant impairments could effectively be detected by combining eye tracking and a verbal task (Freeth, Ropar et al., 2011), or by focusing on the moment-to-moment structure of eye movements (Falck-Ytter, Von Hofsten, Gillberg, & Fernell, 2013;

Nakano et al., 2010). In this study we combine these two approaches, which to our knowledge has not been done before.

### **Pragmatic Language and Social Cue Perception in High-Functioning Adults with ASD**

Individuals with ASD commonly have pragmatic language difficulties, such as a reduced understanding of hidden meanings and conversational nuances (Loukusa & Moilanen, 2009; Martin & McDonald, 2003). It has also been suggested that individuals with ASD often do not catch the social cues expressed in communication situations well enough to resolve the pragmatic content of what is said (Grynszpan & Nadel, 2015; Guillon et al., 2014). These social cues can be verbal or nonverbal, and meanings are often built on the interaction of, or discrepancy between, various cues (Gibbs & Orden, 2012; Roche, Dale, & Caucci, 2012). Theoretical accounts of these difficulties relate to models explaining autism spectrum disorders in general (Loukusa & Moilanen, 2009; Martin & McDonald, 2003; Vulchanova, Saldaña, Chahboun, & Vulchanov, 2015). Some well-known models include the social cognition model, the weak coherence model and the executive functions model. The different variants of the social cognition model propose that pragmatic difficulties rise from a weakened ability to take other interlocutors' perspectives or the general context into account when interpreting communication, i.e. a problem with empathising, mentalising or Theory of Mind processing (Frith & Frith, 2012; Tomasello, Carpenter, Call, Behne, & Moll, 2005; Wheelwright, Baron-Cohen, Goldenfeld, & Delaney, 2006). It also includes the account of a reduced spontaneous interest in social targets leading to a deficit in social cue perception (Senju, 2013). According to the Weak Coherence model, individuals with ASD might have deficits in combining information from many sensorial modalities or cognitive domains

(Happé & Frith, 2006; Happé & Booth, 2008). This could hamper the perception and processing of social cues, which most often are complex and multimodally expressed (Ames & Fletcher-Watson, 2010). The executive function model states that in high-functioning individuals with ASD understanding of complex communication may be compromised due to difficulties with cognitive functions needed for planning, performing and controlling complex actions (Hill, 2004; Loukusa & Moilanen, 2009). The reduction of cognitive flexibility, that is, the ability to adaptively shift attention or change cognitive strategy, is often mentioned as an executive function relevant for following complex communication and compromised in individuals with ASD (Diamond, 2013; Hill, 2004). Cognitive flexibility is needed to track social cues in conversations, in which the cues are typically expressed simultaneously or in quick succession, often by more than one person at a time (Lin & Carlile, 2015; Shintel & Keysar, 2009).

Because of the continuously changing, multiple-source nature of conversations, interlocutors need to direct their attention to the right thing at the right moment in order to perceive the cues needed for appropriate interpretations of the situation (Habets, Kita, Shao, Özyurek, & Hagoort, 2011; Holler & Kendrick, 2015; Lin & Carlile, 2015; Richardson & Dale, 2005). Hence, the unprecise timing of attention allocation during conversations might hamper refined social understanding, especially if the problem occurs repeatedly and the effect cumulates. Atypical timing related to processing of social stimuli has been observed in individuals with ASD in behavioural tasks, eye movements as well as at the neural level (Falter, Elliott, & Bailey, 2012; Nakano et al., 2010; Salmi et al., 2013). It has been suggested that difficulties in basic time processing might be part of the core ASD symptomatology (Allman & Meck, 2012; Falter et al., 2012). On the other hand, atypical timing might be secondarily caused by decreased attention to social cues, deficits in processing multimodal

information or reduced cognitive flexibility (Ames & Fletcher-Watson, 2010; Guillon et al., 2014; Hill, 2004). These models and mechanisms are not necessarily mutually exclusive but they may rather be complementary and affect each other multidirectionally (Ames & Fletcher-Watson, 2010; Loukusa & Moilanen, 2009; Martin & McDonald, 2003).

### **Eye Movements in High-Functioning Individuals with ASD**

Eye tracking is commonly used to investigate how social cues are perceived in the visual mode (Birmingham, Bischof, & Kingstone, 2008; Smith & Mital, 2013). In these studies, potential sources of social cues, such as eyes, faces or humans in general, are often referred to as the *social features* of the stimulus. In the following review on social scene perception in ASD, we mainly include studies in which the participants were adolescents or adults and the participants with ASD were high-functioning individuals. Individuals with ASD and their neurotypical peers differ in how they look at social features, but the group differences are not consistent across conditions (for reviews see Falck-Ytter & von Hofsten, 2011; Guillon et al., 2014). Naturalistic and complex stimuli, such as social scenes, seem to provoke differences more effectively than nonnaturalistic or relatively simple stimuli, such as isolated faces (Hanley et al., 2015; Hanley, McPhillips, Mulhern, & Riby, 2013; Speer, Cook, McMahon, & Clark, 2007). Most studies that used dynamic stimuli, as opposed to static pictures, reported group differences (Bird, Press, & Richardson, 2011; Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Speer et al., 2007).

Results also vary depending on which aspect of eye movement is studied, and it has been suggested that group differences are more likely to be found in studies focusing on the precise timing of gaze shifts than in studies measuring total looked-at times for social features (Falck-Ytter & von Hofsten, 2011; Guillon et al., 2014). Individuals with ASD have

commonly been found to turn their gaze towards the social features of the stimulus later than their neurotypical peers (Freeth et al., 2010; Freeth, Foulsham et al., 2011; Freeth, Ropar et al., 2011; Norbury et al., 2009), although some conflicting or ambivalent findings have been reported (Fletcher-Watson, Leekam, & Benson, 2009; Sasson et al., 2007). Temporally atypical eye movements were also observed in studies that analysed eye movements on a moment-by-moment basis (Falck-Ytter et al., 2013; Nakano et al., 2010). Unlike the other articles reviewed in the present study, participants in the study by Falck-Ytter et al. (2013) were children. The moment-by-moment studies showed that individuals with ASD followed the events of a dynamic social scene less precisely than control participants (Falck-Ytter et al., 2013; Nakano et al., 2010). Nakano et al. (2010) additionally discovered that there was more diversity in the eye movements in the ASD group than in the control group. In studies with neurotypical individuals, this diversity is usually relatively low, that is, participants have similar eye movements (Richardson & Dale, 2005; Smith & Mital, 2013). Common ground knowledge or shared interpretations of a situation may result in similar eye movements of the participants (Shockley, Richardson, & Dale, 2009). On the other hand, it has been shown that participants are more likely to establish a common interpretation of a scene in a video when experimentally manipulated to look at the same targets (Richardson & Dale, 2005). The existence and use of common ground knowledge usually enhances pragmatic understanding in social situations (Gibbs & Orden, 2012). Consequently it has been suggested that the similarity of eye movements may be causally related to social understanding (Richardson & Dale, 2005; Shockley et al., 2009). In line with this notion, similarity in eye movement was negatively correlated to autism spectrum traits, including social difficulties, in the Nakano et al. (2010) study.

In addition to Nakano et al. (2010), many research groups have found links between social skills and eye movements. Both positive (Bird et al., 2011; Hanley et al., 2015; Klin et al., 2002; Speer et al., 2007), and negative (Klin et al., 2002; Norbury et al., 2009) correlations between social skills and atypical eye movements have been reported. The above studies concerned the links to general social skills assessed with standard tools such as the Autism Diagnostic Observational Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 2000). In other studies, researchers have investigated the relation between how participants look at a social stimulus and interpret its content (Freeth, Ropar et al., 2011; Grynszpan & Nadel, 2015; Sasson et al., 2007). The results include positive (Grynszpan & Nadel, 2015), negative (Sasson et al., 2007), and absent correlations (Freeth, Ropar et al., 2011) between social performance and the amount or latency of face watching. Stimulus types, social tasks, as well as outcome measures for eye movements differ between these studies, so the studies cannot be directly compared with each other.

### **The Present Study**

The aim of this study was to investigate how high-functioning adults with ASD watch and interpret a scene which includes pragmatically complex communication. This aim was divided into three subsidiary objectives:

- To study whether individuals with ASD interpret the pragmatic content of a social scene differently from their neurotypical peers.
- To study whether individuals with ASD visually follow the dynamics of a social scene as precisely as their neurotypical peers

- To study whether looking at a person when this person gives a social cue is positively correlated with a pragmatically adequate interpretation of the scene in either or both of the study groups.

We predicted that participants with ASD would have more difficulties than control participants in a pragmatic understanding task that challenged social attention, cognitive flexibility, central coherence and mentalisation processes, because individuals with ASD often have deficits in one or many of these skills (Loukusa & Moilanen, 2009; Martin & McDonald, 2003; Vulchanova et al., 2015). As this was a study with high-functioning individuals with ASD we expected the differences in pragmatic understanding to be relatively subtle. Based on the difficulties these individuals often have in social attention and flexible attention shifts, we predicted that participants with ASD would follow the dynamics of the scene with their gaze in a less precise way than their peers and that this difference would emerge when investigating eye movements moment by moment. We also expected to find some correlations between ways of watching and ways of interpreting the video. More specifically, we predicted that looking at a person in the video when this person did something that could be regarded a social cue would correlate positively with the performance of a task assessing pragmatic understanding.

## Methods

The present study was carried out in collaboration with the University of Oulu, Oulu University Hospital and Oulu Autism Spectrum Disorder Research Center as a part of the multidisciplinary research project *Autism Spectrum Disorders – a follow-up study from childhood to young adulthood*. The present work belongs to a section of the study called *Pragmatic language abilities in adolescents and young adults with autism spectrum disorder*.

The study was approved by the Regional Ethics Committee of the Northern Ostrobothnia Hospital District and conducted in accordance with the Declaration of Helsinki. Participants gave their informed consent to participating in the study and were informed that they could cancel their participation at any time during the project.

## **Participants**

Thirty young adults with ASD (19.4–33.3 years, median 22.7 years) and 33 neurotypical controls (19.3–25.2 years, median 22.4 years) took part in the study. Individuals with ASD originally participated in an epidemiological study in the Northern Ostrobothnia Hospital District area (Mattila et al., 2007, 2011) or a clinical ASD study conducted at Oulu University Hospital (Weiss, Arking, & the Gene Discovery Project of John Hopkins & the Autism Consortium, 2009; see also Kuusikko et al., 2008, 2009), in years 2000–2003. Controls were randomly selected from the epidemiological study (Mattila et al., 2007, 2011), or randomly selected (Kuusikko et al., 2008, 2009) or recruited (Jansson-Verkasalo et al., 2005) from mainstream schools in Oulu.

During the original recruiting processes during 2000–2003, ASD diagnoses were determined by trained clinical psychologists and medical doctors (Kuusikko et al., 2008, 2009; Mattila et al., 2007, 2011; Weiss et al., 2009). In these processes the Autism Diagnostic Interview Revised (ADI-R; Lord, Rutter, & Le Couter, 1995) and the ADOS (Lord et al., 2000) were used to obtain structured information from parents and for semi-structured observation of individuals with ASD (Kuusikko et al., 2008, 2009; Mattila et al., 2007, 2011). The ADI-R

and the ADOS were not used to make diagnostic classifications, that is, the diagnostic algorithms were not used. The diagnoses were clinical best estimates made according to the ICD-10 criteria (WHO, 1993). For the present study, conducted in years 2014–2015, participants completed the Autism-Spectrum Quotient as an online questionnaire (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) and the Wechsler Adult Intelligence Scale-IV in a live situation (WAIS-IV; Wechsler, 2008).

In this study we did not include data from participants whose eye-tracking data were valid for less than 40 % of the total sampling time. Applying this relatively conservative criterion resulted in the exclusion of 14 participants with ASD and 16 neurotypical controls. We additionally excluded one control, because this participant had a WAIS-IV score lower than 70. The final sample consisted of 16 participants with ASD (10 males, 6 females) and 16 control participants (10 males, 6 females). The groups were similar in age and had similar scores on the WAIS-IV (Table 1). There was a significant group difference in scores on the AQ, indicating a difference in ASD symptomatology. Two participants with ASD and two control participants did not accomplish the AQ because of technical issues or other difficulties in completing the online questionnaire.

*TABLE 1 APPROXIMATELY HERE*

### **Stimuli and Verbal Questions**

The stimulus used in the eye-tracking session was a semi-naturalistic video depicting a pragmatically complex communication situation. We used a dynamic, naturalistic, and complex stimulus, because it has been suggested that stimuli demonstrating these features

facilitate the emergence of subtle differences between individuals with ASD and their neurotypical peers (Falck-Ytter & von Hofsten, 2011; Guillon et al., 2014). The video was designed for this study. We chose an unedited clip without any cuts, because stimuli with a lot of cuts tend to increase central biasing, that is, the tendency to fixate on the centre of the scene more than other locations (Dorr, Gegenfurtner, & Barth, 2010; Le Meur, Le Callet, & Barba, 2007). The stimulus video shows a scene with four women of the same age as the participants. In the beginning the women sit around a table having a relaxed conversation. One of the four women offers the others sandwiches, but drops the last sandwich. The woman who drops the sandwich (person 1) apologises briefly. The woman whose sandwich fell (person 2) replies, “*It’s okay,*” but her voice, posture, and facial expression imply that she actually did mind. During the rest of the clip, person 2 looks sad and does not say anything, whereas the others take up the previous discussion in the same relaxed tone as in the beginning. So the main pragmatic and social content of the video is the discrepancy between the verbal and nonverbal cues given by person 2, and furthermore the discrepancy between the reaction of person 2 and the amount of sympathy the others show her. To understand this, participants needed many skills underlying pragmatic understanding. First they had to attend towards verbal and nonverbal social clues which required executive function skills and social attention (Hill, 2004; Senju, 2013). Participants also had to be able to flexibly shift their attention from one potential source of social information to another and to combine information from various sources, demanding cognitive flexibility and good central coherence (Happé & Booth, 2008; Hill, 2004). Lastly the participants had to use the social cues and contextual information they had perceived to make interpretations about social constellations and other persons’ thoughts and feelings, that is, use their social cognition and mentalisation skills (Frith & Frith, 2012; Tomasello et al., 2005). After watching the video clip participants were asked the following questions:

1. What did the woman with short hair (person 2) think?
2. How do you know she thought so?
3. What did the woman with the checkered shirt (person 1) think?
4. How do you know she thought so?
5. What was the thing that fell in the video? (control question)

Answers to all questions were transcribed and scored by one of the authors. A second rater, blind to the group status of the participants, additionally scored the answers for an analysis of inter-rater reliability. Questions 1–4 were scored on a three-point scale. If an answer revealed that the participant did not grasp the basic pragmatic content of the video (questions 1 and 3), or could not explain his or her answer (questions 2 and 4), he or she scored zero on the question. For an adequate but simple answer or explanation the participants got one point. If the answer contained adequate remarks on social implications or underlying causes (questions 1 and 3), or two or more explanations (questions 2 and 4), the participant got two points. Finally, the scores for the four questions were summed so that each participant ended up having one score in the value range 0–8. The term verbal task score will hereinafter refer to this summed score unless stated otherwise. A control question (question 5) was included to ensure that all participants had watched the clip and were aware of the main event at a nonsocial level. Answers to the control question were rated as either correct or incorrect, and not included in the summed verbal task score. A detailed description of the scoring criteria is provided in the Supplemental Material A.

## **Apparatus and Procedure**

Participants' eye movements were recorded at a sampling rate of 300 Hz using a Tobii TX300 remote eye tracker (Tobii Technology, 2014) and Tobii Studio 3.3 software (Tobii Technology, 2015). The eye tracker was calibrated with a five-point calibration procedure before each recording. A detailed description of the eye tracking apparatus and data sampling is provided in the Supplemental Material B.

The video clip used in this study was shown to the participants as a part of a larger set of stimuli. The video was shown in a silent room with only fluorescent light. Once the participants had seen the video, they answered the verbal task questions. Questions were posed and answered orally. Before starting the task, the participants were told that they would be asked questions about the video, but they were not informed about the content of the questions.

## **Data Analysis**

**Verbal task and background variables.** Descriptive and inferential statistics for verbal task scores were calculated using the computer programming software MATLAB (Mathworks, 2015a). Equality of the central tendency was tested with the Mann–Whitney U test. Internal consistency of items in the verbal task was estimated using the unstandardised version of Cronbach's alpha and the inter-rater reliability was estimated with Kendall's tau<sub>b</sub>.

**Eye-tracking data .** The eye-tracking data were analysed in a three-stage approach, described shortly here and in detail in the Supplemental Material C.

We first identified and marked the video elements that would be used as targets, namely the sandwich and the faces of persons 1–4, and calculated the Euclidean distance between the target points and gaze points. We then calculated the Euclidean distance between target points and gaze points and transformed each distance to a value between zero and one, hereinafter referred to as a *gaze value*. A gaze value, in other words, was an index of how close to a target any participant's gaze was at any given moment. At zero distance the gaze value was one, and as the distance increased, the corresponding gaze value decreased according to a coefficient based on a Gaussian distribution.

In the second stage of the eye-tracking data analysis, we performed frame-by-frame tests for group equality of gaze values and to determine a correlation between the gaze values and verbal task scores. We analysed the eye tracking data frame by frame as we wanted to detect any group differences or correlation with a high temporal resolution and analyse them in parallel to the plot of the video. Using this method enabled, for instance, the detection of situations where one group looked at some of the targets later or for a shorter period than the other group. Statistical tests were conducted in series, with one test for each target and frame. Associations between the gaze values and verbal task scores were investigated using version b of Kendall's rank-order correlation coefficient (Kendall's  $\tau_b$ ). Correlation coefficients and corresponding  $p$  values were calculated target-wise for the ASD group and the control group.

In the third stage of the eye-tracking data analysis, we only used data representing seconds 31–53 in the video, because the events most relevant for understanding the pragmatic content took place in this time interval. During these 22 seconds, the sandwich fell and the discrepancy between reactions to the incident became evident. The frame-by-frame approach with tests conducted in series introduced the need for adjusting the  $p$  values for multiple

comparisons. We decided to adjust the  $p$  values for clusters, i.e. moments longer than just a couple of frames, rather than for individual tests to avoid the risk of overcorrection (see the Supplemental Material C for further explanation). We started this stage of the analysis by identifying moments during which an effect, that is, a significant group difference or correlation, was likely. These moments were sequences of at least four frames with  $p$  values below .05 for either group difference or correlation. The number of frames accepted determined the duration of the moment; a sequence of four frames, for instance, equalled a duration of 160 ms. We identified moments separately for each target on the video. To ensure that the target in question actually was looked at when the suggested effect occurred, we excluded frames with small median (group difference) or third quartile (correlation) gaze values (see the Supplemental Material C for further description). We then calculated each participant's average gaze value for each of the moments with possible effects and re-tested the group difference or correlation using these averaged values. The gaze values reported in the results and discussion sections refer to these averaged values, i.e. moment-level gaze values, unless otherwise stated. At the moment level, group differences were tested with the Mann–Whitney  $U$  test and correlations with Kendall's  $\tau_b$ . Finally we adjusted all moment-level  $p$  values for multiple comparisons using the Bonferroni–Holm method (Eichstaedt, Kovatch, & Maroof, 2013; Holm, 1979). Throughout the study, we considered  $p$  values of .05 or less to be statistically significant.

## Results

### Verbal Task Scores

All participants answered the control question correctly. Verbal task scoring was consistent between the items (Cronbach's alpha,  $\alpha = .82$ ) and raters (Kendall's  $\tau_b$ ,  $\tau = .75$ ,  $p < .001$ ).

Verbal task scores were lower in the ASD group (median = 3, IQR = 3.5) than in the control group (median = 5, IQR = 2). The difference was significant with a moderate effect size ( $U = 82$ ,  $n_1 = n_2 = 16$ , one-tailed  $p = .042$ ,  $r = -.31$ ).

### **Group Differences in Moment-Level Gaze Values**

Two moments with potential group differences fulfilled the preselection criteria and were included in the final analyses (Table 2). There was one moment for which the moment-level gaze values for the sandwich were significantly higher in the ASD group ( $U = 62$ ,  $n_1 = n_2 = 15$ , two-tailed  $p = .038$ ) and one in which the moment-level gaze values were significantly lower in the ASD group ( $U = 32$ ,  $n_1 = n_2 = 16$ , two-tailed  $p = .018$ ). The effect sizes for group difference for both re-tested moments were moderate ( $r = .38$  and  $r = -.46$ , respectively).

### **Correlations Between Gaze Values and Verbal Task Scores**

Fourteen moments with potential correlations fulfilled the preselection criteria and were included in the final analyses. In the ASD group there were three moments with positive correlations between moment-level gaze values for the sandwich and verbal task scores, two moments with a positive correlation between moment-level gaze values for person 1 and verbal task scores, one moment with a positive correlation between moment-level gaze values for person 3 and verbal task scores, and one moment with a positive correlation between moment-level gaze values for person 4 and verbal task scores (Table 3, Figure 1). The  $\tau_b$  values indicated strong correlations.

## **Moment-to-Moment Description of Group Differences and Correlations**

In the following section, moments with significant group differences in moment-level gaze values or with correlations between moment-level gaze values and verbal task scores are described in parallel with the plot of the stimulus video. All the correlations mentioned concern the ASD group because no significant correlations were found in the control group. During seconds 31–34, person 1 first offered the plate with sandwiches to person 2, then the sandwich slid off the plate and hit the ground. At the time, there was first a moment during which individuals with ASD looked more at the sandwich than the control participants, and then a moment during which individuals with ASD looked less than control participants at person 3 (Table 2, Figure 1). Temporarily overlapping both of these moments, individuals with ASD showed a positive correlation between the moment-level gaze values for the sandwich and verbal task scores (Table 3, Figure 1). In addition, there was a moment with a positive correlation between moment-level gaze values for person 3 and verbal task scores immediately after the moment during which individuals with ASD looked less at person 3 than the control participants.

When the sandwich hit the ground at around second 34, there was a moment with a significant correlation between the moment-level gaze values for person 1 and verbal task scores in the ASD group (Table 3). After the sandwich fell, person 1 said she was sorry and looked apologetic, but at around seconds 36–37 she took on her previous joyful expression. At this time point there was a moment with a positive correlation between moment-level gaze values for person 1 and verbal task scores (Table 3). During seconds 39–42 person 2 picked up the sandwich and claimed in a sad voice that she did not mind that the sandwich had

fallen. Around second 40 there was a moment with a positive correlation between the moment-level gaze values for the sandwich and verbal task scores in the ASD group (Table 3). During seconds 42–53, the discussion continued between persons 1, 3 and 4, whereas person 2 was silent and looked sad. Around second 48, when person 4 was speaking, the moment-level gaze values for person 4 correlated positively with verbal task scores (Table 3).

FIGURE 1 AND TABLES 2 AND 3 APPROXIMATELY HERE (landscape format)

## Discussion

In the present study we investigated the ways in which individuals with ASD and their peers watched and interpreted a video depicting a complex communication situation. We found significant correlations between moment-level gaze values and verbal task scores in the ASD group, but not among the controls. There were two moments with significant group differences in moment-level gaze values. Individuals with ASD had lower verbal task scores than the controls.

As we expected, the correlation between verbal task scores and moment-level gaze values co-occurred with moments during which the gazed-at target did something that could be interpreted as a social cue. The first correlation between the moment-level gaze values for person 1 and verbal task scores co-occurred with person 1's initial reaction to dropping the sandwich and second one with the moment during which person 1 changed her facial expression from an apologetic to a joyful look. This change of facial expression might have been a social cue that helped the participants notice that person 1 overlooked the feelings of person 2. The correlation between the moment-level gaze values for person 3 and the verbal

task scores, in turn, was socially relevant because it co-occurred at the moment during which person 3 started to turn her head and gaze towards the sandwich that just had dropped. The other targets, i.e. person 4 and the sandwich, whose moment-level gaze values correlated with verbal task scores, did not express anything relevant for understanding the pragmatic content of the scene at the moments the correlations occurred. It is possible that the positive correlation between the moment-level gaze values for these targets and the verbal task scores was due to a difference in general attention: The individuals with high moment-level gaze values for the sandwich or person 4 evidently showed some interest in the events of the video at the moment of positive correlation, whereas the individuals with low moment-level gaze values might have looked at some object in the periphery of the scene. The fact that we did not find any significant correlations between the verbal task performance and the moment-level gaze values in the control group could suggest that most control participants understood the pragmatic content of the video regardless of which target they were observing at any given moment. In the ASD group, on the contrary, decreased attention to social cues or to the scene in general seemed to be related to reduced pragmatic understanding. Similar findings have been reported in previous studies with high-functioning adults with ASD, although methodological differences prevent direct comparisons (Grynszpan & Nadel, 2015; Sasson et al., 2007).

Based on results obtained in previous studies using social videos as stimuli (e.g. Bird et al., 2011; Nakano et al., 2010; Speer et al., 2007), we expected to find some differences in how the individuals with ASD and their neurotypical peers followed the dynamics of the social scene with their gaze. There were, in fact, two moments during which significant group differences were found: first the individuals with ASD looked more at the sandwich and then less at person 3 than the control participants. This result could be interpreted as a implication

of a lower than usual spontaneous interest in social targets, which has been found to be common in high-functioning individuals with ASD (Senju, 2013). On the other hand, the moments with significant group differences only lasted for some milliseconds out of the 22 seconds analysed. Therefore, it seems to be an overextension to claim that there were fundamental differences in the ways in which the individuals with ASD and their neurotypical peers followed the dynamics of the social interaction. In this regard, the results did not meet our expectations. The large intragroup variability and the consequential intergroup overlap of the eye-tracking data partly explain why there were less group differences than expected. High variability of eye movements have previously been found among individuals with ASD (Nakano et al., 2010), but in this study the variability was also present in the control group, which is not typically the case (Nakano et al., 2010; Shockley et al., 2009). For neurotypical participants it has been shown that the variability tends to be higher when participants are performing a task than during free viewing, presumably because then the participants are using different perceptual strategies to solve the task (Smith & Mital, 2013). In the present study the participants were not actively performing a task during the eye tracking, but knew that they would be asked questions after they had seen the video, which might have had a similar effect to performing a task.

The findings of the verbal task were in line with previous studies indicating that even very high-functioning individuals with ASD often have weakened, though not absent, pragmatic language skills (Loukusa & Moilanen, 2009; Vulchanova et al., 2015). In the present study many participants with ASD had some difficulties in understanding the pragmatic content of the video shown, but very few of the individuals with ASD misinterpreted the scene completely, as indicated by the fact that the participants rarely scored zero on the verbal task. However, even relatively mild difficulties of pragmatic understanding can have important

consequences (Hanley et al., 2015; Tobin et al., 2014). If an individual repeatedly does not notice hidden insults or indirectly expressed feelings of offense, he or she probably finds it hard to follow everyday social interactions, which may have negative effects on the social functioning and well-being of the individual.

## **Limitations**

In this study at least 40 % of the eye-movement data from a participant had to be valid for the participant to be included in the study, and applying this relatively conservative criterion resulted in the exclusion of many participants. The excluded participants might have moved their head extensively (Holmqvist et al., 2011), and they possibly had more difficulties than the included participants in directing their attention towards the stimulus. We do not, however, see this as a factor that could undermine the findings concerning the group differences, because the proportion of the excluded participants did not differ between the groups. Even after the preselection described above, the number of participants with valid eye movements varied slightly from one frame to another. We compared the data concerning sample sizes and the results from statistical tests using temporally aligned graphs, and could not detect any relation between decreased or unequal sample sizes and the notion of a statistically significant effect. In order to sharpen the analytic focus, we removed the data representing the beginning and the end of the video. This selection would optimally have been made during the planning stage of the study. The mid-study selection was, however, justified, because it was based on expectations of which time points would be the most relevant for understanding the pragmatic content rather than on results from the early analyses.

Both frame- and moment-level gaze values for all targets were relatively low in general. The low group medians for gaze values were partly a consequence of the high within-group variability discussed above. At the level of the individual participants, low gaze values can to some extent be explained by the scaling procedure that was used when a gaze point was close to two or more targets at once and the sum of the gaze values for all targets exceeded one. In these situations, a participant could have attended to any of the closely located targets, and to account for this we did not assign “full” gaze values to any of the targets, but instead divided the original gaze values for each target by the sum of the values (see the Supplemental Material C for a full description of the scaling procedure). This procedure resulted in a large number of scaled, and thus low, gaze values.

The interpretation of the results in the verbal task is somewhat compromised by the fact that the verbal task questions were open-ended and scored with a system in which answers going beyond the obvious facts received higher scores than adequate but simple answers.

Individuals with ASD might have had a tendency to give the simplest possible answer, even though they would have been able to offer the information required for a full score if asked directly (Simon Baron-Cohen & Barnes, 2012; Loukusa & Moilanen, 2009). On the other hand, the interpretation of what is important in the scene and hence worth describing even when not directly asked for can be seen as a part of the pragmatic understanding.

### **Clinical Implications**

Results from this and previous studies (Grynszpan & Nadel, 2015; Sasson et al., 2007) indicate that the connection between eye movement and social performance is more

pronounced for high-functioning adults with ASD than for their neurotypical peers. It could be that individuals with ASD need to attend to the events of social scenes more closely than their peers in order to make adequate social and pragmatic interpretations (Grynszpan & Nadel, 2015; Vulchanova et al., 2015). If individuals with ASD need to catch more social cues than their neurotypical peers to understand pragmatic language, they may benefit from practicing their ability to perceive these cues (Grynszpan et al., 2009; Hanley et al., 2015). On the other hand, the interventions could focus on how to make social interpretations when only a sparse amount of cues are available or perceived.

## **Conclusions**

In this study participants with ASD had difficulties in understanding the pragmatic content of a complex social situation. The connection between eye movement and pragmatic understanding was more pronounced for participants with ASD than for neurotypical participants. These insights about the links between social perception and interpretation could be used for developing interventions for high-functioning individuals with ASD. Combining moment-by-moment eye tracking and a verbal task pinpointing complex pragmatic skills appeared to be a sensitive way of investigating the characteristics of social understanding in high-functioning individuals with ASD, and this approach may be further developed in future studies.

## **Acknowledgements**

This research was supported by the Academy of Finland, Helsinki, Finland, under grant 276578; the Alma and K. A. Snellman Foundation, Oulu, Finland; the Child Psychiatry

Research Foundation, Finland; the Emil Aaltonen Foundation, Tampere, Finland; Northern Ostrobothnia Hospital District, Finland; Rinnekoti Research Foundation / the Finnish Brain Foundation, Espoo / Helsinki, Finland; the Sigrid Jusélius Foundation, Helsinki, Finland. The authors wish to thank Leena Joskitt (MA) for statistical support and Jasmin Alian (MA) for scoring the verbal task answers for reliability analyses. We particularly want to thank all the participants for taking part in this study. The authors declare no conflict of interest.

### Supplemental Material

The Supplemental Material can be found at the address

### References

- Allman, M. J., & Meck, W. H. (2012). Pathophysiological distortions in time perception and timed performance. *Brain*, *135*, 656–677. doi:10.1093/brain/awr210
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Ames, C., & Fletcher-Watson, S. (2010). A review of methods in the study of attention in autism. *Developmental Review*, *30*, 52–73. doi:10.1016/j.dr.2009.12.003
- Arndt, S., Turvey, C., & Andreasen, N. C. (1999). Correlating and predicting psychiatric symptom ratings: Spearman's r versus Kendall's tau correlation. *Journal of Psychiatric Research*, *33*, 97–104. doi:10.1016/S0022-3956(98)90046-2
- Baez, S., Rattazzi, A., Gonzalez-Gadea, M. L., Torralva, T., Vigliecca, N. S., Decety, J., ... Ibanez, A. (2012). Integrating intention and context: assessing social cognition in adults with Asperger syndrome. *Frontiers in Human Neuroscience*, *6*, 1–21.

doi:10.3389/fnhum.2012.00302

Baron-Cohen, S., & Barnes, J. L. (2012). The big picture: Storytelling ability in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *42*, 1557–1565. doi:10.1007/s10803-011-1388-5

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*, 5–17. doi:10.1023/A:1005653411471

Barracough, J. K. (2011). PhysMo - Video Motion Analysis [Computer Software]. Retrieved March 1, 2015, from <http://physmo.sourceforge.net>

Ben Saïda, A. (2014). Shapiro-Wilk parametric hypothesis test of composite normality (Version 3). [Code file]. Retrieved September 1, 2015, from <http://www.mathworks.com/matlabcentral/fileexchange/13964-shapiro-wilk-and-shapiro-francia-normality-tests/content/swtest.m>

Bird, G., Press, C., & Richardson, D. C. (2011). The role of alexithymia in reduced eye-fixation in Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*, *41*, 1556–1564. doi:10.1007/s10803-011-1183-3

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Gaze selection in complex social scenes. *Visual Cognition*, *16*, 341–355. doi:10.1080/13506280701434532

Blignaut, P. (2009). Fixation identification: The optimum threshold for a dispersion algorithm. *Attention, Perception, & Psychophysics*, *71*, 881–895. doi:10.3758/APP.71.4.881

Cribbie, R. A., Wilcox, R. R., Bewell, C., & Keselman, H. J. (2007). Tests for treatment group equality when data are nonnormal and heteroscedastic. *Journal of Modern*

*Applied Statistical Methods*, 6, 117–132. Retrieved from

<http://digitalcommons.wayne.edu/jmasm/vol6/iss1/12>

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168.

doi:10.1146/annurev-psych-113011-143750

Dorr, M., Gegenfurtner, K. R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision*, 10, 1–17. doi:10.1167/10.10.28

Eichstaedt, K. E., Kovatch, K., & Maroof, D. A. (2013). A less conservative method to adjust for familywise error rate in neuropsychological research: The Holm's sequential Bonferroni procedure. *NeuroRehabilitation*, 32, 693–696. doi:10.3233/NRE-130893

Falck-Ytter, T., & von Hofsten, C. (2011). How special is social looking in ASD: A review. *Progress in Brain Research*, 189, 209–222. doi:10.1016/B978-0-444-53884-0.00026-9

Falck-Ytter, T., Von Hofsten, C., Gillberg, C., & Fernell, E. (2013). Visualization and analysis of eye movement data from children with typical and atypical development. *Journal of Autism and Developmental Disorders*, 43, 2249–2258. doi:10.1007/s10803-013-1776-0

Falter, C. M., Elliott, M. A., & Bailey, A. J. (2012). Enhanced visual temporal resolution in autism spectrum disorders. *PloS One*, 7. doi:10.1371/journal.pone.0032774

Fletcher-Watson, S., Leekam, S., & Benson, V. (2009). Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia*, 47, 248–257. doi:10.1016/j.neuropsychologia.2008.07.016

Freeth, M., Chapman, P., Ropar, D., & Mitchell, P. (2010). Do gaze cues in complex scenes capture and direct the attention of high functioning adolescents with ASD? Evidence from eye-tracking. *Journal of Autism and Developmental Disorders*, 40, 534–547.

doi:10.1007/s10803-009-0893-2

Freeth, M., Foulsham, T., & Chapman, P. (2011). The influence of visual saliency on fixation patterns in individuals with autism spectrum disorders. *Neuropsychologia*, *49*, 156–160.

doi:10.1016/j.neuropsychologia.2010.11.012

Freeth, M., Ropar, D., Mitchell, P., Chapman, P., & Loher, S. (2011). How adolescents with ASD process social information in complex scenes. Combining evidence from eye movements and verbal descriptions. *Journal of Autism and Developmental Disorders*,

*41*, 364–371. doi:10.1007/s10803-010-1053-4

Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology*, *63*, 287–313. doi:10.1146/annurev-psych-120710-100449

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, *141*, 2–18. doi:10.1037/a0024338

Gibbs, R. W., & Orden, V. (2012). Pragmatic choice in conversation. *Topics in Cognitive Science*, *4*, 7–20. doi:10.1111/j.1756-8765.2011.01172.x

Grynszpan, O., & Nadel, J. (2015). An eye-tracking method to reveal the link between gazing patterns and pragmatic abilities in high functioning autism spectrum disorders. *Frontiers in Human Neuroscience*, *8*, 1–6. doi:10.3389/fnhum.2014.01067

Grynszpan, O., Nadel, J., Constant, J., Le Barillier, F., Carbonell, N., Simonin, J., ...

Courseon, M. (2009). A new virtual environment paradigm for high functioning autism intended to help attentional disengagement in a social context: Bridging the gap between relevance theory and executive dysfunction. *2009 Virtual Rehabilitation International Conference, VR 2009*, 51–58. doi:10.1109/ICVR.2009.5174205

doi:10.1109/ICVR.2009.5174205

- 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. doi:10.1016/j.neubiorev.2014.03.013
- Habets, B., Kita, S., Shao, Z., Özyurek, A., & Hagoort, P. (2011). The role of synchrony and ambiguity in speech-gesture integration during comprehension. *Journal of Cognitive Neuroscience*, *23*, 1845–1854. doi:10.1162/jocn.2010.21462
- Hanley, M., McPhillips, M., Mulhern, G., & Riby, D. M. (2013). Spontaneous attention to faces in Asperger syndrome using ecologically valid static stimuli. *Autism*, *17*, 754–761. doi:10.1177/1362361312456746
- Hanley, M., Riby, D. M., Carty, C., McAteer, A. M., Kennedy, A., & McPhillips, M. (2015). The use of eye-tracking to explore social difficulties in cognitively able students with Autism Spectrum Disorder: A pilot investigation. [Special issue]. *Autism*, *19*, 868–873. doi:10.1177/1362361315580767
- Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*, 5–25. doi:10.1007/s10803-005-0039-0
- Happé, F. G. E., & Booth, R. D. L. (2008). The power of the positive: Revisiting weak coherence in autism spectrum disorders. *Quarterly Journal of Experimental Psychology*, *61*, 50–63. doi:10.1080/17470210701508731
- Heller, R., Stanley, D., Yekutieli, D., Rubin, N., & Benjamini, Y. (2006). Cluster-based analysis of fMRI data. *NeuroImage*, *33*, 599–608. doi:10.1016/j.neuroimage.2006.04.233
- Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences*, *8*, 26–32. doi:10.1016/j.tics.2003.11.003

Holler, J., & Kendrick, K. H. (2015). Unaddressed participants' gaze in multi-person interaction: Optimizing reciprocity. *Frontiers in Psychology*, 6, 1–14.

doi:10.3389/fpsyg.2015.00098

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70. doi:10.2307/4615733

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking. A comprehensive guide to methods and measures*. New York: Oxford University Press.

Jansson-Verkasalo, E., Kujala, T., Jussila, K., Mattila, M. L., Moilanen, I., Näätänen, R., ... Korpilahti, P. (2005). Similarities in the phenotype of the auditory neural substrate in children with Asperger syndrome and their parents. *European Journal of Neuroscience*, 22, 986–990. doi:10.1111/j.1460-9568.2005.04216.x

Kendall, M. G. (1970). *Rank correlation methods* (4th ed.). London: Griffin.

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

doi:10.1001/archpsyc.59.9.809

Kuusikko, S., Haapsamo, H., Jansson-Verkasalo, E., Hurtig, T., Mattila, M. L., Ebeling, H., ... Moilanen, I. (2009). Emotion recognition in children and adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39, 938–945.

doi:10.1007/s10803-009-0700-0

Kuusikko, S., Pollock-Wurman, R., Jussila, K., Carter, A. S., Mattila, M. L., Ebeling, H., ... Moilanen, I. (2008). Social anxiety in high-functioning children and adolescents with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, 38,

1697–1709. doi:10.1007/s10803-008-0555-9

Le Meur, O., Le Callet, P., & Barba, D. (2007). Predicting visual fixations on video based on low-level visual features. *Vision Research*, *47*, 2483–2498.

doi:10.1016/j.visres.2007.06.015

Lin, G., & Carlile, S. (2015). Costs of switching auditory spatial attention in following conversational turn-taking. *Frontiers in Neuroscience*, *9*, 1–11.

doi:10.3389/fnins.2015.00124

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2000). *Autism diagnostic observation schedule (ADOS)*. Los Angeles: Western Psychological Services.

Lord, C., Rutter, M., & Le Couter, A. (1995). *Autism diagnostic interview – revised (ADI-R)* (3rd ed.). Los Angeles: Western Psychological Services.

Loukusa, S., & Moilanen, I. (2009). Pragmatic inference abilities in individuals with Asperger syndrome or high-functioning autism. A review. *Research in Autism Spectrum Disorders*, *3*, 890–904. doi:10.1016/j.rasd.2009.05.002

Martin, I., & McDonald, S. (2003). Weak coherence, no theory of mind, or executive dysfunction? Solving the puzzle of pragmatic language disorders. *Brain and Language*, *85*, 451–466. doi:10.1016/S0093-934X(03)00070-1

Mathworks. (2015a). MATLAB Release 2015a. Natick, MA.

Mathworks. (2015b). MATLAB, Statistics and Machine Learning Toolbox Release 2015a. Natick, MA.

Mattila, M.-L., Kielinen, M., Jussila, K., Linna, S.-L., Bloigu, R., Ebeling, H., & Moilanen, I. (2007). An epidemiological and diagnostic study of Asperger syndrome according to four sets of diagnostic criteria. *Journal of the American Academy of Child and*

*Adolescent Psychiatry*, 46, 636–646. doi:10.1097/chi.0b013e318033ff42

Mattila, M.-L., Kielinen, M., Linna, S.-L., Jussila, K., Ebeling, H., Bloigu, R., ... Moilanen,

I. (2011). Autism spectrum disorders according to DSM-IV-TR and comparison with DSM-5 draft criteria: An epidemiological study, *50*, 583–592.

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 of The Royal Society B: Biological Sciences*, 277, 2935–2943. doi:10.1098/rspb.2010.0587

Norbury, C. F., Brock, J., Cragg, L., Einav, S., Griffiths, H., & Nation, K. (2009). Eye-

movement patterns are associated with communicative competence in autistic spectrum disorders. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 50, 834–842. doi:10.1111/j.1469-7610.2009.02073.x

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between

speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29, 1045–1060. doi:10.1207/s15516709cog0000\_29

Roche, J. M., Dale, R., & Caucci, G. M. (2012). Doubling up on double meanings: Pragmatic

alignment. *Language and Cognitive Processes*, 27, 1–24.

doi:10.1080/01690965.2010.509929

Salmi, J., Roine, U., Glerean, E., Lahnakoski, J., Nieminen-von Wendt, T., Tani, P., ... Sams,

M. (2013). The brains of high functioning autistic individuals do not synchronize with those of others. *NeuroImage. Clinical*, 3, 489–97. doi:10.1016/j.nicl.2013.10.011

Sasson, N., Tsuchiya, N., Hurley, R., Couture, S. M., Penn, D. L., Adolphs, R., & Piven, J.

(2007). Orienting to social stimuli differentiates social cognitive impairment in autism and schizophrenia. *Neuropsychologia*, 45, 2580–2588.

doi:10.1016/j.neuropsychologia.2007.03.009

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

Shintel, H., & Keysar, B. (2009). Less is more: A minimalist account of joint action in communication. *Topics in Cognitive Science, 1*, 260–273. doi:10.1111/j.1756-8765.2009.01018.x

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science, 1*, 305–319. doi:10.1111/j.1756-8765.2009.01021.x

Smith, T. J., & Mital, P. K. (2013). Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. *Journal of Vision, 13*, 1–24. doi:10.1167/13.8.16

Speer, L. L., Cook, A. E., McMahon, W. M., & Clark, E. (2007). Face processing in children with autism: Effects of stimulus contents and type. *Autism, 11*, 265–277. doi:10.1177/1362361307076925

Tobii Technology. (2014). Tobii TX300 Eye Tracker. User manual (Revision 2). Retrieved from [www.tobii.com](http://www.tobii.com)

Tobii Technology. (2015). User's manual Tobii Studio (Version 3.3.1). Retrieved from [www.tobii.com](http://www.tobii.com)

Tobin, M. C., Drager, K. D. R., & Richardson, L. F. (2014). A systematic review of social participation for adults with autism spectrum disorders: Support, social functioning, and quality of life. *Research in Autism Spectrum Disorders, 8*, 214–229.

doi:10.1016/j.rasd.2013.12.002

- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675–691. doi:10.1017/S0140525X05000129
- Ullman, M. T., & Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience and Biobehavioral Reviews*, 51, 205–222. doi:10.1016/j.neubiorev.2015.01.008
- Vulchanova, M., Saldaña, D., Chahboun, S., & Vulchanov, V. (2015). Figurative language processing in atypical populations: The ASD perspective. *Frontiers in Human Neuroscience*, 9, 1–11. doi:10.3389/fnhum.2015.00024
- Wechsler, D. (2008). *Manual for the Wechsler Adult Intelligence Scale* (4th ed.). San Antonio, TX: Psychological Corporation.
- Weiss, L. A., Arking, D. E., & the Gene Discovery Project of John Hopkins & the Autism Consortium. (2009). A genome-wide linkage and association scan reveals novel loci for autism. *Nature*, 461, 802–808. doi:10.1038/nature08490
- Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, 68, 130–136. doi:10.2307/2332579
- Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., & Delaney, J. (2006). Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ), 1079, 47–56. doi:10.1016/j.brainres.2006.01.012
- World Health Organization. (1993). *International Classification of mental and behavioural disorders (ICD-10). Diagnostic criteria for research*. Geneva: World Health Organization.

Figure captions

*Figure 1.* Gaze values at frame-level and moments with group differences or correlations.

The panels show results for the sandwich, person 1 and person 3, respectively. The gaze values used for comparing groups are group medians (blue line, ASD; red line, controls), and the gaze values used for correlations are third quartile values (black line, ASD). Third quartile values are not shown for the control group, because there were no significant correlations in their group. Moments with significant group differences are marked with grey boxes and moments with correlations are marked with a sequence of green stars. Significance threshold  $p < .05$  (two-tailed, Bonferroni-Holm adjusted).

*Figure C1.* Gaze value transformation. Black circles = gaze point; X = target point. The distances are Euclidean distances between gaze points and target points expressed in pixels (px). The gaze values are distance-based values between 0 and 1, and they decrease as the distances increase. At a distance of 90 pixels the unscaled gaze value would be .2 (panel A and B) and at a distance of 33 pixels the unscaled gaze value would be .8 (panel B). If the sum of the gaze values for all targets exceeds one at some time point, the gaze values are scaled by dividing each unscaled gaze value with the sum of all the unscaled gaze values (panel B).

MAIN TEXT: 6377 WORDS

WHOLE DOCUMENT 11 170 WORDS

Table 1

*Participant characteristics*

	ASD group			Control group		
	<i>n</i>	<i>Mdn</i>	<i>IQR</i>	<i>n</i>	<i>Mdn</i>	<i>IQR</i>
Age in years	10 / 6	22.5	4.2	10 / 6	22.6	2.0
AQ	9 / 5	23.0	10.0	8 / 6	12.0	3.0
GAI	10 / 6	109.5	25.3	10 / 6	107.5	13.3

*Note.* *n* = Number of participants, male / female; *Mdn* = median; *IQR* = interquartile range;

AQ = Autism-Spectrum Quotient; GAI = General Ability Index on Wechsler Adult

Intelligence Scale-IV.

Table 2

Statistics for moments tested for group differences in moment-level gaze values

<u>Target</u>	<u>Time</u>	ASD group			Control group			Group difference		
		<u>n</u>	<u>Mdn</u>	<u>IQR</u>	<u>n</u>	<u>Mdn</u>	<u>IQR</u>	<u>U</u>	<u>p</u>	<u>r</u>
Sandwich	31.48–31.6	15	0.52	0.58	15	0.11	0.28	62.0	<b>.039</b>	<b>.38</b>
Person 3	32.64–32.76	16	0.01	0.03	16	0.45	0.61	33.0	<b>.019</b>	<b>-.48</b>

Note. Target = the looked-at target in the video in each respective moment; Time = the duration of the moment displayed in seconds; Mdn = median gaze value, IQR = interquartile range; U = Mann–Whitney U statistic; r = effect size. Significance threshold  $p < .05$ , p values are two-tailed and adjusted for multiple comparisons using the Bonferroni–Holm method. Significant values are indicated in **bold** text.

Table 3

Statistics for moments tested for Kendall tau<sub>b</sub> correlation between verbal task scores and moment-level gaze values

<u>Target</u>	<u>Group</u>	<u>Time</u>	<u>n</u>	<u>Q3</u>	<u>IQR</u>	<u>τ</u>	<u>p</u>
Sandwich	ASD	31.44–31.76	16	0.58	0.56	<b>.54</b>	<b>.042</b>
Sandwich	ASD	32.6–33.00	16	0.63	0.63	<b>.61</b>	<b>.021</b>
Sandwich	ASD	39.92–40.36	16	0.77	0.72	<b>.56</b>	<b>.036</b>
Sandwich	ASD	44.84–45.08	14	0.33	0.31	.48	.133
Sandwich	Control	37.12 – 37.36	16	0.69	0.69	-.43	.133
Person 1	ASD	33.96–34.16	13	0.36	0.36	<b>.80</b>	<b>.004</b>
Person 1	ASD	36.56–36.96	15	0.50	0.50	<b>.62</b>	<b>.025</b>
Person 2	ASD	39.76–39.88	16	0.77	0.76	.44	.133

(Continued)

Table 3

*Statistics for moments tested for Kendall tau<sub>b</sub> correlations between verbal task scores and moment-level gaze values (continued)*

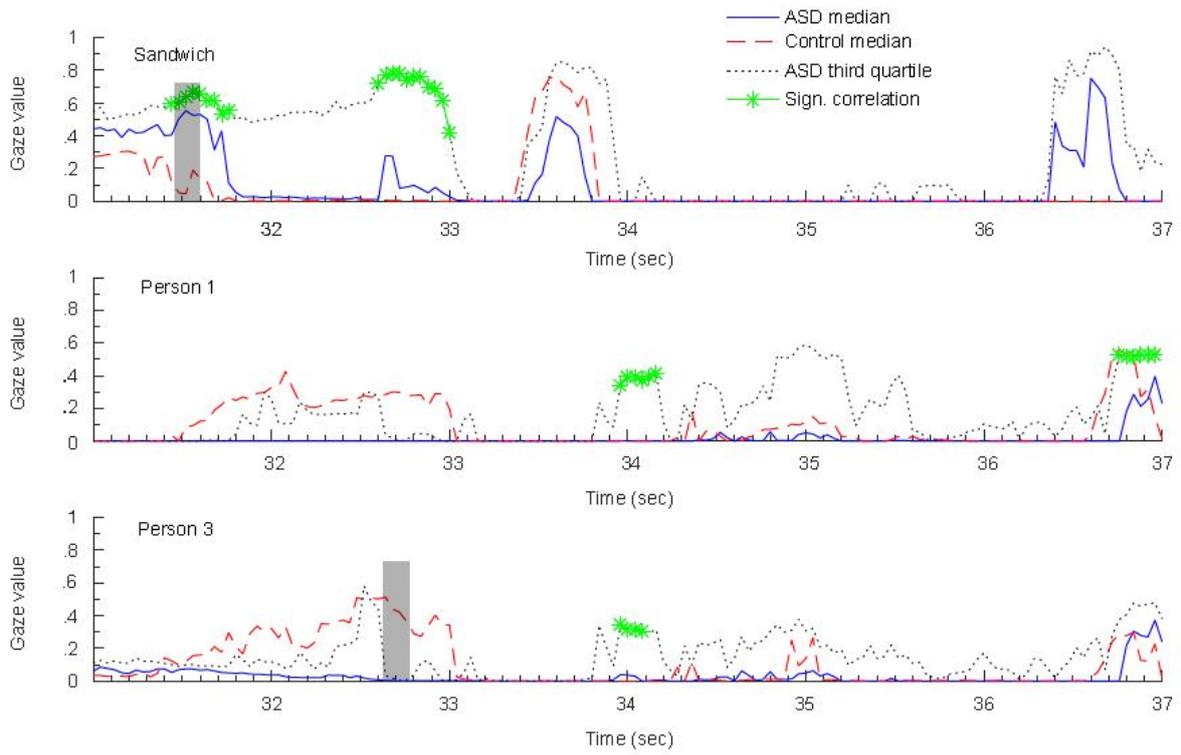
<u>Target</u>	<u>Group</u>	<u>Time</u>	<u><i>n</i></u>	<u><i>Q3</i></u>	<u><i>IQR</i></u>	<u><math>\tau</math></u>	<u><i>p</i></u>
Person 3	ASD	33.96-34.08	13	<b>0.32</b>	<b>0.32</b>	<b>.64</b>	<b>.037</b>
Person 3	ASD	42.04-42.24	16	0.48	0.48	.46	.133
Person 4	ASD	47.72-48	15	<b>0.71</b>	<b>0.65</b>	<b>.62</b>	<b>.025</b>
Person 4	Control	43.96-44.12	15	0.76	0.70	-.42	.133
Person 4	Control	44.2-44.36	14	0.77	0.73	-.50	.133
Person 4	Control	45.08-45.24	15	0.73	0.72	-.45	.133

*Note.* Target = the looked-at target in the video at each respective moment; Group = the group in which the potential correlation was found;

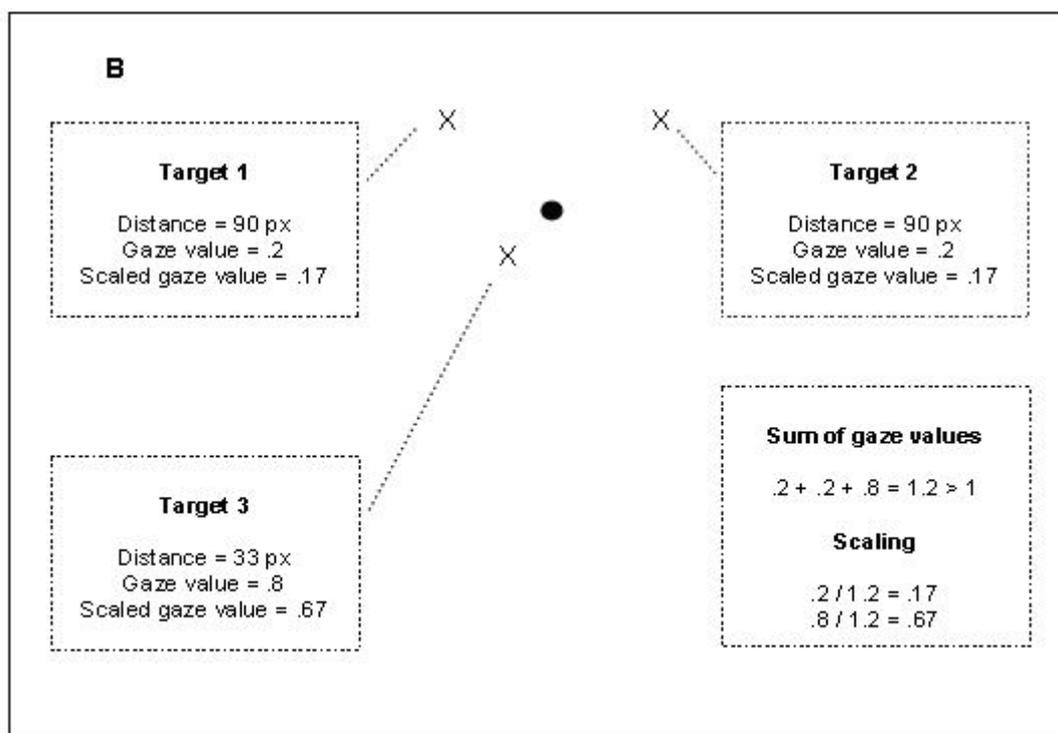
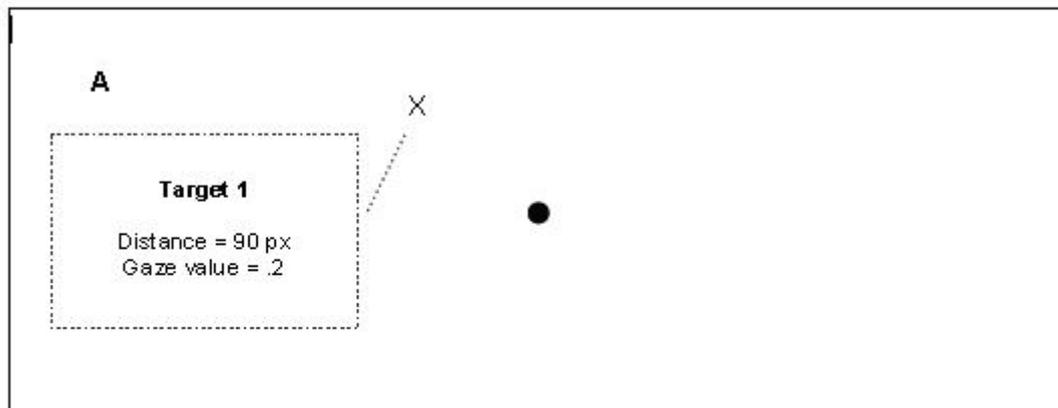
Time = the duration of the moment displayed in seconds; *Q3* = third quartile gaze value; *IQR* = interquartile range;  $\tau$  = Kendall tau<sub>b</sub> statistic

Significance threshold  $p < .05$ , *p* values are two-tailed and adjusted for multiple comparisons using the Bonferroni–Holm method. Significant values are indicated in **bold** text.

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT



ACCEPT