

Exploring the building blocks of social cognition: spontaneous agency perception and visual perspective taking in autism

Jan Zwickel,¹ Sarah J. White,² Devorah Coniston,² Atsushi Senju,³ and Uta Frith^{2,4}

¹Department of Psychology, Ludwig Maximilians University, Munich, Germany, D-80802 Munich, ²Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, ³Centre for Brain and Cognitive Development, Birbeck, University of London, London WC1E 7HX, UK, and ⁴Center of Functionally Integrative Neuroscience, University of Aarhus, 8000 Århus C, Denmark

Individuals with autism spectrum disorders have highly characteristic impairments in social interaction and this is true also for those with high functioning autism or Asperger syndrome (AS). These social cognitive impairments are far from global and it seems likely that some of the building blocks of social cognition are intact. In our first experiment, we investigated whether high functioning adults who also had a diagnosis of AS would be similar to control participants in terms of their eye movements when watching animated triangles in short movies that normally evoke mentalizing. They were. Our second experiment using the same movies, tested whether both groups would spontaneously adopt the visuo-spatial perspective of a triangle protagonist. They did. At the same time autistic participants differed in their verbal accounts of the story line underlying the movies, confirming their specific difficulties in on-line mentalizing. In spite of this difficulty, two basic building blocks of social cognition appear to be intact: spontaneous agency perception and spontaneous visual perspective taking.

Keywords: theory of mind; mental state; social cognition; eye movements; triangle animations

INTRODUCTION

The ability to detect another agent is fundamental to social interaction. Over and above this, successful human interaction depends on the ability to attribute mental states to other agents (for a recent review, see Frith and Frith, 2010). This ability can be studied even when the agents in question are simply animated shapes (Heider and Simmel, 1944). Autism is a neurodevelopmental disorder spanning a whole spectrum of conditions from mild to severe, where the ability to attribute mental states is impaired. This has been studied in high functioning individuals diagnosed as having an autism spectrum condition, and more specifically, Asperger syndrome (AS) using Heider and Simmel-type paradigms (Abell *et al.*, 2000; Klin and Jones, 2006). In contrast, the ability to distinguish between animate and inanimate objects is thought to be intact in autism (Celani, 2002; New *et al.*, 2009). Nevertheless, we still know little about how high functioning autistic individuals extract information when

observing animated agents. Eye movements while watching Heider and Simmel-type shapes might give some insight into the way these agents and their interactions are perceived and this was the purpose of the first of the two experiments we report.

We chose the Frith–Happé animations of two triangles interacting in different scenarios (see below) as they have been used in several previous studies. These stimuli reliably show differences in verbal descriptions between participants with autism and controls (Abell *et al.*, 2000) and are associated with differences in brain activation of the mentalizing systems (Castelli *et al.*, 2000, 2002; Kana *et al.*, 2009). Eye movements differ when neurotypical participants watch these animations according to whether they involve social agents (Klein *et al.*, 2009). However, no study has yet tested whether individuals with autism would also show this pattern of eye movements in response to the different animations. If they did, one could argue that the detection and early processing of social agent information is intact and that their difficulties arise at later stages of processing where information has to be integrated to report a coherent and fitting description of these animations. However, it is equally plausible that not just the later but also the early stages of social information processing are affected in autistic individuals, including the selection of relevant information about social agents.

Whether eye movements reveal difficulties with extracting relevant cues to social agents in autism is of particular interest in light of recent findings that autistic individuals

Received 9 July 2010; Accepted 8 September 2010

Advance Access publication 7 October 2010

The authors thank Idalmis Santiesteban for her help in rating the verbal reports of participants. They would also like to thank Kevin Pelphrey and Antonia Hamilton whose helpful suggestions have been incorporated in this paper. This work is supported within the Deutsche Forschungsgemeinschaft (DFG) excellence initiative research cluster 'cognition for technical systems – CoTeSys', see also <http://www.cotesys.org>, the Joint Medical Research Council UK/Economic and Social Research Council UK Grant (PTA-037-27-0107), the UK Medical Research Council (G0701484), the UK Economic and Social Research Council (RES-063-27-0207), and the University of Aarhus Research Foundation.

Correspondence should be addressed to Jan Zwickel, Ludwig Maximilians University Munich, Department of Psychology, Leopoldstrasse 13, 80802 Munich, Germany. E-mail: jan.zwickel@yale.edu; zwickel@psy.uni-muenchen.de.

are sensitive to social situations as measured by change-blindness paradigms (Fletcher-Watson *et al.*, 2008; Freeth *et al.*, 2010). However, other studies have produced mixed results with respect to differences in eye movements between individuals with and without autism (Klin *et al.*, 2002; Kemner and van Engeland, 2003; Fletcher-Watson *et al.*, 2008; Freeth *et al.*, 2009). Importantly, all of these studies used stimuli that depicted humans; it is therefore unclear whether performance on these tasks is affected by the presence of humans and it remains an open question whether the ability to detect social agent cues in movement alone is abnormal or intact. In the present study, the stimuli did not depict humans and the presence of social agents had to be inferred from movement cues alone. Finding normal eye movements in individuals with high functioning autism/AS would therefore provide compelling evidence that they spontaneously process information about social agents in the same way as neurotypical controls.

With Experiment 2 we investigated another basic building block of social cognition, namely the ability to spontaneously take the visuo-spatial perspective of a social agent into account (Belopolsky *et al.*, 2008; Frischen *et al.*, 2009; Tversky and Hard, 2009; Zwicker, 2009; Samson *et al.*, 2010). If autistic participants would adopt the visuo-spatial perspective of an observed social agent, this would be evident that they not only detect the presence of a social agent but also suggest that they can engage in a first step towards mentalizing (Apperly, 2008), namely stepping into the shoes of the observed social agent.

We already know that another basic building block of social cognition is intact in individuals with high functioning autism: such individuals have been shown to co-represent the task of another person just like healthy controls (Sebanz *et al.*, 2005). In this paradigm, participants reacted to the colour of a ring on a finger that had a spatially irrelevant component (the finger could point either at the participant or at another location). Neurotypical participants showed an influence of the irrelevant pointing direction on their performance only if a second person was sitting at the pointed to location and was concurrently working on the task (Sebanz *et al.*, 2003). Importantly, participants with autism showed the same effect (Sebanz *et al.*, 2005) and thus were not 'other-blind'.

Whilst this work shows that autistic individuals do co-represent the tasks of others, it remains to be seen whether they also co-represent the perspective of others. Individuals with autism certainly have no difficulty performing some explicit level 1 visuo-spatial perspective tasks (Hobson, 1984; David *et al.*, 2010), whereas recent evidence shows that they have difficulty with explicit level 2 visuo-spatial perspective tasks (Hamilton *et al.*, 2009), but it remains to be seen whether implicit level 1 visuo-spatial perspective taking would also occur spontaneously in response to social agent cues.

The Frith–Happé animations

The Frith–Happé animations present short scenarios with two triangles, a small (blue) and a big (red) one, in three different conditions: moving randomly (R), moving in a simple goal-directed fashion (GD) and moving in complex interaction sequences (theory of mind condition, ToM). An example of movements in the R condition would be that the two triangles are floating around from side to side without interacting with each other. In the GD condition the triangles would be engaged in simple scenarios, for instance fighting or dancing, when the movements of the triangles could be interpreted without referring to mental states. This was not the case in the ToM animations, which followed scripts of social interactions. For instance, in the animation 'mocking' the small triangle mimics the big triangle, but stops as soon as the big triangle turns round. In this way the number of cues to social agent behaviour increases from the R to GD to ToM animations. Examples of some of the stimuli can be found at <http://sites.google.com/site/utafrih/research>.

The validity of this classification of the animations has been established with different methods and in various studies. Castelli *et al.* (2000) showed higher activation in brain areas associated with mentalizing for ToM compared with R and GD animations (medial prefrontal cortex, temporo-parietal junction and basal temporal regions), but to a lesser extent in individuals with autism compared with healthy volunteers (Castelli *et al.*, 2002; Kana *et al.*, 2009). In terms of verbal reports, the three types of animations differ systematically in the intentional and mental state language used to describe them by children and adults (Abell *et al.*, 2000; Castelli *et al.*, 2000; Klein *et al.*, 2009) and, whilst using such language in their descriptions, autistic individuals typically do so less discriminately and less appropriately (Abell *et al.*, 2000; Castelli *et al.*, 2002). In terms of eye movements, Klein and colleagues showed that neurotypical adults display an increase in fixation durations from R to GD to ToM animations. This was interpreted as reflecting the greater complexity of information processing involved in watching stories that have more complex scenarios and are also more likely to elicit mental state attribution.

EXPERIMENT 1

In Experiment 1 we examined whether individuals with high functioning autism/AS would show the same pattern of changes in fixation durations and locations across the different types of animation as neurotypical adults. We hypothesised that the early building blocks of agent perception are intact in autism and only later stages of mentalizing are affected. Hence we predicted that autistic adults would show the same eye gaze patterns, but would give different verbal descriptions of the underlying scenarios. As in the study by Klein *et al.* (2009) we measured fixation duration, which was previously shown to differ between GD and ToM compared with R animations. Furthermore, as in a study by Zwicker

and Müller (2009), we evaluated how long gaze stayed on either triangle rather than on background features and how long it stayed on one triangle relative to the other. On the basis of this study we expected that gaze would be directed to the triangles for longer period and be distributed more equally between both the triangles; the more the interaction between the triangles determined the story.

Method

Participants

The study obtained ethical approval from the Joint UCL/UCLH committees on the Ethics of Human Research and informed consent to participate was obtained from each adult. Nineteen participants (mean age: 37 years; mean verbal IQ: 117, Wechsler Adult Intelligence Scale, WAIS III-UK; Wechsler, 1997) were diagnosed with AS. By using this term we do not wish to imply that our results apply exclusively to individuals with AS and not to those with high functioning autism given that differences between clinical categories rest solely on the absence of early language and cognitive delay, but not on current signs and symptoms. There were also 18 neurotypical participants matched for age and general ability (mean age: 39 years; mean verbal IQ: 115). All participants were recruited from the autism database at the Institute of Cognitive Neuroscience, UCL. The participants in the AS group had all received a prior diagnosis from a qualified clinician and all but six also met criteria for an autism-spectrum disorder (ASD)/autism on the Autism Diagnostic Observation Schedule [(ADOS-G); Lord *et al.*, 2000; autism group ADOS-total mean: 8]. The six remaining participants were not excluded as they all had high IQs and their social and communication difficulties were felt to be more obviously evident in their daily lives; furthermore five of them had an Autism-Spectrum Quotient score [(AQ); Baron-Cohen *et al.*, 2001] above the recommended cut-off of 32. As expected, AQ scores for the AS group were significantly greater than for the control group [autism mean: 35; control mean: 17; $t(35) = 7.34$, $P < 0.001$].

Apparatus and stimuli

Eyetracking was enabled using a Tobii (Tobii, Stockholm) T120 Eyetracker, integrated with a 17-inch TFT monitor. Participants were seated on a chair ~50 cm away from the screen. Four R, four GD and four ToM Frith–Happé animations (Abell *et al.*, 2000) were shortened to ~18 s (Klein *et al.*, 2009) and served as the stimuli. All of the animations involved two triangles: a large red one (~3° × 4.5° in width and height) and a smaller blue one (~2° × 2.5° in width and height).

Design and procedure

We asked participants to watch the animations and report after each animation what they had seen. Each animation was presented once in a different randomised order for each participant. After each presentation, participants

reported verbally what they had seen in the prior animation. These reports were written down by the experimenter and coded later. The experimenter only encouraged verbal statements in general but did not give any content-related feedback. Three additional practice animations, one for each stimulus category, were presented before the test trials to familiarize the participants to the procedure.

Data analysis

Verbal reports. Participants' reports were coded by two independent raters (one blind to group, the other D.C.) according to criteria similar to those described in Castelli *et al.* (2000), that is, they were rated for intentionality and appropriateness of expression. A value of 0 on the intentionality scale would indicate that the report contained no intentionality descriptions, e.g., 'the triangles were floating around'. In contrast, a value of 5 would mean that the descriptions contained expressions of high intentionality, e.g., 'the blue triangle wanted to surprise the red one'. It is important to note that the intentionality score provides no measure of the correctness of the description. However, a value of 0 for appropriateness would mean that the descriptions did not fit with the displayed interaction while a value of 5 would indicate that the interaction was adequately described.¹ Inter-rater correlation was satisfactory (0.89) and the two raters' scores were averaged.

Eye movement measures. Fixation was defined as a continuous gaze within a 35 pixel radius (~1.5°). Tobii Studio Software (Tobii, Stockholm) was used to calculate the co-ordinate and duration of each fixation; the *fixation duration*. *Triangle time* was calculated as the time when eye gaze fell within a circle of 3° around the centre (of mass) of either the blue or the red triangle divided by the total length of the animation. This measure thus gives an indication of the importance attributed to the triangles rather than to the background. *Relative time* was used to measure how relevant each triangle was for story understanding. It was defined as the time when gaze was within a circle of 3° around the centre of the blue triangle, divided by the time gaze was within a circle of 3° around the centre of either triangle. A value of 50% would indicate that gaze was divided between both triangles equally, while larger values show a preference for the blue triangle and vice versa.

For each measure, a two-way analysis of variance (ANOVA) with the within subject factor 'animation condition' (R, GD, ToM) and the between subject factor 'group' (AS, control) was calculated to see how the film content influenced gaze parameters and whether the groups differed in these parameters. Significant interactions were followed

¹One difference to the scoring in Castelli *et al.* (2000) should be noted: the range for appropriateness was expanded to a five point scale to mirror the intentionality scale. The five points of the scale were redefined as: 0 = missed the point completely or does not know what has happened; 1 = only a small fraction of the sequence is correct; 2 = partial description of the sequence; 3 = mostly correct, but missing something; 4 = correct, but either imprecise or too much extraneous detail; 5 = correct.

up by t-tests between groups for each animation condition separately. Greenhouse–Geisser corrections were used when necessary but, to reduce complexity, only non-corrected degrees of freedom are reported.

Results

One participant's gaze data could not be tracked reliably and were excluded from analyses.

Verbal reports

Table 1 shows intentionality ratings increased from R to GD to ToM for both groups [$F(2,68) = 253.39$, $MSE = 0.21$, $P < 0.05$]. Group alone had no significant influence [$F(1,34) < 1$], however, a significant interaction between group and animation condition was observed [$F(2,68) = 4.48$, $MSE = 0.21$, $P < 0.05$]. From Table 1 it can be seen that this interaction was due to the participants in the AS group over-attributing intentions in the R animations and under-attributing in the GD and ToM animations compared with controls; this was only significant in the R condition [$t(34) = 2.23$, $P < 0.05$] although there was also a trend towards a group difference for GD [$t(34) = -2.02$, $P < 0.10$] but no difference for ToM [$t(34) = -1.07$, $P > 0.10$].² Appropriateness ratings decreased from R to GD to ToM across the whole sample [$F(2,68) = 70.29$, $MSE = 0.63$, $P < 0.05$] and there was also a significant influence of group with controls showing higher ratings [$F(1,34) = 13.27$, $MSE = 0.76$, $P < 0.05$] but no significant interaction [$F(2,68) < 1$].

Eye gaze measures

For both groups, fixation duration increased from R to GD to ToM (Table 2), which was reflected in a significant main effect of animation condition [$F(2,68) = 76.40$, $MSE = 2840$, $P < 0.05$] but no significant effect of group or interaction between the two [$F(1,34) = 1.23$, $MSE = 22852$, $P > 0.10$; $F(2,68) < 1$]. Triangle time increased similarly in both groups with R being lowest, ToM highest and GD in between [$F(2,68) = 166.17$, $MSE = 74.07$, $P < 0.05$]. Group had no significant influence, nor did it interact with animation condition [$F(1,34) < 1$; $F(2,68) = 1.58$, $MSE = 74.07$, $P > 0.10$]. Finally, relative time showed a decrease from R to GD to ToM [$F(2,68) = 34.96$, $MSE = 37.22$, $P < 0.05$] and was not affected by group, nor did the group show a main effect (all F 's < 1). As can be seen from the 95% confidence intervals in Table 2, only the R and GD but not the ToM condition differed significantly from 50%.

²We note that the control group had very low intentionality and appropriateness scores for the ToM animations, in fact as low as the AS group. Compared with other studies we know that such low performance in a control group is highly unusual (Castelli *et al.*, 2000; Klein *et al.*, 2009). Whatever the reason, it is important to note that the lack of group difference here does not denote unusually high performance on the ToM animations by the AS group; their scores were comparable with those found in other studies (Castelli *et al.*, 2002).

Table 1 Experiment 1—verbal descriptions of the animations coded for intentionality and appropriateness

Animation condition	Intentionality		Appropriateness	
	AS	Control	AS	Control
R	0.46 (0.16–0.76)	0.12 (0.01–0.23)	3.89 (3.35–4.43)	4.66 (4.43–4.90)
GD	1.99 (1.79–2.19)	2.23 (2.09–2.37)	3.48 (3.10–3.86)	4.22 (3.96–4.48)
ToM	2.51 (2.22–2.81)	2.73 (2.42–3.03)	2.01 (1.61–2.40)	2.33 (1.82–2.84)

Mean ratings of intentionality (0–5) and appropriateness (0–5) are reported for each group (AS, control) across the three animation conditions (R, GD, ToM). Confidence intervals of 95% are given in brackets.

Table 2 Experiment 1—means and 95% confidence intervals for eye gaze measures of fixation duration, triangle time and relative time for each group (AS, control) and animation condition (R, GD, ToM)

Animation condition	AS	Control
Fixation duration (ms)		
R	311 (277–345)	332 (286–378)
GD	333 (297–369)	360 (312–408)
ToM	428 (377–479)	477 (410–544)
Triangle time (%)		
R	56.23 (50.22–62.24)	52.82 (45.08–60.56)
GD	66.35 (62.32–70.37)	64.29 (59.87–68.70)
ToM	84.69 (82.31–87.06)	86.25 (83.93–88.57)
Relative time (%)		
R	61.41 (58.48–64.34)	62.37 (57.87–66.87)
GD	54.79 (50.84–58.74)	55.03 (52.19–57.87)
ToM	51.57 (48.76–54.38)	50.71 (48.89–52.54)

Discussion

As in previous studies, high functioning participants with autism differed from control participants in their verbal descriptions of the animations: they attributed higher intentionality to the triangles in the R animations and their reports were generally less appropriate. This fits with the observation of Abell *et al.* (2000) suggesting that the main impairment lies not in a lack of mental state language, which our able participants had mastered, but in providing fitting descriptions. We speculate that our high functioning participants concluded from the experimental setting that mental state language was expected and hence used such language even more than our controls, but also more indiscriminately, extending such language also to the R animations. In support of these mentalizing difficulties, a previous study involving the same participants revealed a lack of spontaneous prediction of another's mental states by those with AS (Senju *et al.*, 2009).

In contrast to mentalizing, none of our eye movement measures differed between the groups. Both groups showed the same increase in mean fixation duration, an indicator of information integration (Klein *et al.*, 2009). Triangle time, an indicator of how much importance was attributed to

the actors, was also similar in both groups; the more the story depended on the actions of the triangles instead of on physical laws, the more time was spent watching the triangles rather than other parts of the display. There was also no difference between the groups in the extent to which they attributed importance to the individual triangles, as reflected in relative time spent on each triangle. In the R and GD animation conditions, where information from one actor was sufficient for story understanding while the other triangle could be followed in the periphery, both groups preferred to focus on the smaller triangle. In contrast, during ToM animations, where the actions of both triangles were equally important and had to be integrated to follow the interaction, the red and blue triangles were fixated for approximately the same amount of time by both groups.

Together, these results suggest that it is one thing to track moving social agents and to understand that an interaction occurs between them and quite another thing to attribute mental states to social agents. Specifically, the AS group could derive from perceptual cues that the triangles were interacting with each other without necessarily understanding the nature of the interaction. This shows that the groups did not differ in *what* visual input they processed but on *how* that input was processed. However, where exactly is the boundary between perceiving social agents and perceiving their mental states? This question was explored in the next experiment.

EXPERIMENT 2

Would individuals with autism/AS also be compelled to spontaneously take into account the visuo-spatial perspective of a social agent in those movies where strong social agent cues are perceived? We investigated this by means of a task developed in a previous study by Zwickel (2009) with neurotypical adults. In this study, the Frith–Happé animations were modified in such a way that at certain points in time a dot appeared next to the red triangle. Neurotypical adults were asked to press either a right or left button to indicate on which side of the screen the dot appeared relative to the triangle. With this method two conditions can be contrasted: the congruent perspective where the dot is on the same side from the viewpoint of both the triangle and the observer and likewise an incongruent perspective where the triangle's 'nose' is turned downwards and the viewpoint of the triangle and observer therefore differ (Figure 1).

The impression that the triangle has a 'nose' and that it consequently has its own left and right side presupposes the perception of the triangle as a social agent. In Zwickel's experiment, observers were slowed down when there was perspective incongruency, but only in those animations that gave strong cues to social agent behaviour. This experiment led to the proposal that spontaneous coding of visual events relative to the observed orientation of an object is closely bound to agency perception as well as to mental state attribution. Similar effects of spontaneous coding of the

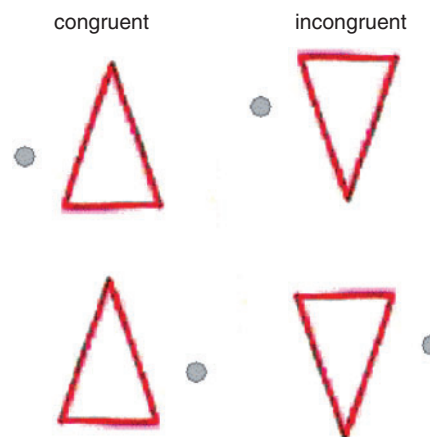


Fig. 1 Illustration of congruent (decisions made from the perspectives of the participant/heading direction of triangle are the same) and incongruent conditions in Experiment 2.

perspective of another merely observed social agent have also been reported recently by a number of other authors (Thomas *et al.*, 2006; Belopolsky *et al.*, 2008; Frischen *et al.*, 2009; Tversky and Hard, 2009; Samson *et al.*, 2010). All these findings provide strong evidence for the proposal that taking another social agent's visuo-spatial perspective into account is a basic spontaneous process. Whether or not this is another basic building block of social cognition that occurs independently of mentalizing remains an open question. Autism however provides a possible way to explore this question. If basic visuo-spatial perspective taking is independent of mentalizing, then we would expect no group difference. However, if it is intrinsic to automatic mentalizing ability, participants with autism should not engage in spontaneous visuo-spatial perspective taking while watching moving social agents.

Method

Participants, stimuli and apparatus

The same participants as in Experiment 1 watched three R (billiard, drifting, tennis) and three ToM (coaxing, mocking surprising) animations.³ These animations were edited to each include six brief (30 ms) presentations of a 0.5° grey dot. These dots occurred 2° to the right or left of the red triangle's centre. In addition, every trial was repeated with right and left presentation reversed. During each film, three of the dots occurred while the triangle was pointing downwards (incongruent) and three while the triangle was pointing upwards (congruent). Each dot presentation was separated by at least 1.5 s. Side of dot presentation and congruency condition was pseudo-randomly controlled and occurred with an equal frequency within the animations.

³Only three of each of the R and ToM movies were found suitable for this experiment as, in these animations, the triangles moved up as well as down frequently enough for both congruent and incongruent conditions to occur.

Design

Animations from the two animation conditions (R and ToM) were presented as two separate pseudo-random blocks of 30 animations, resulting in 10 presentations of each animation in total. The sequence of blocks was balanced across participants.

Procedure

Before each experiment three training trials were run that were identical to the experimental trials except that different animations were used. Participants were instructed to respond with a key press to the dots appearing either to the left or right of the red triangle as quickly as possible. Participants were explicitly instructed that the decision should be made from their perspective. In addition, participants were required to watch the animations attentively to report on the stories after each block. These reports were only used to ensure that participants paid attention to the animation story and were not further analyzed.

Data analysis

Reaction Time (RT) was defined as the interval between the dot presentation and the key press. To ensure that we included only trials in which participants paid attention, the following exclusion criteria were applied: RTs greater than 1500 ms (*no response*); incorrect responses (*wrong responses*); and all responses more than 2 s.d. from the mean of the participant (*unfocused responses*). The remaining RTs were averaged for each participant in each animation and congruency condition separately. A measure of the *congruency effect* was calculated by subtracting the means in the congruent from the means in the incongruent condition. Because a Levene-test suggested that the variance of the congruency effect might be higher in the autism group than the control group in the ToM animation condition these values were first square-root transformed before they were entered as the dependent variable in a 2×2 ANOVA (animation condition \times group). However, the outcome did not differ relative to a significance level of 5% when the raw values were used.

Results

Mean exclusion rates for no, wrong and unfocused responses were 12.33, 9.71 and 3.29% in the AS group and 7.62, 3.21 and 3.60% in the control group. The groups only differed on the wrong responses, with the AS group performing more poorly [no response: $t(35) = 1.59$, $P > 0.10$; wrong response: $t(35) = 2.70$, $P < 0.05$; unfocused response: $t(35) = 0.77$, $P > 0.10$]. Table 3 shows the RTs for each condition and group. In both groups, the difference in RTs was higher in the ToM than R animation condition. Importantly, the size of the congruency effect was similar in both groups. This observation was corroborated by a significant main effect of animation condition [$F(1,35) = 35.29$, $MSE = 11.74$, $P < 0.05$], but no statistical influence of group or interaction

between the two [$F(1,35) = 1.45$, $MSE = 16.76$, $P > 0.10$; $F(1,35) < 1$]. The error rates reflected this pattern of greater difficulty in the incongruent ToM condition, excluding the possibility of a speed-accuracy trade-off. In both groups, the congruency effect was only significant in the ToM condition [95% confidence intervals of Figure 2; autism group: R $t(18) = 1.07$, $P > 0.10$, ToM $t(18) = 4.70$, $P < 0.05$; control group: R $t(17) = 0.37$, $P > 0.10$, ToM: $t(17) = 7.37$, $P < 0.05$].

Discussion

A previous experiment (Zwikel, 2009) found that neurotypical individuals show spontaneous encoding of visual events relative to a social agent's perspective when watching the ToM animations in contrast to the R animations. This finding left open whether this spontaneous coding is part and parcel of mental state attribution or whether it is merely a function of agency attribution, which may be stronger for ToM than for R animations because the former has more social agent cues. The inclusion of individuals with AS here

Table 3 Experiment 2—RT (ms) and errors (%) for congruent and incongruent conditions

Animation condition	Incongruent		Congruent	
	AS	Control	AS	Control
Reaction times (ms)				
R	577 (521–634)	518 (482–555)	566 (506–625)	520 (481–559)
ToM	601 (537–664)	548 (510–586)	559 (502–616)	522 (484–559)
Error rates (%)				
R	15.95 (6.02–25.87)	3.83 (2.43–5.23)	10.47 (4.15–16.80)	3.78 (2.16–5.39)
ToM	11.00 (4.43–17.57)	3.44 (2.02–4.87)	4.21 (1.67–6.75)	1.89 (0.18–3.59)

Means for each group (AS, control) and animation condition (R, ToM) are provided together with their 95% confidence intervals.

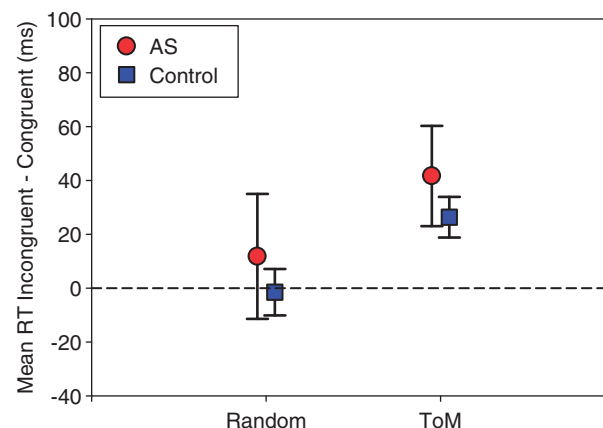


Fig. 2 Difference in Experiment 2 between incongruent and congruent conditions (ms) for random and ToM animations for each group separately. Whiskers indicate 95% confidence intervals.

allowed us to clarify this question. While our first experiment concluded that autistic adults were able to spontaneously detect interactions between social agents and show similar eye movements to control participants, it was still unclear whether detecting a social agent would be linked to further social processing such as visuo-spatial perspective taking.

The results from our second experiment were clear cut: individuals with autism and control participants displayed the same pattern of results. Both groups showed a larger difference in RTs between incongruent and congruent conditions in the ToM compared with the R animations. Therefore, processing of ToM scenarios led to stronger perspective adoption in both groups, presumably because the ToM scenarios contained more cues suggestive of a social agent role of the triangles. Thus, the co-representation of a social agent's perspective in dynamic events seems to occur only with stimuli that are strongly perceived as social agents and this occurs in neurotypical as well as individuals with autism. This indicates that the processes involved in detecting a social agent are intact up to and including level 1 visuo-spatial perspective taking in autistic individuals. While these same cues also trigger spontaneous mentalizing in neurotypical individuals, this is not the case in autistic individuals.

Our findings are reminiscent of another set of data relating to the spontaneous co-representation of another's task (Sebanz *et al.*, 2005). The current study supports and extends these findings. Thus there is not only spontaneous co-representation of another's task but also co-representation of another's viewpoint in autism. Moreover, both studies suggest that these social capacities can be found intact in the absence of spontaneous mentalizing. We strongly agree with Sebanz and colleagues that people with high functioning autism/AS spontaneously take the presence of another person into account. Thus people with autism possess some basic building blocks of social cognition and are neither other-blind nor viewpoint-blind.

GENERAL DISCUSSION

There is a stark contrast between the competency individuals with high functioning autism/AS display in basic social processes, for example, between co-representing another's task (Sebanz *et al.*, 2005) and another's mental state, the former being intact and the latter impaired. With the current experiments we sought to further investigate the building blocks of social interaction that are spared in autistic individuals. In the first experiment we tested whether individuals with and without autism would show similar eye movements in response to various films. We found there were clear differences in eye movement parameters between movies that varied in the number of social agent cues. Importantly, differences in response to the different movie types were similar in both groups, showing that autistic individuals formed similar representations of different degrees of socially

intentional behaviour, as revealed in their eye movements. This is consistent with the findings of New *et al.* (2009) that high functioning individuals with and without autism distinguish along the animate/inanimate dimension. This competency in distinguishing moving agents from moving objects is in marked contrast to their poor performance when asked to explicitly describe the content of the movies.

Having shown that at least a rudimentary form of social agent detection was present in the AS group, in Experiment 2 we asked whether another building block, possibly a further step towards mentalizing—level 1 visuo-spatial perspective taking—was present. Autistic individuals and controls both spontaneously adopted the visuo-spatial perspective of an observed social agent and the more so the stronger the agency cues. Thus, individuals with autism are not only able to do this when explicitly asked (David *et al.*, 2010), but they do so spontaneously when they detect social agent cues. It should be noted that there is a big divide between level 1 and level 2 visual perspective taking. Level 2 visual perspective taking, which is assessed by tasks where the observer explicitly needs to take a meta-cognitive stance 'while I am seeing side A of an object, at the same time the other person is seeing side B', is closely related to mentalizing (Hamilton *et al.*, 2009).

The fact that these abilities were tested with indirect measures ensures that agency processing was truly spontaneous and this makes it extremely unlikely that participants with autism made their responses via different and possibly compensatory processes. Furthermore, the lack of difference between the two groups is not simply a null effect as both groups showed significant differences between the animation conditions, reflecting the ability of both groups of participants to distinguish between films with and without social agents.

Our experiments concerned online perception of dynamic stimuli that triggered perspective taking and, in the case of the neurotypical group, mentalizing simply by their movement. The performance of autistic individuals clearly indicates that the processes underlying mentalizing and level 1 visuo-spatial perspective taking are not one and the same; instead they are likely to require quite different neuro-cognitive processes. This modular interpretation is in line with the recent findings of Bedny *et al.* (2009). At the very least, we can conclude that it is not the requirement for on-line tracking of social agents' perspectives that makes mentalizing hard for people with autism.

At the moment we can only speculate why level 1 visuo-spatial perspective taking is separate from mentalizing. Considering human evolution, it is likely that being able to detect whether an animate being is present and decide whether one is seen by this potential predator was a fundamental survival skill long before attributing mental states to them became relevant. Attributing mental states to a predator most likely became of relevance only later in evolution after predators acquired relatively complex mental states.

One can take this line of thought even further: to distinguish between a hostile and a friendly conspecific, it might be more effective to judge whether a spear is held in the left (friend) or right (foe) hand of the opponent than to engage in a more elaborate mental state analysis. Therefore, detecting a social agent and representing the visuo-spatial perspective of a social agent might both be successful basic survival strategies independent of mental state attribution. Which specific cues might trigger visuo-spatial perspective adoption will be the subject of future studies.

Conflict of Interest

None declared.

REFERENCES

- Abell, F., Happé, F., Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development*, 15(1), 1–16.
- Apperly, I.A. (2008). Beyond simulation-theory and theory-theory: why social cognitive neuroscience should use its own concepts to study “theory of mind”. *Cognition*, 107(1), 266–83.
- 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(1), 5–17.
- Bedny, M., Pascual-Leone, A., Saxe, R.R. (2009). Growing up blind does not change the neural bases of Theory of Mind. *Proceedings of the National Academy of Science USA*, 106(27), 11312–17.
- Belopolsky, A.V., Olivers, C.N.L., Theeuwes, J. (2008). To point a finger: attentional and motor consequences of observing pointing movements. *Acta Psychologica*, 128(1), 56–62.
- Castelli, F., Frith, C., Happe, F., Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, 125(Pt 8), 1839–49.
- Castelli, F., Happé, F., Frith, U., Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage*, 12(3), 314–25.
- Celani, G. (2002). Human beings, animals and inanimate objects: what do people with autism like? *Autism*, 6(1), 93–102.
- David, N., Aumann, C., Bewernick, B.H., Santos, N.S., Lehnhardt, F.G., Vogeley, K. (2010). Investigation of mentalizing and visuospatial perspective taking for self and other in asperger syndrome. *Journal of Autism and Developmental Disorders*, 40(3), 290–99.
- Fletcher-Watson, S., Leekam, S.R., Benson, V., Frank, M.C., Findlay, J.M. (2009). Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia*, 47(1), 248–57.
- Fletcher-Watson, S., Leekam, S.R., Findlay, J.M., Stanton, E.C. (2008). Brief report: young adults with autism spectrum disorder show normal attention to eye-gaze information-evidence from a new change blindness paradigm. *Journal of Autism and Developmental Disorders*, 38(9), 1785–90.
- Freeth, M., Chapman, P., Ropar, D., Mitchell, P. (2009). 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(5), 534–47.
- Freeth, M., Ropar, D., Chapman, P., Mitchell, P. (2010). The eye gaze direction of an observed person can bias perception, memory, and attention in adolescents with and without autism spectrum disorder. *Journal of Experimental Child Psychology*, 105(1–2), 20–37.
- Frischen, A., Loach, D., Tipper, S.P. (2009). Seeing the world through another person’s eyes: simulating selective attention via action observation. *Cognition*, 111(2), 212–18.
- Frith, U., Frith, C. (2010). The social brain: allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1537), 165–76.
- Hamilton, A.F.deC., Brindley, R., Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition*, 113(1), 37–44.
- Heider, F., Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American Journal of Psychology*, 57(2), 243–59.
- Hobson, R.P. (1984). Early childhood autism and the question of egocentrism. *Journal of Autism and Developmental Disorders*, 14(1), 85–104.
- Kana, R.K., Keller, T.A., Cherkassky, V.L., Minshew, N.J., Just, M.A. (2009). Atypical frontal-posterior synchronization of Theory of Mind regions in autism during mental state attribution. *Social Neuroscience*, 4, 135–52.
- Kemner, C., van Engeland, H. (2003). Autism and visual fixation. *American Journal of Psychiatry*, 160(7), 1358–59, author reply 1359.
- Klein, A., Zwickel, J., Prinz, W., Frith, U. (2009). Animated triangles: An eye tracking investigation. *Quarterly Journal of Experimental Psychology*, 62(6), 1189–97.
- Klin, A., Jones, W. (2006). Attributing social and physical meaning to ambiguous visual displays in individuals with higher-functioning autism spectrum disorders. *Brain and Cognition*, 61(1), 40–53.
- 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(9), 809–16.
- Lord, C., Risi, S., Lambrecht, L., Cook, E.H.Jr, Leventhal, B.L., DiLavore, P.C., et al. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3), 205–23.
- New, J.J., Schultz, R.T., Wolf, J., Niehaus, J.L., Klin, A., German, T.C., et al. (2009). The scope of social attention deficits in autism: prioritized orienting to people and animals in static natural scenes. *Neuropsychologia*, 48(1), 51–59.
- Samson, D., Apperly, I.A., Braithwaite, J.J., Andrews, B.J., Bodley Scott, S.E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, 36(5), 1255–66.
- Sebanz, N., Knoblich, G., Prinz, W. (2003). Representing others’ actions: just like one’s own? *Cognition*, 88(3), B11–21.
- Sebanz, N., Knoblich, G., Stumpf, L., Prinz, W. (2005). Far from action blind: Representation of others’ actions in individuals with autism. *Cognitive Neuropsychology*, 22(3/4), 433–54.
- Senju, A., Southgate, V., White, S., Frith, U. (2009). Mindblind eyes: an absence of spontaneous theory of mind in Asperger syndrome. *Science*, 325(5942), 883–5.
- Thomas, R., Press, C., Haggard, P. (2006). Shared representations in body perception. *Acta Psychologica*, 121(3), 317–30.
- Tversky, B., Hard, B. M. (2009). Embodied and disembodied cognition: spatial perspective-taking. *Cognition*, 110(1), 124–9.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale (WAIS) (3rd UK ed.)*. London: Psychological Corporation.
- Zwickel, J. (2009). Agency attribution and visuo-spatial perspective taking. *Psychonomic Bulletin & Review*, 16, 1089–93.
- Zwickel, J., Müller, H. J. (2009). Eye movements as a means to evaluate and improve robots. *International Journal of Social Robotics*, 1(4), 357–66.