Joint attention difficulties in autistic adults: An interactive eye-tracking study

Nathan Caruana[1,2,3], Heidi Stieglitz Ham[4], Jon Brock[1,2,5,6], Alexandra Woolgar[1,2,3], Nadine Kloth[2,7], Romina Palermo[2,7] and Genevieve McArthur[1,2,6]

Abstract

Joint attention – the ability to coordinate attention with a social partner – is critical for social communication, learning and the regulation of interpersonal relationships. Infants and young children with autism demonstrate impairments in both initiating and responding to joint attention bids in naturalistic settings. However, little is known about joint attention abilities in adults with autism. Here, we tested 17 autistic adults and 17 age- and nonverbal intelligence quotient–matched controls using an interactive eye-tracking paradigm in which participants initiated and responded to joint attention bids with an on-screen avatar. Compared to control participants, autistic adults completed fewer trials successfully. They were also slower to respond to joint attention bids in the first block of testing but performed as well as controls in the second block. There were no group differences in responding to spatial cues on a non-social task with similar attention and oculomotor demands. These experimental results were mirrored in the subjective reports given by participants, with some commenting that they initially found it challenging to communicate using eye gaze, but were able to develop strategies that allowed them to achieve joint attention. Our study indicates that for many autistic individuals, subtle difficulties using eye-gaze information persist well into adulthood.

Joint attention is the ability to achieve a common focus of attention with another person during a social interaction and is an important precursor to the development of language and social learning (Adamson et al., 2009; Baron-Cohen, 1995; Charman, 2003; Mundy et al., 1990; Murray et al., 2008; Tomasello, 1995). In a joint attention episode, a person initiates a joint attention bid by intentionally guiding another person's attention towards an object or event. Joint attention is achieved when the other person responds by following the instigator's communicative bid (Bruinsma et al., 2004) and usually involves mutual awareness of the shared experience (Emery, 2000).

Eye gaze is typically the first communicative modality that humans develop and use to experience joint attention with others, which is often accompanied in later life by language and pointing gestures (see Pfeiffer et al., 2013 for a review). In typical development, infants begin to use eye gaze to respond to and initiate joint attention bids at approximately 6 months (Bakeman and Adamson, 1984; Scaife and Bruner, 1975) and 12 months of age, respectively (Bates et al., 1979). However, in autistic children, responding to joint attention (RJA) may not begin to emerge until cognitive development is equivalent to that of 30–36 months of typical development (Mundy et al., 1994), and impairments in initiating joint attention (IJA) often persist well into adolescence (e.g. Charman, 2003; Hobson and Hobson, 2007; Mundy et al., 1986; Sigman and Ruskin, 1999). Difficulties in IJA or RJA are reliable predictors of social communication and social interaction (Lord et al., 2000; Stone et al., 1997) as well as significant predictors of future expressive language development in children on the autism spectrum (Charman, 2003; Dawson et al., 2004; Mundy et al., 1990).

[1]Department of Cognitive Science, Macquarie University, Australia
[2]ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Australia
[3]Perception in Action Research Centre, Macquarie University, Australia
[4]School of Occupational Therapy and Social Work, Curtin University, Australia
[5]Department of Psychology, Macquarie University, Australia
[6]Centre for Atypical Neurodevelopment, Macquarie University, Australia
[7]School of Psychology, University of Western Australia, Australia

Corresponding author:
Nathan Caruana, The Australian Hearing Hub, Macquarie University, 16 University Avenue, Sydney, NSW 2109, Australia.
Email: nathan.caruana@mq.edu.au

To date, joint attention impairments in autism have mostly been investigated in observational studies of very young children in natural and semi-structured social interactions (Charman et al., 1997; Dawson et al., 2004; Loveland and Landry, 1986; Mundy et al., 1990; Osterling and Dawson, 1994; Osterling et al., 2002; Wong and Kasari, 2012). However, observational paradigms often lack the sensitivity to detect joint attention difficulties that may affect older children and adults, nor are they amenable to the experimental manipulation necessary to provide insight into the cognitive or neural mechanisms that underlie joint attention impairments.

A separate group of studies have employed variations on the Posner cueing task to investigate the extent to which individuals reflexively orient their attention to gaze cues. These tasks require participants to respond to a target that is preceded by a gaze cue directing them towards the target or in the opposite direction (e.g. Friesen and Kingstone, 1998). The main outcome measure is the time taken to detect the target's location. Such tasks provide a sensitive, standardised experimental manipulation of the mechanisms underlying the reflexive aspects of gaze processing. However, they fail to capture the truly interactive and intentional nature of joint attention. This may partly explain why studies using the Posner cueing paradigms have failed to provide consistent evidence for impairments in gaze processing in autistic children and adults (see Leekam, 2015; Nation and Penny, 2008 for review).

The aim of this study was to use a paradigm that minimised the limitations of observational and the Posner cueing paradigms to better understand the joint attention behaviours of autistic adults. To this end, we employed a new interactive eye-tracking paradigm developed by Caruana et al. (2015) in which participants played a cooperative game with an animated virtual character (avatar) whom they believed to be controlled by another person (cf. Bayliss et al., 2013; Courgeon et al., 2014; Kim and Mundy, 2012; Schilbach et al., 2010). The goal of the game was to catch a burglar who was hiding inside one of six houses displayed on a screen. Each trial began with a search phase in which both the participant and the avatar searched their allotted houses. Whomever found the burglar had to make their partner aware of his or her location by establishing eye contact and then gazing at the location of the burglar. This procedure created a social condition that (1) elicited intentional RJA and IJA behaviours, (2) informed participants of their social role (i.e. responder or initiator) throughout the course of each trial without overt instruction and (3) required participants to use eye contact as a cue to identify joint attention opportunities. The performance of RJA and IJA trials was compared with the performance in non-social conditions that presented the same task demands but did not require any social interaction (RJAc and IJAc).

Using a number of performance metrics, we investigated whether autistic participants performed the responding and initiating tasks as well as control (i.e. non-autistic)

participants and whether any group differences were specific to the social context or were also observed in the non-social control conditions. We also contrasted performance in the first versus the second block of testing to investigate the ability of participants in both groups to learn and adapt to the new task. This analysis aimed to determine whether autistic individuals were able to overcome any initial difficulties they may have had performing the task.

Method

Ethical statement

This study was approved by the Human Research Ethics Committee at Macquarie University (MQ; reference number: 5201200021) and ratified by the University of Western Australia (UWA). Participants received payment or course credit for their time and provided written consent before participating.

Participants

A total of 18 autistic adults were tested at the UWA (Perth, Australia). All adults reported that they had been formally diagnosed with Autism or Asperger syndrome by a clinical psychologist in line with *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; DSM-IV; American Psychiatric Association, 2000) criteria. As such, they would automatically qualify for an autism spectrum disorder under *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; DSM-5; American Psychiatric Association, 2013). Participants also completed the Ritvo Autism Asperger Diagnostic Scale–Revised (RAADS-R; Ritvo et al., 2011). This is a self-reporting diagnostic measure that we used to provide a uniform diagnostic assessment. All but one participant (score = 60) scored above the diagnostic threshold of 65 on the RAADS-R. This participant was included in the final analyses as they scored close to threshold. The pattern of effects did not change if this participant was excluded.

Nonverbal intelligence quotient (IQ) was assessed using the matrices subtest of the Kaufman Brief Intelligence Test–Second Edition (KBIT-2; Kaufman and Kaufman, 2004). One participant was excluded because their nonverbal IQ score was below 85. This resulted in a final sample of 17 autistic adults (6 females). Their performance was compared to 17 control participants (6 females) with typical development who were tested at MQ (Sydney, Australia). The two groups were matched on gender, age and nonverbal IQ. No control participant scored above threshold on the RAADS-R. Relative to the control group, autistic participants scored higher on the Autism Quotient (AQ; Baron-Cohen, 2008) and the Empathising Quotient (EQ; Wheelright et al., 2006), but not the Systemising Quotient (SQ; Wheelright et al., 2006) questionnaires. Demographic and questionnaire data for each group are shown in

Table 1. Participant details.

	Autism group		Control group		Statistics	
	M	SD	M	SD		
Age	26.47	11.86	26.43	14.53	t(32) = 0.01	p = 0.993
Nonverbal IQ	105.94	13.45	105.70	12.46	t(32) = 0.05	p = 0.958
RAADS-R	126.44	25.47	53.06	25.06	t(31) = 9.87	p < 0.001
AQ	27.81	11.08	9.24	6.31	t(31) = 5.96	p < 0.001
EQ	27.31	10.93	55.35	10.69	t(31) = −7.45	p < 0.001
SQ	66.44	25.47	53.05	25.06	t(31) = 1.52	p = 0.139

*M*: mean; *SD*: standard deviation; IQ: intelligence quotient.
Nonverbal IQ scores were based on the standard score obtained using the KBIT-2 Matrices subtest (Kaufman and Kaufman, 2004). Total raw scores are reported for the Ritvo Autism Asperger Diagnostic Scale–Revised (RAADS-R; Ritvo et al., 2011), Autism Quotient (AQ; Baron-Cohen, 2008), Empathising Quotient (EQ), and Systemising Quotient (SQ; Wheelright et al., 2006).

Table 1. All participants had normal vision and reported no known history of acquired neurological impairment or injury.

Joint attention task

Social conditions (RJA and IJA). In the social conditions, participants played a cooperative game with an avatar that they believed represented the gaze behaviour of another person named 'Alan' who was in a nearby eye-tracking laboratory. Alan was represented by a face, generated using *FaceGen* (Singular Inversions, 2008), that subtended 6.5° of visual angle in the centre of the screen. His eyes could be directed either at the participant or towards one of the six houses that were presented on the screen (see Figures 1 and 2). The houses, which each subtended 4° of visual angle, were arranged in two horizontal rows above and below the avatar. Participants were told that Alan could control the avatar's gaze using live-infrared eye-tracking over a high-speed network. In reality, a gaze-contingent algorithm used the online recordings of the participant's eye movements to program the avatar's responsive behaviour (see Caruana et al., 2015, for a description of this algorithm and a video depicting example trials from each condition).

Search phase. Each trial began with a search phase. During this period, the participant and the avatar (i.e. Alan) were required to search for the burglar. The participant searched houses with blue doors (e.g. bottom row in Figure 1) while the avatar searched houses with red doors (e.g. top row in Figure 1). Each time the participant fixated upon a blue door, it opened to reveal either an empty house or the burglar (Figure 2, first column). However, from the participant's perspective, Alan's doors remained closed as he completed his search. Participants were able to search their houses in any order they chose. On some trials, one or two blue doors were already open at the start of the trial, revealing an empty house. This ensured a different pattern of gaze behaviour on each trial made by the participant and
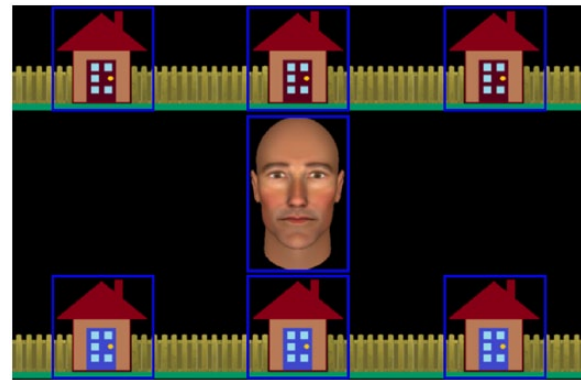
Figure 1. Experimental display showing the central avatar ('Alan') and the six houses in which the burglar could be hiding. Gaze-related areas of interest (GAOIs), represented by blue rectangles, were not visible to participants.

provided a context for non-repetitive patterns of gaze behaviour made by the avatar.

RJA. On RJA trials, the participant opened all the blue doors to find them empty (Figure 2, row 1) and could thus infer that the burglar was hiding in one of Alan's houses. Once the participant fixated back on the avatar's face, he searched 0–2 more houses before establishing mutual gaze with the participant. This ensured that, for a brief interval, participants were required to monitor the avatar's gaze behaviour to determine whether Alan was ready to initiate a joint attention bid. We randomised the location of the house that the avatar searched last to ensure that it was not predictive of the burglar's location. Provided that the participant was still looking at the avatar when the avatar returned eye contact, the avatar then directed his gaze to the burglar's location. The participant was required to follow the avatar's gaze and fixate on the same location. We refer to this eye movement as a 'responding saccade'.

IJA. On IJA trials, the participant found the burglar behind one of the blue doors (Figure 2, row 3). The relevant blue

Figure 2. Schematic representation of trial sequence by condition. The eye symbol represents the fixation required by the participant and was not visible to the participant. Analysis periods for each eye-tracking analysis are indicated in red cells.

door 'closed' to conceal the burglar once the participant fixated back on the avatar's face. Again, the avatar searched 0–2 more houses before making eye contact with the participant. Once eye contact was established, participants were required to initiate joint attention by fixating on the blue door that concealed the burglar. We refer to this eye movement as an 'initiating saccade'. The avatar's gaze was programmed so that it always followed the participant's initiating saccade even if the participant fixated on the incorrect house. The avatar only responded to a participant's initiating saccade after eye contact was established.

Feedback. For 'correct' RJA and IJA trials, the burglar appeared behind bars to indicate that the participant and Alan had succeeded in achieving joint attention to capture the burglar (e.g. Figure 2, seventh column). On 'incorrect' trials, the burglar appeared in red at its true location to provide negative feedback. This occurred if participants (1) spent more than 3 s fixated on the background (i.e. away from the houses or the avatar stimulus), (2) took longer than 3 s to execute a responding or initiating saccade after being guided (RJA trials) or establishing eye contact (IJA trials) or (3) made a responding or initiating saccade to an incorrect location. If participants took longer than 3 s to begin searching their houses at the beginning of the trial,

red text reading 'Failed Search' appeared on the screen and the trial was terminated.

Non-social conditions (RJAc and IJAc). We developed two non-social conditions to control for the non-social task demands involved when responding to (RJAc; Figure 2, second row) and initiating (IJAc; Figure 2, fourth row) joint attention bids in the social conditions (i.e. to control for task complexity, attentional load and number of eye movements required). The differences between the non-social and social conditions were that (1) the avatar's eyes remained closed throughout, (2) a grey fixation point was presented over the avatar's nose until the participant completed their search and fixated upon it, (3) the fixation point turned green when fixated (analogous to the avatar making eye contact) and (4) on RJAc trials, a green arrow subtending 3° of visual angle cued the burglar's location (analogous to the gaze cue on RJA trials).

Procedure

To ensure the testing environments of the two sites (UWA and MQ) were matched as closely as possible, we ensured that the testing rooms were similar in size, had no windows and that the experimenter was positioned behind the

participant during testing. The same experimenter (N.C.) conducted every testing session at each site, and all participants were provided with the same instructions (see Supplementary Material 1). Stimuli at both testing sites were presented at the same visual angle and eye movements were recorded using identical eye-tracking systems and recording parameters (described below).

Eye tracking. Eye movements were recorded with a sampling rate of 500 Hz from the right eye using a desktop-mounted EyeLink 1000 Remote Eye-Tracking System (SR Research Ltd, Ontario, Canada). A chinrest was used to stabilise head movements and standardise viewing distance. We conducted a nine-point eye-tracking calibration and validation at the beginning of each block. Seven gaze-related areas of interest (GAOIs) were used by our gaze-contingent algorithm (depicted by blue rectangles in Figure 1). A GAOI covered each of the six houses and the avatar. Eye movements were monitored online and recalibration was conducted on trials where the participant made at least two consecutive fixations on the borders or outside the GAOIs. The trials requiring recalibration were excluded from all analyses. On average, this accounted for 0.87% of trials from the autism group (standard deviation (SD) = 1.14) and 0.05% of trials from the control group (SD = 0.22).

Joint attention task. The task was presented using *Experiment Builder* 1.10.165 (SR Research, 2004). At the beginning of the experiment, a scripted set of instructions was read aloud to the participant, and a series of cue cards were used to provide a schematic representation of the interactive eye-tracking interface (see Supplementary Material 1). Participants then completed two blocks of trials (Block 1 and Block 2). Each block comprised 27 trials from each condition (i.e. RJA, RJAc, IJA and IJAc). Social (RJA and IJA) and non-social (RJAc and IJAc) trials were presented in clusters of six trials throughout each block. Each cluster began with a 1000 ms cue presented in the centre of the screen that read 'Together' for a social cluster and 'Alone' for a non-social cluster of trials. The randomisation of trial order within and across clusters was constrained to ensure that, within each block, conditions were matched on (1) the frequency that the burglar appeared in each location, (2) the number and location of houses that required searching on each trial and (3) the number of gaze shifts made by the avatar before establishing eye contact.

There were four trial-order protocols that could be completed on each block. Two required the participant to search the upper row of houses (upper blocks), and two required the participant to search the lower row of houses (lower blocks). For each pair of protocols, one began with a social cluster of trials (RJA and IJA) and the other began with a non-social cluster of trials (RJAc and IJAc). Each participant completed only one upper and one lower

protocol. Protocol and cluster order were counterbalanced across participants and matched between the autism and control groups. Participants were not offered any opportunity for practice so that learning effects between blocks could be examined.

Subjective ratings. At the end of testing, we interviewed participants about their subjective experience of the task to determine whether they were convinced that they were interacting with a real person. During the interview, participants rated their subjective experience of the task across a number of dimensions (see Supplementary Material 2 for procedural details and results).

Measures

RJA and RJAc. We measured accuracy as the proportion of trials (excluding trials that required eye-tracking recalibration or where the participant failed to complete their search) where the participant succeeded in catching the burglar. For correct trials, we also measured *saccadic reaction time*, which was the latency (in ms) between the presentation of the orienting cue (gaze for RJA and arrow for RJAc) and the onset of the responding saccade that resulted in a fixation at the correct burglar location (see Figure 2, Analysis Period A).

IJA and IJAc. In addition to trial accuracy, we derived two measures of participants' use of eye contact. *Target dwell time* was the total amount of time (in ms) between finding the burglar and saccading back to the avatar's face (see Figure 2, Analysis Period B). It is analogous to the saccadic reaction time (RJA and RJAc) insofar as it represents the time between the participant learning of the burglar's location and making the next appropriate saccade. *Premature initiating saccades* was the proportion of trials on which participants made a saccade from the avatar to the burglar location before he had established eye contact (IJA) or the fixation point had turned green (IJAc; see Figure 2, Analysis Period C).

Statistical analyses

Joint attention task. For each measure, we conducted an analysis of variance (ANOVA) using the ezANOVA (ez) package in R (Lawrence, 2013), reporting the generalised eta squared (η_G^2) measure of effect size. Group (autism versus control) was the between-subjects factor, and condition (social versus non-social) and block (Block 1 versus Block 2) were within-subjects factors. Significant interactions were followed up with ANOVAs and Welch's two sample unequal variances *t*-tests as appropriate (Welch, 1947). Full details of these analyses, including syntax and data screening, can be found in Supplementary Material 3. For reaction-time measures (i.e. saccadic reaction time and
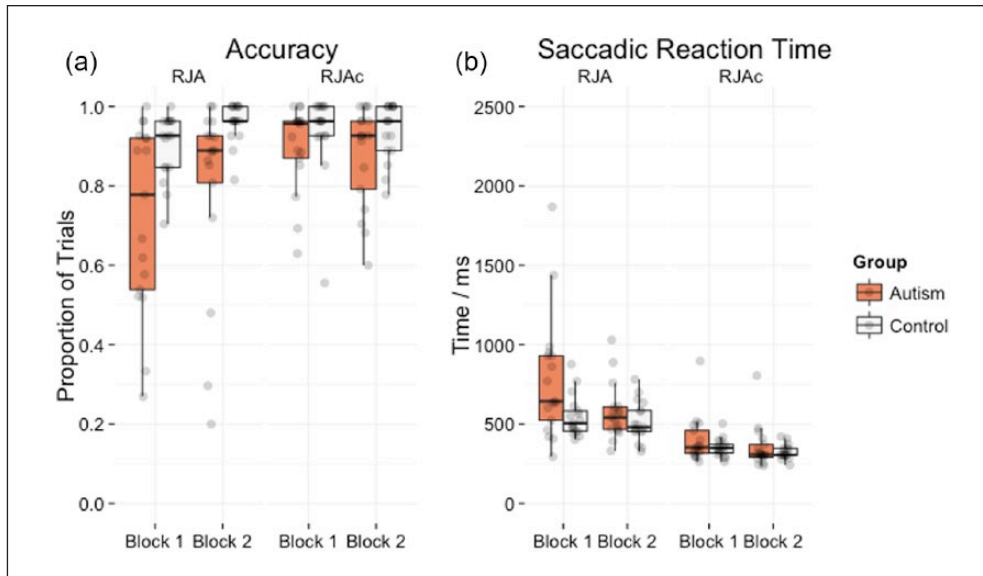
Figure 3. Box plots displaying (a) average proportion of correct trials and (b) average saccadic reaction times in RJA and RJAc conditions, separated by block and group. Data points represent individual participant means.

target dwell time), we report analyses of the mean reaction time, having excluded trials with dwell times less than 150 ms or more than 3000 ms (as trials timed out after 3000 ms in the RJA condition). We also re-analysed all reaction-time data taking the median of the untrimmed data (see Supplementary Material 3). This did not change the pattern of effects for any of the analyses.

Results

For each analysis, we report the main effects of condition and group to determine whether there were behavioural differences between the social and non-social conditions and between autistic and control participants, respectively. In addition, we report the group $\times$ condition and group $\times$ condition $\times$ block interaction effects since we were primarily interested in exploring whether differences between autistic and control participants were specific to the social conditions and whether these changed with increased exposure to the task (see Supplementary Material 3 for a full summary of the ANOVA output).

RJA

Accuracy data are shown in Figure 3(a). Participants made significantly more errors on RJA trials than RJAc trials ($F(1, 32)=7.04$, $p=0.012$, $\eta_G^2=0.06$). Autistic adults made significantly more errors than control participants ($F(1, 32)=9.04$, $p=0.005$, $\eta_G^2=0.13$). Importantly, we found a significant group $\times$ condition interaction ($F(1, 32)=5.60$, $p=0.024$, $\eta_G^2=0.04$). Posthoc tests revealed that autistic adults made significantly more errors than the control group on RJA trials ($t(16.97)=-3.08$, $p=0.007$),

but not on RJAc trials ($t(26.49)=-1.58$, $p=0.127$). There was no significant group $\times$ condition $\times$ block effect ($F(1, 32)=0.50$, $p=0.485$, $\eta_G^2<0.01$).

Saccadic reaction-time data are presented in Figure 3(b). Participants were significantly slower to respond to RJA trials than RJAc trials ($F(1, 32)=73.65$, $p<0.005$, $\eta_G^2=0.33$). The main effect of group ($F(1, 32)=3.96$, $p=0.055$, $\eta_G^2=0.07$) and the group $\times$ condition interaction ($F(1, 32)=3.57$, $p=0.068$, $\eta_G^2=0.02$) failed to reach significance. However, there was a significant group $\times$ condition $\times$ block interaction ($F(1, 32)=4.65$, $p=0.039$, $\eta_G^2=0.13$) indicating different group $\times$ condition effects in the two blocks. For Block 1, there was a significant group $\times$ condition interaction ($F(1, 32)=4.96$, $p=0.033$, $\eta_G^2=0.05$) with the autistic participants being significantly slower than controls to respond to RJA trials ($t(19.41)=2.36$, $p=0.029$), but not on RJAc trials ($t(20.45)=1.25$, $p=0.226$). For Block 2, there was no significant group by condition interaction ($F(1, 32)=0.37$, $p=0.546$, $\eta_G^2<0.01$) and no significant group effect for either RJA ($t(28.86)=1.12$, $p=0.272$) or RJAc ($t(19.65)=0.87$, $p=0.392$).

IJA

Accuracy data for IJA and IJAc are shown in Figure 4(a). There was no significant effect of condition ($F(1, 32)=2.61$, $p=0.116$, $\eta_G^2=0.02$). Autistic adults made significantly more errors than controls ($F(1, 32)=6.38$, $p=0.017$, $\eta_G^2=0.07$). We tested the effect of group in both conditions separately and found that autistic adults made more errors than controls on both IJA ($t(30.58)=2.45$, $p=0.020$) and IJAc trials ($t(16.65)=2.33$, $p=0.033$).
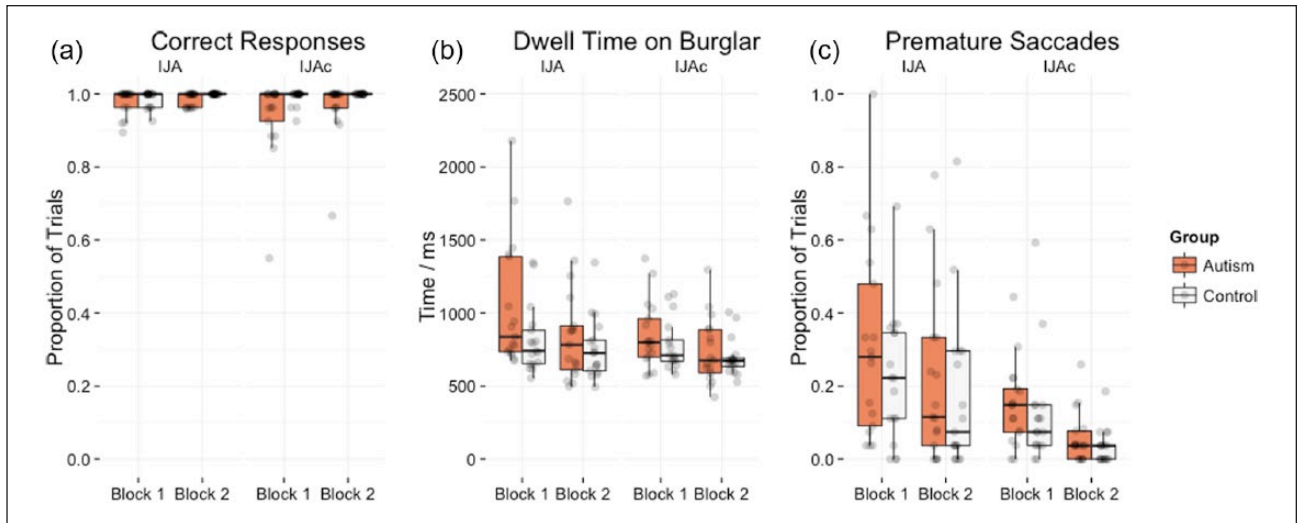
Figure 4. Box plots displaying (a) average proportion of correct trials in the IJA and IJAc conditions, (b) average dwell times on the burglar before establishing eye contact (IJA) or looking back at the fixation point (IJAc) and (c) average proportion of trials containing a saccade from the central stimulus to the burglar before the avatar made eye contact (IJA) or the fixation point turned green (IJAc), separated by block and group. Data points represent individual participant means.

However, as depicted in Figure 4(a), the majority of autistic participants performed at ceiling, and so these differences are largely driven by three individuals in the sample. There was no group × condition interaction ($F(1, 32) = 3.54$, $p = 0.069$, $\eta_G^2 = 0.02$) and no group × condition × block interaction ($F(1, 32) = 0.88$, $p = 0.355$, $\eta_G^2 < 0.01$).

Target dwell-time data are presented in Figure 4(b). Participants spent significantly more time fixated on the burglar before establishing eye contact on IJA trials relative to analogous eye movements on IJAc trials ($F(1, 32) = 10.68$, $p = 0.003$, $\eta_G^2 = 0.04$). There was no main effect of group ($F(1, 32) = 2.06$, $p = 0.161$, $\eta_G^2 = 0.05$), and no group × condition ($F(1, 32) = 2.79$, $p = 0.104$, $\eta_G^2 = 0.01$) or group × condition × block interactions ($F(1, 32) = 1.99$, $p = 0.168$, $\eta_G^2 < 0.03$).

Data for prematurely initiated saccades are presented in Figure 4(c). Participants made significantly more premature attempts at IJA on IJA trials relative to analogous eye movements on IJAc trials ($F(1, 32) = 20.76$, $p < 0.005$, $\eta_G^2 = 0.14$). There was no significant main effect of group ($F(1, 32) = 1.02$, $p = 0.321$, $\eta_G^2 = 0.02$), no group × condition interaction ($F(1, 32) = 0.52$, $p = 0.478$, $\eta_G^2 < 0.01$) and no group × condition × block interaction ($F(1, 32) = 0.64$, $p = 0.429$, $\eta_G^2 < 0.01$).

Discussion

Difficulty establishing joint attention is a cardinal feature of autism (American Psychiatric Association, 2013). However, little is known about joint attention abilities in older children or adults, most likely due to a lack of sensitive and ecologically valid experimental paradigms. In this

study, we addressed this issue using a novel interactive eye-tracking paradigm and provide the first evidence that joint attention impairments also affect autistic adults.

RJA

Compared to controls, autistic adults were less accurate at responding to the joint attention bid of an avatar. They also responded more slowly during the first block of testing. However, the autistic participants showed a significant improvement in response speed and, by the second block, were indistinguishable from control participants. Importantly, these group differences were specific to the social (RJA) condition: autistic and non-autistic individuals did not differ in their responses to arrow cues in the non-social (RJAc) condition. Thus, the reduced and delayed ability to respond to joint attention exhibited by autistic participants cannot be explained by differences in oculomotor control, attention orienting or executive function demands, which were equivalent in the RJA and RJAc conditions. Instead, the interaction between group and condition indicates that the difficulties of participants in the autism group were specific to the condition involving eye-gaze cues.

One possible explanation for the difference between groups may be that autistic individuals have different sensitivities to the low-level perceptual properties that unavoidably differ between gaze cues and non-social arrow cues. However, existing empirical studies of autistic children and adults show little evidence of specific difficulties in processing eye gaze as compared to non-social attention cues (see Leekam, 2015; Nation and Penny, 2008). Importantly, there is a key difference between our RJA

condition and conventional gaze-following tasks. In this study, the virtual partner made multiple eye movements during each trial and participants had to differentiate eye movements that were preceded by eye contact and thus signalled the intent to initiate joint attention from eye movements that were merely a continuation of their partner's search (cf. Caruana et al., 2016). This 'intention monitoring' component is an important feature of real-life gaze-based interactions (Cary, 1978) but is absent from more conventional measures of gaze processing in which a single unambiguous gaze shift is presented on each trial. Poor performance in our RJA condition may, therefore, reflect difficulties determining the social relevance of different eye-gaze cues rather than a deficit in eye gaze processing per se.

Following this interpretation, our findings are consistent with the idea that joint attention impairments in autism reflect a difficulty in evaluating the meaning of particular gaze cues (i.e. what they tell us about the perspectives and intentions of others) rather than an inability to effectively discriminate and orient to gaze cues (Baron-Cohen, 1995; Senju and Johnson, 2009). They are also consistent with evidence that autistic individuals are less effective in using eye contact to understand the goals and actions of others (Phillips et al., 1992) or to assess the relevance of an upcoming gaze shift (Böckler et al., 2014).

IJA

On average, autistic participants made more errors than control participants in the initiating conditions. However, in contrast to the responding conditions, group differences were not specific to the social version of the task, but were evident for both IJA and IJAc. This finding indicates a difficulty with one or more of the task components that were common to both conditions, such as oculomotor control, attentional demands or the requirement to remember the burglar's location (recall that the burglar disappeared once the participant made a saccade back to the avatar). That said, it is important to note that the majority of participants in both groups performed at or close to ceiling in terms of successful trial completion (see Figure 4(a)). Our accuracy measure may, therefore, have lacked sensitivity to detect subtle group differences. However, we also considered two eye-tracking measures of *how* participants were completing the task – the length of time between finding the burglar and saccading back to the avatar, and the number of premature saccades. Again, there were no significant group differences, despite much greater individual variation.

These findings are at odds with previous studies of joint attention which suggest that IJA difficulties, unlike RJA difficulties, tend to persist into later development (Mundy et al., 1994). It has been suggested that IJA impairments in autism may be related to a reduced motivation to engage in social interactions (Chevallier et al., 2012). This idea is consistent with neuroimaging studies which associate the achievement of joint attention, following IJA behaviour, with activation in the ventral striatum, a region associated with social reward processing (Schilbach et al., 2010). It is possible that IJA difficulties were not observed in this study because IJA behaviours were externally motivated by the goals defined by the task, rather than the participant's intrinsic motivation to share a social experience with another person.

There were some interesting differences between the IJA and IJAc conditions that were common to both groups of participants. First, having found the burglar during the search phase, participants took longer to saccade back towards the avatar's eyes in the IJA condition than the central fixation point in the IJAc condition. Second, they were more likely to make a premature guiding saccade to capture the burglar in the IJA condition than they were to make analogous saccades in the IJAc condition. Both findings were unexpected and may reflect the fact that participants expect a certain degree of flexibility from a human partner that they know not to expect from a computer. That is, participants may have expected Alan to follow their guiding gaze even when they did not wait to intentionally establish eye contact. Future studies could test this explanation by investigating whether participants are faster to establish eye contact, and make fewer premature initiating saccades, when they believe a virtual character is controlled by a computer rather than a human.

Furthermore, the fact that most participants attempted to initiate joint attention before establishing eye contact raises questions about the phenomenology of optimal joint attention behaviour and how it ought to be measured and assessed. Specifically, it calls into question whether establishing eye contact should be considered a mandatory aspect of IJA or simply an adaptive behaviour that may facilitate the achievement of joint attention under certain conditions. Naturalistic studies of genuine face-to-face interactions are needed to better characterise the role of eye contact during successful joint attention experiences between adults with typical development. This work will also inform the design and future implementation of joint attention paradigms.

Subjective experiences

At the end of testing, we interviewed participants about their subjective experience of the task. Only two participants, both in the autism group, claimed to have suspected that Alan was not real. However, prior to being told that Alan was computer-controlled, neither participant had given any indication that the deception had been unsuccessful. For instance, when asked whether they preferred completing the task with Alan or on their own, one participant commented 'Together … more accurate because you can see the other person's perspective'.

The comments made by autistic individuals during the debriefing session also provide some intriguing insights into the difficulties they faced while completing the task. Six autistic adults explicitly stated that they found it challenging to complete a task that required them to establish and use eye contact. Different individuals commented that

> The eyes were harder to figure out.

> Alone [i.e. non-social condition) was easier to complete because you didn't have to catch his eye to tell him where to go.

> When they [eyes] were closed I didn't have to worry about him and what he wants. Didn't have to have the patience to wait for him.

> I felt a bit anxious during the together task [i.e. social condition]. The alone task [i.e. non-social condition] was easier because it was clear what the dot and arrow meant […] I don't normally look at peoples' eyes […] In the game I had to look at the eyes […] Then I thought, 'Why are eyes harder than arrows?' So I decided to treat the eyes like arrows.

None of the control participants reported difficulties processing the eyes.

These comments demonstrate an awareness on the part of many autistic adults that establishing eye contact and using gaze as a communicative technique was challenging for them. This is consistent with a larger body of literature suggesting that autistic individuals find it difficult to use eye contact to understand others and regulate social interactions (e.g. Pelphrey et al., 2011; Senju and Johnson, 2009). Specifically, this difficulty in establishing eye contact may also explain why some autistic adults demonstrated markedly more premature saccades and took longer to establish eye contact with the avatar on IJA trials. For example, one participant spent up to 30 s fixating on the burglar before establishing eye contact on IJA trials (median 3 s). Another spent up to 6 s fixating on the burglar (median 2 s) and made premature saccades on 81% of IJA trials. While this was not representative of the entire autism group, this reluctance to establish eye contact could hinder the achievement of joint attention during the fast-paced social interactions of real life. Further investigation is needed to elucidate the factors that contribute to the inter-individual variation in joint attention behaviour and experiences for autistic individuals. For instance, one focus for future work could be to investigate the relationship between individual differences in social anxiety and joint attention behaviour (Kuusikko et al., 2008).

Some autistic participants also indicated that while they preferred to complete the task alone than with Alan, they also preferred the virtual interaction over real-life face-to-face interactions. They indicated that the computer interface provided a less anxiety-provoking social interaction: 'I don't like dealing with people so this was better. Feels like you're socialising, but not. Feels more relaxed'. Another autistic participant preferred real-life interactions, but only if eye contact could be avoided: '[Virtual interface] makes it more comfortable … I am an audio person. I like to ask things if they're not clear. So I would prefer real life. Not face-to-face, but side-by-side'. Others noted that the virtual interface allowed them to focus on specific aspects of their social interaction without being overwhelmed by multiple cues: (1) 'Easier to segment the task and interaction … only focus on one thing' and (2) 'I can interact but don't have too many things to think about'.

Conclusion

To our knowledge, this is the first study designed to use an ecologically valid, objective, quantified and experimentally-controlled measure to test the ability to respond to and initiate joint attention bids in autism. Our data indicates that autistic adults experience significant difficulties in RJA bids. Some autistic individuals also experienced difficulties in IJA, but this was not consistent across our entire sample. These findings encourage further work investigating the individual characteristics that may account for the heterogeneity of joint attention abilities in autism. In particular, there is a need for further studies that apply these paradigms across larger samples, with individuals across the autism spectrum, and at different stages of development. Ideally, future studies would obtain additional detailed measures of individuals' social functioning in daily life situations in order to determine which (if any) aspects of daily social functioning are associated with joint attention impairments. Virtual interaction paradigms could also be used in neuroimaging studies to investigate the neural correlates of atypical joint attention (cf. Caruana et al., 2015).

This study also highlights the potential for interactive computer-based tasks to guide the training of social information processing and communication skills among individuals with autism. Preliminary findings using virtual reality video games, albeit from a third person perspective, have already revealed promising gains in social cognitive skills in autistic children (Didehbani et al., 2016). Tasks like ours, which support real-time social interaction, may be used to identify the precise aspects of face-to-face social interactions that autistic people find difficult and provide strategies that are likely to make social communication more effective and possibly less stressful. The minimal 'social' environment of a computer interface may allow individuals to become gradually accustomed to one aspect of social communication at a time (e.g. eye gaze), while other aspects of the social interaction and environment are controlled. This could provide a less perceptually overwhelming context for autistic individuals to develop

their skills in social information processing and communication. This approach has the potential to make social interactions more pleasant and less-intimidating for autistic individuals while improving opportunities for social learning, language acquisition and fostering the development of social relationships (Adamson et al., 2009; Baron-Cohen, 1995; Charman, 2003; Mundy et al., 1990; Murray et al., 2008; Tomasello, 1995).

Funding

References

Adamson LB, Bakeman R, Deckner DF, et al. (2009) Joint engagement and the emergence of language in children with autism and Down syndrome. *Journal of Autism and Developmental Disorders* 39(1): 84–96.

American Psychiatric Association (2000) *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington, DC: American Psychiatric Association.

American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association.

Bakeman R and Adamson LB (1984) Coordinating attention to people and objects in mother-infant and peer-infant interactions. *Child Development* 55(4): 1278–1289.

Baron-Cohen S (1995) *Mindblindness*. Cambridge, MA: MIT Press.

Baron-Cohen S (2008) Autism, hypersystemizing, and truth. *Quarterly Journal of Experimental Psychology* 61(1): 64–75.

Bates E, Benigni L, Bretherton I, et al. (1979) *The Emergence of Symbols: Cognition and Communication in Infancy*. New York: Academic Press.

Bayliss AP, Murphy E, Naughtin CK, et al. (2013) Gaze leading: initiating simulated joint attention influences eye movements and choice behavior. *Journal of Experimental Psychology: General* 142(1): 76–92. Available at: http://doi.org/10.1037/a0029286

Böckler A, Timmermans B, Sebanz N, et al. (2014) Effects of observing eye contact on gaze following in high-functioning autism. *Journal of Autism and Developmental Disorders* 44(7): 1651–1658.

Brock J (2011) Commentary: complementary approaches to the developmental cognitive neuroscience of autism – reflections on Pelphrey et al. (2011). *Journal of Child Psychology and Psychiatry* 52(6): 645–646.

Bruinsma Y, Koegel RL and Koegel LK (2004) Joint attention and children with autism: a review of the literature. *Mental Retardation and Developmental Disabilities Research Reviews* 10(3): 169–175.

Caruana N, Brock J and Woolgar A (2015) A frontotemporoparietal network common to initiating and responding to joint attention bids. *Neuroimage* 108: 34–46.

Caruana N, McArthur G, Woolgar A, et al. (2016) Detecting communicative intent in a computerised test of joint attention. *PeerJ Preprints* 4: e2410v1. Available at: https://doi.org/10.7287/peerj.preprints.2410v1

Cary MS (1978) The role of gaze in the initiation of conversation. *Social Psychology* 41(3): 269–271.

Charman T (2003) Why is joint attention a pivotal skill in autism? *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 358: 315–324.

Charman T, Swettenham J, Baron-Cohen S, et al. (1997) Infants with autism: an investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology* 33(5): 781–789.

Chevallier C, Kohls G, Troiani V, et al. (2012) The social motivation theory of autism. *Trends in Cognitive Sciences* 16: 231–239.

Courgeon M, Rautureau G, Martin J-C, et al. (2014) Joint attention simulation using eye-tracking and virtual humans. *IEEE Transactions on Affective Computing* 5(3): 238–250.

Dawson G, Toth K, Abbott R, et al. (2004) Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental Psychology* 40(2): 271–283.

Didehbani N, Allen T, Kandalaft M, et al. (2016) Virtual reality social cognition training for children with high functioning autism. *Computers in Human Behavior* 62: 703–711. Available at: http://dx.doi.org/10.1016/j.chb.2016.04.033

Emery NJ (2000) The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews* 24(6): 581–604.

Friesen C and Kingstone A (1998) The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review* 5(3): 490–495.

Hobson J and Hobson P (2007) Identification: the missing link between joint attention and imitation? *Development and Psychopathology* 19(02): 411–431.

Kaufman AS and Kaufman NL (2004) *Kaufman Brief Intelligence Test*. 5th ed. Bloomington, MN: Pearson, Inc.

Kim K and Mundy P (2012) Joint attention, social-cognition, and recognition memory in adults. *Frontiers in Human Neuroscience* 6. Available at: http://doi.org/10.3389/fnhum.2012.00172

Kuusikko S, Pollock-Wurman R, Jussila K, et al. (2008) Social anxiety in high-functioning children and adolescents with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders* 38: 1697–1709.

Lawrence MA (2013) ez: easy analysis and visualization of factorial experiments. R package version 4.2–2. Available at: http://CRAN.R-project.org/package=ez

Leekam S (2015) Social cognitive impairment and autism: what are we trying to explain? *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 371(1686): 20150082.

Lord C, Risi S, Lambrecht L, 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): 205–223.

Loveland KA and Landry SH (1986) Joint attention and language in autism and developmental language delay. *Journal of Autism and Developmental Disorders* 16(3): 335–349.

Mundy P, Sigman M and Kasari C (1990) A longitudinal study of joint attention and language development in autistic children. *Journal of Autism and Developmental Disorders* 20(1): 115–128.

Mundy P, Sigman M and Kasari C (1994) Joint attention, developmental level, and symptom presentation in autism. *Development and Psychopathology* 6(03): 389–401.

Mundy P, Sigman M, Ungerer J, et al. (1986) Defining the social deficits of autism: the contribution of non-verbal communication measures. *Journal of Child Psychology and Psychiatry* 27(5): 657–669.

Murray DS, Creaghead NA, Manning-Courtney P, et al. (2008) The relationship between joint attention and language in children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities* 23(1): 5–14.

Nation K and Penny S (2008) Sensitivity to eye gaze in autism: is it normal? Is it automatic? Is it social? *Development and Psychopathology* 20(1): 79–97.

Osterling J and Dawson G (1994) Early recognition of children with autism: a study of first birthday home videotapes. *Journal of Autism and Developmental Disorders* 24(3): 247–257.

Osterling J, Dawson G and Munson JA (2002) Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development and Psychopathology* 14(2): 239–251.

Pelphrey KA, Shultz S, Hudac CM, et al. (2011) Research review: constraining heterogeneity: the social brain and its development in autism spectrum disorder. *Journal of Child Psychology and Psychiatry* 52(6): 631–644.

Pfeiffer UJ, Vogeley K and Schilbach L (2013) From gaze cueing to dual eyetracking: novel approaches to investigate the neural correlates of gaze in social interaction. *Neuroscience and Biobehavioral Reviews* 37(10): 2516–2528.

Phillips W, Baron-Cohen S and Rutter M (1992) The role of eye contact in goal detection: evidence from normal infants and children with autism or mental handicap. *Development and Psychopathology* 4(3): 375–383.

Ritvo RA, Ritvo ER, Guthrie D, et al. (2011) The Ritvo Autism Asperger Diagnostic Scale-Revised (RAADS-R): a scale to assist the diagnosis of autism spectrum disorder in adults: an international validation study. *Journal of Autism and Developmental Disorders* 41: 1076–1089.

Scaife M and Bruner JS (1975) The capacity for joint visual attention in the infant. *Nature* 253(5489): 265–266.

Schilbach L, Wilms M, Eickhoff SB, et al. (2010) Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience* 22(12): 2702–2715. Available at: http://doi.org/10.1162/jocn.2009.21401

Senju A and Johnson MH (2009) The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences* 13(3): 127–134. Available at: http://dx.doi.org/10.1016/j.tics.2008.11.009

Sigman M and Ruskin E (1999) Continuity and change in the social competence of children with autism, Down syndrome and developmental delays. *Monographs of the Society for Research in Child Development* 64(1): 1–114.

Singular Inversions (2008) *FaceGen Modeller* (Version 3.3) [Computer Software]. Toronto, ON: Singular Inversions.

SR Research (2004) *Experiment Builder* (version 1.10.165). Ottawa, ON, Canada: SR Research.

Stone WL, Ousley OY and Littleford CD (1997) Motor imitation in young children with autism: what's the object? *Journal of Abnormal Child Psychology* 25: 475–485.

Svensson E (2001) Guidelines to statistical evaluation of data from rating scales and questionnaires. *Journal of Rehabilitation Medicine* 33(1): 47–48.

Tomasello M (1995) Joint attention as social cognition. In: Moore C and Dunham PJ (eds) *Joint Attention: Its Origins and Role in Development*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp.1–14.

Welch BL (1947) The generalization of 'student's' problem when several different population variances are involved. *Biometrika* 34: 28–35.

Wheelwright S, Baron-Cohen S, Goldenfeld N, et al. (2006) Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ). *Brain Research* 1079(1): 47–56. Available at: http://dx.doi.org/10.1016/j.brainres.2006.01.012

Wong C and Kasari C (2012) Play and joint attention of children with autism in the preschool special education classroom. *Journal of Autism and Developmental Disorders* 42(10): 2152–2161.