ORIGINAL PAPER

Self-Monitoring of Gaze in High Functioning Autism

Ouriel Grynszpan · Jacqueline Nadel · Jean-Claude Martin ·
Jérôme Simonin · Pauline Bailleul · Yun Wang · Daniel Gepner ·
Florence Le Barillier · Jacques Constant

Abstract Atypical visual behaviour has been recently proposed to account for much of social misunderstanding in autism. Using an eye-tracking system and a gaze-contingent lens display, the present study explores self-monitoring of eye motion in two conditions: free visual exploration and guided exploration via blurring the visual field except for the focal area of vision. During these conditions, thirteen students with High Functioning Autism Spectrum Disorders (HFASD) and fourteen typical individuals were presented naturalistic and interactive social stimuli using virtual reality. Fixation data showed a weaker modulation of eye movements according to the conditions in the HFASD group, thus suggesting impairments in self-monitoring of gaze. Moreover, the gaze-contingent lens induced a visual behaviour whereby social understanding scores were correlated with the time spent gazing at faces. The device could be useful for treating gaze monitoring deficiencies in HFASD.

O. Grynszpan (✉) · J. Nadel · P. Bailleul · D. Gepner
CNRS USR 3246, Université Pierre et Marie Curie, Bâtiment de
La Force, Hôpital de La Salpêtrière, 47 Boulevard de l'Hôpital,
75013 Paris, France
e-mail: ouriel.grynszpan@upmc.fr

J.-C. Martin · Y. Wang
LIMSI-CNRS, Université Paris Sud, Orsay, France

J. Simonin · D. Gepner
LORIA-CNRS-INRIA, Vandoeuvre-lès-Nancy, France

F. Le Barillier · J. Constant
La Maison pour les Personnes Autistes, Hôpitaux de Chartres,
Chartres, France

Introduction

Gaze peculiarities are now listed as part of the specific dysfunctions in Autism Spectrum Disorders (ASD). They are considered as altering both the processing of social and non-social stimuli (Samson et al. in press). Notwithstanding, they are largely examined in light of their social outcomes such as face scanning deficits and weak social attention (Spezio et al. 2007). Some bottom-up approaches even hypothesize that social deficits themselves would originate from gaze atypicalities. For instance, the 'amygdala hyperactivity' model posits that heightened aversive sensitivity to social stimuli explains the reduced gaze fixations on faces, which in turn result in the hypoactivation of the fusiform gyrus attributed to ASD (Corden et al. 2008; Dalton et al. 2005). Another model, named the Fast Track Modulator model (FTM), stems from neuroimaging studies evidencing that the capacity for eye contact is linked to a fast subcortical pathway that modulates face processing, such that atypical eye contact reported in ASD could hamper the development of the cortical social brain circuitry (Senju and Johnson 2009a, b). As eye-to-eye contact plays a critical role in face-to-face communication, atypical visual exploration is postulated to be associated with poor social attention and communication. This corpus of research underscores the necessity of investigating the mechanisms such as self-monitoring that subserve gaze control in social contexts.

Our daily social life drives us to experience how much the eyes can serve as a powerful nonverbal communication channel. Baron-Cohen et al. (1997) term this communicative function of gaze the "language of the eyes" (p. 311). They emphasize the difficulties encountered by individuals with High Functioning Autism Spectrum Disorders (HFASD) in interpreting emotions expressed by the eyes. A trend of research attempts to tackle issues regarding the

misinterpretation of social interactions using eye-tracking technology (Boraston and Blakemore 2007). Most studies tracking eye movements on static pictures of emotional facial expressions show a tendency to omit core features of the face and especially the eyes (Corden et al. 2008; Dalton et al. 2005; Pelphrey et al. 2002; Spezio et al. 2007), although there are contradicting reports suggesting that abnormal gaze behaviour in HFASD is due mostly to social interactions rather than facial stimuli per se (Speer et al. 2007; van der Geest et al. 2002). Few studies employed more ecologically valid stimuli depicting animated social characters. Klin et al. (2002a) report that participants with HFASD, while watching a film featuring dense social interactions, tended to focus less on the eyes and more on the mouth, body and surrounding objects than controls. Moreover, their fixation time on objects was negatively correlated with standardized measures of their social aptitudes. Speer et al. (2007) found similar results only when the film featured more than one character. Riby and Hancock (2009) confirmed the tendency in ASD to omit faces when watching films and extended the finding to cartoon movies. In a case study, Klin et al. (2002b) highlighted how the atypical visual behaviour of an individual with HFASD could account for much of his misunderstanding of the film. For instance, when a comedian would speak to another one, the participant with HFASD would not look as much to the listener's facial expressions as would the typical participant, thus missing out the listener's reactions and failing to understand the social dynamics of the conversation. Despite preserved intellectual abilities and often well-developed vocabulary, individuals with HFASD exhibit profound difficulties in pragmatics, which refers to the ability to use language in order to achieve effective communication in social contexts (Tager-Flusberg 2000). They have a tendency to interpret speech literally rather than in reference to a context (Jolliffe and Baron-Cohen 1999), failing to understand irony, metaphors and deception. In a previous study (Grynszpan et al. 2008), we emphasized the difficulties experienced by adolescents with HFASD in using facial expressions as context to interpret dialogues in a non-literal manner.

As conceived by Baron-Cohen et al. (1997), the language of the eyes depends on the ability to interpret the expression conveyed by another person's eyes; however, the ability to monitor one's own eye movements can arguably hold a critical role as well. Sasson et al. (2007) showed that individuals with HFASD who were to assess the emotional content of static social scenes where faces were either included or digitally erased, failed to modulate the orientation of their gaze differently when faces were present. Convincing evidence from the literature shows that the monitoring of eye movements is a key component of the human visual system (Lindner et al. 2006). Monitoring

of action is classically described as dependent on the 'forward model'; an internal model that predicts the sensory consequences of motor commands (Blakemore et al. 2002; Wolpert and Miall 1996) and thus enables self-monitoring. This concept has been conducive to the understanding of eye motion (Robinson et al. 1986).

The present study seeks to test the monitoring of the eyes' movements in a conversational context. Given that social interactions represent the archetypical situation in which individuals with HFASD have difficulties, we devised that such a context would be propitious for detecting a possible dysfunction of self-monitoring. We designed a task requiring participants to look at the face of an animated virtual character, while provided with visual feedback on the direction of their gaze. The task was specially conceived to assess performances in non-literal interpretation of speech and should therefore allow to discriminate between typical and HFASD participants. Given the work of Riby and Hancock (2009) using cartoon movies, we assumed that the abnormalities in gaze behaviour would remain, even though characters were not real. Moreover, we relied on virtual reality tools that enable precise control over the design of facial expressions and their synchronicity with speech, in characters holding highly realistic features. According to Russell and Hill (2001), action-monitoring "refers to the mechanisms that ensure that agents always know, without self-observation, (1) for which changes in perceptual input they are responsible and (2) what they are currently engaged in doing" (p. 317). With regards to (1), we tested whether participants could detect that the visual feedback was controlled by their gaze and whether they behaved accordingly.

Method

Participants

Fourteen adolescents and young adults with ASD and 14 typical individuals participated in this study. The ASD diagnoses conducted by child psychiatrists, based on DSM IV criteria, were confirmed with the Autism Diagnostic Interview (ADI) (Le Couteur et al. 1989) and the Childhood Autism Rating Scale (CARS) (Schopler et al. 1980). One participant with ASD was excluded due to a failure in our apparatus. We administered the Raven's Progressive Matrices (Raven and Court 1986) to the remaining participants and their Verbal Intelligence Quotient was assessed with the Wechsler Adult Intelligence Scale, 3rd edition (Wechsler 1997). As can be seen in Table 1, this group was high functioning. They comprised 1 female and 12 male and their ages ranged from 13 to 31 years. Note should be

Table 1 Participants details

	HFASD ($n = 13$)	Typical ($n = 14$)
Gender		
Female	1	7
Male	12	7
Vision		
Normal	5	8
Corrected-to-normal	8	6
Chronological age		
Mean	20.19	23.18
SD	4.73	6.04
Range	13–31	13–31
Raven's matrices		
Mean	50.26	
SD	5.95	
Range	36–56	
Verbal IQ		
Mean	90.5	
SD	16.6	
Range	68–119	

Table 2 An animated scene sample: the upper part of the table presents each utterance of the virtual character associated with its simultaneous facial expression

Animated scene's script	
Utterance	Facial expression
It's lunch time, I'm hungry	Joy
They are serving spinach	Disgust
What a great surprise	Disgust
We usually have French fries	Sadness
Choice of answers for the first question, i.e. "How does John feel?"	
He is disgusted	
He is happy	
He is afraid	
Choice of answers for the second question, i.e. "How do you know that?"	
John does not like spinach	
John likes spinach	
The cook was late in serving food	

The key sentence inducing an ambiguity is in italic font. The possible choices for the two subsequent questions are shown beneath. The correct answers are in italic font. This example is an English translation of a French script that was actually used

taken that vision control is considered mature and similar to adult vision in this age range (Slater 1999). Each typical participant was interviewed before the experiment. Participants with neurological and/or psychiatric history and psychotropic treatment were excluded. Due to ethical considerations, the IQ of typical participants were not assessed, however we checked that their academic level was coherent with their age. They had the same age range as the HFASD group. A *t* test on the participants' ages showed no significant difference between the two groups ($t = 1.42$, $p = 0.17$). Although the gender ratio of the typical group did not match with the HFASD group, it was consistent with the gender ratio in the general population. This research was reviewed and approved by the regional ethics committee. An informed consent was obtained from each participant.

Apparatus and Task

Participants had to complete a task designed to induce a behaviour whereby they would look at the face of a virtual character addressing them. The character reported a situation he or she was experiencing and, while doing so, uttered a sentence that could be interpreted in two distinct ways according to the context. The context was provided by the character's facial expressions that enabled disambiguating this key sentence and therefore understanding the whole message (see example in Table 2). After each such animated scene, the participants were asked two closed-choice questions in order to evaluate their understanding. The first was about the feeling of the virtual character; the second was about the basis for the participant's perception. Participants had three possible choices for each question: the correct interpretation coherent with the key sentence and the facial expressions; an out-of-context interpretation coherent with the key sentence left alone, but incoherent with the facial expressions; an erroneous interpretation incoherent with both. A female and a male virtual character were designed with Poser Pro software (Smith Micro Software, Inc.). Each appeared in half of the animated scenes. The wording used had been carefully constructed to be easily understandable by young adolescents. The facial expressions associated with the key sentence in each animated scene were chosen among five basic emotions (disgust, joy, fear, anger and sadness) which individuals with HFASD are reported to recognize better than subtler emotions (Baron-Cohen et al. 1997). These expressions were designed based on Ekman's model (Ekman 2003) and the stimuli were validated in a preliminary study (Buisine et al. 2010) including 23 typical participants. The virtual characters were embedded in videos of real life settings, thus providing a naturalistic context. To avoid possible biases induced by the virtual character's voice, intonation was reduced to the minimum by using synthesized speech created with Virtual Speaker software (Acapela Group, Inc.). The virtual characters' lips movements and speech were synchronized with Poser Pro. Previous research in HFASD supports the ecological validity of virtual
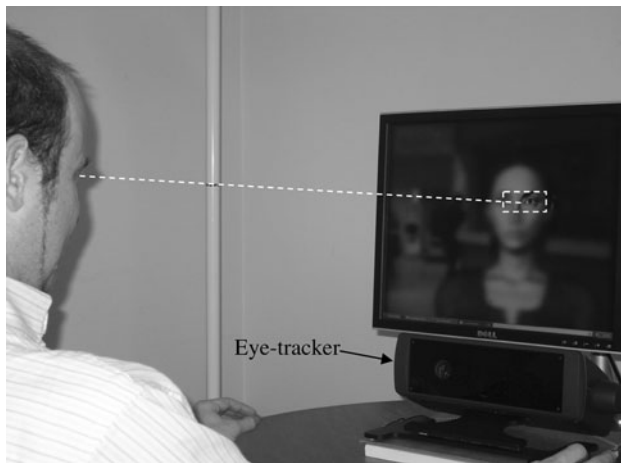
Fig. 1 The gaze-contingent prototype: an eye-tracking remote camera enables the participant to control a lens on the visual display. The whole display is blurred except for the lens centred on the focal point of the participant

characters, even when social behaviours are approximated (Parsons et al. 2004).

During the experiment, the participant's gaze was tracked by a remote infrared camera (model D6-HS Remote from Applied Science Laboratories) placed under the screen. The eye-tracker was combined with a gaze-contingent prototype that provided participants with real-time feedback about the location of their gaze on the screen. The prototype processed the graphic display in real time so that it was entirely blurred, except for an area centred on the focal point of the participant. This system can be seen as simulating a gaze-contingent lens (Fig. 1). The animations were displayed on a 19 inch (377 × 302 mm^2) screen with a resolution of 1,280 × 1,024 pixels. The lens was a rectangle with rounded angles measuring 233 × 106 pixels, large enough to see the virtual character's two eyes with eyebrows. The processing time required by the gaze-contingency system induced a delay of approximately 100 ms between the movement of the eye and the positioning of the lens. This delay was expected to induce a behaviour whereby participants would stabilize their gaze. The gaze-contingent lens would therefore serve as an indicator of the ability to self-monitor one's gaze. This prototype could be switched on or off. In the latter case, the eye-tracker would still function, but would merely record the direction of the gaze.

Procedure

The experimental protocol followed an ABA design: visual exploration was first free (baseline condition), then the gaze-contingent lens was switched on (experimental condition) and finally switched off, enabling free exploration again (final condition). We created sixty different animated scenes, twenty scenes per condition. Animated scenes lasted 18 s on average and ranged from 13 to 24 s. The order of the scenes was randomly counterbalanced across participants so that any given scene was presented under various conditions and the total duration of animations was the same in every condition. Special care was taken to make sure participants understood the task. Before starting the experiment, participants had to read the instructions for the task. The experimenter then reviewed the instructions with the participant and, in case of doubt on the participant's understanding, repeated them. In particular, participants were told that sometimes only some parts of the visual display would remain clear, but they were not told why. They then completed a standard calibrating procedure for the eye-tracker. The experiment started with a demo animation where both virtual characters introduced themselves and provided the instructions for the task. This demo also served as an example: It was followed by two closed-choice questions that were similar to those used subsequently except that they did not involve non-literal interpretation. The experimental and final baseline conditions also started with a short demo showing both virtual characters cheering the participants. The purpose was to give them some time for adapting to the new condition and reward their efforts for continuing the experiment. The data collected during the starting animations was not used for statistical analysis. Just before the experimental condition, a written instruction appeared on the screen that explicitly encouraged the participant to look at the facial expressions (e.g. "Think about looking at the characters' faces to understand what they feel"). This instruction was intended to prepare the participants for the gaze-contingent display and exhort them to behave consistently even though their vision would be constrained.

We also sought to examine whether participants became eventually aware that they were controlling the lens in the experimental condition. Throughout the experiment, participants were left uninformed about the gaze-contingent system although they knew that their gaze was measured by the eye-tracker. At the end of the experiment, they were asked a question that translates into English as follows: "You noticed that in some videos, there were blurred areas and clear areas. What causes the clear areas?". The answers were recorded and analysed by two independent judges to determine whether participants had noticed that they controlled the lens. There was complete agreement between judges.

Data Analyses

Eye-tracking data and scores on the task were automatically recorded during the experiment. Each of the two closed-choice questions that followed every animated
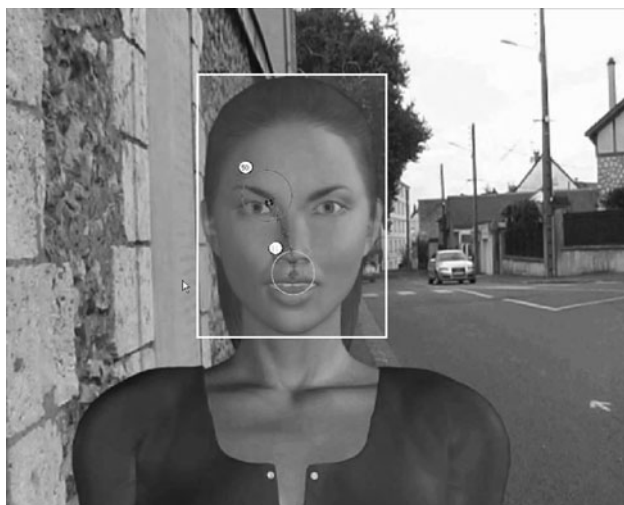
Fig. 2 An example of an Area of Interest (AOI) used for aggregating gaze fixations on the face. The *circles* represent consecutive fixations linked by lines depicting the visual path

scenes was worth one point. As there were 20 scenes in each condition, the computed scores ranged from 0 to 40. The sampling rate of the eye-tracker was 50 Hz, thus providing a Point-Of-Gaze (POG) every 20 ms. Fixations were computed with a proprietary algorithm conceived by the provider of the eye-tracker (Applied Science Laboratories) on the basis of clusters of POGs remaining in 1° of visual angle for at least 100 ms. The number of POGs collected during a fixation provides a measure of the fixation's duration equal to the duration in seconds divided by the sampling rate. The gaze data was analysed with a software prototype (Gepner et al. 2007), adapted for the present study, that could handle eye-tracking on dynamic visual displays. This prototype enabled aggregating gaze data on pre-defined rectangular Areas Of Interest (AOI). In each animated scene, we define an AOI that was circumscribed around the face (Fig. 2) and another AOI, named "no face", that encompassed the rest of the screen.

Given that the experimental procedure involved repeated measures that were conceivably highly correlated and that the group with HFASD could potentially yield very heterogeneous results, we conducted analyses of variance (ANOVA) based on a mixed-design associated with an unstructured residual covariance matrix and unequal variances. Post hoc *t* tests were performed using the Tukey adjustment procedure; the p-values provided hereunder are adjusted values. The task was intended to test the understanding of the animated scenes and the ability to focus on faces. To validate this task and ascertain that it could discriminate between the typical and HFASD groups, participants' performances were assessed based on the scores, the total viewing time spent on fixations (sum of fixation durations), and the number of fixations. We first checked

for possible influences of age and gender on these variables during the baseline condition. Pearson correlation coefficients were used for age and *t* tests (female versus male) for gender. Scores were analysed using an ANOVA with Condition as the within-subjects variable and Group (HFASD vs. typical) as the between-subjects variable. As performances depended on whether fixations were on the face, we added the within-subjects variable AOI (face vs. no face) in the ANOVA used for the sum of fixation durations and the number of fixations. We checked whether the participants' answers were linked to the time spent on viewing faces by calculating the correlation between the scores and the sum of fixation durations on faces for every condition in each group. Correlations being computed for each group separately, we used the Spearman coefficient given the small sample sizes. To test changes in gaze behavioural strategies due to the gaze-contingent lens, we sought to evaluate whether the gaze would be more stable during the experimental condition. We thus computed the average duration of fixations. The gaze-contingent lens was active on the entire screen and did not depend on whether the participant was looking at the face or not. Therefore, the AOI were not relevant for this ANOVA. It was computed with Condition as the within-subjects variable and Group as the between-subjects variable.

Results

Performances

Correlations between age and performance variables (scores, number of fixations, sum of fixation durations) were not significant, nor were the *t* tests comparing female and male. The ANOVA on the score variable yielded a main effect of Group [$F(1, 25) = 15.40$ $p = 0.0006$] with a mean score of 26.72 (SD = 7.13) for the HFASD group and 34.86 (SD = 4.61) for the typical group. There was a main effect of Condition [$F(2, 50) = 9.04$ $p = 0.0004$]; the mean scores were 30.41 (SD = 7.04) in the baseline condition, 30.04 (SD = 7.40) in the experimental condition and 32.37 (SD = 7.19) in the final condition. Post hoc comparisons showed that scores in the final condition were significantly higher than in both the baseline condition [$t(50) = 2.53$ $p = 0.0385$] and experimental condition [$t(50) = 2.84$ $p = 0.0177$].

The ANOVA for the sum of fixation durations yielded a significant Group × Condition interaction [$F(2, 47) = 7.28$ $p = 0.0018$]. Post hoc comparisons of conditions within each group revealed that the sum of fixation durations decreased significantly from the experimental condition to the final condition for the typical group [$t(47) = 4.14$ $p = 0.0019$] with a differences' mean of 39.16

Table 3 Spearman's correlations between the scores and the sum of fixation durations

	Condition					
	Baseline		Experimental		Final	
	r	p	r	p	r	p
HFASD group	0.014	NS	0.620	0.0239	0.33	NS
Typical group	−0.421	NS	0.091	NS	−0.089	NS

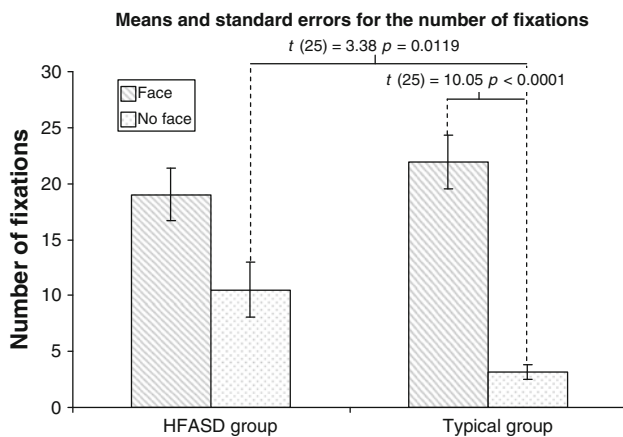NS not significant, HFASD high functioning autism spectrum disorders

Means and standard errors for the number of fixations

$t (25) = 3.38 \ p = 0.0119$

$t (25) = 10.05 \ p < 0.0001$

Fig. 3 Number of fixations per group per Area Of Interest (AOI), illustrating the Group × AOI interaction. *HFASD* High Functioning Autism Spectrum Disorders

Means and standard errors for the number of fixations

$t (47) = 3.87 \ p = 0.0010$
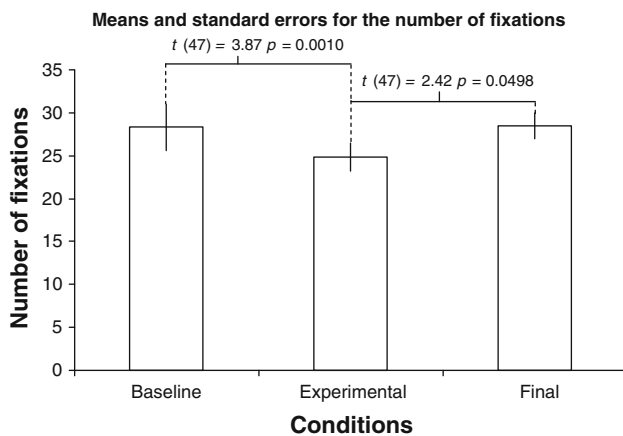
$t (47) = 2.42 \ p = 0.0498$

Fig. 4 Number of fixations per condition, illustrating the main effect of condition

(SD = 91.68). This variable also showed a main effect of AOI [$F(1, 25) = 53.43 \ p < 0.0001$] with more time spent on the "face" (mean = 499.96 SD = 239.92) than on the "no face" AOI (mean = 107.41 SD = 117.18). Spearman correlation coefficients between the scores and the sum of fixation durations were significant only in the experimental condition for the HFASD group (Table 3). This correlation was of medium amplitude (classically defined as ranging from 0.5 to 0.8) and the difference with the correlation in the baseline condition was nearly significant ($p = 0.0559$).

The ANOVA for the number of fixations revealed a significant Group × AOI interaction [$F(1, 25) = 6.03 \ p = 0.0213$]. Post hoc comparisons revealed that the group with HFASD made more fixations on the "no face" AOI than the typical group (Fig. 3). Moreover, contrasting with the HFASD group, the typical group had significantly more fixations on the "face" than on the "no face" AOI. Notwithstanding, there was a main effect of AOI [$F(1, 25) = 40.53 \ p < 0.0001$] showing more fixations on the "face" (mean = 20.51 SD = 8.89) than on the "no face" AOI (mean = 6.70 SD = 7.20). A nearly significant main effect of Group was observed [$F(1, 25) = 4.17 \ p = 0.0519$] with a mean of 12.54 (SD = 11.55) for the typical group and 14.75 (SD = 9.48) for the HFASD group. There was a main effect of Condition [$F(2, 47) = 7.72 \ p = 0.0013$] showing a decrease during the experimental condition compared to the baseline and final conditions (Fig. 4).

Gaze Stability

The ANOVA for the average duration of fixations yielded a significant Group × Condition interaction [$F(2, 47) = 4.07 \ p = 0.0234$]. Post hoc comparisons showed that this variable increased significantly from the baseline to the experimental condition for the typical group but not for the HFASD group (Fig. 5). It decreased significantly from the experimental to the final condition for the HFASD group. There was a main effect of Condition [$F(2, 47) = 6.94 \ p = 0.0023$]. According to post hoc comparisons, the experimental condition differed significantly from the baseline condition [$t(47) = 3.65 \ p = 0.0019$] and from the final condition [$t(47) = 3.65 \ p = 0.0019$]. This variable was higher in the experimental condition (mean = 29.87 SD = 14.47) than in either the baseline condition (mean = 19.37 SD = 10.45) or the final condition (mean = 22.35 SD = 9.01). There was no significant difference between the baseline and final conditions.

To the question asked at the end of the experiment, seven typical participants and one participant with HFASD responded being aware of controlling the lens in the
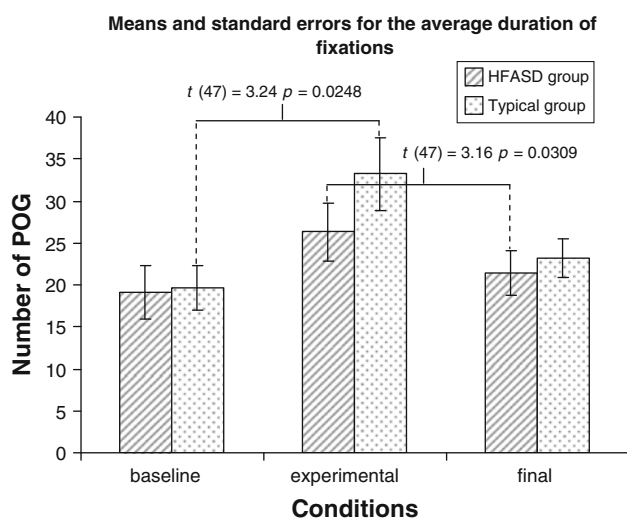
Means and standard errors for the average duration of fixations

Fig. 5 The average duration of fixations per condition per group, illustrating the Group × Condition interaction. *HFASD* High Functioning Autism Spectrum Disorders

experimental condition. The difference between groups was significant [$\chi^2 = 5.79$ $p = 0.0329$]. Detailed observation of the average duration of fixations in participants who noticed they controlled the lens showed that they increased the stability of their gaze from the baseline to the experimental condition.

Discussion

Overall, the results support our hypothesis of an impaired self-monitoring of gaze in conversational contexts in HFASD. The Group × Condition interaction found for the average duration of fixations indicates that the two groups had different stabilization behaviours when the gaze-contingent lens was introduced. The main effects of condition showed that the number of fixations decreased while their average duration increased during the experimental condition, thus supporting our assumption that the gaze-contingent lens caused gaze to stabilize. Yet, participants with HFASD showed weaker adaptation of gaze between the baseline and the experimental conditions and their average duration of fixations fell back to the baseline level when the gaze-contingent lens was removed in the final condition. The eyes' motor reactions in regard to the gaze-contingent lens highlight the high inter-individual heterogeneity among participants with HFASD and provide some evidence for an alteration in self-monitoring of eye movements in at least a sub-group. This interpretation is further corroborated by the answers to the question asked at the end of the experiment. Half of the typical participants noticed that their gaze was controlling the lens, compared to only one participant with HFASD. This lack of

awareness on the part of participants with HFASD actually supports a deficiency in the sense of agency (the awareness of being the author of one's action). Moreover, participants who noticed they controlled the lens also stabilized their gaze during the experimental condition, thus drawing a link between adapting to the lens and being aware of controlling it. An impaired sense of agency in gaze may preclude adequate eye language during social interactions. Deficient self-monitoring of gaze is most likely to entail maladaptive gaze control, thus hampering relevant social attention. Such findings should help refine current bottom-up models in ASD, such as the 'amygdala hyper-activity' (Corden et al. 2008; Dalton et al. 2005) or FTM (Senju and Johnson 2009b) models mentioned earlier.

Alterations in the monitoring of action have been suspected in ASD (Russell 1996; Russell and Jarrold 1998, 1999), though recent attempts to find evidence have not been conclusive (Russell and Hill 2001; Williams and Happé 2009). The monitoring of action is traditionally considered as belonging to the realm of executive functions, which are associated with impairments in HFASD (Hill 2004). Russell and Jarrold (1998) initially suspected a specific disorder in action-monitoring and devised an experiment where participants were to adjust a ballistic action that they had initiated for reaching a target. The authors interpreted the poorer performances of children with ASD compared to matched controls as an inability to generate a correct visual prediction of their motor command. In an ensuing experiment (Russell and Jarrold 1999), participants with ASD were playing a picture lotto game with an experimenter where they alternatively placed cards on a grid. Failures in remembering after the game which cards had been placed by them were interpreted as signalling impairment in monitoring their own action while placing cards. Russell and Hill (2001) attempted to provide direct evidence for action-monitoring impairments in ASD by testing online discrimination in a task where participants had to detect which of a number of dots moving on a computer screen was controlled by their hand hidden from their sight. Results yielded no signs of action-monitoring impairments. Williams and Happé (2009) tried to replicate the picture lotto experiment of Russell and Jarrold (1999) and the online discrimination task of Russell and Hill (2001). The outcomes contrasted with the results found by Russell and Jarrold (1999) while supporting the conclusion of Russell and Hill (2001) that there was no evidence of action-monitoring impairments in ASD. Williams and Happé (2009) suggest that the sense of agency is unaltered in HFASD. A similar view was expressed by David et al. (2008) who tested agency in HFASD with an online discrimination task of visual stimuli that were either produced by the participants moving a joystick or automatically generated. These studies relied on experimental paradigms

where participants were asked to match intended motor commands with perceptual input. Our paradigm differs in that participants did not know in advance that they controlled the lens and, therefore, were not necessarily aware of their intention to move it. To our knowledge, the present experiment is the first to provide direct evidence in favour of impairments in action-monitoring and agency.

The present study also yields informative outcomes concerning performances on the task. As expected, scores showed that the task discriminated HFASD from typical participants. The number of fixations in the HFASD group was nearly significantly superior to the typical group, although this difference may stem from their failure to adapt gaze during the experimental condition. Main effects showing more fixations on faces in terms of time and number suggest a general compliance with the task. Actually, the Group × Condition interaction for the sum of fixation durations suggests that typical participants paid less attention towards the end of the experiment, presumably because of habituation with a task that was not so challenging for them. Not surprisingly however, the HFASD group had more fixations on areas other than the face compared to the typical group, thus converging with findings by Klin et al. (2002a) and Riby and Hancock (2009). Speer et al. (2007) drew different conclusions when using stimuli with isolated characters. However, note should be taken that participants in our experiment were addressed face-to-face by the virtual character, contrasting with previous studies where participants were viewing social scenes in which they were not involved. This argument may also explain why the HFASD group did not spend less time than the typical group looking at faces, in contrast with reports by Klin et al. (2002a) and Riby and Hancock (2009). Interestingly, the time spent on fixating the face only correlated with the scores of the HFASD group in the experimental condition. The gaze-contingent lens constrained the visual field of the participants and presumably hindered compensatory strategies based on lateral vision that are often mentioned in clinical and anecdotal reports on ASD. Mottron et al. (2007) report that lateral glances are the most frequent atypical visual behaviour observed among young children with ASD. The gaze-contingent lens induced a visual behaviour whereby performances and time spent looking at faces were correlated, thus yielding a setting that might prove highly beneficial for educational purposes. Actually, scores improved significantly in the final condition, although participants were never given feedback about their answers.

The present study holds several limitations. First, the relatively small sample size in this study may obliterate some significant effects, such as a Group × Condition interaction for the number of fixations that would seem logical given the group differences in dealing with the gaze-contingent lens. Second, this study involves adolescents and adults with HFASD and additional studies would be required to extend the investigation to younger or lower functioning individuals. Third, only half of the typical participants noticed that they were controlling the gaze-contingent lens. The reason for this limited judgement of agency in the typical population should be further explored, as it could foster new relevant hypotheses for autism.

References

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition, 4*(3), 311–331.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences, 6*(6), 237–242.

Boraston, Z., & Blakemore, S. J. (2007). The application of eye-tracking technology in the study of autism. *The Journal of Physiology, 581*(Pt 3), 893–898. doi:10.1113/jphysiol.2007.133587.

Buisine, S., Wang, Y., & Grynszpan, O. (2010). Empirical investigation of the temporal relations between speech and facial expressions of emotion. *Journal on Multimodal User Interfaces, 3*(4), 263–270. doi:10.1007/s12193-010-0050-4.

Corden, B., Chilvers, R., & Skuse, D. (2008). Avoidance of emotionally arousing stimuli predicts social-perceptual impairment in Asperger's syndrome. *Neuropsychologia, 46*(1), 137–147. doi:10.1016/j.neuropsychologia.2007.08.005.

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience, 8*(4), 519–526. doi:10.1038/nn1421.

David, N., Gawronski, A., Santos, N. S., Huff, W., Lehnhardt, F.-G., Newen, A., et al. (2008). Dissociation between key processes of social cognition in autism: impaired mentalizing but intact sense of agency. *Journal of Autism and Developmental Disorders, 38*(4), 593–605. doi:10.1007/s10803-007-0425-x.

Ekman, P. (2003). *Emotions revealed*. London, UK: Weidenfeld & Nicolson.

Gepner, D., Simonin, J., & Carbonell, N. (2007). Gaze as a supplementary modality for interacting with ambient intelligence environments. In *Proceedings of human-computer interaction 2007* (pp. 848–857). Berlin: Springer, LNCS Series.

Grynszpan, O., Martin, J. C., & Nadel, J. (2008). Multimedia interfaces for users with high functioning autism: An empirical investigation. *International Journal of Human-Computer Studies, 66*(8), 628–639.

Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences, 8*(1), 26–32.

Jolliffe, T., & Baron-Cohen, S. (1999). A test of central coherence theory: Linguistic processing in high-functioning adults with autism or Asperger syndrome: Is local coherence impaired? *Cognition, 71*(2), 149–185.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002a). Visual fixation patterns during viewing of naturalistic social

situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59*(9), 809–816.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002b). Defining and quantifying the social phenotype in autism. *The American Journal of Psychiatry, 159*(6), 895–908.

Le Couteur, A., Rutter, M., Lord, C., Rios, P., Robertson, S., Holdgrafer, M., et al. (1989). Autism diagnostic interview: A standardized investigator-based instrument. *Journal of Autism and Developmental Disorders, 19*(3), 363–387.

Lindner, A., Haarmeier, T., Erb, M., Grodd, W., & Thier, P. (2006). Cerebrocerebellar circuits for the perceptual cancellation of eye-movement-induced retinal image motion. *Journal of Cognitive Neuroscience, 18*(11), 1899–1912. doi:10.1162/jocn.2006.18.11.1899.

Mottron, L., Mineau, S., Martel, G., Bernier, C. S. T., Berthiaume, C., Dawson, M., et al. (2007). Lateral glances toward moving stimuli among young children with autism: Early regulation of locally oriented perception? *Development and psychopathology, 19*(01), 23–36.

Parsons, S., Mitchell, P., & Leonard, A. (2004). The use and understanding of virtual environments by adolescents with autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 34*(4), 449–466.

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders, 32*(4), 249–261.

Raven, J. C., & Court, J. H. (1986). *Raven's progressive matrices and Raven's coloured matrices*. London: HK Lewis.

Riby, D., & Hancock, P. J. (2009). Looking at movies and cartoons: Eye-tracking evidence from Williams syndrome and autism. *Journal of Intellectual Disability Research, 53*(2), 169–181.

Robinson, D. A., Gordon, J. L., & Gordon, S. E. (1986). A model of the smooth pursuit eye movement system. *Biological Cybernetics, 55*(1), 43–57.

Russell, J. (1996). *Agency: Its role in mental development*. UK: Taylor & Francis.

Russell, J., & Hill, E. L. (2001). Action-monitoring and intention reporting in children with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(3), 317–328.

Russell, J., & Jarrold, C. (1998). Error-correction problems in autism: Evidence for a monitoring impairment? *Journal of Autism and Developmental Disorders, 28*(3), 177–188.

Russell, J., & Jarrold, C. (1999). Memory for actions in children with autism: Self versus other. *Cognitive Neuropsychiatry, 4*(4), 303–331. doi:10.1080/135468099395855.

Samson, F., Mottron, L., Soulières, I., & Zeffiro, T. A. (in press). Enhanced visual functioning in autism: An ALE meta-analysis. *Human Brain Mapping*. doi:10.1002/hbm.21307.

Sasson, N., Tsuchiya, N., Hurley, R., Couture, S. M., Penn, D. L., Adolphs, R., et al. (2007). Orienting to social stimuli differentiates social cognitive impairment in autism and schizophrenia. *Neuropsychologia, 45*(11), 2580–2588. doi:10.1016/j.neuropsychologia.2007.03.009.

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders, 10*(1), 91–103.

Senju, A., & Johnson, M. H. (2009a). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences, 13*(3), 127–134. doi:10.1016/j.tics.2008.11.009.

Senju, A., & Johnson, M. H. (2009b). Atypical eye contact in autism: Models, mechanisms and development. *Neuroscience and Biobehavioral Reviews, 33*(8), 1204–1214. doi:10.1016/j.neubiorev.2009.06.001.

Slater, A. (1999). *Perceptual development: Visual, auditory and speech perception in infancy*. UK: Psychology press.

Speer, L. L., Cook, A. E., McMahon, W. M., & Clark, E. (2007). Face processing in children with autism: Effects of stimulus contents and type. *Autism, 11*(3), 265–277. doi:10.1177/1362361307076925.

Spezio, M. L., Adolphs, R., Hurley, R., & Piven, J. (2007). Abnormal use of facial information in high-functioning autism. *Journal of Autism and Developmental Disorders, 37*(5), 929–939. doi:10.1007/s10803-006-0232-9.

Tager-Flusberg, H. (2000). Language and understanding minds: Connections in autism. In Baron-Cohen S, Tager-Flusberg H, Cohen DJ (Eds.), *Understanding other minds, 2nd edn* (pp. 124–149). New York: Oxford University Press.

van der Geest, J. N., Kemner, C., Verbaten, M. N., & van Engeland, H. (2002). Gaze behavior of children with pervasive developmental disorder toward human faces: a fixation time study. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 43*(5), 669–678.

Wechsler, D. (1997). *Wechsler adult intelligence scale-III (WAIS-III)*. San Antonio, TX: The Psychological Corporation.

Williams, D., & Happé, F. (2009). Pre-conceptual aspects of self-awareness in autism spectrum disorder: The case of action-monitoring. *Journal of Autism and Developmental Disorders, 39*(2), 251–259. doi:10.1007/s10803-008-0619-x.

Wolpert, D. M., & Miall, R. C. (1996). Forward Models for Physiological Motor Control. *Neural Networks: The Official Journal of the International Neural Network Society, 9*(8), 1265–1279.