

## Tools of the Trade

# It's in your eyes—using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience

Marcus Wilms,<sup>1,\*</sup> Leonhard Schilbach,<sup>2,\*</sup> Ulrich Pfeiffer,<sup>2</sup> Gary Bente,<sup>3</sup> Gereon R. Fink,<sup>1,4</sup> and Kai Vogeley<sup>2</sup>

<sup>1</sup>Institute of Neurosciences and Medicine, Research Centre Juelich, Juelich, <sup>2</sup>Department of Psychiatry, <sup>3</sup>Department of Social Psychology, and <sup>4</sup>Department of Neurology, University of Cologne, Cologne, Germany

**The field of social neuroscience has made remarkable progress in elucidating the neural mechanisms of social cognition. More recently, the need for new experimental approaches has been highlighted that allow studying social encounters in a truly interactive manner by establishing 'online' reciprocity in social interaction. In this article, we present a newly developed adaptation of a method which uses eyetracking data obtained from participants in real time to control visual stimulation during functional magnetic resonance imaging, thus, providing an innovative tool to generate gaze-contingent stimuli in spite of the constraints of this experimental setting. We review results of two paradigms employing this technique and demonstrate how gaze data can be used to animate a virtual character whose behavior becomes 'responsive' to being looked at allowing the participant to engage in 'online' interaction with this virtual other in real-time. Possible applications of this setup are discussed highlighting the potential of this development as a new 'tool of the trade' in social cognitive and affective neuroscience.**

**Keywords:** magnetic resonance imaging-compatible interactive eyetracking; truly interactive paradigms; gaze feedback; social cognition

### INTRODUCTION

Social neuroscience has helped to shed light upon the neural mechanisms underlying our ability to understand other minds under the headings of 'theory of mind' or 'mentalizing', commonly understood as the ability to represent other people's mental states (Frith and Frith, 2003, 2008). An increasing number of functional neuroimaging studies suggests that two large-scale neural networks are involved, namely the so-called mirror neuron system, comprising essentially the parietal and premotor cortices and the so-called 'social brain', comprising essentially the medial prefrontal, the temporopolar, the temporoparietal cortices and the amygdala (Keysers and Gazzola, 2007; Lieberman, 2007).

Most of these studies, however, rely on paradigms in which participants are asked to merely observe others ('offline' mentalizing; Schilbach *et al.*, 2006), while mentalizing during 'online' social interaction has only been studied

by a minority of studies (e.g. Montague *et al.*, 2002; Eisenberger *et al.*, 2003), often making use of game theory paradigms from economics (e.g. Sanfey, 2007). Consequently, the need to develop 'interactive mind' paradigms that could provide a platform to systematically study the neural mechanisms of social interaction in an ecologically valid manner has been pointed out recently (Singer, 2006). 'Online' interaction crucially involves 'closing the loop' between interaction partners and establishing reciprocal relations where actions feed directly into the communication loop and elicit reactions which, in turn, may subsequently lead to reactions of the initiator and so forth. This has been referred to as adopting a 'second-person-perspective' (2PP; Reddy, 2003) which can be taken to suggest that awareness of mental states results from being psychologically engaged with someone and being an active participant of reciprocal interaction thereby establishing a subject-subject ('Me-You') rather than a subject-object ('Me-She/He') relationship. Paradigms that permit the systematic investigation of the reciprocity of interactions as well as the involvement of implicit and explicit processes will substantially enrich our knowledge of the neurobiology of social cognition (Frith and Frith, 2008).

The challenge for social neuroscience here will be twofold: a suitable experimental platform should allow real-time, 'online' interactions between participants and the social

Received 20 November 2009; Accepted 17 February 2010

Advance Access publication 11 March 2010

We are grateful to our colleagues from the MR group at the Institute for Neurosciences and Medicine, in particular Ralph Weidner and Sharam Mirzazade, for their assistance in acquiring eyetracking data inside the scanner. The study was supported by a grant of the German Ministry of Research and Education (BMBF) on 'social gaze' (G.B. and K.V.). L.S. was also supported by the Koeln Fortune Program of the Faculty of Medicine, University of Cologne.

Correspondence should be addressed to Leonhard Schilbach, Department of Psychiatry, University Hospital Cologne, Kerpener Str. 62, 50924 Cologne, Germany. E-mail: leonhard.schilbach@uk-koeln.de

\*These authors contributed equally to this work.

stimuli while also providing means for experimental control over changes of the social stimuli. Here, we suggest that measurements of participants' gaze behavior inside the MR environment could be used to influence a virtual character's gaze behavior making it 'responsive' to the participant's gaze allowing to engage in 'online' interaction with the virtual other in real-time.

From a theoretical standpoint, gaze is known to be an important social cue in dyadic interaction indicating interest or disinterest, regulating intimacy levels, seeking feedback and expressing emotions (e.g. Argyle and Cook, 1976; Emery, 2000). In addition, gaze can also influence object perception (Becchio, *et al.*, 2008) by means of establishing triadic relations between two observers and an object onto which the interactors can look 'together' and thereby establish 'joint attention' (Moore and Dunham, 1995). From a methodological standpoint, gaze behavior represents one of the few ways in which participants can interact with stimuli naturally in spite of the movement constraints when lying inside an MR scanner. Gaze is a socially most salient non-verbal behavior, which can be reliably measured even within an functional magnetic resonance imaging (fMRI) setting and can, therefore, be used as input for interactive paradigms in neuroimaging studies.

This leads to the further question of how gaze input can be used to automatically control contingent behavior of a social stimulus to create the illusion of real-time interaction in such a setting. Pre-programmed and strictly controlled visual presentation of nonverbal behaviors in general and of gaze cues in particular, can be established by using anthropomorphic virtual characters. Such computer-generated characters have been suggested as a valuable tool for social neuroscience (e.g. Sanchez-Vives and Slater, 2005) as they convey social information to human observers (e.g. Bente *et al.*, 2001; Bente *et al.*, 2002, 2008a; Bailenson *et al.*, 2003) and cause reactions strikingly similar to those in real human interactions (e.g. Slater *et al.*, 2006). An increasing number of behavioral as well as fMRI studies has now used such stimuli to study different aspects of social cognition including gaze perception (e.g. Pelphrey *et al.*, 2005; Spiers *et al.*, 2006; Schilbach *et al.*, 2006; Bente *et al.*, 2007, 2008b; Park *et al.*, 2009; Kuzmanovic *et al.*, 2009).

To establish a paradigm that actively engages participants in the 2PP as opposed to being a 'detached' observer of social stimuli (from a 'third-person-perspective'; 3PP), we present a new technique that makes use of eyetracking data obtained from participants inside an MR scanner to control a virtual character's gaze behavior in real-time making it 'responsive' to the human observer's gaze (Schilbach *et al.*, in press). In this setup the eye movements of the participant become a means to 'probe' the behavior of the virtual other similar to real-life situations. Importantly, this also seems compatible with an 'enactive' account of social cognition which understands social cognition as bodily experiences resulting from an organism's adaptive actions upon salient and self-relevant

aspects of the environment (e.g. Klin *et al.*, 2003) that feed back into the social interaction process.

Consequently, our setup promises to allow the exploration of the neural basis of processes of interpersonal alignment and the reciprocity inherent to social interaction, i.e. whether and how social cues are detected as contingent upon one's own behavior and how interaction partners initiate and respond to each other's actions (Schilbach *et al.*, in press). Both aspects seem highly relevant to make substantial progress in the field of social cognitive neuroscience and may lead to a reconsideration of the current emphasis on similarities between self- and other-related processing.

To implement these different eyetracking setups were tested in the fMRI setting to produce gaze-contingent stimuli. We review results of these different approaches which use eyetracking measurements overtly or covertly to drive MR-compatible experimental paradigms and underline the usability of this technique. Furthermore, we give examples for the applications of these interactive, eyetracking-based paradigms. Given the importance of gaze behavior during real-life social interaction, this approach, we suggest, provides a much needed, new 'tool of the trade' for the study of real-time 'online' interaction in social neuroscience.

## METHODS

### Interactive eyetracking setups

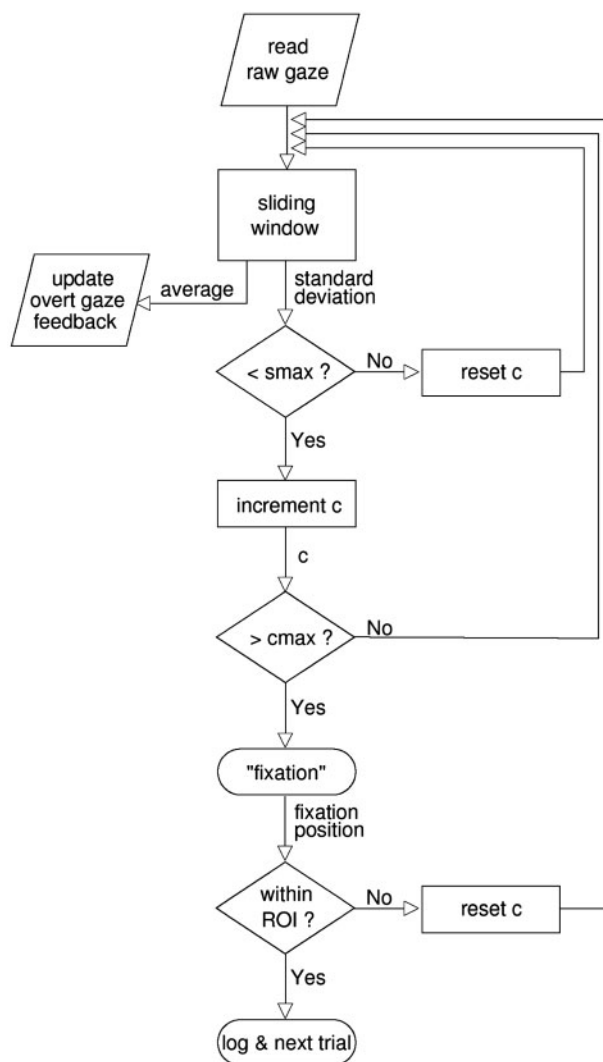
'Interactive eyetracking' relies on an MR-compatible eyetracking system that allows real-time data transmission to a visual stimulation controller. The controller receives the ongoing gaze data and adapts the visual stimulation according to preset task conditions and the volunteer's current gaze position on screen.

For stimulus delivery, different presentation devices were tested employing either a TFT screen or two different goggle systems. First, a custom-built, shielded TFT screen was used for the stimulus presentation at the rear end of the scanner ( $14^\circ \times 8^\circ$  horizontal  $\times$  vertical viewing angle, screen distance from volunteer's eyes: 245 cm). Volunteers watched the stimuli via a mirror mounted on the head coil. Volunteers' eye movements were monitored by means of an infrared camera (Resonance Technology, CA, USA). The camera and infra-red light source were mounted on the head coil using a custom-built gooseneck that allowed easy access to the volunteer's eyes without interfering with the visual stimulation (setup A). Second, stimuli were presented using MR-compatible goggles. Volunteers' eye movements were monitored by means of an infrared camera that was built into the goggles. In a 3T MR environment we used a VisuaStim<sup>TM</sup> system ( $30^\circ \times 22.5^\circ$  horizontal  $\times$  vertical viewing angle; Resonance Technology, CA, USA; setup B<sub>1</sub>) whereas in a 1.5T environment we tested a Silent Vision<sup>TM</sup> ( $25.5^\circ \times 18^\circ$  horizontal  $\times$  vertical viewing angle; AvoTec, FL, USA; setup B<sub>2</sub>). The raw analog video signals of all setups were digitized at a frame rate of 60 Hz on a dedicated PC running a gaze extraction software

(iViewX<sup>TM</sup>, SMI, Germany, and Clearview<sup>TM</sup>, Tobii Technology AB, Sweden, respectively) which produced real-time gaze position output. Careful eyetracking calibration was performed prior to each training or data acquisition session in order to yield valid gaze positions in a stimulus-related coordinate system. Via a fast network connection, gaze position updates were transferred and, thus, made available to another PC running the software which controlled the stimulation paradigm (Presentation<sup>TM</sup>, <http://www.neurobs.com>).

### Interactive eyetracking with overt feedback

This version was established so that study participants could engage in cognitive tasks using their eye movements only while receiving visual feedback. To this end we coupled the volunteer's current gaze position to the location of a cursor-like object on the screen (hereafter: gaze cursor). Using their eyes, the volunteers could, thus, voluntarily move the gaze cursor according to the demands of the tasks. For the automatic detection of gaze fixations on screen targets in real-time the following computer-based algorithm was devised: Using Presentation<sup>TM</sup> software, gaze positions were transformed into stimulus screen coordinates (pixels). A continuously proceeding 'sliding window' average of the preceding 60 gaze positions was calculated throughout the whole stimulus presentation (Figure 1). In effect, the gaze cursor marked the volunteers' average gaze position within the preceding 1 s time window providing the observer with a smooth gaze-contingent visual stimulus to which the volunteers quickly adapted despite a brief temporal lag. In particular, this procedure lessened blinking artifacts, averaged out fixational eye movements (2–120 arcmin; Martinez-Conde *et al.*, 2004), and attenuated the impact of erroneous gaze estimates caused by intermittent residual imaging artifacts in the eye video signal. Each 'sliding window' average was tested for being part of a coherent fixation period or not and was accepted by the algorithm as part of an ongoing fixation, if the standard deviation of the sliding window gaze elements was below a pre-specified threshold, in which case a counter was incremented. If the standard deviation criterion was not fulfilled, the counter was reset to zero. This procedure was repeated until a fixation period of a pre-specified length, i.e. a pre-specified number of consecutive sliding window averages, was detected. This procedure reliably recognized effective fixations from gaze behavior without prior knowledge of fixation coordinates. Fixations were subsequently tested for being within one of a set of predefined region-of-interests (ROIs) on the stimulus screen. If this was not the case, the algorithm searched for another fixation. This cycle was repeated until either a fixation was found that was within one of the predefined ROIs or the maximum duration of the current task was reached. Time stamps as well as coordinates of detected fixations were stored in a text file for offline data analysis.

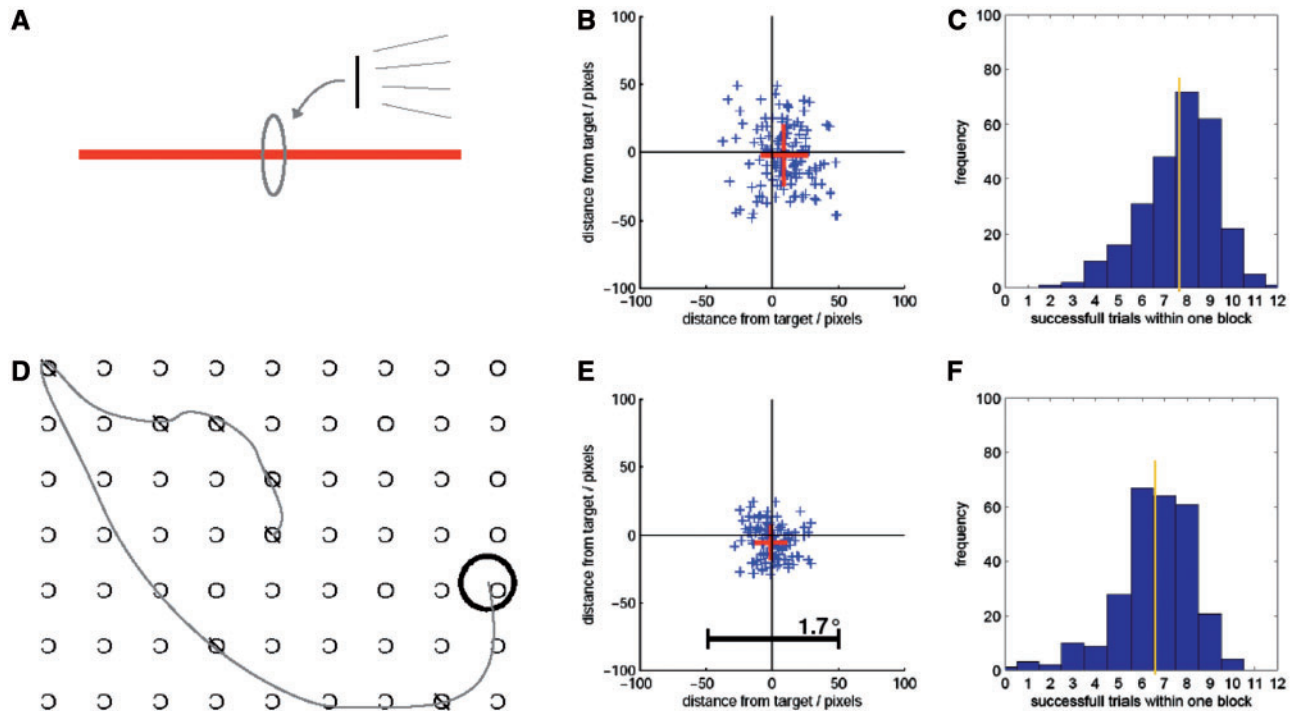


**Fig. 1** Flow diagram of gaze data processing. Raw gaze data is stored in a sliding window vector whose average value is shown as a gaze cursor in overt gaze feedback paradigms. A counter ( $c$ ) increments for each consecutive updated sliding window data vector whose standard deviation is below a prespecified threshold ( $smax$ ). The gaze status is identified as 'fixation' if a prespecified number ( $cmax$ ) of consecutive sliding windows is reached. The average gaze position is thereafter tested for being within one of a given set of ROIs. The described procedure is typically run until a fixation was found within one of the given ROIs triggering a step forward in the experimental paradigm, e.g. the presentation of a new visual stimulus or the change of the gaze direction of a virtual character.

### Interactive eyetracking with covert feedback

As during interactive eyetracking with overt feedback, here, participants engage in and 'drive' an experimental paradigm by looking at different locations on the screen. In this version of the setup, however, they do not receive visual feedback in form of a gaze cursor.

In conjunction with a virtual character whose gaze behavior could be made contingent upon fixations detected in ROIs this was done to generate an ecologically valid setting in which the gaze behavior of the virtual other could change in response to the human observer's gaze position on the



**Fig. 2** Overt gaze feedback tasks. (A) While being scanned using ‘dummy’ fMRI scans, a subject was instructed to bisect a horizontal line with a vertical line which is locked to her gaze. In effect, the subject performed the task by fixating the perceived line center. The screen positions of the horizontal lines were randomized. (B) Example of the spatial precision of one subject performing the bisection task. Each of the blue symbols represents one bisection position relative to the true line center. The red cross denotes the average and standard deviation of the spatial bisection error. (C) This histogram shows the frequency distribution of the number of line bisections that subjects were able to perform within one task/block length (21.9 s). On average over 18 subjects, each performing the tasks 15 times, 7.7 line bisections were successfully performed within one block length revealing a rather fluent task performance. (D) In a second task, the subject was asked to cancel targets (‘O’) among distractors by centering the black circle over each target until marked as cancelled. Since the circle’s position was locked to the subject’s gaze she only had to find and fixate the targets one by one. (E) The spatial precision of the same subject as in (B) performing the cancellation task. (F) On average over 18 subjects, 6.6 cancellations were successfully performed within one block.

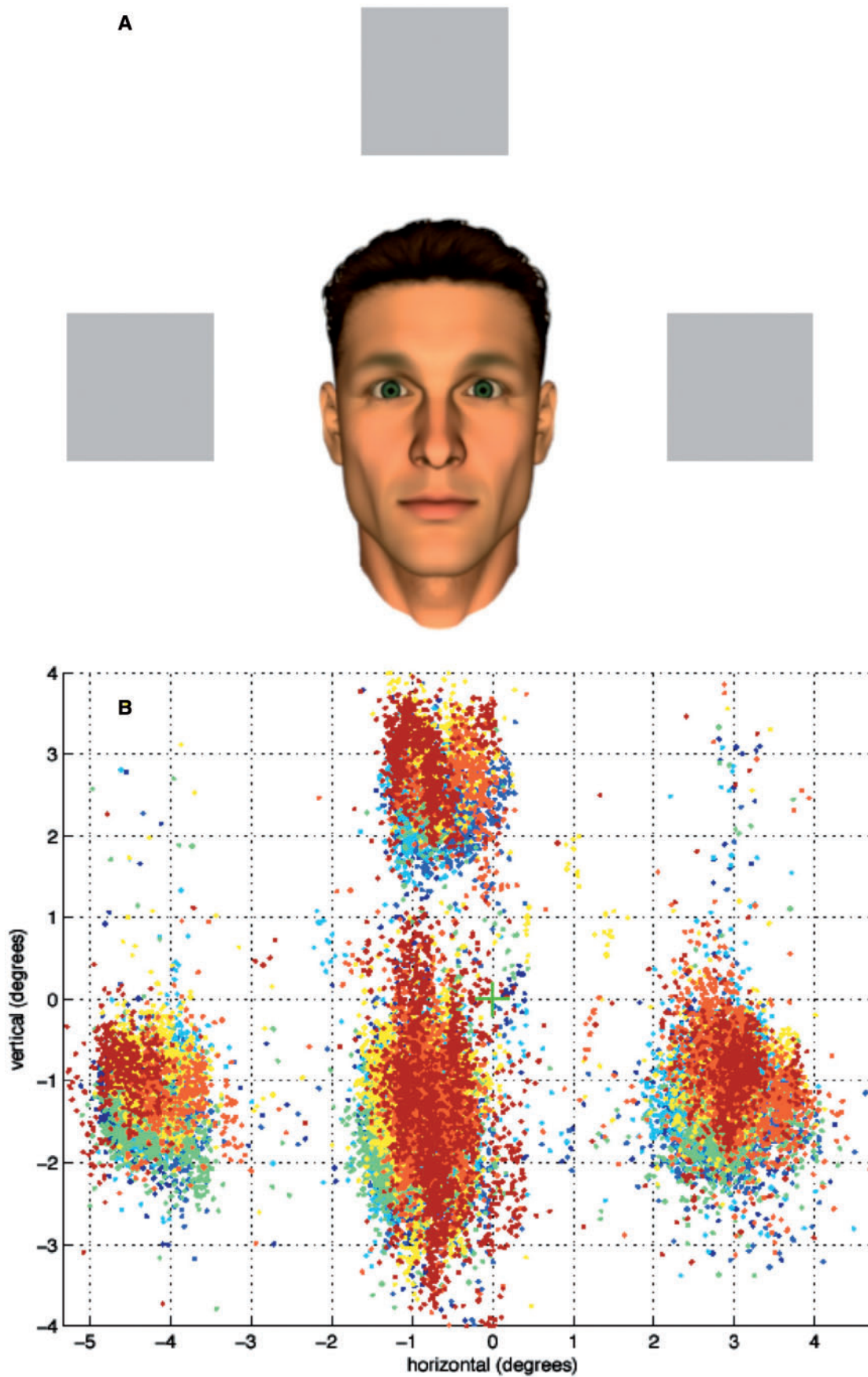
stimulus screen. In this setup the temporal delay between a relevant fixation and the reaction of the virtual character needed to be small to successfully induce a fluent experience of reciprocal exchange between the participant and the virtual character. Not providing continuous visual feedback via a gaze cursor precluded the possibility of participants being distracted by the gaze cursor but also to adapt to possible measurement errors, even if minimal in size. For example, participants’ head movements or variations in eye illumination could invalidate the initial eye tracking calibration and lead to displaced gaze coordinates. This type of technical problem could be met by defining larger ROIs that were less sensitive to distortions in gaze estimations. Despite minimal offset errors in the gaze coordinates, thus, meaningful reactions of the virtual character were still possible.

### The tasks inside the MR environment

To make use of the *overt* interactive eyetracking mode we implemented two tasks resembling clinical bedside tests for visuo-spatial neglect. In a line bisection task participants had to bisect a horizontal line by fixating it centrally thereby moving a gaze cursor (a vertical line) on the screen into the desired position (Figure 2). In a target cancellation task volunteers had to search for and single out randomly

distributed targets among distractors by fixating them one by one thereby moving a gaze cursor (a circle) over each of the detected targets (see online supplementary data for video of task performance). Participants were informed about the paradigm prior to entering the scanner room. Both overt interactive eyetracking tasks were tested with visual stimulation delivered via the TFT screen (setup A) and via goggles (setups B<sub>1</sub> and B<sub>2</sub>). After careful calibration of the eyetracking, volunteers were allowed to get adjusted to the procedure and then went on to perform the task while lying inside the scanner. During this period we ran ‘dummy’ EPI sequences with MRI parameters identical to those in standard imaging experiments. We thereby introduced EPI artifacts in the eyetracking data to test that the devised algorithm would be able to successfully cope with the added noise.

In order to test the *covert* interactive eyetracking mode we made use of a task in which test subjects were asked to respond to or probe the gaze behavior shown by an anthropomorphic virtual character on screen (Figure 3; Schilbach *et al.*, in press). Before participation test subjects were instructed that the gaze behavior shown by the virtual character on screen was actually controlled by a real person who was also participating in the experiment outside the scanner. Likewise, their own gaze behavior was said to be visualized



**Fig. 3** Covert gaze feedback task. (A) Screenshot (as seen by participants) depicting anthropomorphic virtual character and three objects. (B) Illustration of gaze samples obtained for one exemplary participant during the experiment.

for the other participant outside the MR scanner, so that both participants could engage in gaze-based 'online' interaction. During functional neuroimaging participants were instructed to direct the gaze of the other person towards one of three objects by looking at it. Alternatively, they were asked to respond to gaze shifts of the virtual character by either following or not following them to then look at an object. The gaze behavior of the other was made contingent to the participant's gaze and systematically varied in a  $2 \times 2$  factorial design (joint attention *vs* non-joint attention; self-initiation *vs* other-initiation; see online supplementary data for video of task performance). The neural correlates of task performance were investigated employing fMRI in 21 participants.

In this task, stimuli were presented to the participants lying inside the MR scanner using setup A. Due to the screen's distance from the volunteers' eyes and the corresponding narrow field of view; changes of the virtual character's gaze behavior were easily observable while focusing on one of the three objects. Functional MRI (fMRI) data was acquired on a Siemens Trio 3T whole-body scanner (Erlangen, Germany) using blood-oxygen-level-dependent (BOLD) contrast (Gradient-echo EPI pulse sequence, TR = 2.304 s, slice thickness 3 mm, 38 axial slices, in-plane resolution  $3 \times 3$  mm). Additional high-resolution anatomical images (voxel size  $1 \times 1 \times 1$  mm<sup>3</sup>) were acquired using a standard T<sub>1</sub>-weighted 3D MP-RAGE sequence. The neuroimaging data was preprocessed and analyzed using a general linear model (GLM) as implemented in SPM5 (for further details see: Schilbach *et al.*, in press).

## RESULTS

Interactive eyetracking under the constraints of fMRI was successfully installed in all setups. The quality depended on the length of the sliding window which in turn depended on the raw gaze data variance, the main source of which were residual imaging artifacts. Fixational eye movements were the other important source of raw gaze data jitter. The amplitude of such eye movements are in the range of 2–120 arcmin (Martinez-Conde *et al.*, 2004) and their size in screen pixel coordinates scales with the eye's distance to the screen display. This rendered the fixational eye movements' impact particularly disadvantageous in setup A where a screen distance of 245 cm translated the range of 2–120 arcmin to 2–120 pixels on the screen (given a  $800 \times 600$  pixels resolution). On the other hand, setup A did also include important advantages regarding the handling of the eyetracking camera compared to setups using goggles. The distance as well as the angle with which the camera was positioned in relation to the test subject's eyes region could be adjusted more easily when the eyetracking camera was mounted to the head coil as in the TFT-based setup. However, goggle systems that allowed a precise eyetracking with their built-in camera did hardly need any camera adjustments (e.g. setup B<sub>2</sub>). After the feasibility of

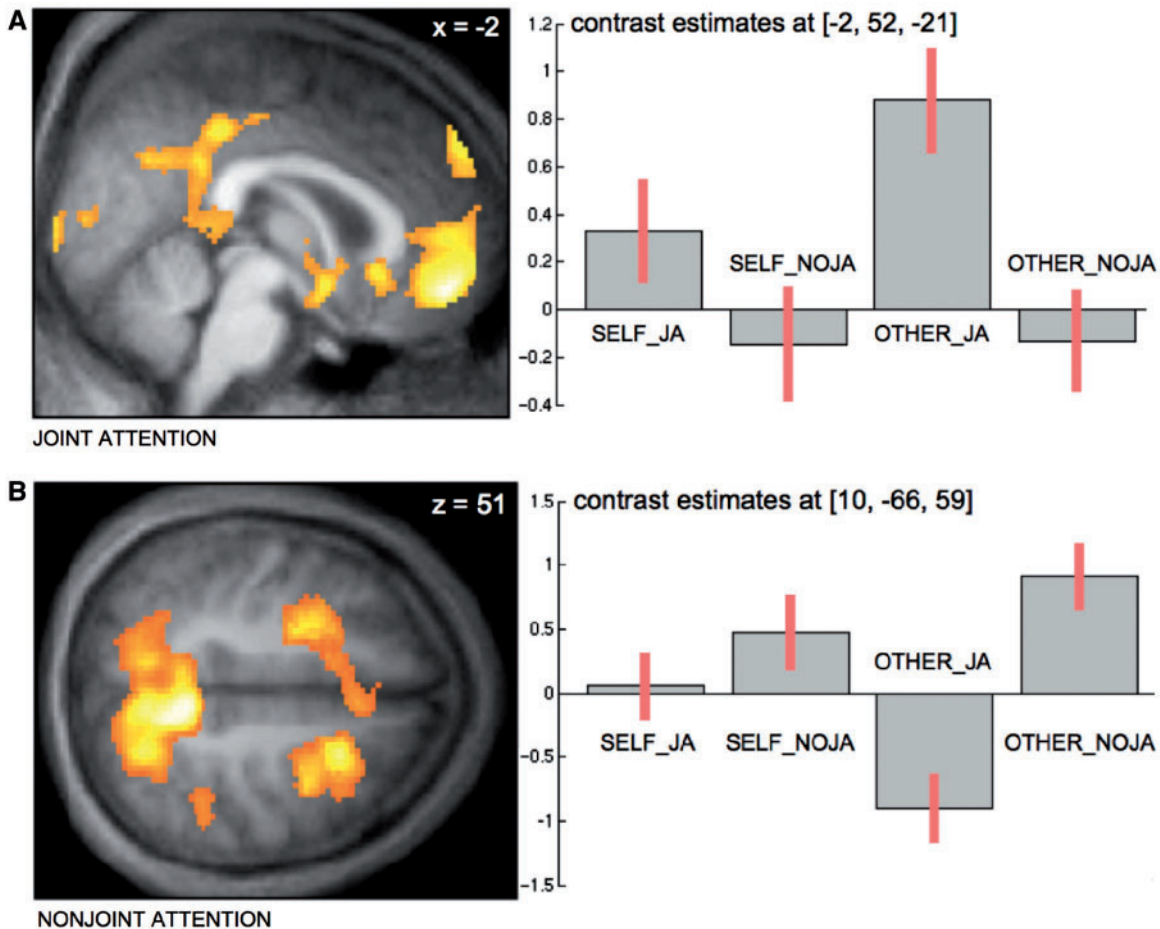
all setups was secured we chose to use our 3 T MRI system and opted for the TFT-based visual stimulation setup for data acquisition. We nevertheless stress, that the other setups allowed running the paradigms as well.

### Interactive eyetracking with overt feedback (line bisection and target cancellation)

During the task in which overt feedback was given, participants were able to use the gaze cursor which they could move across the screen in concordance with their eye movements to execute the task. For the bisection task this amounted to subjects completing an average of 7.7 line bisections in blocks of 21.9 s ( $n = 18$  subjects), whereas during the cancellation task subjects were able to cancel an average of 6.6 targets within the same time (Figure 2). Subjects needed an average of 2.8 s to judge the center of a given line and position the gaze cursor in the respective position. An average additional 0.5 s was needed to search, find, and position the gaze cursor in cancellation tasks (average time between cancellations: 3.3 s). The time subjects needed to choose targets generally depended on the length of the sliding window. If the sliding window length was too small, increased residual gaze cursor jittering made it hard for volunteers to 'focus' a target, whereas too long a sliding window increased the temporal lag of the gaze cursor, which reduced the intuitive usability of the gaze feedback. Apart from this, the spatial precision of target choices was increased when the eye tracking calibration was optimal and subjects were well adjusted to the temporal lag of the gaze cursor (Figure 2).

### Interactive eyetracking with covert feedback (joint attention)

Having been able to use overt gaze feedback successfully to drive an experiment, we went on to perform the joint attention task which included running regular fMRI measurements. During this task subjects were not given visual feedback in the form of a gaze cursor because we wanted to create a naturalistic as possible setup which allowed for an immersive experience during which participants could interact with the virtual other similarly to how one might interact with another person by means of gaze behavior in real life. In spite of the absence of continuous visual feedback participants were able to fulfill the task in which they had been asked to engage: they were able to establish 'eye-contact' with the virtual character and to respond to the virtual character either by following or not following its gaze to either fixate one of three visible objects 'together with' the virtual character (other-initiated joint attention: OTHER\_JA) or not (other-initiated nonjoint attention: OTHER\_NOJA; Figure 3A). Conversely, they were also able to establish 'eye-contact' and subsequently direct the virtual character's gaze towards one of the three objects (self-initiated joint attention: SELF\_JA). In an equal number of occasions subjects were unable to do so as the character would 'react' by fixating an object other than the one chosen by the



**Fig. 4** Neural correlates of joint attention task. (A) Differential increase of neural activity in MPFC, PCC as well as ventral striatum and anterior temporal poles (latter not illustrated here) for main effect of joint attention. (B) Differential increase of neural activity in medial and lateral parietal as well as frontal cortex bilaterally for main effect of nonjoint attention (taken from: Schilbach *et al.*, in press).

participant (self-initiated nonjoint attention: SELF-NOJA). On average, this procedure amounted to approximately four object fixations per ‘interaction segment’ (i.e. a block of 18 s duration) for all conditions (SELF\_JA: 4.10 ( $n=21$ , s.d. = 0.68), SELF\_NOJA: 4.06 ( $n=21$ , s.d. = 0.79), OTHER\_JA: 3.96 ( $n=21$ , s.d. = 0.88), OTHER\_NOJA: 4.03 ( $n=21$ , s.d. = 0.88); see Figure 3B for exemplary gaze data; for more details see Schilbach *et al.*, in press). fMRI results demonstrated, firstly, that interpersonal gaze coordination and ‘joint attention’ (main effect of JA) resulted in a differential increase of neural activity in the medial prefrontal cortex (MPFC) and posterior cingulate cortex (PCC) as well as the anterior temporal poles bilaterally. While this activation pattern bears some resemblance with the ‘default mode of brain function’ (Raichle *et al.*, 2001; Schilbach *et al.*, 2008), activations in ventral and dorsal MPFC—at voxel-level correction for multiple comparisons—have been related to outcome monitoring and the understanding of communicative intent (Amodio and Frith, 2006) as well as representations of triadic relations (Saxe, 2006). Conversely, looking at an object different from the one attended by the virtual character—regardless of whether or not this was self-initiated

(main effect of NOJA)—recruited a bilateral fronto-parietal network known to be involved in attention and eye-movement control (Schilbach *et al.*, in press; Figure 4).

## DISCUSSION

Here, we have presented a method by which eyetracking data obtained from study participants lying inside an MR scanner can be processed in real-time in order to directly influence visual stimulus material in spite of the electromagnetic noise associated with fMRI measurements. This can be realized by using a form of *overt* gaze feedback, i.e. a gaze cursor which subjects can control with their eye movements to carry out a task. Alternatively, this can be done by means of *covert* gaze feedback where gaze data is used to systematically manipulate and drive the visual stimulation unbeknownst to the participant. In combination with the presentation of anthropomorphic virtual characters whose behavior can be made responsive to the participants’ fixations, the latter technique can be used to allow participants to engage in reciprocal ‘online’ interaction with a virtual other similar to instances of interpersonal gaze coordination in real-life social

encounters. Even though similar approaches of using gaze-contingent stimuli have been used in other areas of research (e.g. Duchowski *et al.*, 2004), the development of an MR-compatible version of the technique as presented here is of crucial importance for social cognitive and affective neuroscience as it allows to experimentally target the neural underpinnings of processes during 'online' interaction which have so far been largely inaccessible due to the technical constraints of the MR environment.

Different possible paradigms come to mind which could benefit from making use of the here described method. Given the scope of this article, we will limit our description to interactive paradigms in which gaze behavior is exchanged between a human observer and a virtual character. We will focus here on *dyadic interaction* between two interactors ('Me-You'), but also on *triadic interaction* where two interactors relate to an object in the environment ('Me-You-This'; Saxe, 2006).

Within dyadic interaction gaze is known to have important regulatory functions impacting on a wide range of cognitive, affective and motivational processes (Argyle and Cook, 1976; Emery, 2000). Furthermore, gaze is known to influence our social perception and evaluation of others (e.g. Macrae *et al.*, 2002; Mason *et al.*, 2005) as it conveys the direction of an agent's attention and has been suggested to be closely related to mentalizing, i.e. our ability to understand other people's mental states (Nummenmaa and Calder, 2009). Importantly, gaze is also known to 'connect' human beings in everyday life situations by means of a 'communication loop' in which interactors impact reciprocally on each others' behavior (e.g. Frith, 2007, p. 175). This procedural dimension of 'social gaze' in 'online' interaction has only recently begun to be systematically investigated (e.g. Senju and Csibra, 2008) and promises to allow radically new insights into the temporal dynamics of implicit interpersonal 'alignment'.

As many previous studies concerning the social effects of gaze on person perception have used static, non-interactive stimuli, it may be important to revisit these paradigms by making use of this new technique to validate whether the findings actually result from social communicative processes or not. Here, it is important to note, however, that our setup in its current version does not allow to investigate real social interaction (as in the setup used by Montague *et al.*, 2002), but uses anthropomorphic virtual characters in conjunction with a cover story to generate the impression of interacting with a 'mindful' agent. While this can be seen as a limitation of our setup, it is important to note that using gaze feedback has the important benefit of enabling the systematic study of a major component of real interpersonal interaction as it provides a naturalistic way to engage participants. Furthermore, future research could explicitly address how variations of the temporal and stochastic characteristics of a virtual character's gaze behavior made contingent upon the human observer's gaze impact on a human

observer's perception of the nature of the agent ('social' Turing test).

Apart from aspects related to dyadic interaction, gaze is also known to contribute to the establishment of triadic relations between two interactors who can look at an object together and engage in (gaze-based) joint attention (Moore and Dunham, 1995). Apart from the convergence of gaze directions, this, importantly, also requires mutual awareness of being intentionally directed towards the same aspect of the world which may result directly from the process of interaction. Therefore, joint attention can be construed as an interactively constituted phenomenon whose different facets can only be explored by making use of an interactive paradigm (e.g. Schilbach *et al.*, in press). Interestingly, it has been suggested that being actively engaged in triadic interaction may have an impact both on the perception of the other person (e.g. his/her trustworthiness and attractiveness) as well as on the perception of an object (e.g. its value) that may be jointly attended (Heider, 1958).

To the best of our knowledge, there are no neuroimaging studies which have targeted the neural correlates of the perception of jointly attended objects. Such investigations might be extremely informative, however, by allowing the empirical investigation of the neural correlates of different formats or varying degrees of shared intentionality and could help to investigate the complex relationship of implicit and explicit processes involved in social cognition (Frith and Frith, 2008). Further applications of the method could include investigations of how interactive gaze cues shown by a virtual character impact on object-related decision-making or memory performance. Finally, gaze-based triadic interaction could also help to disentangle differences between 'online' and 'offline' social cognition, i.e. social cognition from a second- or third-person-perspective, by realizing interactions between two virtual agents and a human observer while introducing systematic differences in social responsiveness of the agents making them more or less likely to actually engage with the participant (Eisenberger *et al.*, 2003).

Apart from investigations of dyadic and triadic interaction in healthy adults, we also see great potential in using the above described method to investigate social cognition in development and diagnostic groups. In spite of the importance of joint attention in ontogeny the neural correlates of this significant phenomenon are only incompletely understood. Given that our paradigm does not rely upon verbal information and higher-order reasoning about others' mental states, but relies upon naturally occurring social behavior, it might prove to be particularly helpful for the study of the neurofunctional substrates of the development of social cognitive and perceptual abilities during ontogeny.

Specific alterations of social cognition are known to be characteristic of psychiatric disorders such as autism and schizophrenia. In the former case, a dissociation between implicit and explicit processes underlying social cognition



has recently been emphasized (Senju *et al.*, 2009). Also, it has been suggested that autistic individuals might be more sensitive to perfect, non-social as compared to imperfect, social contingencies in the environment (Gergely, 2001; Klin *et al.*, 2009). We suggest that the investigation of the neural mechanisms underlying these clinically relevant differences in high-functioning autism will benefit substantially from the method described here (see also Boraston and Blakemore, 2007).

## CONCLUSIONS

Taken together, we have shown that the interactive eyetracking method here presented allows to generate gaze-contingent stimuli during fMRI in spite of the electromagnetic noise introduced by such measurements. Used in conjunction with anthropomorphic virtual characters whose behavior can be made 'responsive' to the participant's current gaze position, this method has the potential to substantially increase our knowledge of the neural mechanisms underlying social cognition by making psychological processes accessible for empirical investigation that are involved in the interpersonal coordination of gaze behavior, both in dyadic and triadic interaction. Making use of this new 'tool of the trade', we suggest, could open up an entire new avenue of research in social cognitive and affective neuroscience.

## SUPPLEMENTARY DATA

Supplementary data are available at SCAN online.

## REFERENCES

- Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Review Neuroscience*, 7(4), 268–77.
- Argyle, M., Cook, M. (1976). *Gaze and Mutual gaze*. Cambridge: Cambridge University Press.
- Bailenson, J.N., Blascovich, J., Beall, A.C., Loomis, J.M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, 29(7), 819–33.
- Becchio, C., Bertone, C., Castiello, U. (2008). How the gaze of others influences object processing. *Trends in Cognitive Science*, 12(7), 254–8.
- Bente, G., Krämer, N.C., Petersen, A., de Ruiter, J.P. (2001). Computer animated movement and person perception: methodological advances in nonverbal behavior research. *Journal of Nonverbal Behavior*, 25(3), 151–66.
- Bente, G., Krämer, N.C. (2002). Virtual gestures: Analyzing social presence effects of computer-mediated and computer-generated nonverbal behaviour. In: Gouveia, Feliz, R., editors. *Proceedings of the Fifth Annual International Workshop PRESENCE, October, 2002, Porto, Portugal*. [http://www.temple.edu/ispr/prev\\_conferences/proceedings/2002/Final%20papers/Bente%20and%20Kramer.pdf](http://www.temple.edu/ispr/prev_conferences/proceedings/2002/Final%20papers/Bente%20and%20Kramer.pdf) (15 Dec 2009, date last accessed).
- Bente, G., Eschenburg, F., Aelker, L. (2007). Effects of simulated gaze on social presence, person perception and personality attribution in avatar-mediated communication. *Proceedings of the 10th Annual International Workshop on Presence 2007*. [http://www.temple.edu/ispr/prev\\_conferences/proceedings/2007/Bente,%20Eschenburg,%20and%20Aelker.pdf](http://www.temple.edu/ispr/prev_conferences/proceedings/2007/Bente,%20Eschenburg,%20and%20Aelker.pdf) (15 Dec 2009, date last accessed).
- Bente, G., Krämer, N.C., Eschenburg, F. (2008a). Is there anybody out there? Analyzing the effects of embodiment and nonverbal behavior in avatar-mediated communication. In: Konijn, E., Utz, S., Tanis, M., Barnes, S., editors. *Mediated Interpersonal Communication*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 131–57.
- Bente, G., Senozkchieva, M., Pennig, S., Al Issa, A., Fischer, O. (2008b). Deciphering the secret code: A new methodology for the cross-cultural analysis of nonverbal behavior. *Behavior Research Methods*, 40(1), 269–77.
- Boraston, Z., Blakemore, S.J. (2007). The application of eye-tracking technology in the study of autism. *Journal of Physiology*, 581(Pt 3), 893–8.
- Duchowski, A.T., Cournia, N., Murphy, H. (2004). Gaze-contingent displays: a review. *Cyberpsychol Behav*, 7(6), 621–34.
- Eisenberger, N.I., Lieberman, M.D., Williams, K.D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science*, 302(5643), 290–2.
- Emery, N.J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience Biobehavior Review*, 24(6), 581–604.
- Frith, C.D. (2007). *Making up the mind*. Blackwell Publishing.
- Frith, C.D., Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, 60(3), 503–10.
- Frith, U., Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, 358(1431), 459–73.
- Gergely, G. (2001). The obscure object of desire: 'nearly, but clearly not, like me': contingency preference in normal children versus children with autism. *Bull Menninger Clin*, 65(3), 411–26.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Keysers, C., Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends in Cognitive Science*, 11(5), 194–6.
- Klin, A., Jones, W., Schultz, R., Volkmar, F. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, 358(1430), 345–60.
- Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, 459(7244), 257–61.
- Kuzmanovic, B., Georgescu, A.L., Eickhoff, S.B., et al. (2009). Duration matters: dissociating neural correlates of detection and evaluation of social gaze. *Neuroimage*, 46(4), 1154–63.
- Lieberman, M.D. (2007). Social cognitive neuroscience: a review of core processes. *Annual Review of Psychology*, 58, 259–89.
- Macrae, C.N., Hood, B.M., Milne, A.B., Rowe, A.C., Mason, M.F. (2002). Are you looking at me? Eye gaze and person perception. *Psychological Science*, 13(5), 460–4.
- Martinez-Conde, S., Macknik, S. L., Hubel, D.H. (2004). The role of fixational eye movements in visual perception. *Nature Review of Neuroscience*, 5(3), 229–40.
- Mason, M.F., Tatlow, E.P., Macrae, C.N. (2005). The look of love: gaze shifts and person perception. *Psychological Science*, 16(3), 236–9.
- Montague, P.R., Berns, G.S., Cohen, J.D., et al. (2002). Hyperscanning: simultaneous fMRI during linked social interactions. *Neuroimage*, 16(4), 1159–64.
- Moore, C., Dunham, P.J. (1995). *Joint Attention: Its Origin and Role in Development*. Hillsdale: Lawrence Erlbaum.
- Nummenmaa, L., Calder, A.J. (2009). Neural mechanisms of social attention. *Trends in Cognitive Science*, 13(3), 135–43.
- Park, K.M., Kim, J.J., Ku, J., et al. (2009). Neural basis of attributional style in schizophrenia. *Neuroscience Letters*, 459(1), 35–40.
- Pelphrey, K.A., Morris, J.P., McCarthy, G. (2005). Neural basis of eye gaze processing deficits in autism. *Brain*, 128, 1038–48.
- Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L. (2001). A default mode of brain function. *Proceedings of National Academy of Sciences United States of America*, 98(2), 676–82.
- Reddy, V. (2003). On being the object of attention: implications for self-other consciousness. *Trends in Cognitive Sciences*, 7(9), 397–402.
- Sanchez-Vives, M.V., Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Review of Neurosciences*, 6(4), 332–9.
- Saxe, R. (2006). Uniquely human cognition. *Current Opinion of Neurobiology*, 16, 235–9.
- Sanfey, A.G. (2007). Social decision-making: insights from game theory and neuroscience. *Science*, 318(5850), 598–602.

- Schilbach, L., Wohlschlaeger, A.M., Kraemer, N.C., et al. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, 44(5), 718–30.
- Schilbach, L., Eickhoff, S.B., Rotarska-Jagiela, A., Fink, G.R., Vogeley, K. (2008). Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the “default system” of the brain. *Conscious Cognition*, 17(2), 457–67.
- Schilbach, L., Wilms, M., Eickhoff, S.B., et al. (in press). Minds made for sharing. Initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience*, [Epub-ahead of print; 25 November 2009]; doi: 10.1162/jocn.2009.21401.
- Senju, A., Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, 18(9), 668–71.
- Senju, A., Southgate, V., White, S., Frith, U. (2009). Mindblind eyes: an absence of spontaneous theory of mind in Asperger syndrome. *Science*, 325(5942), 883–5.
- Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neurosci Biobehaviour Review*, 30(6), 855–63.
- Slater, M., Pertaub, D.P., Barker, C., Clark, D.M. (2006). An experimental study on fear of public speaking using a virtual environment. *Cyberpsychol Behaviour*, 9(5), 627–33.
- Spiers, H.J., Maguire, E.A. (2006). Spontaneous mentalizing during an interactive real world task: an fMRI study. *Neuropsychologia*, 44(10), 1674–82.