

Minds Made for Sharing: Initiating Joint Attention Recruits Reward-related Neurocircuitry

Leonhard Schilbach¹, Marcus Wilms², Simon B. Eickhoff^{2,3},
Sandro Romanzetti², Ralf Tepest¹, Gary Bente¹, N. Jon Shah²,
Gereon R. Fink^{1,2}, and Kai Vogeley¹

Abstract

■ The ability and motivation to share attention is a unique aspect of human cognition. Despite its significance, the neural basis remains elusive. To investigate the neural correlates of *joint attention*, we developed a novel, interactive research paradigm in which participants' gaze behavior—as measured by an eye tracking device—was used to contingently control the gaze of a computer-animated character. Instructed that the character on screen was controlled by a real person outside the scanner, 21 participants interacted with the virtual other while undergoing fMRI. Experimental variations focused on leading versus following the gaze of the character when fixating one of three objects also shown on the screen. In concordance with our hypotheses, results demonstrate, firstly, that following someone else's gaze

to engage in joint attention resulted in activation of anterior portion of medial prefrontal cortex (MPFC) known to be involved in the supramodal coordination of perceptual and cognitive processes. Secondly, directing someone else's gaze toward an object activated the ventral striatum which—in light of ratings obtained from participants—appears to underlie the hedonic aspects of sharing attention. The data, therefore, support the idea that other-initiated joint attention relies upon recruitment of MPFC previously related to the “meeting of minds.” In contrast, self-initiated joint attention leads to a differential increase of neural activity in reward-related brain areas, which might contribute to the uniquely human motivation to engage in the sharing of experiences. ■

INTRODUCTION

Gaze behavior is a crucial element of social interactions (Argyle & Cook, 1976) and helps to establish triadic relations between self, other, and the world in *joint attention* (JA; Moore & Dunham, 1995). Ontogenetically, JA has been considered an important precursor for the emergence of social cognitive capacities (Charman, 2003) while impairments thereof have been associated with autistic spectrum disorders (Dawson et al., 2002). Here, an important distinction has been made between *responding* to JA bids of others as compared to the *initiation* of JA (Mundy & Newell, 2007). Although autistic individuals can respond to JA bids of others, they show impairment in its initiation (Mundy, 2003). Consistently, it has been suggested that it might be the motivation to spontaneously engage in triadic relations which constitutes a unique element of (typically developing) human cognition and influences cognitive development by promoting engagement in shared, social realities (Moll & Tomasello, 2007).

Despite its significance, the neurobiological correlates of JA remain incompletely characterized as previous fMRI investigations have relied on third-party observation para-

digms, which cannot inform us about motivational factors inherent to the reciprocal nature of JA as well as their potentially distinct neural correlates (e.g., Materna, Dicke, & Thier, 2008; Williams, Waiter, Perra, Perrett, & Whiten, 2005).

To realize a truly interactive JA paradigm, we developed an eye tracking setup which allows to track the participants' gaze position calibrated to the stimulus screen and to contingently control the gaze behavior of a computer-animated character visible on the screen. The participants were, however, told that the virtual character's gaze corresponded to that of a real person whose eye movements were tracked outside the scanner and to whom their own gaze behavior was similarly visualized (Figures 1 and 2). During neuroimaging, participants were instructed to *direct* the gaze of the other person toward one of three objects by looking at it (hereafter: SELF). The gaze behavior of the other was made responsive to the participant's gaze and varied to either follow it, inducing a sense of JA, or to look elsewhere, inducing a sense of *nonjoint attention* (NOJA). Alternatively, participants were asked to *respond* to the other by looking at the same or at another object (OTHER; see Figure 3 for all condition-specific interaction sequences). Using this 2 × 2 factorial design, we investigated the neural correlates of JA employing 3-T fMRI in 21 participants.

¹University of Cologne, Germany, ²Research Centre Jülich, Germany, ³University of Aachen, Germany

We hypothesized that JA would rely upon the recruitment of medial prefrontal cortex (MPFC) previously related to the “meeting of minds” (Amodio & Frith, 2006; Saxe, 2006; Schilbach et al., 2006). Furthermore, we expected differential neural mechanisms depending upon whether JA is self- or other-initiated: Although gaze-following (OTHER_JA) was thought to increase neural activity in anterior MPFC in line with suggestions concerning this region’s involvement in stimulus-oriented attending (Burgess, Dumontheil, & Gilbert, 2007), SELF_JA was thought to rely upon recruitment of reward-related neurocircuitry, which might help explain why human beings have a propensity to engage in the sharing of experiences (Mundy & Newell, 2007; Tomasello & Carpenter, 2007).

METHODS

Participants

Twenty-one right-handed, healthy male volunteers, aged 18 to 30 years, with no record of neurologic or psychiatric illness, participated in the fMRI study. All volunteers were naive with respect to the experimental task as well as to the purpose of the study. Handedness was confirmed using the Edinburgh Handedness Questionnaire (Oldfield, 1971). The study was approved by the local ethics committee of the University of Cologne, Germany.

Stimulus Material, Instructions, and Study Design

Before participation, all participants received standardized instructions and were familiarized with the task. They were told that they would be asked to engage in an “interactive game” with two other participants (1 man, 1 woman) located outside the scanner. During the game, they would be asked to either probe or respond to the gaze behavior

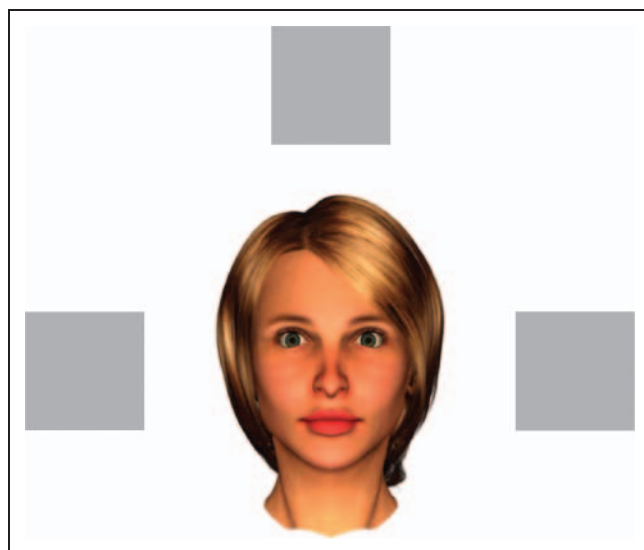


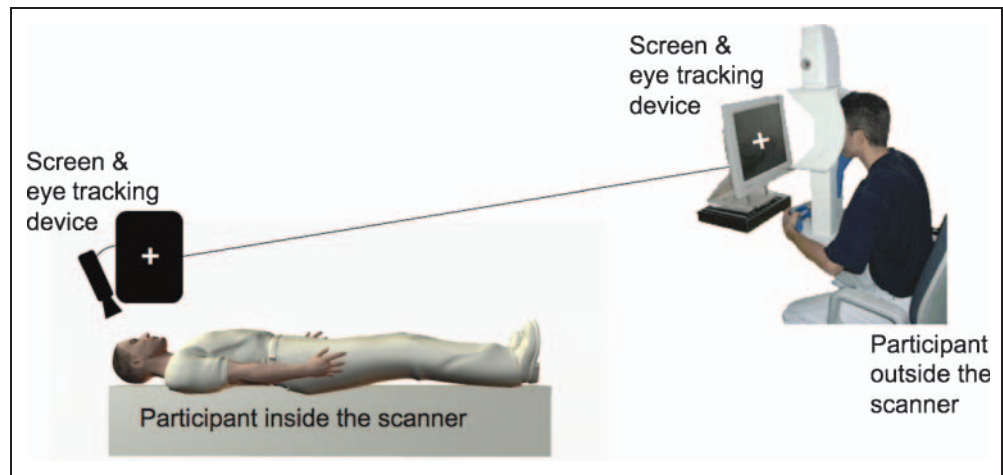
Figure 1. Screen shot of an exemplary virtual character and the three objects shown on the stimulus screen inside the scanner.

of an anthropomorphic virtual character shown on the stimulus screen (Figure 1). The gaze behavior shown would correspond to the gaze behavior of either one of the other, real participants having their eye movements tracked outside the scanner (Figure 2). Likewise, the participant’s own gaze behavior would be tracked inside the scanner and visualized for the other participants outside the scanner to allow for real-time interaction. To increase the likelihood of participants accepting this “cover story,” they were personally introduced to the other participants and were shown one of the eye tracking devices that would be used outside the scanner. Furthermore, participants were instructed that each of the “interaction phases” (i.e., 18-sec blocks) would be introduced by a word cue on the screen signaling how to interact with the other participant (Figure 3):

During the SELF-task, participants were instructed to “begin by establishing eye contact” with the other participant by looking directly at him or her to “make sure that he or she is ready to start.” Having done so, they were asked to choose and gaze steadily at one of the three objects also shown on the screen. Participants were prepared that the object would change color from gray to blue upon fixation and that they should maintain their gaze until the color changed back to gray. Unknown to the participant, this always occurred after 1500 msec to exert experimental control over the temporal spacing and number of fixations performed by the participants during each block. During sustained fixation of the object, they would be able to (peripherally) observe the reaction of the other participant occurring in response to their own fixation. Having done this successfully, participants were asked to reinitiate “eye contact” with the other participant and to choose a new object to probe the other participant anew. They were asked to continue the interaction in this self-paced manner until the end of the block to achieve several repetitions. Participants were given no information about possible factors influencing the response selection by the other participants represented by a respective virtual character, but were told that it was their own task and the “idea of the game” was to find out about the response tendencies of the other participants by means of interacting. Instead of corresponding to the gaze behavior of a real person—as suggested by the instructions given to participants—the gaze behavior shown by the virtual character was, in fact, made contingent upon the participant’s own gaze and systematically varied on a block-by-block basis to either consistently follow the participant’s gaze for the duration of a block, thereby inducing a sense of JA, or to consistently look elsewhere, thereby inducing an experience of NOJA.

During the OTHER-task, participants were instructed to also begin by establishing “eye contact” with the other participant. Afterward, they were told that the other participant would “choose to look at one of the objects,” and it would be their task “to react to this.” They were further instructed to adjust their response behavior on a block-by-block basis according to the OTHER-cue shown before the block. During a block following a *green* OTHER-cue,

Figure 2. Experimental setup as depicted in the instructions.



they were asked to always react congruently to their partner's gaze behavior and to look at whichever object the other had chosen to look at, responding positively to the JA bid. Doing this correctly would change the object's color from gray to blue. Again, participants were asked to look at the object until the color changed back to gray. Afterward, they were asked to reinitiate "eye contact" whereupon their partner would choose again to look at an object to continue the interaction. In this iterative, self-paced manner, participants were asked to proceed until the end of the block. During blocks following a *red* OTHER-cue, they were instructed to always react incongruently to the gaze

behavior shown for the entire duration of the block and to look at one of the other two objects, thereby engaging in NOJA. Again, focusing on an object not being looked at by the other party would change the object's color. Sustained fixation of the object would lead to the reversal of the color change, indicating that the participants should re-establish "eye contact" and continue until the end of the block.

During the baseline condition (not illustrated), participants were shown the exact same visual stimuli except for the face, which was shown with its eyes closed. Participants were told that this baseline was inserted "due to

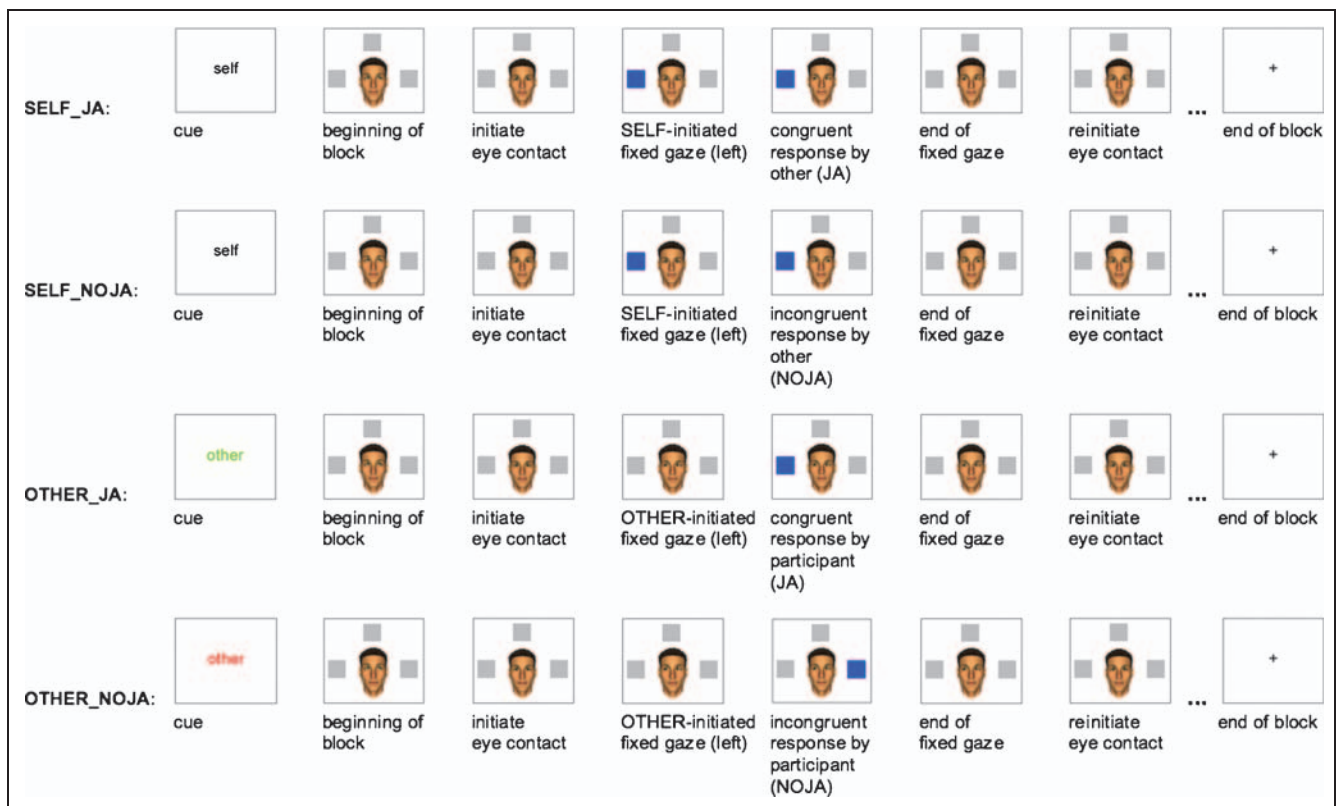


Figure 3. Illustration of all condition-specific gaze-based interaction sequences.

methodological reasons,” but also to give them “a break from the interaction.” Participants were told that, indicated by the closed eyes of the virtual character, the gaze behavior of the interactors was “not broadcast” and no interaction could, therefore, occur during this time.

Stimuli were presented to the participants lying inside the scanner on a custom-built, shielded TFT screen at the rear end of the scanner ($14^\circ \times 8^\circ$ horizontal \times vertical viewing angle). Volunteers watched the stimuli via a mirror mounted on the headcoil. Due to the screen’s distance from the volunteers’ eyes, changes of the virtual character’s gaze behavior were easily observable while focusing on one of the three objects. The participant’s head was placed into an MR-compatible vacuum hood inside the headcoil (Vacuform hood by B. & W. Schmidt GmbH, Germany). Air was extracted from the hood once the participant had been made comfortable in order to provide maximum stabilization of the head and to minimize head movements. During each of the three sessions, 20 blocks of 18 sec duration (4 repetitions of each of the four target conditions and the baseline condition) were shown in a randomized order. The appearance of male and female characters was balanced across conditions and runs.

“Interactive Eye Tracking” during fMRI

“Interactive eye tracking,” that is, the delivery of gaze-contingent stimuli during fMRI, relied on an MR-compatible eye tracking system allowing real-time gaze data transmission to a visual stimulation controller. The controller received the ongoing gaze data and adapted the visual stimulation according to pre-set task conditions and the volunteer’s current gaze position on screen. Participants’ eye movements were monitored by means of an infrared camera (Resonance Technology, CA, USA). The camera and infrared light source were mounted on the headcoil using a custom-built gooseneck that allowed easy access to the participants’ eyes without interfering with the visual stimulation. The raw analog video signal of the eye tracking camera was digitized at 60 Hz frame rate (iViewX, SMI, Germany) on a computer dedicated to this task and the gaze extraction software (iViewX, SMI) produced real-time gaze position output. Eye tracking calibration was performed prior to each data acquisition session in order to yield gaze positions in a stimulus-related coordinate system. Via a fast network connection, gaze position updates were transferred, and thus, made available to another computer running the software which controlled the visual stimulation (Presentation, Version 9.9, Neurobehavioral Systems, Albany, CA). For detecting fixations on screen targets (face, 3 objects), we devised the following procedure: Using Presentation software, gaze positions were transformed to stimulus screen coordinates (pixels). A continuous “sliding window” average of the preceding 10 gaze positions was calculated throughout the whole stimulus presentation. This averaging technique is necessary to cope with the electromagnetic noise and its detrimental effect on eye tracking

measurements, which is introduced by the use of standard EPI sequences during fMRI. These fixations were subsequently tested to see if they occurred within one out of four regions of interest (ROIs). 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 block length of the task was attained. In our paradigm, the virtual character’s face and the three objects shown on the stimulus screen were entered as ROIs for the on-line analysis to make the virtual character’s gaze behavior “responsive” to and contingent upon the participant’s fixations. Detected fixations on targets as well as measurements of pupil size were logged in a file for later data analysis. Comparisons of the individual mean were performed by means of repeated measures analyses of variance (ANOVAs) as implemented in SPSS Version 12 (SPSS, Chicago, IL).

Functional Magnetic Resonance Imaging

Functional magnetic resonance images were acquired on a Siemens Trio 3-T whole-body scanner (Erlangen, Germany) using blood oxygen level dependent (BOLD) contrast (gradient-echo, EPI pulse sequence, TR = 2.304 sec, slice thickness 3 mm, 38 axial slices, in-plane resolution 3×3 mm). Additional high-resolution anatomical images (voxel size = $1 \times 1 \times 1$ mm³) were acquired using a standard T1-weighted 3-D MP-RAGE sequence. Images were analyzed using SPM5 (www.fil.ion.ucl.ac.uk/spm) as follows. The EPI images were corrected for head movement between scans by an affine registration (Ashburner & Friston, 2003). For realignment, we used a two-pass procedure, by which images were initially realigned to the first image of the time series and subsequently re-realigned to the mean of all images after the first step. After completing the realignment, the mean EPI image for each subject was computed and spatially normalized to the MNI single-subject template (Collins, Neelin, Peters, & Evans, 1994) using the “unified segmentation” function in SPM5. This algorithm is based on a probabilistic framework that enables image registration, tissue classification, and bias correction to be combined within the same generative model. The resulting parameters of a discrete cosine transform, which define the deformation field necessary to move the subject’s data into the space of the MNI tissue probability maps (Evans, Kamber, Collins, & MacDonald, 1994) were then combined with the deformation field transforming between the latter and the MNI single-subject template. The ensuing deformation was subsequently applied to the individual EPI volumes as well as to the T1 scan, which was coregistered to the mean of the realigned EPIs beforehand. All images were transformed into standard stereotaxic space. The normalized images were spatially smoothed using an 8-mm FWHM Gaussian kernel to meet the statistical requirements of the general linear model (GLM) and to compensate for residual macroanatomical variations across subjects.

The data were analyzed using a general linear model as implemented in SPM5 (Kiebel & Holmes, 2003). Each experimental condition was modeled using a boxcar reference vector convolved with a canonical hemodynamic response function and its first-order temporal derivative. Low-frequency signal drifts were filtered using a cutoff period of 128 sec. Parameter estimates were subsequently calculated for each voxel using weighted least squares to provide maximum likelihood estimators based on the temporal autocorrelation of the data (Penny & Holmes, 2003) in order to get identical and independently distributed error terms. No global scaling was applied. For each subject, simple main effects for each experimental condition were computed by applying appropriate baseline contrasts. The experimental conditions were modeled as 18-sec blocks in the first-level analysis. Blocks which did not include at least two successful object fixations, that is, fixations which led to condition-specific changes of the virtual character's gaze behavior, were modeled using a separate regressor of no interest. Each experimental condition was contrasted with the baseline condition and these first-level individual contrasts were then entered into a second-level group analysis using an ANOVA (factor: condition; blocking factor: subject) employing a random-effects model (Worsley et al., 1996). Incorporating the "high-level" baseline on the first level was done in order to further constrain the analysis and to selectively target the social cognitive processes of JA. In the modeling of variance components, we allowed for violations of sphericity by modeling non-independence across parameter estimates from the same subject and allowing unequal variances both between conditions and subjects using the standard implementation in SPM5. On the second level, the main effect of JA [(SELF_JA + OTHER_JA) > (SELF_NOJA + OTHER_NOJA)], as well as the main effect of NOJA, was calculated [(SELF_NOJA + OTHER_NOJA) > (SELF_JA + OTHER_JA)]. Furthermore, we analyzed the main effect of self-initiated interaction (SELF) [(SELF_JA + SELF_NOJA) > (OTHER_JA + OTHER_NOJA)] as well as the main effect of responding to other-initiated actions (OTHER) [(OTHER_JA + OTHER_NOJA) > (SELF_JA + SELF_NOJA)]. To test for statistical interactions between the main effects, that is, relative activation for JA \times SELF [(SELF_JA > OTHER_JA) > (SELF_NOJA > OTHER_NOJA)] and relative activation for JA \times OTHER [(OTHER_JA > SELF_JA) > (OTHER_NOJA > SELF_NOJA)], appropriate contrasts were calculated. The resulting SPM(*T*) maps were then interpreted by referring to the probabilistic behavior of Gaussian random fields (Worsley et al., 1996) and thresholded at $p < .05$ (cluster-level corrected for multiple comparisons). The cluster-forming threshold was set to $p_{uc} < .001$. In light of significant interactions of the main effects and in order to detect differential effects driven by whether JA was self- or other-initiated, we also analyzed the relevant simple contrasts (SELF_JA > SELF_NOJA and OTHER_JA > OTHER_NOJA) which were inclusively masked by the respective interaction at $p < .0001$, uncorrected, and then thresholded at a family-wise

error (FWE) corrected significance level of $p < .05$ using an extent threshold of 10 voxels.

Functional activations were anatomically localized by using the SPM anatomy toolbox (Eickhoff et al., 2007) employing a maximum probability map. This map (Eickhoff, Heim, Zilles, & Amunts, 2006) denotes the most likely anatomical area at each voxel of the MNI single-subject template based on probabilistic cytoarchitectonic maps derived from the analysis of cortical areas in a sample of 10 human postmortem brains, which were subsequently normalized to the MNI reference space. If no cytoarchitectonic maps were available, the macroanatomical labels are provided based on the Automated Anatomic Labeling atlas (Tzourio-Mazoyer et al., 2002) or the labels of the MNI single-subject template brain were used in comparison with the mean structural image of the analyzed subjects after normalization.

Debriefing Scores

During a debriefing period, we asked the participants to fill out a questionnaire which assessed aspects of their experience and their attitude toward the task. This short questionnaire consisted of nine questions, which participants were asked to answer on a 5-point scale (Table 1). They were also given the opportunity to write down comments concerning the study. Furthermore, participants were notified of the "cover story" they had been exposed to and were given the opportunity to withdraw their individual dataset at this point. None of our participants decided to do so. We used the paired-sample *t* test procedure as implemented in SPSS Version 12 (SPSS) to perform comparisons of those questions which target the pleasantness as part of the experience of JA as compared to NOJA, as well as the perceived difficulty of sustaining JA as compared to NOJA.

Combining Debriefing Scores and fMRI Data

To assess whether the BOLD signal change in the ventral striatum was correlated with the subjective ratings of the pleasantness of JA as compared to NOJA, we conducted a bivariate correlation analysis as implemented in SPSS Version 12 of the relevant questionnaire scores and the BOLD signal, that is, the individual values for the modeled effects, extracted at the principally activated voxel in the left ventral striatum (MNI coordinates: $x = -12$, $y = 14$, $z = -19$). Given that the pleasantness ratings of looking at the same object were significantly nonnormal according to the Kolmogorov-Smirnov test [$D(21) = 0.241$, $p < .05$], we used nonparametric correlations as implemented in SPSS Version 12 (Spearman's rho). To test the specificity of the correlation, the pleasantness ratings were entered as a covariate in the second-level analysis of the neuroimaging data. This covariate was then used to perform a conjunction analysis (global null hypothesis, thresholded at 0.01, FWE corrected) with those contrasts that target JA being self- or other-initiated ([SELF_JA > SELF_NOJA \cap covariate] and [OTHER_JA > OTHER_NOJA \cap

Table 1. Debriefing Questionnaire Used in the Study and Mean Responses to All Questionnaire Items

Question	Mean Response
“How did you like the study?”	1.38 (0.67)
“Was the visualization of the gaze behavior by the other participants satisfactory?”	1.67 (0.80)
“Did you always have the impression that the other participant reacted promptly to your own gaze behavior?”	1.95 (0.59)
“Were there times during the study when you felt that the gaze behavior by the computer face was <i>not</i> controlled by a real person?”	4.19 (0.60)
“Was it always clear whether the other participant was looking towards the object which you were also looking at (or vice versa)?”	2.05 (0.92)
“Did you find it rather pleasant (as opposed to rather unpleasant) when the other participant looked at the same object as you did?”	2.19 (0.75)
“Did you find it rather pleasant (as opposed to rather unpleasant) when the other participant looked away towards a different object than you did?”	3.62 (0.92)
“Did you find it rather easy (as opposed to rather difficult) to look at the same object which was not being viewed by the other participant?”	1.86 (0.73)
“Did you find it rather easy (as opposed to rather difficult) to look at an object which was not being viewed by the other participant?”	3.52 (1.47)

Scores were obtained on a 5-point scale with 1 signifying agreement and 5 signifying disagreement (*SD* in brackets).

covariate], respectively). This procedure was performed to assess which brain regions show a differential effect related to JA being self- or other-initiated and also to show a correlation with the postscan pleasantness rating of JA.

In order to assess a possible relationship between the participants' reports of task difficulty and differences in BOLD signal for JA and NOJA, the task difficulty ratings of JA and NOJA were used as covariates in the second-level data analysis. These covariates were then used to perform conjunction analyses to directly test whether an increase in neural activity is related to the measure of perceived task difficulty ($JA > NOJA \cap \text{covariate}_{\text{difficulty_JA}}$; $NOJA > JA \cap \text{covariate}_{\text{difficulty_NOJA}}$; global null hypothesis, .01, FWE corrected).

Follow-up Behavioral Study

In light of the neuroimaging data suggesting that the pleasantness of JA differs depending upon whether or not JA is self- or other-initiated, that is, a result of leading or following someone else's gaze, we performed an additional behavioral study employing a between-subject design to test this hypothesis. In this study, two groups of subjects participated in a shortened version of the paradigm which either involved engagement in SELF_JA and SELF_NOJA (Group 1) or OTHER_JA and OTHER_NOJA blocks (Group 2). Twenty-four healthy volunteers aged 21 to 35 years, with no record of neurologic or psychiatric illness, participated in the behavioral study. All volunteers were naive with respect to the experimental task as well as to the purpose of the study. A Tobii eye tracking system (Tobii, Sweden) was used to acquire the gaze position data during the experiment and to produce gaze-contingent stimuli. After participation, sub-

jective ratings were obtained from the participants by means of a questionnaire to investigate the perceived pleasantness of engaging in JA or NOJA on a 7-point scale (Table 2).

Within-group comparisons of the individual mean responses to the questionnaire items pertaining to the experience of JA and NOJA were performed by using the paired-sample *t* test procedure as implemented in SPSS Version 12 (SPSS). Between-group comparisons for responses to each question were performed by using an independent-sample *t* test procedure.

RESULTS

Debriefing Scores

After scanning, participants completed a debriefing questionnaire (Table 1) to assess their belief in the “cover story” (treatment check) and to document their subjective experience of participation. Scores obtained on a 5-point scale (1 = high, 5 = low) indicated that participants enjoyed being included in our study, resulting in a mean response of 1.38 for question 1 ($n = 21$, $SD = 0.67$) (see Table 1 for all mean responses). On a number of subsequent questions, they demonstrated their engagement in the task and assessed different aspects of the interaction; participants gave high ratings for the quality of the “broadcast” used to visualize the gaze behavior, resulting in a mean response of 1.67 for Question 2 ($n = 21$, $SD = 0.80$) and indicated that the other participant had reacted promptly during the interaction [mean response of 1.95 for Question 3; $n = 21$, $SD = 0.59$]. Most importantly, results also showed that the participants had been receptive to the cover story, resulting in a low score when asked whether they had experienced moments during which they had felt that the

Table 2. Debriefing Questionnaires Used in the Follow-up Study and Mean Responses to Questionnaire Items

Question	Mean Response
<i>Group 1 (SELF_JA and SELF_NOJA)</i>	
“Did you find it rather pleasant (as opposed to rather unpleasant) when the other participant looked at the same object as you did?”	1.5 (0.67)
“Did you find it rather pleasant (as opposed to rather unpleasant) when the other participant looked away towards a different object than you did?”	5.75 (0.87)
<i>Group 2 (OTHER_JA and OTHER_NOJA)</i>	
“Did you find it rather pleasant (as opposed to rather unpleasant) when you looked at the same object as the other participant?”	2.33 (1.23)
“Did you find it rather pleasant (as opposed to rather unpleasant) you looked away towards a different object than the other participant did?”	5.08 (1.24)

Scores were obtained on a 7-point scale with 1 signifying agreement and 7 signifying disagreement (*SD* in brackets).

gaze behavior shown by the virtual character had *not* been controlled by a real person (mean response of 4.19 for Question 4; $n = 21$, $SD = 0.60$). Participants also indicated that, in all cases, it had been quite clear which object the other participant had been focusing on [mean response of 2.05 for Question 5; $n = 21$, $SD = 0.92$]. The pleasantness of looking at the same object as the other participant provided a score of 2.19 ($n = 21$, $SD = 0.75$). Conversely, when asked about the pleasantness of looking at a different object, participants gave a lower rating of 3.62 ($n = 21$, $SD = 0.92$). Being asked for an evaluation of how easy or difficult it had been to look at the same object resulted in a mean response of 1.86 ($n = 21$, $SD = 0.73$), whereas the same question asked with regard to looking at different objects resulted in a mean response of 3.52 ($n = 21$, $SD = 1.47$). The statistical comparison of participants' ratings concerning the pleasantness as well as the perceived difficulty of the experience of engaging in JA as compared to NOJA demonstrated that the mean responses to the two pairs of questions differed significantly. Subjects' reports described engaging in JA as significantly more pleasant than engaging in NOJA [$t(20) = -7.52$, $p < .001$, $r = .86$], and engaging in JA as significantly less difficult to do than engaging in NOJA [$t(20) = -4.62$, $p < .001$, $r = .72$].

Fixations on Target

Eye tracking measurements obtained from study participants during fMRI measurements were successfully used

in real time to control and drive the experimental paradigm and also served as dependent measures of attention and arousal. Across conditions, we recorded the number of “successful” fixations, that is, those detected by our computer algorithm in the face and object regions on the stimulus screen and translated into real-time changes of the visual stimulation, thereby providing for the continuation of the experiment. Due to the self-paced manner of the task and the temporal spacing of fixations resulting from the change in color of the objects, the mean number of successful object fixations per 18-sec block was comparable across all target conditions [SELF_JA: 4.10 ($n = 21$, $SD = 0.68$); SELF_NOJA: 4.06 ($n = 21$, $SD = 0.79$); OTHER_JA: 3.96 ($n = 21$, $SD = 0.88$); OTHER_NOJA: 4.03 ($n = 21$, $SD = 0.88$); see Table 3]. Comparisons of these results across conditions employing a two-way repeated measures ANOVA, as implemented in SPSS Version 12 (SPSS), did not yield any significant difference [effect of SELF: $F(1, 20) = 0.26$, $p = .62$; effect of OTHER: $F(1, 20) = 0.05$, $p = .83$; statistical interaction: $F(1, 20) = 1.21$, $p = .29$].

Pupil Size

Apart from gaze coordinates, we also recorded changes of pupil size measured in degrees of screen range [deg] at the eye tracking device's sampling rate of 60 Hz. Given that we employed a blocked design in which condition-specific behavior was repeatedly shown during each block—as indicated by the number of successful fixations on target—pupil

Table 3. Mean Values of Pupil Size (in Degrees of Screen Range [deg]) and Mean Number of Object Fixations per Block across All Experimental Conditions (*SD* in Brackets)

	Experimental Conditions			
	SELF_JA	SELF_NOJA	OTHER_JA	OTHER_NOJA
Pupil size	1.349 (0.274)	1.359 (0.289)	1.330 (0.276)	1.353 (0.328)
Fixations on target	4.10 (0.68)	4.06 (0.79)	3.96 (0.88)	4.03 (0.88)

Table 4. Neural Correlates (Main Effects, Statistical Interactions, and Simple Contrasts at $p < .05$, Cluster-level Corrected for Multiple Comparisons; MNI Coordinates of Principally Activated Voxels for Each Cluster Are Given)

<i>Brain Region</i>	<i>x</i>	<i>y</i>	<i>z</i>	<i>k</i>	<i>T</i>
<i>(A) Common Activations of Joint Attention (JA > NOJA)</i>					
Left superior frontal gyrus ^a	-10	56	37	387	5.63*
Rectal gyrus	-2	52	-21	2314	6.67*
Left middle temporal gyrus	-56	-2	-21	180	4.50
Right hippocampus ^a	22	-24	-5	239	4.56
Middle cingulate cortex	4	-36	39	1052	4.92
Right calcarine gyrus ^a	8	-88	13	682	5.41*
<i>(B) Common Activations of Nonjoint Attention (NOJA > JA)</i>					
Right middle frontal gyrus	38	36	33	378	4.64
Left middle frontal gyrus	-36	26	33	196	4.64
Right superior frontal gyrus	26	8	53	1606	6.76*
Left middle frontal gyrus ^a	-24	0	57	970	6.96*
Left inferior parietal lobule ^a	-38	-38	37	227	4.40
Right inferior parietal lobule ^a	40	-40	45	991	5.97*
Right precuneus	10	-66	59	3477	8.29*
Right middle temporal gyrus	46	-74	25	309	4.61
<i>(C) Common Activations of Self-sustained Interaction (SELF > OTHER)</i>					
Left superior frontal gyrus	-32	54	19	749	7.92*
Right middle frontal gyrus	28	12	59	5792	6.54*
Left inferior frontal gyrus ^a	-46	12	5	1033	5.28*
Right pallidum	20	6	1	1092	5.82*
Right inferior parietal lobule ^a	46	-38	41	3387	6.74*
Left inferior parietal lobule ^a	-48	-42	41	1684	5.62*
Right cerebellum	38	-62	-29	410	4.99
Left cerebellum	-30	-68	-25	576	5.90*
Left precuneus	-6	-74	59	446	4.31
<i>(D) Common Activations of Other-sustained Interaction (OTHER > SELF)</i>					
Left middle frontal gyrus ^a	-8	-10	49	5696	6.70*
Right Rolandic operculum ^a	38	-16	21	1293	6.89*
Left posterior cingulate cortex	-12	-46	9	287	4.86
Right inferior occipital gyrus ^a	24	-92	-1	394	5.53*
<i>(E) Common Activations of Statistical Interaction JA × SELF</i>					
Left middle frontal gyrus	-34	56	19	179	5.00

Table 4. (continued)

<i>Brain Region</i>	<i>x</i>	<i>y</i>	<i>z</i>	<i>k</i>	<i>T</i>
Right insula ^a	38	18	5	647	5.22*
Middle cingulate cortex	-2	14	41	647	4.87
Left insula	-36	14	1	1169	5.83*
Right superior frontal gyrus ^a	26	-6	67	1527	6.06*
Right thalamus	10	-12	13	169	4.40
Left inferior parietal lobule	-46	-34	41	161	4.64
Right inferior parietal lobule	40	-40	45	597	4.79
Left cerebellum	-48	-54	-35	184	4.57
Right lingual gyrus	24	-64	-1	458	5.82*
Right superior parietal lobule	16	-74	63	3721	6.42*
<i>(F) Common Activations of Statistical Interaction JA × OTHER</i>					
Left superior frontal gyrus	-18	32	43	166	5.03
<i>(G) Common Activations of SELF_JA > SELF_NOJA (Incl. Masked by JA × SELF)</i>					
Right anterior ventral striatum	10	16	-15	14	5.89*
Left anterior ventral striatum	-12	14	-19	80	6.28*
<i>(H) Common Activations of OTHER_JA > OTHER_NOJA (Incl. Masked by JA × OTHER)</i>					
Medial frontal gyrus	2	56	-15	149	6.74*

^aAnatomy assigned by the Anatomy Toolbox.*Also significant at $p < .05$, FWE voxel-level corrected.

size data were averaged for each participant across all blocks belonging to a particular condition. Two individual pupil size datasets could not be entered into the analysis due to technical difficulties during data acquisition. Across the remaining 19 subjects, mean pupil size for SELF_JA was 1.349 ($n = 19$, $SD = 0.274$), for SELF_NOJA 1.359 ($n = 19$, $SD = 0.289$), for OTHER_JA 1.330 ($n = 19$, $SD = 0.276$), and for OTHER_NOJA 1.353 ($n = 19$, $SD = 0.328$) (see Table 3). Comparisons of these results by means of a two-way repeated measures ANOVA as implemented in SPSS Version 12 (SPSS) demonstrated a significant main effect indicating significantly larger pupil size during NOJA [$F(1, 18) = 13.91$, $p = .002$].

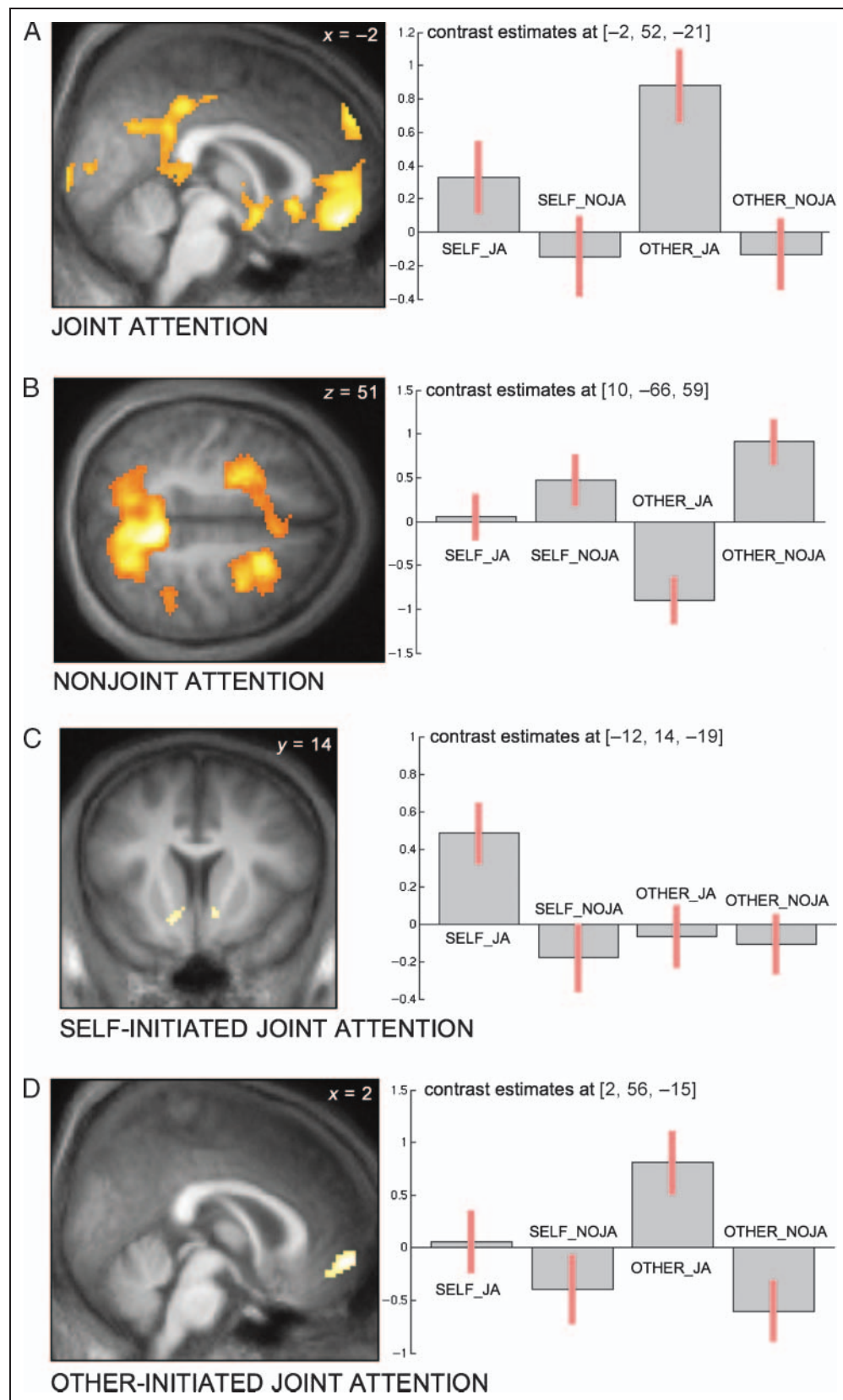
Neural Correlates

Neural correlates of all main effects and interactions are summarized in Table 4. Figure 4 shows the SPMs of supra-threshold clusters of the main effect of JA (Figure 4A) as well as NOJA (Figure 4B) as overlay images onto an averaged normalized 3-D dataset of the participants of the study and also includes depictions of the contrast estimates for

principally activated voxels within clusters of interest across all experimental conditions. Figure 4 also demonstrates activations of simple contrasts highlighting the differential effects of JA being self- (Figure 4C) or other-initiated (Figure 4D).

The results demonstrated that engagement in JA, that is, the main effect of JA [(SELF_JA + OTHER_JA) > (SELF_NOJA + OTHER_NOJA)], resulted in the recruitment of a neural network comprising the dorsal and ventral portions

Figure 4. A–D: Neural correlates of the main effect of *joint* attention (A) and *nonjoint* attention (B), as well as simple contrasts inclusively masked by the respective statistical interaction effect highlighting the differential neural effects related to *self-initiated* (C) and *other-initiated* joint attention (D).



of MPFC (DMPFC, VMPFC), medial orbito-frontal and subgenual cingulate cortices extending into the ventral striatum, posterior cingulate cortex, the calcarine gyrus, the right hippocampus, and anterior temporal cortex. Although the differential increase of neural activity in VMPFC was driven by the effect of gaze-following (OTHER_JA), the activation of the ventral striatum was most pronounced during self-initiated JA (see plot in Figure 4). Similarly to the activation in VMPFC, the differential increase of activity in the left anterior middle temporal gyrus was most pronounced for gaze-following. The main effect of NOJA [(SELF_NOJA + OTHER_NOJA) > (SELF_JA + OTHER_JA)] demonstrated a differential increase of neural activity in the precuneus, superior parietal cortex, the superior and middle frontal gyrus bilaterally, the inferior parietal lobule bilaterally, the right temporo-parietal junction area, and the right middle temporal gyrus.

Initiating interaction—regardless of the type of responses this elicited, that is, the main effect of SELF [(SELF_JA + SELF_NOJA) > (OTHER_JA + OTHER_NOJA)], recruited a bilateral fronto-parietal network more pronounced in the right hemisphere, including inferior, middle, and superior frontal gyrus, as well as activations in inferior parietal lobule, angular gyrus, medial and superior parietal cortex. Furthermore, the results showed a differential increase of neural activity in the basal ganglia and the cerebellum. Conversely, responding to the gaze shifts produced by the virtual other regardless of the kind of response shown, that is, the main effect of OTHER [(OTHER_JA + OTHER_NOJA) > (SELF_JA + SELF_NOJA)], resulted in a differential increase of neural activity in the left middle frontal gyrus, the right Rolandic operculum, left posterior cingulate cortex, and right inferior occipital gyrus.

A statistical interaction between the main effects was noted for JA \times ME [(SELF_JA > OTHER_JA) > (SELF_NOJA > OTHER_NOJA)] in inferior and superior parietal cortex, middle cingulate cortex, as well as insular cortex. Furthermore, the results demonstrated a differential involvement

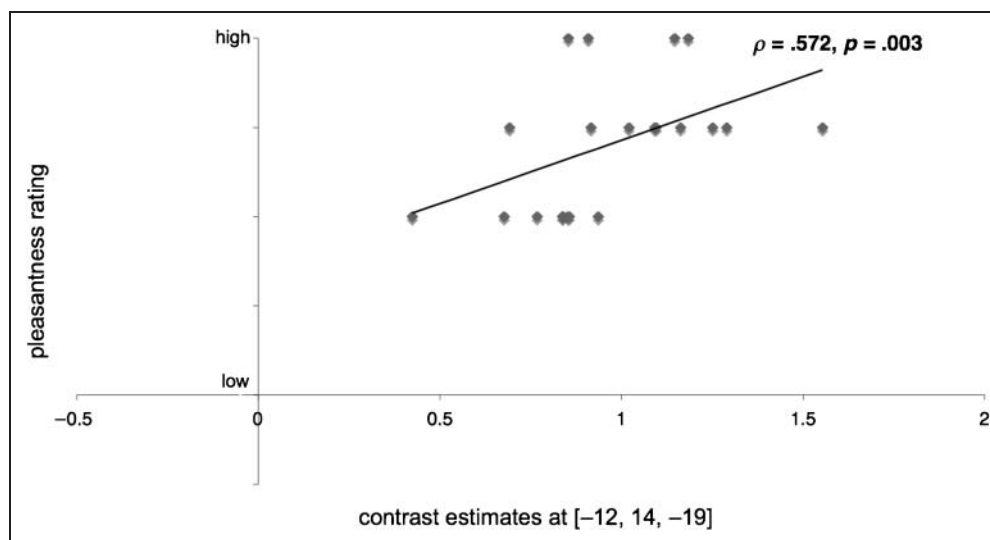
of the middle and superior frontal gyrus, the cerebellum, the right lingual gyrus, and the dorsal thalamus. The second interaction, JA \times OTHER [(OTHER_JA > SELF_JA) > (OTHER_NOJA > SELF_NOJA)], showed differential activity in the left superior frontal lobe.

In light of significant interactions between the main effects and to further assess the differential effects of JA being self- or other-initiated, we also analyzed simple contrasts which were inclusively masked with the respective interaction effect. This analysis demonstrated that *self-initiated* JA, that is, the single contrast of SELF_JA > SELF_NOJA, resulted in a differential increase of neural activity in the anterior ventral striatum bilaterally (Figure 4C). Conversely, *other-initiated* JA, that is, the single contrast of OTHER_JA > OTHER_NOJA, resulted in the recruitment of VMPFC (Figure 4D).

Combining Debriefing Scores and fMRI Data

In order to investigate the relationship of subjective rating scores (obtained by a poststudy debriefing questionnaire concerning the pleasantness of experiencing JA), bivariate correlation analyses between the responses and the BOLD signal acquired in the ventral striatum were performed. This analysis showed a positive and significant correlation of the fMRI signal changes related to being engaged in SELF_JA and the subjective ratings of pleasantness ($\rho = .572, p = .003$) (Figure 5). No other comparison between the pleasantness ratings and fMRI signal changes pertaining to any of the three remaining experimental conditions showed a significant correlation. Consistently, the conjunction analysis of the simple contrast for SELF_JA > SELF_NOJA and the covariate of the pleasantness rating showed activation of the ventral striatum bilaterally (Figure 2; MNI: $-26, 8, -8$ and $24, 10, -15$) and anterior cingulate cortex (MNI: $4, 36, 18$), whereas the same procedure for the contrast OTHER_JA > OTHER_NOJA did not show any supra-threshold clusters of activation.

Figure 5. Scatterplot showing the relation between the reported pleasantness of engaging in joint attention (y -axis) and ventral striatal neural activity during self-initiated joint attention (x -axis).



Using the participants' ratings of task difficulty to explore their possible relationship to activation results observed for the main effects of JA and NOJA by means of conjunction analysis did not yield any suprathreshold voxels of activation, which suggests that the activation results obtained for the main effects cannot be explained in terms of task-unrelated processes (such as perceived differences in task difficulty), but can be regarded as specifically task-related.

Follow-up Behavioral Study

Results of the follow-up behavioral study show an average rating of 1.5 for SELF_JA ($n = 12$, $SD = 0.67$) and of 5.75 for SELF_NOJA ($n = 12$, $SD = 0.87$). Results further demonstrate an average rating of 2.33 ($n = 12$, $SD = 1.23$) for OTHER_JA and of 5.08 ($n = 12$, $SD = 1.24$) for OTHER_NOJA (Table 2). Statistical comparison of these responses by means of the paired-samples t test procedure revealed that SELF_JA was rated to be significantly more pleasant than SELF_NOJA [SELF_JA vs. SELF_NOJA: $t(11) = -4.213$, $p = .001$, $r = .79$] and OTHER_JA to be significantly more pleasant than OTHER_NOJA within the respective groups [OTHER_JA vs. OTHER_NOJA: $t(11) = -10.851$, $p < .001$, $r = .95$].

As hypothesized, the statistical comparison of the ratings between groups by means of an independent-sample t test demonstrated that the mean responses to SELF_JA and OTHER_JA differed in that participants rated engagement in SELF_JA as significantly more pleasant than in OTHER_JA [$t(22) = 2.057$, $p = .03$ one-tailed, $r = .40$]. Comparison of the ratings of SELF_NOJA and OTHER_NOJA did not show a statistically significant result.

DISCUSSION

Our study highlights the potential of combining computer-animated characters and real-time processing of gaze data to generate truly interactive research paradigms (Singer, 2006). Such paradigms may be helpful in expanding the investigation of the neurobiology of social cognition, by taking into account the reciprocal nature of social interaction, and could help to target the differential contribution of implicit and explicit processes in social cognition under ecologically valid conditions (Frith & Frith, 2008). In contrast to theories that either emphasize similarity or dissimilarity between self and other, making use of such a "second-personal" approach may help to address how engagement in on-line interaction pragmatically changes and informs the understanding of other minds (Legrand & Iacoboni, in press; Gallagher, 2008; Reddy, 2003). The paradigm used in this study reliably produced gaze-contingent stimuli, which allowed participants to engage in JA with someone while leading or following the social interaction. Future studies could expand the use of this experimental platform by introducing real-life objects which may help to further increase ecological validity and to study aspects

related to object choice and preference as well as person perception (Becchio, Bertone, & Castiello, 2008).

As one of two key findings, our analysis of the fMRI data demonstrates that JA results in the recruitment of MPFC previously implicated in the "meeting of minds" (Amodio & Frith, 2006). Recruitment of DMPFC, in particular, has been related to the processing of communicative intent (Kampe, Frith, & Frith, 2003). It has been proposed that in order to understand communicative intent, both sender and receiver must recognize that the receiver attributes a mental state to the sender and that signals are sent with the intention of altering the receiver's mental state. In this way, sender and receiver engage in a triadic interaction by exchanging intentional actions for communicative purposes (Frith, 2007). We suggest that (gaze-based) sharing of attention can be thought of as a paradigmatic case of this process. Consistent with previous neuroimaging studies targeting the understanding of communicative intent (Kampe et al., 2003), JA recruited DMPFC in our study.

Our results also demonstrate activation of the ventral portion of MPFC (VMPFC) during JA. We suggest that this might be related to the monitoring of one's own emotions and the interaction's outcome (Amodio & Frith, 2006). Consistent with its anatomical connections to sensory association areas, VMPFC might be involved in higher-order integration of stimulus-dependent, sensory input with stimulus-independent information accrued as part of an adaptive response to the stimulus (Klin, Jones, Schultz, & Volkmar, 2003). VMPFC could, thus, contribute to the processing of aspects which are not "intrinsic" to a given stimulus but may require supramodal integration of self-generated information (Shamay-Tsoory, Tomer, Berger, Goldsher, & Aharon-Peretz, 2005). Such self-generated information may also result from activity in retrosplenial cortex, known to be involved in episodic memory (Hassabis, Kumaran, & Maguire, 2007), as well as the anterior temporal pole, previously associated with socioemotional knowledge (Zahn et al., 2007), also observed for the main effect of JA.

The inverse contrast targeting the neural correlates of NOJA, that is, looking at an area of the screen not attended to by the other participant, demonstrated recruitment of a fronto-parietal network comprising activations in the medial and inferior parietal lobe and superior as well as middle frontal gyrus bilaterally. These activations are most likely related to the voluntary control of attention and inhibitory processes related to the control of eye movements, suggesting that participants had to resist an urge to follow someone else's gaze in order to look in a different direction instead (Pierrot-Deseilligny, Milea, & Muri, 2004). Consistent with this suggestion, participants' subjective reports demonstrated that they found it significantly more difficult to look at an object which was not looked at by the virtual other. This experiential difference was also reflected by measurements of autonomic arousal demonstrating significantly larger pupil sizes during NOJA.

In concordance with our hypotheses, the results further demonstrate differential effects for *self*- as compared to

other-initiated JA, as indicated by brain regions that show corresponding statistical interactions. Confirming our first hypothesis, *other*-initiated JA resulted in differential recruitment of anterior MPFC with significantly higher increase of activity during OTHER_JA > OTHER_NOJA than during SELF_JA > SELF_NOJA. Notably, the activation observed for this contrast is more dorsal and more anterior than the activation seen for the main effect of JA. Similar to suggestions of VMPFC being involved in coordinating perceptual and cognitive operations, the fronto-polar region has been implicated in “sensory gating,” possibly contributing to the coordination of “stimulus-oriented” as compared to “stimulus-independent processing.” This mechanism might be particularly relevant for “open-ended situations,” imposing a requirement for self-determined and self-maintained behavior, which corresponds to the task demands of our interactive experimental setup (Burgess et al., 2007). In this context, involvement of anterior MPFC during gaze-following could be interpreted as part of an orienting response to gaze shifts of others potentially relevant for the adoption of behavioral strategies from others. Results from studies in clinical groups support this interpretation. In particular, it has been shown that autistic individuals do not exhibit spontaneous responses to gaze shifts (Dawson et al., 2002), and alterations of anterior MPFC have, indeed, been implicated as a possible neurofunctional substrate of autism (Gilbert, Bird, Brindley, Frith, & Burgess, 2008; Williams et al., 2005; Waiter et al., 2004).

Consistent with our second hypothesis, JA resulted in a differential increase of neural activity in motivation-related brain regions, namely, the ventral striatum, which is more pronounced for self-initiated JA (SELF_JA > SELF_NOJA) than it is for other-initiated JA (OTHER_JA > OTHER_NOJA), consistent with a statistical interaction for this brain region. Ventral striatal neurocircuitry has been implicated during different phases of reward processing and can be closely related to the motivational and hedonic aspects of experiencing reward (Rolls, Grabenhorst, & Parris, 2008; Liu et al., 2007). Although this brain area was also recruited for the main effect of JA—consistent with behavioral data of both SELF_JA and OTHER_JA being rated as significantly more pleasant than SELF_NOJA and OTHER_NOJA—additional findings demonstrate that SELF_JA, in particular, engages reward-related brain areas: Consistent with a differential increase of activity in the ventral striatum for SELF_JA, correlation analyses of the BOLD signal in this region for each condition and ratings of the pleasantness of JA indicated a significant and positive correlation for SELF_JA, but not for any other experimental condition. This finding was supported by results of further analyses using the pleasantness rating as a covariate which demonstrated a differential increase of activity in the ventral striatum for SELF_JA, but not for OTHER_JA. Importantly, differences in neural activity also seem to be consistent with participants’ ratings as data from a follow-up behavioral study demonstrate SELF_JA to be rated significantly more pleasant than OTHER_JA. Taken together, these findings suggest

that although engagement in JA does rely upon reward-related neurocircuitry, being able to elicit a congruent response from another participant *oneself* specifically recruits the ventral striatum.

We suggest that this differential increase might be related to a higher degree of uncertainty and subsequent outcome monitoring for the performance of *initiative* actions as compared to *reactions* to someone else’s gaze as the former case includes the possibility of one’s own action not being reciprocated, whereas in the latter case one’s own action constitutes the reciprocation and may not need to be assessed in this respect. Another intriguing, yet more speculative, explanation would be that the differential increase of activity in the ventral striatum might be related to a higher degree of control over the other participant during self-initiated, reciprocated actions. It has been suggested that people universally seek control—or at least prediction—and that other people are core targets of people’s efforts to control their environments. Apart from affiliative motives, control can also be an important motivation for engagement with other people and for trying to make sense of them (e.g., Fiske & Dépret, 1996). Consequently, our findings may point toward a putative relationship of social motivation resulting from the interpersonal *coordination* of behavior and social influence, that is, *control* one has over the behavior of another person. Although the former has been indicated to impact on person perception, it is more controversial whether and by which mechanisms the latter influences social cognition (e.g., Boksem, Smolders, & De Cremer, 2009; Fiske, 1993). In our experiment, visual feedback was provided in all conditions to ensure that participants would always experience having an effect on the visual display. Likewise, a response by the virtual character—albeit an incongruent one half of the time—could always be observed. Consequently, we interpret the differential effect for SELF_JA in the ventral striatum with regard to participants being able to elicit a specific response from the other, namely, a congruent one. In contrast to our findings obtained from healthy controls which demonstrate activation of reward-related neurocircuitry for the *initiation* of JA, it has been suggested that autistic individuals’ established deficits in JA may result from alterations of its underlying motivational foundations (Mundy & Newell, 2007; Mundy, 1995). Making use of our paradigm to investigate autistic individuals may, therefore, help to investigate these matters further.

In summary, our results parallel recent findings which show that social referencing can activate reward-related brain regions known to realize strong influences on cognition and behavior (Fliessbach et al., 2007). JA also engages these mechanisms which—in light of behavioral data—might be contributory to an intrinsic motivation to engage in the interpersonal coordination of perspectives. We suggest that this could be closely related to the phenomenon’s impact on human cognitive development by contributing to the uniquely human motivation to engage in shared, social realities (Tomasello & Carpenter, 2007). Furthermore,

reward-related neurocircuitry appears to be differentially activated depending upon one's own role—performing *initiative* as compared to *responsive* actions—during joint attentional engagements. This, we believe, highlights the importance of exploring the reciprocal nature of social interaction and could help to explain why we so much enjoy the reciprocation of our own actions, particularly when this leads to a sharing of experience.

Acknowledgments

We thank all the colleagues in the MR and Cognitive Neurology group at the Research Centre Jülich for their support. In particular, we would like to thank Barbara Elghahwagi and Dorothe Krug for their help during fMRI data acquisition. L. S. is also grateful to Nicole David, Shaun Gallagher, Stefan Heim, Rüdiger Ilg, Bojana Kuzmanovic, Anna Rotarska-Jagiela, Tobias Schlicht, David Sharp, and Ralph Weidner for their helpful advice and stimulating comments.

The study was supported by the German Ministry for Education and Research, the Volkswagen Foundation, and by a personal grant to L. S. by the “Kompetenzzentrum NeuroNRW” at the Ministry of Innovation, Science, Research and Technology of North-Rhine Westfalia, Germany.

Reprint requests should be sent to Leonhard Schilbach, Department of Psychiatry, University of Cologne, Kerpener Str. 62, 50924 Cologne, Germany, or via e-mail: leonhard.schilbach@uk-koeln.de.

REFERENCES

- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*, 268–277.
- Argyle, J. M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge, MA: Cambridge University Press.
- Ashburner, J., & Friston, K. J. (2003). Rigid body registration. In R. S. Frackowiak, K. J. Friston, C. Frith, R. Dolan, C. J. Price, S. Zeki, et al. (Eds.), *Human brain function* (2nd ed., pp. 635–655). London: Academic Press.
- Becchio, C., Bertone, C., & Castiello, U. (2008). How the gaze of others influences object processing. *Trends in Cognitive Sciences*, *12*, 254–258.
- Boksem, M. A., Smolders, R., & De Cremer, D. (2009). Social power and approach-related neural activity. *Social Cognitive and Affective Neuroscience*. Epub ahead of print.
- Burgess, P. W., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in Cognitive Sciences*, *11*, 290–298.
- 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.
- Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, *18*, 192–205.
- Dawson, G., Webb, S., Schellenberg, G. D., Dager, S., Friedman, S., Aylward, E., et al. (2002). Defining the broader phenotype of autism: Genetic, brain, and behavioral perspectives. *Development and Psychopathology*, *14*, 581–611.
- Eickhoff, S. B., Heim, S., Zilles, K., & Amunts, K. (2006). Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. *Neuroimage*, *32*, 570–582.
- Eickhoff, S. B., Paus, T., Caspers, S., Grosbras, M. H., Evans, A. C., Zilles, K., et al. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage*, *36*, 511–521.
- Evans, A. C., Kamber, M., Collins, D. L., & MacDonald, D. (1994). An MRI based probabilistic atlas of neuroanatomy. In S. Shorvon, D. Fish, F. Andermann, & G. M. Bydder (Eds.), *Magnetic resonance scanning and epilepsy* (pp. 263–274). New York: Plenum Press.
- Fiske, S. T. (1993). Controlling other people. The impact of power on stereotyping. *American Psychologist*, *48*, 621–628.
- Fiske, S. T., & Dépret, E. (1996). Control, interdependence, and power: Understanding social cognition in its social context. In W. Stroebe & M. Hewstone (Ed.), *European review of social psychology* (pp. 31–61). New York: Wiley.
- Fliebsbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., et al. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science*, *318*, 1305–1308.
- Frith, C. D. (2007). The social brain? *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *362*, 671–678.
- Frith, C. D., & Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, *60*, 503–510.
- Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition*, *17*, 535–543.
- Gilbert, S. J., Bird, G., Brindley, R., Frith, C. D., & Burgess, P. W. (2008). Atypical recruitment of medial prefrontal cortex in autism spectrum disorders: An fMRI study of two executive function tasks. *Neuropsychologia*, *46*, 2281–2291.
- Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis of episodic memory. *Journal of Neuroscience*, *27*, 14365–14374.
- Kampe, K. K., Frith, C. D., & Frith, U. (2003). “Hey John”: Signals conveying communicative intention toward the self activate brain regions associated with “mentalizing,” regardless of modality. *Journal of Neuroscience*, *23*, 5258–5263.
- Kiebel, S., & Holmes, A. P. (2003). The general linear model. In R. S. Frackowiak, K. J. Friston, C. Frith, R. Dolan, C. J. Price, S. Zeki, et al. (Eds.), *Human brain function* (2nd ed., pp. 725–769). London: Academic Press.
- 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 of London, Series B, Biological Sciences*, *358*, 345–360.
- Legrand, D., & Iacoboni, M. (in press). Intersubjective intentional actions. In F. Grammont, D. Legrand, & P. Livet (Eds.), *Naturalizing intention in action. An interdisciplinary approach*. The MIT Press and ENS Editions.
- Liu, X., Powell, D. K., Wang, H., Gold, B. T., Corbly, C. R., & Joseph, J. E. (2007). Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. *Journal of Neuroscience*, *27*, 4587–4597.
- Materna, S., Dicke, P. W., & Thier, P. (2008). Dissociable roles of the superior temporal sulcus and the intraparietal sulcus in joint attention: A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *20*, 108–119.
- Moll, H., & Tomasello, M. (2007). Cooperation and human cognition: The Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *362*, 639–648.
- Moore, C., & Dunham, P. J. (1995). *Joint attention: Its origin and role in development*. Hillsdale, NJ: Erlbaum.
- Mundy, P. (1995). Joint attention, social-emotional approach in children with autism. *Development and Psychopathology*, *7*, 63–82.

- Mundy, P. (2003). The neural basis of social impairments in autism: The role of the dorsal medial-frontal and anterior cingulate system. *Journal of Child Psychology & Psychiatry*, *44*, 793–809.
- Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, *16*, 269–274.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- Penny, W. D., & Holmes, A. P. (2003). Random effects analysis. In R. S. Frackowiak, K. J. Friston, C. Frith, R. Dolan, C. J. Price, S. Zeki, et al. (Eds.), *Human brain function* (2nd ed., pp. 843–850). London: Academic Press.
- Pierrot-Deseilligny, C., Milea, D., & Muri, R. M. (2004). Eye movement control by the cerebral cortex. *Current Opinion in Neurology*, *17*, 17–25.
- Reddy, V. (2003). On being the object of attention: Implications for self–other consciousness. *Trends in Cognitive Sciences*, *7*, 397–402.
- Rolls, E. T., Grabenhorst, F., & Parris, B. A. (2008). Warm pleasant feelings in the brain. *Neuroimage*, *41*, 1504–1513.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, *16*, 235–239.
- Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., et al. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, *44*, 718–730.
- Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., Goldsher, D., & Aharon-Peretz, J. (2005). Impaired “affective theory of mind” is associated with right ventromedial prefrontal damage. *Cognitive and Behavioral Neurology*, *18*, 55–67.
- Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews*, *30*, 855–863.
- Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10*, 121–125.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, *15*, 273–289.
- Waiter, G. D., Williams, J. H., Murray, A. D., Gilchrist, A., Perrett, D. I., & Whiten, A. (2004). A voxel-based investigation of brain structure in male adolescents with autistic spectrum disorder. *Neuroimage*, *22*, 619–625.
- Williams, J. H., Waiter, G. D., Perra, O., Perrett, D. I., & Whiten, A. (2005). An fMRI study of joint attention experience. *Neuroimage*, *25*, 133–140.
- Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, *4*, 58–74.
- Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *104*, 6430–6435.