



Published in final edited form as:

*Neuroimage*. 2010 April 1; 50(2): 693–700. doi:10.1016/j.neuroimage.2009.12.101.

## BRAIN MECHANISMS FOR REPRESENTING WHAT ANOTHER PERSON SEES

Ratha D. Heyda, Steven R. Green, Brent C. Vander Wyk, James P. Morris, and Kevin A. Pelphrey

Yale Child Study Center, Yale University

### Abstract

We used functional magnetic resonance imaging (fMRI) and a naturalistic joint attention scenario to evaluate two, alternative hypotheses concerning the social brain. The first, *Content Specific Attribution* hypothesis, was that core regions previously identified as being involved in social cognition also participate in representing the *contents* of another mind. The second, *Dual Role* hypothesis, was that extrastriate, category-specific visual regions respond to a visible stimulus of a specific category *and* to the same stimulus occluded, but when it appears to be the focus of another person's visual attention. Participants viewed category-specific stimuli (Place and Body images) to localize the extrastriate body area (EBA) and parahippocampal place area (PPA). Then, they observed a computerized character viewing each stimulus category, occluded from the participant's view. In support of the *Content Specific Attribution* hypothesis, whole-brain analyses revealed that viewing someone else looking at an occluded picture of a body activated brain regions previously associated with components of social cognition more than viewing someone else looking at an occluded picture of a place. Counter to the *Dual Role* hypothesis, functional region of interest (ROI) analyses revealed that the EBA and PPA were not clearly involved in representing what the character was seeing.

### Keywords

fMRI; body; place; occlusion; social perception

## INTRODUCTION

In order to navigate a social world successfully, a person must recognize the basic fact that the people around them *have* minds; they must attribute mental states to them, such as beliefs, desires, goals, intentions, and so on. However, they must also populate those attributed minds with *content*. The ability to make contentful attributions of mental states to others has been termed 'theory of mind' (ToM) (Premack & Woodruff, 1978) and is argued to be one basis by which people may predict the future actions and understand prior actions of other people.

Neuroimaging studies to date have identified a number of cortical regions involved in aspects of social cognition. For example, the posterior superior temporal sulcus (STS) region has been shown to respond to eye gaze (Pelphrey, Singerman, Allison, & McCarthy, 2003; Puce, Allison,

---

Author contact: Kevin Pelphrey, Yale University, Yale Child Study Center, 230 South Frontage Road, New Haven, CT 06520, kevin.pelphrey@yale.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Bentin, Gore, & McCarthy, 1998; Wicker, Michel, Henaff, & Decety, 1998) and is sensitive to the intentions conveyed by and the context within which such biological motions as eye gaze shifts are observed (Pelphrey et al., 2003; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004). The temporoparietal junction (TPJ) is involved in attributing and reasoning about others' mental states (Saxe & Kanwisher, 2003), and the medial prefrontal cortex (MpFC) responds to joint attention in triadic relations and when processing socially relevant information about others (Saxe, 2006; Williams, Waiter, Perra, Perrett, & Whiten, 2005).

While these areas are clearly engaged whenever an attribution of another mind is necessary, it is unclear to what extent these regions participate in representing the specific content of others' minds. One possibility, which we will call the *Content Specific Attribution* hypothesis, is that these regions, or some of these regions, directly represent the specific content of another person's mind, as well as participating in representing the attribution that the other person has mental states. A second possibility is that these core regions do not represent the content themselves, but rather coordinate their activity with other regions of the brain that do so. Although there may be many candidates for such a set of regions, we appealed to previous work that describes category selective regions within the extrastriate visual cortex for the content-specific encoding role in the present study. Thus, the logic of the *Dual Role* hypothesis dictates that the very same regions that participate in the first-order encoding of specific information when a person is thinking about or perceiving that kind of stimuli may also play a role in representing the content of attributed mental states of others.

One important way that humans evaluate others' interests and intentions is by following and interpreting their direction of eye gaze, thereby coordinating attention with others. The ability to partake in 'joint attention' permits the observer to know what another perceives, thereby gaining insight into another's intentions, and anticipating his or her actions (Tomasello, Carpenter, Call, Behne, & Moll, 2005). In the present study, participants viewed a realistic virtual avatar shift his gaze to pictures of bodies or places. Importantly, the participants knew the contents of the pictures, but the pictures were occluded to the participant's view while the avatar was looking at them.

Since the current study employs a scenario involving a form of joint attention via gaze perception and invites the attribution of "seeing" to the observed character, we hypothesized that the STS, TPJ, and MpFC would be engaged by our stimulus paradigm. If, as dictated by the Content Specific Attribution hypothesis, these core regions of social cognition are the sole areas responsible for encoding the content of others' minds, then we expected to see differences in activation as a function of stimulus type in these regions and not see differences elsewhere.

To evaluate the Dual Role hypothesis, we gained leverage from the fact that previous functional magnetic resonance imaging (fMRI) studies have identified category-specific brain regions in extrastriate visual cortex that selectively respond to images of faces, bodies, or places, including the fusiform face area (FFA) (Kanwisher, McDermott, & Chun, 1997; Puce, Allison, Asgari, Gore, & McCarthy, 1996), extrastriate body area (EBA) (Downing, Jiang, Shuman, & Kanwisher, 2001), and parahippocampal place area (PPA) (Aguirre, Zarahn, & D'Esposito, 1998; Epstein & Kanwisher, 1998). In formulating our approach, we noted that category-specific visual regions are also active when participants mentally visualize (O'Craven & Kanwisher, 2000) or view either partially occluded (Hulme & Zeki, 2007) or invisible stimuli of a region's 'preferred' category (Moutoussis & Zeki, 2002). The Dual Role hypothesis for these regions suggests that the mechanism for representing the contents of another person's visual field involves stimulus-specific activation in the participant's own category-specific regions of extrastriate cortex in response to seeing another person look towards an occluded stimulus belonging to a particular category. If differential activity is detected within category-specific regions during periods in which the avatar is looking at the occluded objects and is

not detected in the core social cognition regions (STS, TPJ, and MpFC), this would be consistent with the Dual Role hypothesis and not the Content Specific Attribution hypothesis.

## MATERIALS AND METHODS

### Participants

Fifteen healthy adults (ages 20–31, 14 right-handed, 1 left-handed) participated. Nine participants were female and six were male. All had normal or corrected-to-normal vision. Data from three of these 15 participants (all right handed, 2 male, 1 female) were excluded from analysis because of excessive (> 4 mm) participant movement (two participants) or scanner technical problems leading to low signal-to-noise ratios (one participant). All participants provided written informed consent to participate. Participants were recruited from the community surrounding Duke University in Durham, North Carolina, USA. The participants were paid \$50 for their participation in this study.

### Tasks

The experimental design consisted of three parts which we label here: *fMRI localizer*, *fMRI task*, and *eye tracking*. Stimuli for the fMRI task and eye tracking were similar, whereas the fMRI localizer used different, independent stimuli. Each component is detailed below.

**fMRI localizer**—All participants were scanned while viewing independent localizer stimuli following the task scan. There were two localizer runs for each participant, and each run consisted of 16 blocks of images. Each block showed grayscale photographic images of a single stimulus category, namely places (including indoor and outdoor scenes), faces, bodies (without heads), or flowers. Each block consisted of 24 images shown for 500 ms each over the course of 12 s. Between each block, a fixation cross was visible in center screen for 12 s to reestablish a baseline. Overall, 4 blocks of each stimulus category were shown per run.

**fMRI task**—Two experimental conditions were animated using the Poser software program (Curious Labs Inc., Santa Cruz, California). As illustrated in Figure 1, in both conditions, a virtual male figure (head, neck, and shoulders visible) was flanked on either side by two suspended cards. One card (the body card) showed a grayscale image of a whole body (with head and face intact), and the other (the place card) showed a grayscale image of a place (an indoor or outdoor scene). Before each trial began, two animated parts were shown. First, the fronts of the body card and place card were visible for 10 s, allowing the participant to see which card showed a body and which showed a place. Then, the cards flipped over, showing the identical backs of the cards for 12 s, providing a baseline period. The trial began when the card on the right or the left side moved down to a location directly in front of the virtual male figure. In the Place condition, the place card moved down, and in the Body condition, the body card moved down. The character broke eye contact with the participant to look down at the card, which was facing him. The male figure could thus ‘see’ the picture on the card (either of body or place), while the participant could only see the back of the card. The body or place card remained in front and was ‘visible’ to the virtual male figure for 10 s. Then, the card returned to its original position, completing one trial. Two more trials followed, creating a set of three trials in all. Sets of three trials were employed to allow us to remind the participant of the locations (right or left side) of the place and body cards after three trials. A 10-s intertrial interval separated each of the three trials in a set. After a set, the cards were again in their original position at the sides, with body and place images visible to the participant, reestablishing a baseline. A run consisted of 4 sets (12 trials), and there were 6 runs in all. Each run used a different pair of place and body images on the cards, and their respective positions on the right and left alternated with each run. On each trial, the card that came down could be either a place card or body card and could come down from the right or left side of the virtual

male figure. Crossing the card type with the side from which the card came down thus created four possible stimulus configurations. We pseudorandomized the sequence of presentation for these four conditions within blocks and within runs, so that the number of conditions was counterbalanced within and across runs. In total, there were 36 Body trials and 36 Place trials.

**Eye tracking**—Following the scan, 11 of the 12 participants whose data were used in the fMRI analyses participated in an eye-tracking task, using a subset of the fMRI task stimuli. During eye tracking, participants were seated at a comfortable viewing distance from a monitor, upon which the stimuli were presented in 3 separate runs. A remote infrared pupil-corneal reflection eye-movement monitoring system (Tobii 1750, <http://www.tobii.com>) was used to record the participant's point-of-regard data. First, participants viewed one of the same runs used during the fMRI scan. Second, participants viewed a short segment of another run. When the short segment was complete, participants were asked to identify whether the man in the animation was looking at a body or a place, and the answer was recorded by the experimenter. They were not warned in advance that they would be asked this question. Third, participants viewed another of the fMRI task runs and were instructed to press P for place and B for body, depending on what the character was viewing.

### Imaging Parameters

Participants were scanned in a GE Signa EXCITE HD 3.0 Tesla scanner (General Electric, Waukesha, Wisconsin, USA). High-resolution anatomical images were first acquired for each participant. Sixty-eight high-resolution images were acquired using a 3D fast SPGR pulse sequence (FOV = 24 cm; image matrix = 2562; voxel size =  $0.9375 \times 0.9375 \times 1.9$  mm) and used for coregistration with the functional data. These structural images were aligned in a near-axial plane defined by the anterior and posterior commissures. Whole-brain functional images were acquired using echo-planar imaging sensitive to blood oxygenation level dependent (BOLD) contrast (TR = 2000 ms; TE = 27 ms; FOV = 24 cm; voxel size =  $3.75 \times 3.75 \times 3.8$  mm; 34 near axial slices). The functional images were aligned identically to the structural images.

### Data Analysis

Image preprocessing was performed using SPM 99 (Wellcome Department of Cognitive Neurology, Queen Square, London, United Kingdom) modules and custom programs written in MATLAB. Center-of-mass measurements were used to detect participant motion. Images were time-slice adjusted to compensate for an interleaved slice acquisition and realigned to the tenth image to correct for head movements between scans. For some analyses, the realigned scans were spatially normalized to the Montreal Neurological Institute (MNI) template. The functional data were high-pass filtered and spatially smoothed with an 8 mm isotropic Gaussian kernel prior to the statistical analysis. Except where otherwise noted, these normalized and smoothed data were used in the analysis procedures described below.

**fMRI localizer**—For individual-level analysis of the localizer scans, we conducted a time-point-by-time-point *t*-test analysis comparing the responses to Body versus Faces to localize the EBA, and Places versus Bodies to localize the PPA. For this analysis, we used the acquisition-aligned and motion-corrected, smoothed, unnormalized imaging data. Using this data, overlaid on each participant's own anatomical images, we identified, on a participant-by-participant basis, regions of activation within the expected, likely anatomical locations for the EBA and PPA that (1) exhibited significantly ( $t = 1.96, p < .05$ , uncorrected) greater activity to pictures of places (for the PPA) or bodies (for the EBA) compared respectively to Bodies or Faces averaging across the 6-16 s period following the block onset, and (2) encompassed an area greater than eight functional voxels. We then examined the response from the

independently defined EBAs and PPAs to the Body and Place trials from the fMRI-task portion of the experiment.

For group-level analysis of the data from the localizer scan, we conducted random effects analyses of the differences between the hemodynamic responses (HDR) during Bodies, Places, and Faces blocks, beginning 6–16 s after block onset using the normalized data. This analysis consisted of the following steps: (1) The epoch of image volumes beginning 2 images before (–4.0 s) and 8 images after (16 s) the onset of each trial type was excised from the continuous time series of volumes and selectively averaged into one of three bins (Places, Bodies, or Faces). (2) The average intensity of each of the four resulting average HDRs over the post-onset period of 6–16 s was computed for each voxel. A *t*-statistic was then computed at each voxel within the brain to quantify the HDR differences among selected stimulus types (Bodies > Faces and Place > Body). This process was performed separately for each participant. (3) The individual *t*-maps created in the preceding step were then subjected to a random-effects analysis that assessed the significance of differences across-participants, thereby creating group-average Bodies > Faces and Places > Bodies whole-brain statistical maps.

**fMRI task**—Likewise, for some group-level analyses of the fMRI-task data we conducted random effects analyses of the differences between Body and Place trials at the expected peak of the HDR. This analysis consisted of the following steps: (1) The epoch of image volumes beginning 2 images before (–4.0 s) and 6 images after (12 s) the onset of each trial type (time locked to the movement of the animated character’s head) was excised from the continuous time series of volumes and selectively averaged into one of two bins (Place, Body). (2) The average intensity of each of the two resulting average HDRs over the post-onset period of 4–8 sec was computed for each voxel. A *t*-statistic was then computed at each voxel within the brain to quantify the HDR differences between Body and Place trials. This process was performed separately for each participant. (3) The individual *t*-maps created in the preceding step were then subjected to a random-effects analysis that assessed the significance of differences across-participants, thereby creating group-average Body > Place and Place > Body statistical maps.

## RESULTS

### fMRI localizer

Figure 2a shows the results of a random effects analysis contrasting, at a group level, the response to Places vs. Bodies (to identify the PPA) and Bodies vs. Faces (to identify the EBA) from 6–16 s post block onset ( $p < .01$ , cluster size > 8 functional voxels). As can be seen, a robust PPA (blue to light blue color map) was identified in the expected location based on prior reports concerning the PPA (e.g., Aguirre et al., 1998; Epstein & Kanwisher, 1998). Similarly, activation localized to a region consistent with the expected location of the EBA (e.g., Downing et al., 2001) was identified at the group level. Unexpectedly, we identified a portion of cortex in and around the right inferior parietal lobule that responded more strongly to bodies than to any other stimulus class shown during the localizer. Inspection of the waveforms presented in Figure 2b and 2c revealed the expected pattern of responses. The EBA responded more strongly to bodies than to the other three classes of stimuli. The PPA responded most strongly to places compared to all other classes of stimuli. However, the EBA exhibited somewhat less specificity than the PPA in our data set.

We also conducted individual participant analyses. We were able to localize the PPA in 10 of 12 participants in both the right and left cerebral hemispheres, and the EBA in 9 of 12 participants. In those 9 participants with an EBA, 3 had only a right EBA and 6 had both a right and left EBA.

These findings indicate that, in full agreement with the prior literature, we were able to localize regions of extrastriate visual cortex that were specific for their preferred category of visual stimulus. These were apparent at both the group and individual level of analysis. This process of independently localizing the PPA and EBA allowed us to explore how these very same regions behaved in the context of the task that involved a person looking at places and bodies that were occluded from the participant's view.

### fMRI task

Figure 3 shows waveforms time-locked to the onset of the character moving his head to look at the card. These waveforms are the across-participants average HDR from the individual participant-defined PPAs and EBAs. For the PPA (Figure 3a), activation was greatest in the Place condition, but this difference was not statistically significant. A more posterior region in the lingual gyrus did show activation significantly greater for the Place condition (Figure 3c). This region was also selectively active to Places in the localizer. The EBA (Figures 3b) also showed significantly greater activation to the Place condition, not the Body condition. The HDR waveforms from the group average defined PPA and EBA (the less sensitive analysis) did not show statistically significant differences between the two stimulus categories.

We also conducted a whole-brain random effects analysis to identify regions that responded more strongly for Place versus Body trials and Body versus Place trials during the fMRI-task portion of the experiment. A description of these regions is provided in Table 1 in the form of weighted centers of the listed regions of activation. The Place > Body and Body > Place contrast maps are shown in Figure 4 ( $p < .01$ , cluster size  $\geq 8$  functional voxels).

Regions active for the Body > Place contrast included the left inferior parietal lobule, and areas along the anterior and posterior STS bilaterally. Also active were regions in the left precentral gyrus, left superior frontal gyrus (SFG), MpFC, and left insula/ventromedial prefrontal cortex (Ins). Regions active for the Place > Body contrast included left postcentral gyrus, right intraparietal sulcus (IPS), right inferior temporal gyrus (ITG) and right lingual gyrus (LG), right superior frontal sulcus (SFS), and bilateral inferior frontal gyri (IFG).

### Eye tracking

Figure 5 shows the pattern of point-of-regard data during Place and Body conditions across all participants over the first eye-tracking run. During trials when the character is looking at the occluded card, participants' point of regard alternated rapidly between the character's face and the occluded card. When accounting for all data points within the run, 21% of them were on the place card, 21% on the body card, 25% on the character's face during the Place condition, and 18% on the character's face during the Body condition. This difference in looking at the face trended towards significance ( $t = 2.10$ ,  $p = .06$ ). During the Place condition, 7% of data points were outside either the character's face or the occluded card, and during the Body condition, 8% were similarly outside. The results indicate that participants' attention was almost equally divided between the male character and the occluded card.

Following Run 2 of the eye-tracking component of the study, all participants answered the unanticipated question regarding the category of image on the back of the card correctly, indicating that even when not explicitly instructed to remember whether the character was viewing a Body or Place, participants were maintaining an implicit awareness of this information. Ten out of 12 participants completed the button press task in Run 3 with no errors, and 2 out of 12 made one error. This indicates that when so instructed, participants had little difficulty maintaining the contents of the cards in short-term memory.

## DISCUSSION

Whole-brain analyses revealed that brain regions previously associated with components of social cognition activated more strongly to viewing someone else looking at an occluded picture of a body more than looking at an occluded picture of a place. The fMRI results from the present study show that the EBA and PPA are not necessarily involved in representing the contents of what another person is seeing. These findings are consistent with the Content Specific Attribution hypothesis that the representation of the specific content of others' minds is done by the same core network of brain regions that has been implicated in mental state attribution and social cognition more abstractly.

An important departure from prior studies was our naturalistic joint attention/occlusion scenario. Its goal was to identify neural circuitry involved in spontaneous, automatic representation of the contents of another's visual field, not the more top-down, consciously directed processes that have been the objective of prior studies (e.g., O'Craven & Kanwisher, 2000). Prior to the scan, participants were instructed only to "pay attention to what the man in the movie is seeing." They were not directed to create any mental images or to examine mentally such images. Therefore, our study examined whether EBA and PPA are involved in neural processes relatively free from direct perception or strategically directed imagery. We observed a somewhat paradoxical finding of greater EBA activity in the Place condition. This finding might be attributable to the trend showing that participants spent more time looking at the virtual character's face and body during Place trials than during Body trials. Nonetheless, overall our results indicate that these regions are not clearly involved in representing the contents of another's visual field.

Within the core social cognition network, the dorsal MpFC, posterior STS, and IPL showed greater activation to the Body condition than the Place condition. The direction of the contrast may be related to social cognition processes evoked by seeing the character looking at another person. Dorsal MpFC has been implicated in various social cognition tasks, including representation of triadic relations (reviewed in Saxe, 2006). Williams and colleagues (2005) conducted an elegant joint attention study which showed MpFC activation when the participant and an animated character both viewed the same moving object. In our study, both Body and Place conditions involved joint attention, and we did not find dorsal MpFC activation in common between the conditions. However, our Body condition created a more specific kind of social triadic attention, a person viewing an interaction between two other persons (albeit, a picture of a person on a card and an animated cartoon figure). Similarly, Walter and colleagues (2004) found dorsal MpFC activation when participants observed collaboration between two cartoon characters. Furthermore, the pSTS and IPL have been implicated in studies involving Heider and Simmel (1944) type animations of simple geometric shapes that move in such a way as to evoke mentalistic attributions and the perception of social interactions (e.g., Castelli, Frith, Happé, & Frith, 2002).

We also found stimulus-dependent activity in other regions beyond the core social network. These differences might be related to mechanisms of attention and working memory. Eye tracking revealed that when the character was viewing the occluded card, participants' point of regard repeatedly alternated between the character and the card for both the Body and Place conditions. The joint attention situation thus split the participant's overt visual attention between the character and the card. Previous findings, discussed below, might help to illuminate how this distribution of attention might have impacted activity in other regions.

The two conditions evoked unique activation patterns in portions of the frontal cortices. Specifically, the Place > Body contrast revealed activity in the right SFS and the IFG bilaterally. The Body > Place condition activated the left precentral gyrus, dorsal MpFC, and left Ins. The

maintenance of attention to the occluded body or place likely recruited working memory processes. Participants' correct answers to the unexpected question during eye tracking suggested that although they were not asked to actively remember or rehearse whether the character was looking at Body or Place, representations of the task conditions were maintained in working memory.

Extensive evidence points to dissociable neural systems subserving two types of attention-directed visual processing: the dorsal, occipitoparietal space-based stream which specifies location ('where'), and the ventral, occipitotemporal object-based stream which specifies object identity ('what') (Haxby et al., 1991; Shomstein & Behrmann, 2006; Yantis & Serences, 2003). Attention can thus be directed to spatial properties (location, orientation, and layout) or object-based properties (identity, contour). We suggest that spatial properties of the Place card differentially (Place > Body) activated the right IPS, a region known to be involved in spatial attention. In contrast, object-based properties of the Body card differentially (Body > Place) activated unique regions of the occipitotemporal object-based attention stream in the anterior STS region. Non-human primate single-cell recordings reveal the existence of cell populations around the anterior STS that respond to static images of faces and bodies, as well as to implied past or future movement (Oram & Perrett, 1996; Perrett, Smith, Mistlin et al., 1985; Perrett, Smith, Potter et al., 1985; Wachsmuth, Oram, & Perrett, 1994). Further, Baker et al. (2001) have described cell populations in the monkey anterior STS active in conditions in which a person is occluded from view after having been previously seen. They propose that anterior STS may be involved in maintaining object permanence for socially meaningful stimuli. Our observed anterior STS activity may have served a similar purpose.

Analogous to the dissociation of spatial and object-based attention streams, attempts have been made to identify a similar division of labor in working memory. For example, a meta-analysis found evidence of a left/right hemispheric dissociation, with object working memory localized more to the lateral left frontal lobe and spatial working memory more to the right SFS (D'Esposito et al., 1998). Our Place > Body contrast activated right SFS, in agreement with previous studies of spatial working memory. It also activated the IFG bilaterally. In previous studies, the right IFG was active in object working memory (Haxby, Ungerleider, Horwitz, Rapoport, & Grady, 1995) and in spatial working memory (Jonides et al., 1993). Courtney and colleagues (1996) have suggested that this region may contain neurons involved in both object and spatial working memory processes. It is possible that individual components of the scene in our Place cards could activate object working memory, while the scene layout could activate spatial working memory as well, as our study was not designed to dissociate these two aspects of working memory.

In summary, forming and maintaining a representation of another person's mental state requires making an attribution that they have mental states. But to be useful, these attributed mental states must be filled with content. In the simple naturalistic joint attention paradigm, we found differences in activity within core regions of the social cognition network that was dependent on the content being attributed to the avatar. This is consistent with the Content Specific Attribution hypothesis that these regions not only make attributions of mental state, but also encode the specific content of those states. Our finding was also consistent with this hypothesis and inconsistent with a Dual Role hypothesis, in that that the category-specific visual processing regions like EBA and PPA are not automatically involved in representing the contents of what another person is seeing, even though these regions are clearly involved in forms of perception, early forms of stimulus specific processing, and mental imagery.

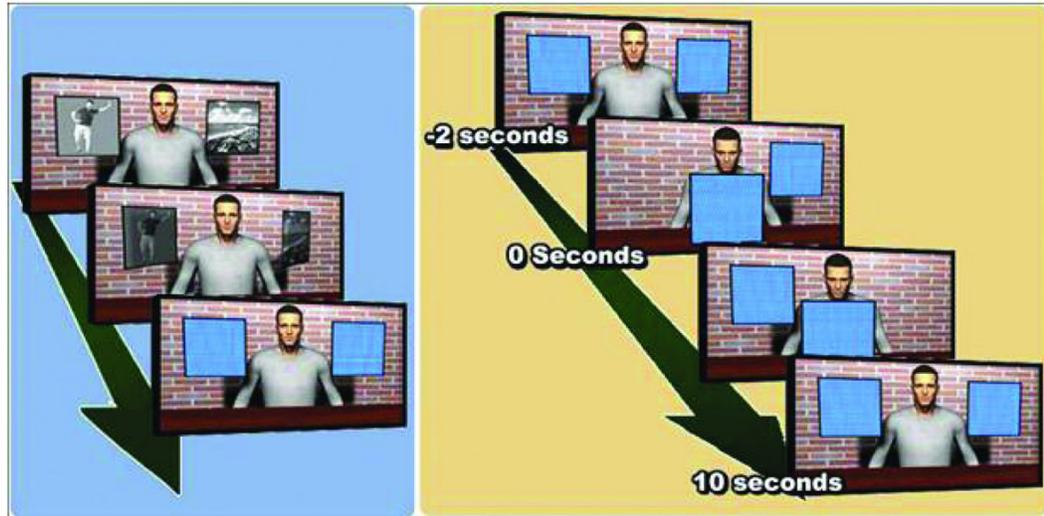
## Acknowledgments

We thank Susan Music for her assistance with this research. This research was supported by the NINDS, NIMH, the John Merck Scholars Fund, and Autism Speaks. A NIH Roadmap Fellowship supported Ratha Heyda. Kevin Pelphrey was supported by a Career Development Award from the NIMH (MH071284).

## References

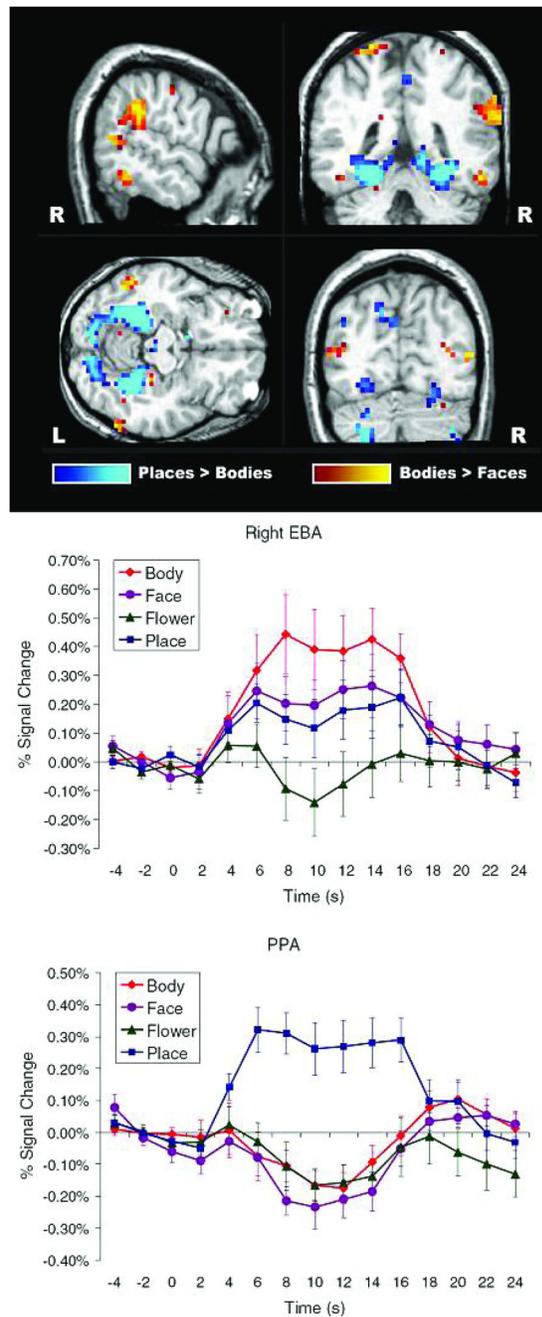
- Aguirre GK, Zarahn E, D'Esposito M. An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron* 1998;21(2):373–383. [PubMed: 9728918]
- Baker CI, Keysers C, Jellema T, Wicker B, Perrett DI. Neuronal representation of disappearing and hidden objects in temporal cortex of the macaque. *Experimental Brain Research* 2001;140(3):375–381.
- Castelli F, Frith C, Happé F, Frith U. Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* 2002;125(8):1839–1849. [PubMed: 12135974]
- Courtney SM, Ungerleider LG, Keil K, Haxby JV. Object and spatial visual working memory activate separate neural systems in human cortex. *Cerebral Cortex* 1996;6(1):39–49. [PubMed: 8670637]
- D'Esposito M, Aguirre GK, Zarahn E, Ballard D, Shin RK, Lease J. Functional MRI studies of spatial and nonspatial working memory. *Brain Research, Cognitive Brain Research* 1998;7(1):1–13. [PubMed: 9714705]
- Downing PE, Jiang Y, Shuman M, Kanwisher N. A cortical area selective for visual processing of the human body. *Science* 2001;293(5539):2470–2473. [PubMed: 11577239]
- Epstein R, Kanwisher N. A cortical representation of the local visual environment. *Nature* 1998;392(6676):598–601. [PubMed: 9560155]
- Haxby JV, Grady CL, Horwitz B, Ungerleider LG, Mishkin M, Carson RE, et al. Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy Sciences U S A* 1991;88(5):1621–1625.
- Haxby JV, Ungerleider LG, Horwitz B, Rapoport SI, Grady CL. Hemispheric differences in neural systems for face working memory: A PET-rCBF study. *Human Brain Mapping* 1995;3:68–82.
- Heider F, Simmel M. An experimental study of apparent behaviour. *American Journal of Psychology* 1944;57:243–259.
- Hulme OJ, Zeki S. The sightless view: neural correlates of occluded objects. *Cerebral Cortex* 2007;17(5):1197–1205. [PubMed: 16844722]
- Jonides J, Smith EE, Koeppe RA, Awh E, Minoshima S, Mintun MA. Spatial working memory in humans as revealed by PET. *Nature* 1993;363:623–625. [PubMed: 8510752]
- Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience* 1997;17(11):4302–4311. [PubMed: 9151747]
- Morris JP, Pelphrey KA, McCarthy G. Face processing without awareness in the right fusiform gyrus. *Neuropsychologia* 2007;45(13):3087–3091. [PubMed: 17643452]
- Moutoussis K, Zeki S. The relationship between cortical activation and perception investigated with invisible stimuli. *Proceedings of the National Academy of Sciences U S A* 2002;99(14):9527–9532.
- O'Craven KM, Kanwisher N. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience* 2000;12(6):1013–1023. [PubMed: 11177421]
- Oram MW, Perrett DI. Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology* 1996;76(1):109–129. [PubMed: 8836213]
- Pelphrey KA, Singerman JD, Allison T, McCarthy G. Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia* 2003;41(2):156–170. [PubMed: 12459214]
- Perrett DI, Smith PAJ, Mistlin AJ, Chitty AJ, Head AS, Potter DD, et al. Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: A preliminary report. *Behavioural Brain Research* 1985;16:153–170. [PubMed: 4041214]

- Perrett DI, Smith PAJ, Potter DD, Mistlin AJ, Head AS, Milner AD, et al. Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London B Biological Sciences* 1985;223:293–317.
- Premack D, Woodruff G. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* 1978;4:515–526.
- Puce A, Allison T, Asgari M, Gore JC, McCarthy G. Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *Journal of Neuroscience* 1996;16(16):5205–5215. [PubMed: 8756449]
- Puce A, Allison T, Bentin S, Gore JC, McCarthy G. Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience* 1998;18(6):2188–2199. [PubMed: 9482803]
- Saxe R. Uniquely human social cognition. *Current Opinion in Neurobiology* 2006;16(2):235–239. [PubMed: 16546372]
- Saxe R, Kanwisher N. People thinking about thinking people. The role of the temporo-parietal junction in “theory of mind”. *Neuroimage* 2003;19(4):1835–1842. [PubMed: 12948738]
- Saxe R, Xiao DK, Kovacs G, Perrett DI, Kanwisher N. A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia* 2004;42(11):1435–1446. [PubMed: 15246282]
- Shomstein S, Behrmann M. Cortical systems mediating visual attention to both objects and spatial locations. *Proceedings of the National Academy of Sciences U S A* 2006;103(30):11387–11392.
- Tomasello M, Carpenter M, Call J, Behne T, Moll H. Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences* 2005;28(5):675–691. discussion 691–735. [PubMed: 16262930]
- Wachsmuth E, Oram MW, Perrett DI. Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. *Cerebral Cortex* 1994;4(5):509–522. [PubMed: 7833652]
- Walter H, Adenzato M, Ciaramidaro A, Enrici I, Pia L, Bara BG. Understanding intentions in social interaction: The role of the anterior paracingulate cortex. *Journal of Cognitive Neuroscience* 2004;16(10):1854–1863. [PubMed: 15701234]
- Wicker B, Michel F, Henaff MA, Decety J. Brain regions involved in the perception of gaze: a PET study. *Neuroimage* 1998;8(2):221–227. [PubMed: 9740764]
- Williams JH, Waiter GD, Perra O, Perrett DI, Whiten A. An fMRI study of joint attention experience. *NeuroImage* 2005;25(1):133–140. [PubMed: 15734350]
- Yantis S, Serences JT. Cortical mechanisms of space-based and object-based attentional control. *Current Opinion in Neurobiology* 2003;13(2):187–193. [PubMed: 12744972]



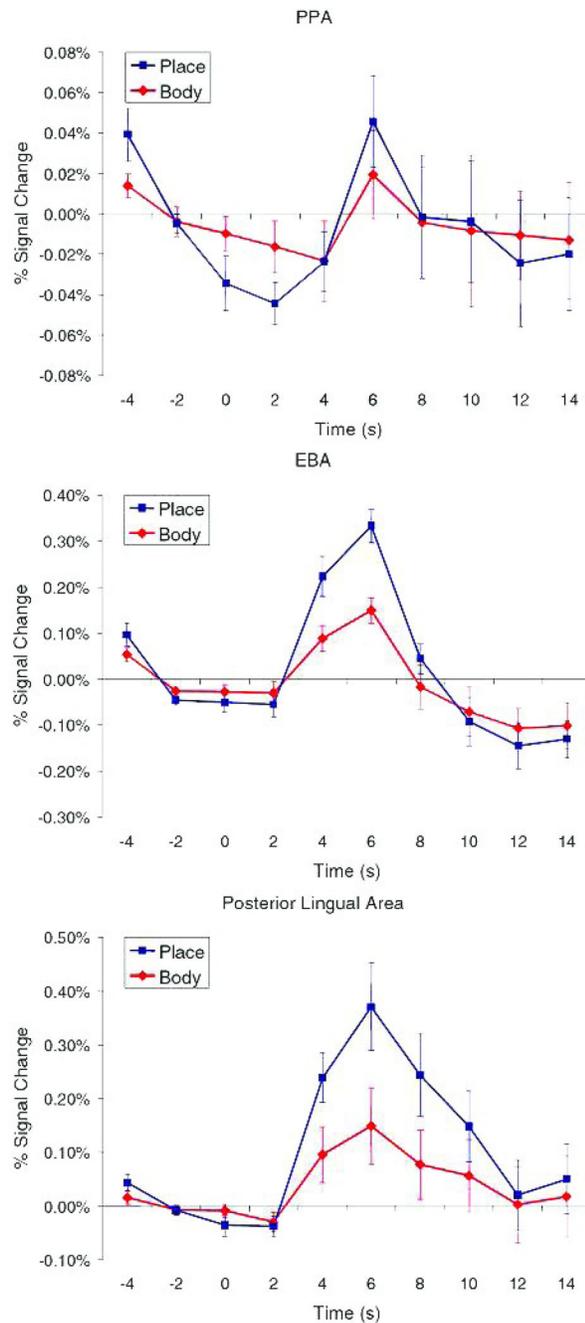
**Figure 1.**

The task paradigm used in the scanner. A virtual reality animation portrayed a male character flanked by 2 cards showing images of a Body and Place. The backs of the cards had identical blue patterns. Initially, both images were visible, and then the cards flipped, so only the backs were visible. In a trial, the card on the R or L moved down in front of the character, so he could see the image, while the participant could only see the card's back. Then the card returned to its original location. After a 3 trial set, the card images were again visible at the sides. There were 6 runs, each made up of 4 sets (12 trials per run).

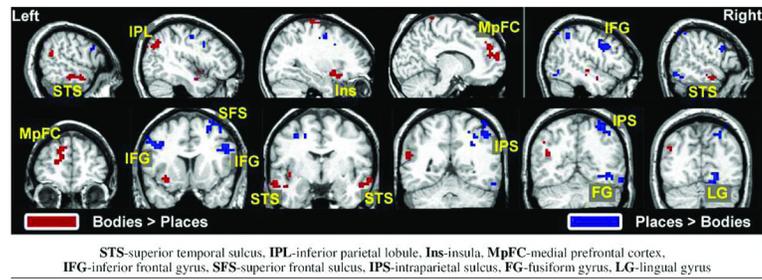


**Figure 2.**

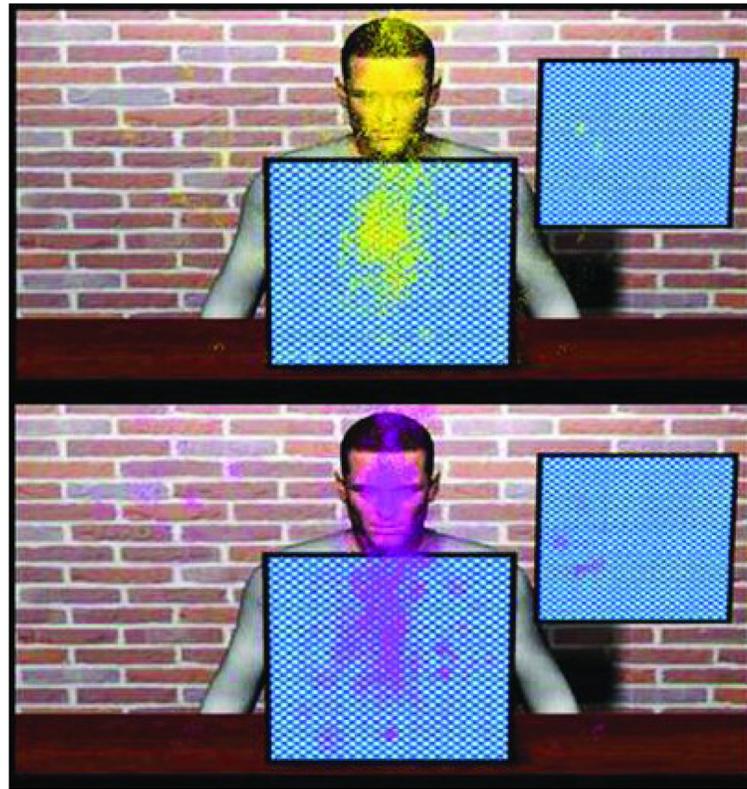
Two group-level random-effects contrasts were performed using BOLD responses to the two Localizer runs averaged across 12 participants ( $p < .01$ , uncorrected,  $k \geq 8$  functional voxels). (a) The Places > Bodies contrast revealed robust PPA activity with extension into posterior lingual cortex. The Bodies > Faces contrast showed EBA activity and unexpected inferior parietal lobule activity. (b) BOLD time courses in right EBA (averaged across hemispheres) for the four localizer conditions. (c) BOLD time courses in PPA for the four localizer conditions.



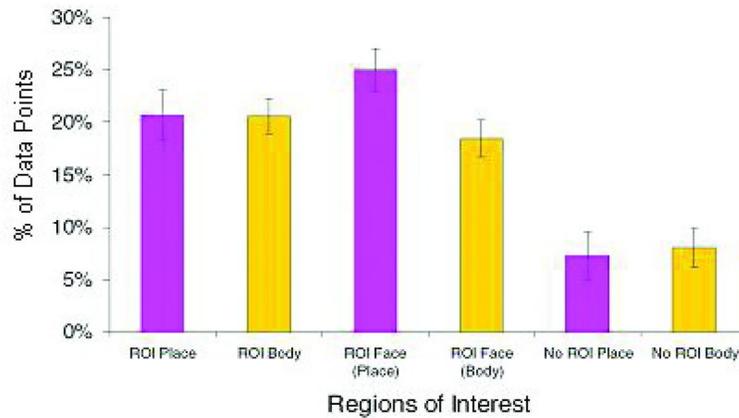
**Figure 3.** Comparison of task waveforms from the individual participant defined regions of interest representing the PPA, EBA, and posterior lingual area. (a) Unexpectedly, EBA showed greater activation bilaterally in the Place condition. (b) PPA showed slightly greater activation to Place condition than to Body condition bilaterally, but the waveforms did not differ significantly. (c) An area slightly posterior to the expected location of the PPA showed greater activation to Place condition. This area also responded to the Place localizer.



**Figure 4.** Whole-brain random-effects analysis. This analysis identified regions responding more strongly for Place > Body and Body > Place contrasts ( $p < .01$ , uncorrected, cluster size  $\geq 8$  functional voxels). The contrast maps revealed distinctly different activation patterns for the two conditions.



Eye-tracking Fixations



**Figure 5.**

Eye-tracking task images showing all participants' point of regard across Run 1. (a) Point-of-regard data points with the character viewing the Body card shown in yellow. (b) Data points with the character viewing the place card shown in purple. (c) Average percentage of data points for all participants for specific ROIs. The "ROI Place" bar shows the percentage of data points on the place card while it was viewed by the animated character. "ROI Body" shows the percentage of data points on the body card while it was viewed by the character. "ROI Face (Place)" shows data points on the character's face during Place condition, and "ROI Face (Body)" shows data points on the character's face during the Body condition. "No ROI Place" and "No ROI Body" show data points outside of the ROIs (character's Face and attended card) during each condition.

**Table 1**  
 Summary of Observed Regions of Places > Bodies and Bodies > Places Activation

<b>Region</b>	<b>Side</b>	<b>X</b>	<b>Y</b>	<b>Z</b>	<b>Nvox</b>	<b>BA</b>
<b>Places &gt; Bodies</b>						
Cingulate Gyrus	R	13	20	32	12	32
Inferior Frontal Gyrus	L	-60	11	31	149	9
Inferior Parietal Lobule	R	49	-49	53	52	40
Inferior Parietal Lobule	R	60	-56	45	452	40
Lingual Gyrus	R	8	-71	3	194	18
Middle Frontal Gyrus	L	-20	-6	49	8	6
Middle Frontal Gyrus	R	32	7	60	108	6
Middle Frontal Gyrus	R	49	7	34	106	9
Precentral Gyrus	L	-36	9	49	140	6
Superior Parietal Lobule	R	49	-63	57	156	7
Superior Parietal Lobule	R	31	-54	41	20	7
<b>Bodies &gt; Places</b>						
Clastrum/Insula	L	-34	5	-5	19	NA
Inferior Parietal Lobule	L	-44	-68	39	88	39
Lentiform Nucleus, Putamen	L	-20	8	-4	148	NA
Middle Frontal Gyrus	L	-21	63	6	40	9
Middle Temporal Gyrus	R	55	-12	-7	8	22
Middle Temporal Gyrus	L	-57	-17	-9	331	21
Parahippocampal Gyrus	L	-19	-29	-4	44	27
Parahippocampal Gyrus	R	21	-29	-4	312	27
Postcentral Gyrus	L	-31	-28	67	332	3
Superior Frontal Gyrus	L	-15	50	19	24	9
Superior Parietal Lobule	L	-37	-71	49	8	7
Supramarginal Gyrus	L	-55	-52	27	72	40

Nvox = number of voxels in the ROI. X, Y, and Z refer to the stereotaxic coordinates of the center of activation within an ROI in the coordinate system of Talairach and Tournoux. R = right hemisphere. L = left hemisphere. BA = Brodmann's Area. All values are reported at a threshold of  $p < .01$ , uncorrected; cluster size  $\geq 8$  functional voxels.