ELSEVIER

Neuropsychologia 44 (2006) 1674–1682

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

# Spontaneous mentalizing during an interactive real world task: An fMRI study

Hugo J. Spiers *, Eleanor A. Maguire

*Wellcome Department of Imaging Neuroscience, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, UK*

Received 9 December 2005; received in revised form 8 February 2006; accepted 20 March 2006
Available online 9 May 2006

**Abstract**

There are moments in everyday life when we need to consider the thoughts and intentions of other individuals in order to act in a socially appropriate manner. Most of this mentalizing occurs spontaneously as we go about our business in the complexity of the real world. As such, studying the neural basis of spontaneous mentalizing has been virtually impossible. Here we devised a means to achieve this by employing a unique combination of functional magnetic resonance imaging (fMRI), a detailed and interactive virtual reality simulation of a bustling familiar city, and a retrospective verbal report protocol. We were able to provide insights into the content of spontaneous mentalizing events and identify the brain regions that underlie them. We found increased activity in a number of regions, namely the right posterior superior temporal sulcus, the medial prefrontal cortex and the right temporal pole associated with spontaneous mentalizing. Furthermore, we observed the right posterior superior temporal sulcus to be consistently active during several different subtypes of mentalizing events. By contrast, medial prefrontal cortex seemed to be particularly involved in thinking about agents that were visible in the environment. Our findings show that it is possible to investigate the neural basis of mentalizing in a manner closer to its true context, the real world, opening up intriguing possibilities for making comparisons with those who have mentalizing problems.


*Keywords:* Theory of mind; Mentalizing; Neuroimaging; Verbal report; Thought; fMRI; Virtual reality; Superior temporal sulcus; Medial prefrontal cortex

## 1. Introduction

Our survival depends on considering the actions and intentions of others. Being able to predict what they might do or expect of us can help direct our own behaviour appropriately. This ability to attribute mental states to others is known as having 'theory of mind' (ToM) (Premack & Woodruff, 1978) or 'mentalizing' (Frith, Morton, & Leslie, 1991). It has been specifically linked with a number of brain regions, in particular the posterior superior temporal sulcus (pSTS), the temporo-parietal junction, the medial prefrontal cortex (mPFC) and the temporal poles (Frith & Frith, 2003; Gallagher & Frith, 2003). Evidence for their involvement in ToM has been provided by neuroimaging and neuropsychological studies in which subjects must make inferences about the mental state or intentions of other individuals.

These individuals, or agents, have been presented through a variety of methods including stories (Fletcher et al., 1995; Rowe, Bullock, Polkey, & Morris, 2001; Saxe & Kanwisher, 2003), static cartoons (Brunet, Sarfati, Hardy-Bayle, & Decety, 2000; Gallagher et al., 2000), animations of interacting shapes (Castelli, Happe, Frith, & Frith, 2000; Schultz, Friston, O'Doherty, Wolpert, & Frith, 2005) and movies of people or simulated characters making intentional actions (Grezes, Frith, & Passingham, 2004; Pelphrey, Singerman, Allison, & McCarthy, 2003; Pelphrey, Viola, & McCarthy, 2004). Converging evidence for the involvement of these brain regions has come from a different approach, using interactive tasks to tease out components of our mentalizing capacity during cooperative or competitive games (Gallagher, Jack, Roepstorff, & Frith, 2002; McCabe, Houser, Ryan, Smith, & Trouard, 2001; Ramnani & Miall, 2004; Rilling et al., 2002; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003; Stuss, Gallup, & Alexander, 2001). While these studies have given insight into the brain regions underlying mentalizing, they often involve simplified repetitive tasks that do not speak

* Corresponding author.
*E-mail addresses:* h.spiers@fil.ion.ucl.ac.uk (H.J. Spiers), e.maguire@fil.ion.ucl.ac.uk (E.A. Maguire).

0028-3932/$ – see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2006.03.028

directly to the experience of spontaneously mentalizing in everyday life. We are all confronted by other individuals on a daily basis, but not all our time is consumed with considering their mental states. At particular moments their actions might give us reason to consider their mental states or we might in turn think about how our own behaviour might affect their future choice of actions.

Examining the brain regions associated with spontaneous mentalizing during our everyday experiences represents a huge challenge. Not only must we capture the richness of external world, but subjects must be free to interact with it as they wish. One means of creating interactive experiences is to use virtual reality (VR). While VR has been used successfully in combination with neuroimaging to examine the neural correlates of large-scale navigation (Aguirre, Detre, Alsop, & D'Esposito, 1996; Gron, Wunderlich, Spitzer, Tomczak, & Riepe, 2000; Hartley, Maguire, Spiers, & Burgess, 2003; Maguire et al., 1998), the VR environments employed have been simplistic, unfamiliar and uninhabited, making them unsuitable for studying ToM. Furthermore, standard behavioural measures are insufficient for identifying when subjects engage in mentalizing or other thought processes during ongoing real world experiences.

We set out to overcome these limitations by combining functional neuroimaging with an accurate interactive virtual simulation of a bustling central London (UK), comprising people and traffic. We also employed a novel means of 'reading' participants' thoughts while they moved around the city. To determine when subjects were engaged in different thought processes, immediately post-scan and without prior warning, subjects watched a video replay of their performance and were interviewed using a verbal report protocol that was carefully developed from pilot studies (Ericsson & Simon, 1980). Simply put, this involved getting subjects to review their performance and report on what they had been thinking while they had been doing the task in the scanner. The transcribed thoughts of the subjects could then be used to model every second of the functional magnetic resonance imaging (fMRI) time series. Analysing the brain activity during the moments when subjects reported engaging in spontaneous mentalizing, we were able to reveal the neural correlates underlying these events.

## 2. Methods

### 2.1. Subjects

Twenty healthy right-handed male licensed London taxi drivers participated in the experiment (mean age 49.8 years, S.D. 8.5 years, range: 27–59 years). Licensed taxi drivers were chosen in order to ensure a consistent pattern of movement around the city as a platform for our analyses. All subjects were naïve to the stimuli used in the experiment and gave informed written consent in accordance with the local research ethics committee.

### 2.2. The virtual environment

The video game 'The Getaway' (©Sony Computer Entertainment Europe 2002) run on a Sony Playstation2 (©Sony Computer Games Inc.) was used to present subjects with a ground-level first person perspective view of a simulation of central London, UK (see Fig. 1). The game designers decided to truly recreate the city and a large team of photographers walked the streets of central London for 2 years recording many streets, shops, other details. Over 110 km (70 miles) of driveable roads have been accurately recreated from Ordinance Survey map data, covering 50 km$^2$ (20 miles$^2$) of the city centre. Breaking all speed limits and ignoring all red traffic lights, it takes 15 min to travel between the furthest points east to west. The one-way systems, working traffic lights, the busy London traffic, and an abundance of Londoners going about their business are all included. The simulated drivers and pedestrians followed the traffic regulations, and reacted to the subjects' movements appropriately, e.g. avoiding being run over and giving way at junctions. The 'Free Roaming' mode of the game was used, permitting free navigation with the normal game scenarios suspended. Subjects moved through the environment with a normal ground-level first person perspective, in a London taxi cab, controlled using a modified MRI-compatible game controller, consisting of two joysticks providing analogue control of acceleration, braking and steering left and right. To avoid constant collisions with other vehicles in the environment, Action Replay Max software (©Datel Design and Development Ltd. 2003) provided a 'cheat' modification to the game, permitting subjects to drive through other vehicles if necessary. Subjects were instructed to drive 'legally' as they would in actual London. All of the taxi drivers confirmed that the game was very reminiscent of their experience of driving in central London.

### 2.3. Pre-scan training and familiarization

Two weeks prior to scanning subjects were given 2 h of practice with the game controls by asking them to navigate to various locations in areas of envi-

(a) (b)

Fig. 1. Example views from within the virtual simulation of London: 'The Getaway' ©Sony Entertainment Europe 2002. These images are reproduced with the kind permission of Sony Computer Entertainment Europe. (a) Piccadilly Circus and (b) Trafalgar Square. The rear view of a London black taxi cab in the lower middle of each image was the vehicle that the subjects drove during the experiment.

ronment not used in the experimental task. To avoid crashes with other vehicles and waiting for long periods at red traffic lights subjects were familiarized with being able to drive through cars and red traffic lights, but were otherwise required to comply with all other road traffic regulations in the UK. Thirty minutes before the scan subjects were again given further practice in an area not used in experimental tasks. During this practice session subjects were trained to respond to a set of recorded customers' requests to take them to destinations in London. They also heard navigationally irrelevant statements which they had to listen to carefully but did not need to respond to. Finally inside the MRI scanner the subjects were given practice in an area of London not tested in the experimental tasks and with the MRI-compatible game control for between 2 and 3 min prior to the start of the experimental task. They were also given experience of hearing voices of customers over the noise of the scanner through head phones worn during the scan. Prior to scanning subjects were told the locations they would be starting from in the experimental tasks, but not the order.

### 2.4. Experimental tasks

During fMRI scanning, navigation was tested in blocks where subjects responded to customers' requests (heard via head phones) by delivering them to their destinations. During each block one route was tested. When the game came on the screen, subjects were given between 3 and 5 s to orient themselves in the environment. Following this they heard a customer request a destination (mean duration 2.0 s). For all routes, at some point during navigation the subjects heard customers request a change of destination (mean duration 3.0 s) and on a separate occasion make a navigationally irrelevant statement (mean duration 2.0 s). For three of the routes an additional request to avoid a location or go via a location was made by the customer (mean duration 3.7 s). Seven routes were tested. Two subjects completed only four routes, in one case due to discomfort, the other due to a technical problem. Each block of navigation ended when either the subject reached the destination or when a predetermined period of time elapsed. Each block of navigation was separated by a period of rest in which the subjects viewed a blank white screen for 60 s. Total mean functional scanning time was 31 min 35 s (S.D. 4 min 9 s).

### 2.5. Video recording

In order to debrief subjects and create an independent record of eye-tracking two videos were recorded during the scan. Video output from the Playstation2 was split three ways: (1) to a projector presenting stimuli in the MRI scanner (view angle of 27.6°), (2) to a VHS video recorder for debriefing and (3) to a video mixer to create an eye-tracking video. Video output going into the video mixer was mixed with camera footage of the scan console and a stopwatch manually synchronized with the time stamp on debriefing video. By noting the time on the stopwatch when the scanner finished it was possible to convert the time in the debriefing video into time from the onset of scanning. Gaze position cross hairs collected via an ASL504LRO infrared eye-tracker (Applied Science Laboratories, Bedford, MA) were overlaid onto the video sent to the mixed video recording. Accurately calibrated eye-gaze tracking was achieved in nine subjects.

### 2.6. Verbal report protocol

Immediately post-scan the subjects were taken from the scan room to a separate room where they were given a surprise debriefing with a verbal report protocol (Ericsson & Simon, 1980). In this debriefing, subjects watched the video of their performance during scanning. They were carefully instructed to describe what they remembered thinking, step-by-step, during their original performance. The interview proceeded at a pace determined by the subject, with the video being paused and rewound by the interviewer where necessary to capture the details provided by the subject. A new copy of the original video was recorded during the interview with the voices of the subject and interviewer collected by a microphone overlaid. In accordance with the methods described by Ericsson and Simon (1980) and others (Jack & Roepstorff, 2003) the interviewer followed a predetermined protocol during the interview. The subject's report was interrupted as little as possible, the interviewer intervening only to improve the subject's specification of the onset and duration of thoughts where possible, and on occasion where clarification was required to later aid analysis. The mean duration of the collection of the verbal reports was 108.9 min (S.D. 16.9 min).

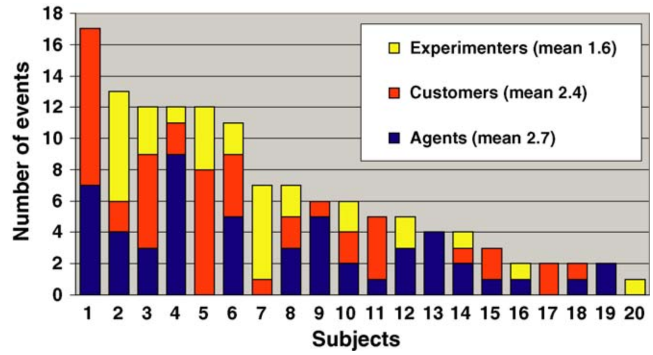


Fig. 2. Number of events of each subcategory for each subject (rank ordered by total number of events). The events comprising the subcategory ToM experimenters/customers are separated into thoughts relating to either the customers or experimenters to show the composition of this category.

### 2.7. Analysis of the verbal reports

Anonymized audio information from the verbal report interviews was transcribed by a professional transcription agency who were blind to the purpose of the experiment. By comparing the transcript with the time stamp from the original performance video, information about the timing of the thoughts was incorporated into the transcripts and any errors or unclear statements rectified. Each statement in the transcript was classified into one of a set of categories, and its onset and duration recorded to create a segmented timeline of the subject's experiences in order to model the fMRI data. Unambiguous categories were predetermined by analysis of common repeated statements in the verbal reports of four subjects who took part in an in-depth pilot study outside of the scanner. These categories included the category 'ToM events', comprising statements concerning the subject's thoughts about other individual's thoughts or intentions (see Fig. 2 and Table 1). Since some of the ToM events related to thoughts about observed agents in the environment and others related to thoughts about unseen individuals consisting of either the customers or experimenters, ToM events was further classified into the subcategories 'ToM agents' and 'ToM customers/experimenters' (see Table 1 for definitions and examples). The verbal reports of thoughts relating to customers and experimenters were deemed sufficiently similar to be included in one subcategory (see examples in Table 1). Another category of thoughts common in the verbal reports was 'coasting', where subjects were actively driving and moving through the city, but did not have any directed thoughts. This served as the baseline task. Other categories of thoughts concerning spatial navigation were also identified and are reported elsewhere (Spiers & Maguire, in press). The independent eye-tracking video was used to aid the identification of onsets and durations where the subjects reported looking at fixed features and moving agents in the environment, and served as an external measure to validate the procedure.

### 2.8. fMRI image acquisition and analysis

T2-weighted echo planar (EPI) images with blood oxygen level dependent (BOLD) contrast were acquired on a 1.5 T Siemens Sonata MRI scanner. We used standard scanning parameters to achieve whole brain coverage: 44 slices, 2 mm thickness (1 mm gap), TR 3.96 s. The first four volumes from each session were discarded to allow for T1 equilibration effects. A T1-weighted structural MRI scan was acquired for each subject. Images were analyzed in a standard manner using the statistical parametric mapping software SPM2 (www.fil.ion.ucl.ac.uk/SPM). Spatial preprocessing consisted of realignment, unwarping, normalization to a standard EPI template in MNI space with a resampled voxel size of 3 mm × 3 mm × 3 mm and smoothing using a Gaussian kernel with full width at half maximum of 10 mm. Following preprocessing, statistical analysis was performed using the general linear model. We used events sampled

Table 1
Definitions and examples from the verbal reports for the ToM subcategories and coasting

| | |
|---|---|
| ToM agents: Thinking about the intentions or thoughts of agents in the environment | |
| Example 1 | Subject 7: [Subject cuts in front of another vehicle] "I cut him right up there, he wouldn't have been happy . . .. It's just we've got rules of the road and he had to stop sharply then and I thought that was a bit of a liberty." |
| Example 2 | Subject 12: [a car pulls out in front of him and the subject has to break to avoid a collision] "I thought what's he doing? Why would he do that? There's a line there, he's got to stop, you know, and that . . .. I've got preference." |
| Example 3 | Subject 19: [Subject knocks into a pedestrian] "And then he suddenly jumped out, didn't he, at the last second. Most of them don't normally jump into your lane . . .. And I thought is he's going to chase me now." |
| Example 4 | Subject 5: [predicts the future movements of a bus he has spotted] "I thought I can't go through there, if I go and overtake him, he'll move. |
| Example 5 | Subject 12: [Car suddenly stops in front of him as it turns into a road] "Even though there's two white lines, you don't have to stop like that dipstick did. I was annoyed with him." |
| Example 6 | Subject 3: "I thought oh great that's a police car and maybe he'll chase me, give me a bit of fun." |
| Example 7 | Subject 15: "I thought about that person on the left there, it looked like the person was just about to walk across the road for some reason, from that angle." |
| Example 8 | Subject 11: [predicts the future actions of a driver ahead of him] "I thought he was going to come across, which he did." |
| ToM customers/experimenters: Thinking about the thoughts of the customers (heard through headphones) or the thoughts of the experimenters | |
| Example 1 | Subject 11: [The subject is near the destination after a long detour] "I'm happy now. I'm thinking are the people happy in the back? The way I've gone." |
| Example 2 | Subject 3: [subject nearly collides with some railings] "Well I just kept thinking about you and I thought don't hit the railings, oh I bet he's gone "*oh that was close*"." |
| Example 3 | Subject 7: [subject makes a mistake with the controls and swerves] "I was worried what you lot would think, I thought you'd be laughing, ha ha." |
| Example 4 | Subject 2: "Also I thought this . . .. I reckon that she's going to change her mind." |
| Example 5 | Subject 5: [After the customer tells him he has run out of time] "I was amused. I thought, well, you're stupid. You're out of time. I know he's been in there two hours but . . .. That was . . . that was a silly remark that was." |
| Example 6 | Subject 11: [Customer has previously requested to go via a further location] "I thought to myself well, you see you can't be wanting me to go down Regent Street and back because that would be a stupid thing to do." |
| Example 7 | Subject 11: [spots a ramp he could drive up] "I thought oh, I wonder what will happen if I go up that ramp, that's what I thought. And I thought no, no, that's not what they want." |
| Example 8 | Subject 15: [concerned about how the customer might react to a large fare] "Hopefully the woman won't give a monkey's about how much money is on the clock [taxi meter]." |
| Coasting events (condition for comparison): Driving automatically, without directed thoughts | |
| Example 1 | Subject 15: "It's just automatic down the Haymarket." |
| Example 2 | Subject 3: "I'm not thinking much." |

in periods of coasting as the control condition for ToM events in order to discern what brain areas were engaged over and above those associated with basic factors such as movement in virtual London, visual stimulation, optic flow and use of the games console that were also present during coasting. The coasting events were specified at the middle of each period of coasting rather than at the start or end to avoid temporal correlation of the other events/epochs in the other categories. Importantly, all the categories, including auditory events, were sufficiently uncorrelated in the time series to model them as separate regressors. To compare ToM events with coasting events a model was generated in which the onsets of these categories were modelled with a stick function and convolved with the canonical haemodynamic response function to create regressors of interest. To examine the subcategories ToM agents and ToM customers/experimenters a second model was generated with these two categories and coasting events entered as regressors. In both models, the other spatial navigation categories as well as turning (left and right combined) and rest periods were modelled with boxcar functions for categories consisting of epochs and stick functions for categories consisting of events. These were then convolved with the canonical haemodynamic response function to create regressors of no interest. Subject-specific parameter estimates pertaining to each regressor (betas) were calculated for each voxel. The parameter estimates were entered into a second level random-effects analysis using *t*-tests. We report results in a priori regions of interest at $P < 0.001$ uncorrected for multiple comparisons, with an extent threshold of >5 contiguous voxels. These regions were identified from previous neuroimaging studies of theory of mind (Gallagher & Frith, 2003). Activations in other regions are reported if they survive whole brain correction for multiple comparisons at $P < 0.05$.

## 3. Results

Aspects of the findings from this experiment relating to spatial navigation have been reported elsewhere (Spiers & Maguire, in press). We now report new analyses form this rich dataset focussed specifically on theory of mind.

### 3.1. Spontaneous ToM thoughts

All subjects reported thinking about other individuals' thoughts or intentions during the task. The average number of reported ToM events was 6.7 (standard deviation = 4.6). All 20 subjects reported some kind of ToM thought, and 14 subjects reported thoughts in both ToM agents and ToM customers/experimenters categories. The distribution of classified ToM events across subjects (see Fig. 2) reveals that the category of ToM is not dominated by one particular subcategory. Examination of the examples of verbal reports provided in Table 1 gives insight into the nature of the spontaneous mentalizing during an everyday experience (see Table 1). The category ToM agents consists of thoughts relating to the future intentions of other agents, for example, thoughts about being challenged by the traffic police, concerns about colliding with other vehicles or pedestrians, and the implication that they might be annoyed, hurt or even killed. The category ToM experimenters/customers consists of thoughts about what the experimenters or the customers might be thinking, for example, what will they think if a bad route is chosen.

### 3.2. Neuroimaging results

Comparison of all ToM events with the coasting events with a correction for multiple comparisons revealed significantly greater activity in only one region, the right pSTS (peak voxel: $x = 60$, $y = -36$, $z = 0$; $Z$-score = 5.80; see Fig. 3). When an uncorrected threshold of $P < 0.001$ was employed, activation was also observed in other regions, including regions predicted a priori: the mPFC (peak voxel: $x = 3$, $y = 45$, $z = 18$; $Z$-score = 3.60) and the right temporal pole (peak voxel: $x = 39$, $y = 18$, $z = -33$; $Z$-score = 3.99).

Having observed an effect with the overall category of ToM we then compared each of the subcategories to coasting events, and to each other. The right pSTS was significantly more active when either ToM agents (peak voxel: $x = 54$, $y = -36$, $z = -3$; $Z$-score = 3.89) or ToM customers/experimenters (peak voxel: $x = 60$, $y = -42$, $z = -3$; $Z$-score = 3.72) were compared with coasting events. Significantly greater activation was observed in the mPFC (peak voxel: $x = 6$, $y = 54$, $z = 27$; $Z$-score = 3.51) when ToM agents were compared with coasting events, but not when ToM customers/experimenters was compared with coasting events. Direct comparison of ToM agents with ToM customers/experimenters revealed no differences in activity in either direction. Examination of the parameter estimates for the peak voxel in the right pSTS region (see Fig. 4) confirms that the effect was similar for the two ToM subcategories. The parameter estimates for the mPFC reveal a different pattern. The effect was greatest during the ToM agents events, while in contrast a large degree of variance was evident during the ToM customers/experimenters condition.

## 4. Discussion

There are moments in everyday life when we think about the thoughts and intentions of other individuals. Unlike most previous neuroimaging studies in which mentalizing was manip-

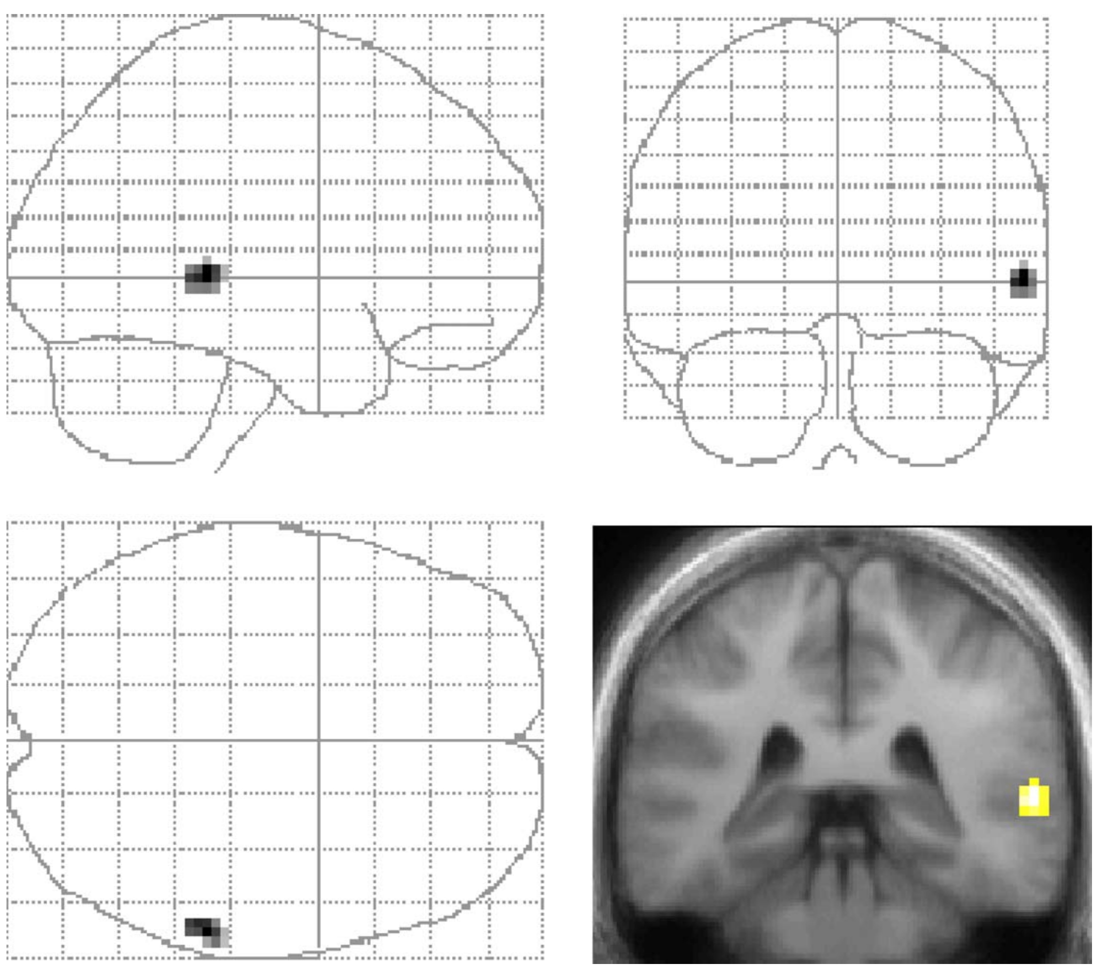

Fig. 3. Significantly greater activation in the right pSTS for the comparison of all ToM thoughts combined compared with coasting events, thresholded at $P < 0.05$ corrected for multiple comparisons. Activation is presented on the glass brain, except for the bottom right panel where the activation is overlaid on the mean structural MRI scan of the 20 subjects.

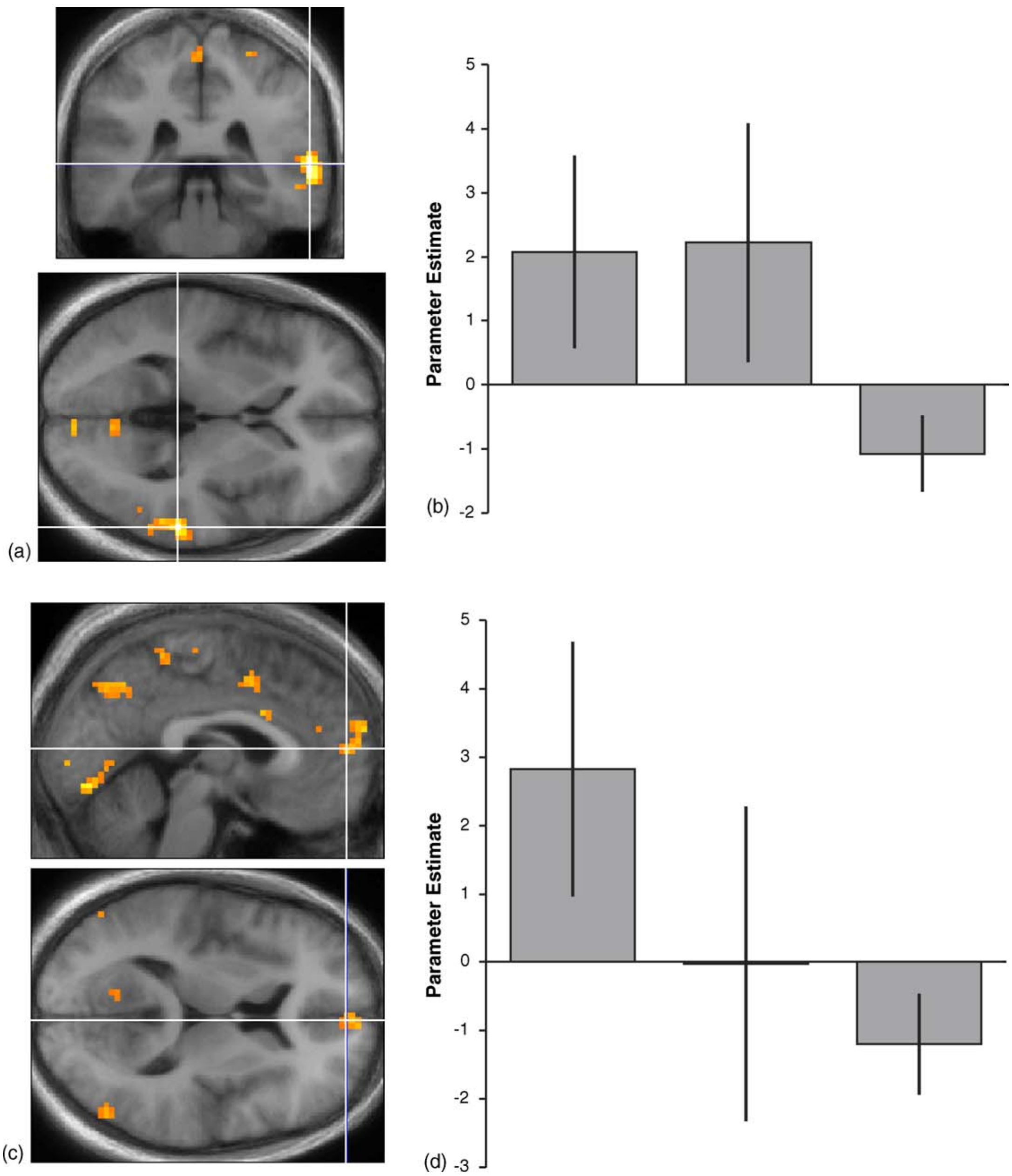


Fig. 4. Comparison of ToM events with coasting events, with ToM agents and ToM customers/experimenters events modelled separately in order to examine their individual parameter estimate plots, and effect comparisons between the two subtypes. The activations shown are from the contrast of ToM (ToM agents + ToM customers/experimenters) vs. coasting. (a) Activation thresholded at $P < 0.001$ uncorrected in the right pSTS overlaid on the mean structural MRI scan of the 14 subjects who reported both subtypes of ToM thoughts. Cross hairs show the location of the peak voxel at $x = 60$, $y = -36$, $z = 0$; $Z$-score = 4.99. (b) Parameter estimates (in arbitrary units) for the peak voxel in the right pSTS for the conditions: (1) ToM agents, (2) ToM customers/experimenters and (3) coasting events. (c) Activation thresholded at $P < 0.001$ uncorrected in the mPFC, overlaid on the mean structural image of the 14 subjects. Cross hairs show the location of the peak voxel at $x = 3$, $y = 57$, $z = 6$; $Z$-score = 3.73. (d) Parameter estimates (arbitrary units) from the peak voxel in the mPFC, with the same order of conditions as shown in (b). See Appendix A for a list of all peak coordinates and $Z$-scores (at a threshold of $P < 0.001$ uncorrected).

ulated by an explicit condition in the study design (see Gallagher & Frith, 2003), we were able to examine spontaneous mentalizing that occurred during ongoing real world experiences. By combining fMRI with an interactive and detailed virtual reality city, and a retrospective verbal report protocol, we were able to provide insight into the content of such thoughts and identify the brain regions that underlie them. We found increased activity in a number of regions, namely the right pSTS, the mPFC and the right temporal pole when we compared events in which subjects described spontaneous mentalizing compared with events where subjects were also actively driving through the city but report no specific thoughts (coasting). Of these regions the right pSTS was found to be the most consistently active and showed increased activation during either of the two subcategories, ToM agents, ToM customers/experimenters, while mPFC seemed to be particularly involved in thinking about agents that were visible in the environment.

### 4.1. The right posterior superior temporal sulcus

Activation of the pSTS has been observed in numerous neuroimaging studies examining mentalizing (Frith & Frith, 2003; Gallagher & Frith, 2003), yet its precise function is still not well understood. The pSTS is a multimodal region receiving input from both the ventral and dorsal streams, and is interconnected with many structures in the limbic system (Barnes & Pandya, 1992). Its involvement, along with nearby temporal lobe structures, in processing auditory information has been noted by previous studies (Calvert, Campbell, & Brammer, 2000). More selectively, it has also been associated with the processing of faces and biological motion (for reviews, see Allison, Puce, & McCarthy, 2000; Puce & Perrett, 2003). For example, activation has been seen in the pSTS in relation to facial expressions (Kilts, Egan, Gideon, Ely, & Hoffman, 2003; Narumoto, Okada, Sadato, Fukui, & Yonekura, 2001), eye-movements or eye-gaze (Calder et al., 2002; Hoffman & Haxby, 2000; Wicker, Michel, Henaff, & Decety, 1998), viewing moving body parts or static images that imply movement (Beauchamp, Lee, Haxby, & Martin, 2003; Gallagher & Frith, 2004; Grezes, Costes, & Decety, 1998, 1999; Puce, Allison, Bentin, Gore, & McCarthy, 1998; Senior et al., 2000; Thompson, Clarke, Stewart, & Puce, 2005) and viewing point light displays of body motion (Bonda, Petrides, Ostry, & Evans, 1996; Grezes et al., 1998, 2001; Grossman et al., 2000).

Based on such findings it has been suggested to that one function of the STS is to signal the observed actions and intentions of another individual (Allison et al., 2000). Consistent with this view, many of the thoughts described by our subjects in the category ToM agents relate to the observation of actions that imply intentions, such as a pedestrian waiting on curb before crossing the street (see Table 1). However, not all the reported ToM thoughts are of this kind. The ToM customers/experimenters thoughts do not appear to relate to the observation of actions or intentions, suggesting that a broader explanation of the role of the pSTS region identified in this study is required. A common feature of the reported thoughts is that they concern the mental states of other individuals, either their intentions or consideration of the outcome of behaviour (e.g. will the customer be annoyed with me for my actions). Thus, our findings are compatible with a broader view of the pSTS in which it not only participates in the detection of intentional behaviour, but is also engaged by the analysis of the goals and outcomes of intentional behaviour (Frith & Frith, 1999).

Two features of the anatomical location of the pSTS activation we report deserve mention. Firstly, the activation is strongly right lateralized, even when the individual subcategories were examined separately. This finding is consistent with the observation of right pSTS lateralization in previous studies exploring mentalizing and the processing of intentions (e.g. Castelli et al., 2000; Gallagher et al., 2000; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004), leading to the suggestion that the right pSTS, in particular, is important for mentalizing (Gallagher & Frith, 2003).

Secondly, the activation location we have observed is in the more anterior portion of the pSTS identified in previous studies examining mentalizing and biological motion (Frith & Frith, 2003). Our pSTS activation lies in a similar location to a right lateralized region reported by Saxe et al. (2004), who labelled it the pSTS-A (where A refers to action). In their study subjects viewed video clips of an actor walking across a room, becoming occluded by a bookshelf and then reappearing after either a short or a long delay. The right pSTS was found to be the only region showing more activation in response to the long delay than the short delay condition. Following Allison et al. (2000), Saxe et al. (2004) suggest that this particular right pSTS region may play a role in the representation of an agent's intentional action, but argue that it is modulated by the environmental context. Since most of the spontaneous mentalizing thoughts reported in our experiment have an environmental context due to the task they were performing (e.g. a pedestrian on a curb waiting to step out, or wondering what the customer might make of their driving) the current results are compatible with such a view. However, since there is increased activation in the same pSTS region when ToM customers/experimenters are compared with coasting events, it would appear that action observation is not a necessary requirement to elicit activity in this region. This view is supported by evidence that a similar right pSTS region was found to be more active when subjects make trustworthiness judgements than age judgements during the viewing of static faces (Winston, Strange, O'Doherty, & Dolan, 2002), a finding attributed to the processing of information relating to the potential intentions of someone. Thus, our results are consistent with the view that the processing of intentions or possible mental states, either from observation of action or from other cues, involves the right pSTS.

### 4.2. The medial prefrontal cortex

The mPFC, and in particular the anterior paracingulate region, has also been associated with mentalizing in numerous neuroimaging studies (see Frith & Frith, 2003). We found that the anterior paracingulate region within the mPFC was more active for ToM agents compared with coasting events, but showed no consistent response during the ToM customers/experimenters events. It has been suggested that this region is important for determining an agent's mental state (Gallagher & Frith, 2003), by facilitating the decoupling of our representations of stimuli from reality (Frith & Frith, 2003). Support for the involvement of the mPFC in the decoupling hypothesis has come from studies where, for example, the amount of mPFC activity present during a task period was found to be significantly correlated with ratings of the frequency of stimulus-independent thoughts occurring during that task period (McGuire, Paulesu, Frackowiak, & Frith, 1996). Using a paradigm based in a real world setting we have been able identify thoughts that go beyond those classified as stimulus-oriented or stimulus-independent. We find that during an ongoing interactive task the mPFC is more active when subjects report considering the mental states of agents that they observe going about their business in the world.

It is interesting that increased activity was observed in mPFC when subjects considered the thoughts or intentions of observed agents but not when subjects considered the thoughts of the customers and experimenters. One possibility is that the increased activity in mPFC in the ToM agents condition reflects a greater demand to interpret the actions, or predict the future actions of individuals based on prior experience (e.g. example 2 of ToM agents in Table 1: "why would he do that?"). Such a possibility is supported by the observation of selectively increased mPFC activity when subjects attempted to guess, or second guess, the responses of another person with whom they were playing an interactive game, than when playing against a computer, though in fact in both cases they were playing against a computer (Gallagher et al., 2002).

### *4.3. The temporal pole and the temporo-parietal junction*

The temporo-parietal junction (TPJ) and the temporal pole were also predicted to be more active on the basis of previous studies. While the temporal pole was more active when all ToM events were compared with coasting events, this activity was absent when either of the two ToM subcategories was compared with coasting events. The TPJ showed no greater response in any contrast. Different processes have been ascribed to these regions, with the temporal pole suggested as generating a wider semantic or emotional context for processing stimuli (Frith & Frith, 2003; Gallagher & Frith, 2003), and the TPJ involved in attributing mental states during the comprehension of narratives (Fletcher et al., 1995; Gallagher et al., 2000; Saxe & Kanwisher, 2003). One possibility is that tasks involving, for example, narratives, require more semantic processing than spontaneous mentalizing in a more real world context, and hence areas such as temporal pole and TPJ are more engaged.

## 5. Conclusions

In summary, the present study provides support for the suggestion that right pSTS and the mPFC are involved in mentalizing about other individuals' thoughts or intentions. Our findings extend previous studies by showing that these regions are active in the context of spontaneous mentalizing during an interactive real world task that did not explicitly call for mentalizing. This was made possible by a unique combination of neuroimaging, a detailed simulation of a bustling familiar city, and the use of a retrospective verbal report protocol. Now that we have shown such an approach is possible, in the future will be interesting to use interactive situations to explore how different aspects of spontaneous mentalizing affect the activity in the relevant neural regions, in healthy volunteers and also those with mentalizing problems.

## Acknowledgements

This work was supported by a Wellcome Trust senior research fellowship in basic biomedical science to E.A.M. We are grateful to the pilot and scan participants for their time, patience and good humour. We thank Chris Frith for helpful discussions. We also thank the major licensed London taxi companies, publications, depots and cafes for facilitating subject recruitment. Thanks to P. Aston, E. Featherstone, C. Freemantle, R. Davis, O. Josephs, C. Hutton, J. Hocking, D. Kumaran, K. Friston and the FIL Methods Group, and the FIL Functional Technologies Team, for advice and technical assistance.

## Appendix A. Peak coordinates and *Z*-scores for the contrast ToM agents and ToM customers/experimenters > coasting events

| Brain regions | *Z*-score | Coordinates of peak activation | | |
|---|---|---|---|---|
| | | *x* | *y* | *z* |
| R posterior superior temporal sulcus | 4.99 | 60 | −36 | 0 |
| L anterior medial frontal gyrus (mPFC) | 3.73 | 3 | 57 | 6 |
| R lingual gyrus | 4.57 | 6 | −84 | −12 |
| R middle frontal gyrus | 4.14 | 60 | 24 | 27 |
| R central sulcus | 4.12 | 30 | −39 | 60 |
| L medial superior frontal gyrus | 4.00 | −6 | 30 | 36 |
| L middle frontal gyrus | 3.91 | −45 | 36 | 30 |
| R superior frontal gyrus | 3.91 | 18 | 15 | 63 |
| L middle occipital gyrus | 3.86 | −51 | −72 | 15 |
| L inferior occipital gyrus | 3.85 | −39 | −81 | −18 |
| L cerebellum | 3.81 | −33 | −60 | −27 |
| L middle temporal gyrus | 3.81 | −60 | −3 | −27 |
| L inferior frontal gyrus | 3.77 | −42 | 27 | −15 |
| L superior parietal gyrus | 3.77 | −18 | −66 | 63 |
| R calcarine sulcus | 3.73 | 6 | −93 | 0 |
| R cingulate sulcus | 3.72 | 6 | 6 | 42 |
| R cingulate gyrus | 3.67 | 3 | 12 | 27 |
| L superior frontal gyrus | 3.67 | −27 | 45 | 36 |
| R superior parietal gyrus | 3.66 | 30 | −57 | 60 |
| R middle occipital gyrus | 3.59 | 51 | −72 | 9 |
| R inferior occipital gyrus | 3.58 | 48 | −63 | −15 |
| R anterior cingulate sulcus | 3.52 | 12 | 39 | 18 |
| L cuneus | 3.39 | −3 | −87 | 27 |
| L calcarine sulcus | 3.26 | −12 | −69 | 6 |

L: left; R: right; mPFC: medial prefrontal cortex. All activations are listed statistically significant at $P < 0.001$. For brevity, each region is listed only once. MNI coordinates are listed.

## References

Aguirre, G. K., Detre, J. A., Alsop, D. C., & D'Esposito, M. (1996). The parahippocampus subserves topographical learning in man. *Cerebral Cortex*, *6*(6), 823–829.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Science*, *4*(7), 267–278.

Barnes, C. L., & Pandya, D. N. (1992). Efferent cortical connections of multimodal cortex of the superior temporal sulcus in the rhesus monkey. *The Journal of Comparative Neurology*, *318*(2), 222–244.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). FMRI responses to video and point-light displays of moving humans and manipulable objects. *Journal of Cognitive Neuroscience*, *15*(7), 991–1001.

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *Journal of Neuroscience*, *16*(11), 3737–3744.

Brunet, E., Sarfati, Y., Hardy-Bayle, M. C., & Decety, J. (2000). A PET investigation of the attribution of intentions with a nonverbal task. *NeuroImage*, *11*(2), 157–166.

Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I., et al. (2002). Reading the mind from eye gaze. *Neuropsychologia*, *40*(8), 1129–1138.

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*(11), 649–657.

Castelli, F., Happe, F., Frith, U., & Frith, C. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage*, *12*(3), 314–325.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215–251.

Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., et al. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition*, *57*(2), 109–128.

Frith, C. D., & Frith, U. (1999). Interacting minds—A biological basis. *Science*, *286*(5445), 1692–1695.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 459–473.

Frith, U., Morton, J., & Leslie, A. M. (1991). The cognitive basis of a biological disorder: Autism. *Trends in Neurosciences*, *14*(10), 433–438.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Science*, *7*(2), 77–83.

Gallagher, H. L., & Frith, C. D. (2004). Dissociable neural pathways for the perception and recognition of expressive and instrumental gestures. *Neuropsychologia*, *42*(13), 1725–1736.

Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11–21.

Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage*, *16*(3 Pt 1), 814–821.

Grezes, J., Costes, N., & Decety, J. (1998). Top–down effect of strategy on the perception of biological motion: A PET investigation. *Cognitive Neuropsychology*, *15*, 553–583.

Grezes, J., Costes, N., & Decety, J. (1999). The effects of learning and intention on the neural network involved in the perception of meaningless actions. *Brain*, *122*(Pt 10), 1875–1887.

Grezes, J., Fonlupt, P., Bertenthal, B., Delon-Martin, C., Segebarth, C., & Decety, J. (2001). Does perception of biological motion rely on specific brain regions? *NeuroImage*, *13*(5), 775–785.

Grezes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: An fMRI study. *NeuroImage*, *21*(2), 744–750.

Gron, G., Wunderlich, A. P., Spitzer, M., Tomczak, R., & Riepe, M. W. (2000). Brain activation during human navigation: Gender-different neural networks as substrate of performance. *Nature Neuroscience*, *3*(4), 404–408.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., et al. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, *12*(5), 711–720.

Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (2003). The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron*, *37*(5), 877–888.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*(1), 80–84.

Jack, A. I., & Roepstorff, A. (2003). *Trusting the subject:* Vol. 1, (Vol. 1, Exeter, UK: Imprint Academic.

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage*, *18*(1), 156–168.

Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S., Frith, C. D., & O'Keefe, J. (1998). Knowing where and getting there: A human navigation network. *Science*, *280*(5365), 921–924.

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of the Sciences of the United States of America*, *98*(20), 11832–11835.

McGuire, P. K., Paulesu, E., Frackowiak, R. S., & Frith, C. D. (1996). Brain activity during stimulus independent thought. *Neuroreport*, *7*(13), 2095–2099.

Narumoto, J., Okada, T., Sadato, N., Fukui, K., & Yonekura, Y. (2001). Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Brain Research. Cognitive Brain Research*, *12*(2), 225–231.

Pelphrey, K. A., Singerman, J. D., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: The influence of context. *Neuropsychologia*, *41*(2), 156–170.

Pelphrey, K. A., Viola, R. J., & McCarthy, G. (2004). When strangers pass: Processing of mutual and averted social gaze in the superior temporal sulcus. *Psychological Sciences*, *15*(9), 598–603.

Premack, D., & Woodruff, G. (1978). Chimpanzee problem-solving: A test for comprehension. *Science*, *202*(4367), 532–535.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, *18*(6), 2188–2199.

Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 435–445.

Ramnani, N., & Miall, R. C. (2004). A system in the human brain for predicting the actions of others. *Nature Neuroscience*, *7*(1), 85–90.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., & Kilts, C. (2002). A neural basis for social cooperation. *Neuron*, *35*(2), 395–405.

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage*, *22*(4), 1694–1703.

Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). Theory of mind" impairments and their relationship to executive functioning following frontal lobe excisions. *Brain*, *124*(Pt 3), 600–616.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, *300*(5626), 1755–1758.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *NeuroImage*, *19*(4), 1835–1842.

Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, *42*(11), 1435–1446.

Schultz, J., Friston, K. J., O'Doherty, J., Wolpert, D. M., & Frith, C. D. (2005). Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron*, *45*(4), 625–635.

Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E. T., Brammer, M., et al. (2000). The functional neuroanatomy of implicit-motion perception or representational momentum. *Current Biology*, *10*(1), 16–22.

Spiers, H. J., & Maguire, E. A. (in press). Thoughts, behavior, and brain dynamics during navigation in the real world. *NeuroImage*.

Stuss, D. T., Gallup, G. G., Jr., & Alexander, M. P. (2001). The frontal lobes are necessary for 'theory of mind'. *Brain*, *124*(Pt 2), 279–286.

Thompson, J. C., Clarke, M., Stewart, T., & Puce, A. (2005). Configural processing of biological motion in human superior temporal sulcus. *Journal of Neuroscience*, *25*(39), 9059–9066.

Wicker, B., Michel, F., Henaff, M. A., & Decety, J. (1998). Brain regions involved in the perception of gaze: A PET study. *NeuroImage*, *8*(2), 221–227.

Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, *5*(3), 277–283.