ACADEMIC
PRESS

Available online at www.sciencedirect.com
SCIENCE @ DIRECT®

NeuroImage

NeuroImage 19 (2003) 1835–1842

www.elsevier.com/locate/ynimg

Rapid Communication

People thinking about thinking people
The role of the temporo-parietal junction in "theory of mind"

R. Saxe[a,]* and N. Kanwisher[a,b]

[a]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA
[b]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, USA

Received 15 January 2003; revised 12 March 2003; accepted 14 April 2003

Abstract

Humans powerfully and flexibly interpret the behaviour of other people based on an understanding of their minds: that is, we use a "theory of mind." In this study we distinguish theory of mind, which represents another person's mental states, from a representation of the simple presence of another person per se. The studies reported here establish for the first time that a region in the human temporo-parietal junction (here called the TPJ-M) is involved specifically in reasoning about the contents of another person's mind. First, the TPJ-M was doubly dissociated from the nearby extrastriate body area (EBA; Downing et al., 2001). Second, the TPJ-M does not respond to false representations in non-social control stories. Third, the BOLD response in the TPJ-M bilaterally was higher when subjects read stories about a character's mental states, compared with stories that described people in physical detail, which did not differ from stories about nonhuman objects. Thus, the role of the TPJ-M in understanding other people appears to be specific to reasoning about the content of mental states.

Keywords: fMRI; Social cognitive neuroscience; False belief; Mentalising; Superior temporal sulcus; EBA

The remarkable human facility with social cognition depends on a fundamental ability to reason about other people. Specifically, we predict and interpret the behaviour of people based on an understanding of their minds: that is, we use a "theory of mind."[1] In this study we show that a region of human temporo-parietal junction is selectively involved in reasoning about the contents of other people's minds.

Brain regions near the temporo-parietal junction (TPJ) have been implicated in a broad range of social cognition tasks (Allison et al., 2000; Gallagher and Frith, 2003; Green and Haidt, 2003). Regions near the TPJ have preferential responses to human faces (e.g., Hoffman and Haxby, 2000), bodies (e.g., Downing et al., 2001) and biological motion (e.g., Grossman et al., 2000). There is also some evidence that regions within human TPJ are involved in theory of mind (ToM). A number of studies have reported increased responses in the TPJ when subjects read verbal stories or see pictorial cartoons that require inferences about a character's (false) beliefs, compared with physical control stimuli (Fletcher et al., 1995; Brunet et al., 2000; Gallagher et al., 2000; Castelli et al., 2000; Voegely et al., 2001. A number of other brain regions have also been implicated in theory of mind; see reviews by Gallagher and Frith, 2003, and Greene and Haidt, 2003).

What is the role of the TPJ in these tasks? ToM reasoning depends upon at least two kinds of representation: a representation of another person per se and a representation of that other person's mental states (see Leslie, 1999). While a representation of a person per se is a likely prerequisite for

---

* Corresponding author. Department of Brain and Cognitive Sciences, MIT, NE20-464, 77 Massachussetts Avenue, Cambridge, MA 02139, USA. Fax: +1-617-258-8654.

E-mail address: saxe@mit.edu (R. Saxe).

[1] The term "theory of mind" has a more restricted sense, referring to the suggestion that the structure of knowledge in the mind is analogous to a scientific theory (e.g., Carey, 1985; Wellman and Gelman, 1992). For discussions about the so-called theory-theory, see Carruthers and Smith, 1996, and Malle et al., 2001. In this study, we use the term theory of mind in a broader sense, to refer to any reasoning about another person's representational mental states (also called "belief-desire psychology," e.g., Bartsch and Wellman, 1995).

1053-8119/03/$ – see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S1053-8119(03)00230-1

ToM, achieving a representation of others' mental states is the core responsibility of a ToM. Some authors suggest that the TPJ is involved only in the preliminary stages of social cognition that "aid" ToM, not in ToM reasoning itself (e.g., Gallagher and Frith, 2003). We here provide evidence against this suggestion, and argue on the contrary that a region of the TPJ is selectively involved in representation other peoples' mental states.

Neuroimaging studies have followed developmental psychology in using "false belief" stories as the prototypical problem for ToM reasoning (Fletcher et al., 1995; Gallagher et al., 2000; see also Vogeley et al., 2001). In these scenarios, a character's action is based on the character's false belief (Wimmer and Perner, 1983). False beliefs provide a useful behavioural test of a ToM, because when the belief is false, the action predicted by the belief is different from the action that would be predicted by the true state of affairs (Dennett, 1978). Note, though, that everyday reasoning about other minds, by adults and children, depends on attributions of mostly true beliefs (e.g., Dennett, 1996; Bartsch and Wellman, 1995).

Previous investigations of the neural correlates of ToM (Fletcher et al., 1995; Gallagher et al., 2000) have compared false belief ("theory of mind") stories with two control conditions: "non-theory of mind stories," which describe actions based on the character's true beliefs, and "control" stories, consisting of unrelated sentences. These authors found that the TPJ response was high during theory of mind stories, but was also high during non-theory of mind stories. They concluded (see also Gallagher and Frith, 2003) that the TPJ is not selectively involved in ToM. This conclusion does not follow. Because the non-theory of mind stories invite inferences about the character's (true) beliefs, a region involved in reasoning about other minds should show a high response to these stories, as well as to the so-called theory of mind stories. (For an argument against the use of unrelated sentences as the baseline condition, see Ferstl and von Cramon, 2002.)

We propose two basic tests for a region selectively involved in ToM reasoning. First, it must show increased response to tasks/stimuli that invite ToM reasoning (about true or false beliefs) compared with logically similar non-social controls. Second, the region must respond not just when a person is present in the stimulus, but specifically when subjects reason about the person's mental states. Below, we provide evidence that a subregion of the TPJ, here called the TPJ-M, passes both these criteria for a selective role in ToM.

Experiment 1

We devised a new version of the false belief stories task (Fletcher et al., 1995) to compare reasoning about true and false beliefs to reasoning about non-social control situations. ToM stories described a character's action caused by his/her false belief. Descriptions of human actions required analysis of mental causes, in the absence of false beliefs. We compared these conditions to two non-social control conditions, (1) mechanical inference control stories, which required the subject to infer a hidden physical (as opposed to mental) process, such as melting or rusting (for examples, see Appendix 1), and (2) descriptions of nonhuman objects.

Unlike previous studies, we did not cue or instruct subjects to attend specifically to mental states. With this design we were able to look for regions of cortex in individual subjects that are selectively and spontaneously involved in understanding the mental (as opposed to physical) causes of events.

To test whether the response to ToM stories was a response to the presence of a person in the stimulus, we presented still photographs of people, and nonhuman objects. Downing et al. (2001) reported a bilateral region near the posterior superior temporal cortex that responds preferentially to the visual appearance of human bodies, compared with a range of control objects (the extrastriate body area, EBA). We tested directly the functional and anatomical relationship between the EBA and the (proposed) TPJ-M.

Methods

Twenty-five healthy right-handed adults (12 women) volunteered or participated for payment. All subjects had normal or corrected-to-normal vision and gave informed consent to participate in the study.

Subjects were scanned in the Siemens 1.5-(9 subjects) and 3.0-T (16 subjects) scanners at the MGH-NMR center in Charlestown, MA, using a head coil. Standard echoplanar imaging procedures were used [TR = 2 s, TE = 40 (3 T) or 30 (1.5 T) ms, flip angle 90°]. Twenty 5-mm-thick near-coronal slices (parallel to the brainstem) covered the occipital lobe and the posterior portion of the temporal and parietal lobes.

Stimuli consisted of short center-justified stories, presented in 24-point white text on a black background (average number of words = 36). Stories were constructed to fit four categories: false belief, mechanical inference, human action, and nonhuman descriptions (Appendix 1). Each story was presented for 9500 ms, followed by a 500-ms interstimulus interval. Each scan lasted 260 s: four 40-s epochs, each containing four stories (one from each condition), and 20 s of fixation between epochs. The order of conditions was counterbalanced within and across runs. Subjects were asked to press a button to indicate when they had finished reading each story. Subjects read a total of 8 (4 subjects) or 12 (21 subjects) stories per condition.

Fourteen of the subjects from Experiment 1 (7 women) were also scanned on an EBA localizer in the same scan session, all at 3.0 T. Stimuli consisted of 20 grayscale photographs of whole human bodies (including faces) in a range of postures, standing and sitting, and 20 photographs

Table 1
Experiment 1[a]

Region	MNI coordinate (max voxel)	Z	No. of voxels ($P < 0.05$, corrected)
LTPJ-M	[−54 −60 21]	5.88	63
LaSTS	[−57 −27 −12]	5.40	55
Prec	[−9 − 51 33]	5.20	41
R TPJ-M	[51 −54 27]	5.10	10
R aSTS	[66 −18 −15]	4.91	2

[a] Five regions showed increased signal during theory of mind, compared with mechanical inference, stories (random effects, $n = 25$, $P < 0.05$): left and right temporo-parietal junction (TPJ-M), left and right anterior superior temporal sulcus (aSTS), and precuneus (Prec). All coordinates are according to the Montreal Neurological Institute standard brain.

of easily recognizable inanimate objects (e.g., car, drum, tulip). (Two other conditions, cropped faces and scrambled objects, were included in the scan but were not analyzed here).

Image presentation followed the blocked design described in Tong et al. (2000; Experiment 1) except that images were presented at a rate of one every 800 ms (stimulus duration = 500 ms, interstimulus interval = 300 ms), and each scan lasted 336 s. Subjects performed a one-back matching task (Tong et al., 2000).

MRI data were analyzed using SPM 99, FS-fast, and in-house software.

Results

Average reading times for theory of mind and mechanical inference stories did not differ significantly (ToM = 6.4 s, MI = 6.5 s, $P > 0.2$).

Random effects analyses of 25 subjects revealed five loci of greater activation during the theory of mind compared with mechanical inference stories ($P < 0.05$ corrected for multiple spatial hypotheses): left and right TPJ-M, left and right anterior superior temporal sulcus (aSTS), and precuneus (Table 1, Fig. 1). [Consistent with many previous studies (e.g., Gusnard and Raichle, 2001; Raichle et al., 2000) the precuneus was deactivated (BOLD signal less than fixation baseline) during all of our story conditions. The ToM stories deactivated the precuneus less than mechanical inference stories. It was therefore unclear whether this effect should be considered a response to ToM or to mechanical inference stories, and the precuneus response was not analyzed further.]

The same pattern of results was apparent in individual subjects (fixed effects $P < 0.0001$, uncorrected for all results reported here). Voxels more responsive during ToM than mechanical inference stories were observed at the TPJ in 22 of 25 subjects (bilaterally in 14, left in 5, and right in 3 subjects). The aSTS activation was significant at this level in 10 of 25 subjects. Because the TPJ-M was most consistent across subjects and was the focus of our prior hypotheses, we concentrated on this region in the subsequent analyses.

We defined TPJ-M regions of interest (ROI) in the left and right TPJ in each individual subject as contiguous voxels in each hemisphere that were more active ($P < 0.0001$) during false belief than mechanical inference stories. The TPJ-M bilaterally generalized beyond false beliefs, responding significantly more to human action (HA) stories than to nonhuman descriptions [N-H D; paired samples t tests, right: HA average percent signal change from fixation (PSC): 0.22, N-H D average PSC: 0.02, $P < 0.0001$; left: HA average PSC: 0.35, N-H D average PSC: 0.10, $P < 0.0001$].

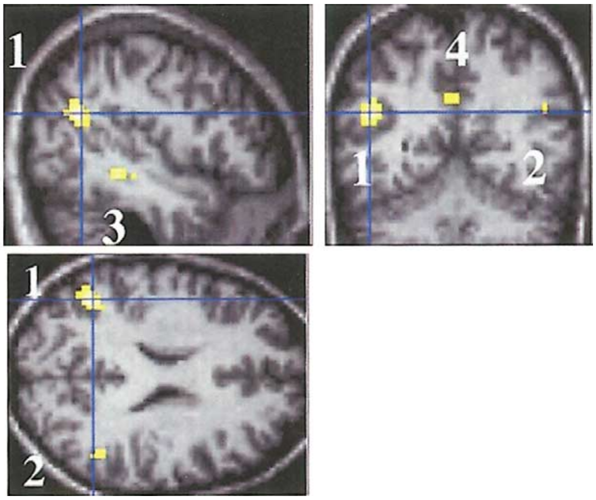
In the 14 subjects who also had an EBA localizer scan, EBA ROIs were defined as the cluster of contiguous voxels in extrastriate cortex (bilaterally in 13 subjects and right-only in 1 subject) that was more active ($P < 0.0001$) during pictures of human bodies than during pictures of nonhuman objects in each individual subject (following Downing et al., 2001). Both right and left EBA ROIs failed to discriminate between any story conditions (paired samples t tests, all $P > 0.4$, all story PSCs below 0.01; Fig. 2).

TPJ-M response to photographs was lateralized. The left TPJ-M did not discriminate between photographs of people (PSC: −0.04) and of objects (PSC: −0.09, paired samples t test, $P > 0.4$). The right TPJ-M showed a trend toward a greater response to photographs of people (PSC: 0.24) than of objects (PSC: 0.10, paired samples t test, $P < 0.10$). A repeated-measures ANOVA of content (person versus object) by stimulus modality (stories versus photograph) by hemisphere (right versus left) revealed a main effect of person > object ($P < 0.001$) and of stories > photographs ($P < 0.05$) modulated by an interaction between stimulus modality and hemisphere (response to photographs only on the right, $P < 0.005$) and a trend toward a three-way interaction (the right TPJ-M response distinguishes photographs of bodies and objects more than the left TPJ-M, $P < 0.1$; Fig. 2).

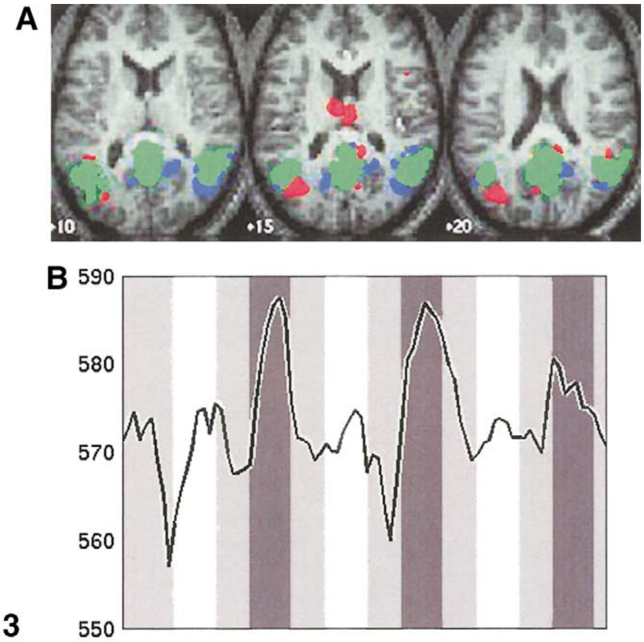
Discussion

Experiment 1 thus shows an increased BOLD response in a region of the TPJ bilaterally, here called the TPJ-M, during ToM compared with mechanical inference stories. This activation is robust and reliable across individual subjects. This finding replicates the earlier reports with a new set of stimuli, a less biased task (no cues), and with more stringent statistical tests (both individual subject analyses and random effects group analyses). Our results confirm that the TPJ-M response to verbal descriptions generalizes to human actions based on true beliefs.

Importantly, we distinguished the TPJ-M from its neighbour, the EBA, which did not respond to any verbal story conditions. However, the TPJ-M response to nonverbal so-



1



3

Fig. 1. Experiment 1. Random effects analysis, $P < 0.05$, corrected, $n = 25$. Theory of mind > mechanical inference stories. Crosshair marks the most significant voxel in the left TPJ (1). Also visible are activations in right TPJ (2), left aSTS (3), and precuneus (4). TPJ, temporo-parietal junction; aSTS, anterior superior temporal sulcus.

Fig. 3. (a) Experiments 1 and 2. Activation overlap within an individual subject showing bilateral temporo-parietal junction (bilateral TPJ) and precuneus regions (fixed effects, $P < 0.001$). Red = theory of mind > mechanical inference (Exp. 1). Blue = false belief > false photo (Exp. 2). Green = both. (b) Single subject time course of response during Experiment 2 to false belief (dark gray) and false photograph (white) stories in the same subject's TPJ-M, independently defined by a greater response to theory of mind than to mechanical inference stories in Experiment 1; $P < 0.0001$, uncorrected. Medium gray indicates fixation. Time course averaged over four runs.

cial stimuli appeared to be lateralized. The left TPJ-M response was selective for verbal descriptions, while the right TPJ-M activation may generalize to nonverbal stimuli, such as photographs.

Experiment 2

The results of Experiment 1 established that bilateral regions near the TPJ show a greater increase in BOLD signal when subjects reason about others' mental states, than when they reason about nonhuman objects. However, in Experiment 1, stories involving people and mental states were compared with stories that involve neither people nor mental states. In Experiment 2, we asked which of these two components was responsible for the observed activation. We directly compared the response of the TPJ-M to stories about people that did (desires) or did not (physical people) require inferences based on mental states.

Also, while they were controlled for difficulty and causal structure, the logical structure of the ToM stories used in Experiment 1 (and previous studies) differed systematically from the control stories: only the false belief stories require the notion of a false representation, in this case a false belief. This confounding factor was perceived by developmental psychologists, who invented its solution: "false pho-

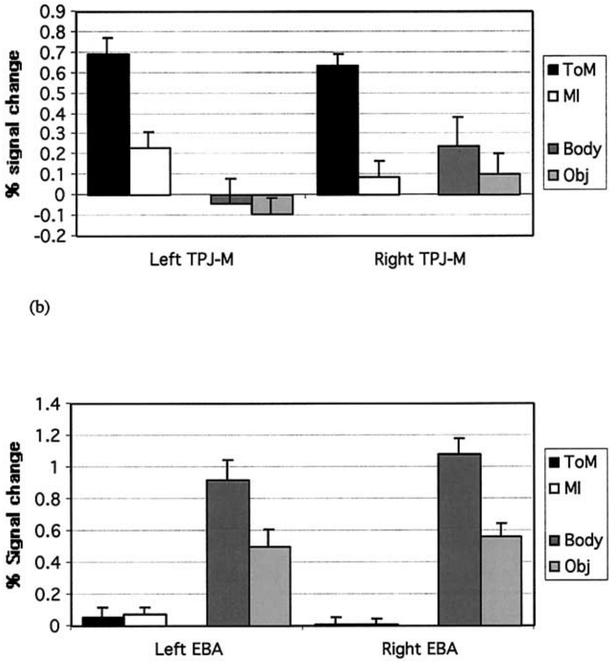


Fig. 2. Experiment 1. Average percent signal change from fixation in (a) left and right TPJ-M and (b) left and right EBA, defined in individual subjects ($n = 14$). The EBA consisted of contiguous voxels in bilateral extrastriate cortex that responded significantly more to pictures of human bodies than pictures of nonhuman objects ($P < 0.0001$, uncorrected). The TPJ-M consisted of contiguous voxels near the temporo-parietal junction that responded significantly more to theory of mind (ToM) stories than to mechanical inference (MI) stories ($P < 0.0001$, uncorrected). (Response magnitudes for the conditions that were used to define the regions of interest are illustrative only.) The EBA did not respond to story stimuli. The right TPJ-M differentiated between pictures of bodies and of objects ($P < 0.05$, paired samples t test), but the left TPJ-M did not. ToM = theory of mind (false belief) stories; MI = mechanical inference stories, Body = photographs of human bodies, Obj = photographs of nonhuman objects; EBA, extrastriate body area.

tograph" stories (Zaitchik, 1990), which require subjects to represent the (false) content of a physical representation such as a photograph or map.

For Experiment 2, we therefore created five new sets of stories (for examples, see Appendix 2): (1) false belief stories, (2) false photograph stories, (3) desires, (4) inanimate descriptions, and (5) physical people. Desire stories described a character's goals or intentions and thus rely on ToM. Nonhuman description stories consisted of short descriptions of nonhuman objects such as plants, cars, or planets. Physical people stories were short descriptions of people from a purely physical perspective: clothing, hair colour, facial markings, and so on.

We predicted that regions specifically involved in ToM should have a equally low response in the nonhuman description and (critically) physical people conditions, and a higher BOLD response in the desire condition. By contrast, regions involved in processing any other representation of other people would show a high BOLD signal for the physical people condition.

Methods

Twenty-one naive right-handed subjects (11 women) were scanned at 1.5 T, using twenty 5-mm-thick axial slices that covered the whole brain. An additional 7 subjects from Experiment 1 (4 women) also participated in part of Experiment 2. All were scanned at 3.0 T using twenty 5-mm-thick near-coronal slices (parallel to the brainstem) covering most of the occipital lobe and the posterior portion of the temporal and parietal lobes.

Story stimuli consisted of 70 stories (12 each of false belief, false photograph, desire, physical description, and nonhuman description, average number of words = 32; see Appendix 2). After each story a two-alternative forced choice "fill-in-the-blank" question was presented for 4 s. The question consisted of a single sentence with a word missing, presented above two alternative completions on the left and right side of the screen. Subjects pressed the left-hand response button if the word on the left completed the sentence to fit the story, and the right-hand button to choose the word on the right. Fifty percent of the false belief, false photograph, and desire story questions probed the character's mental states; the other 50% probed the actual outcome, to prevent formulaic response preparation. Subjects were given three practice trials before going into the scanner: two false belief trials, and one false photograph trial.

Fourteen subjects (including the 7 from Experiment 1) were tested on only false belief and false photograph stories. For these subjects, each run lasted 204 s and consisted of six blocks [each containing 1 story (10 s) and 1 question (4 s)], alternating between the two conditions; there were three blocks per condition per run. The remaining 14 subjects were tested on all five conditions. Each run lasted 272 and consisted of 10 blocks [each containing 1 story (10 s) and 1 question (4 s)]. There were two blocks per condition per run.

Fixations of 12 s were interleaved between blocks. The order of conditions was counterbalanced across runs. Behavioural data were collected during the scan.

Results

Subjects were slower when responding to questions about false photograph than false belief stories (FB: 2.6 vs. FP: 2.8 s, $P < 0.01$), making it unlikely that false belief inferences were simply more difficult.

As predicted, a random effects analysis on the 21 subjects who underwent whole brain scanning revealed regions of increased BOLD signal to false belief compared with false photograph stories ($P < 0.0001$, uncorrected) at the TPJ bilaterally [right: (54 −51 **18**), left: (−48 −63 33)], precuneus/posterior cingulate [(3 −54 30)], right anterior superior temporal sulcus [(54 −18 −15)], and in medial superior frontal gyrus [(6 57 18)] in the frontal pole (Fig. 2). Medial prefrontal cortex has repeatedly been implicated in ToM processing, both in neuroimaging and in lesion studies (e.g., Rowe et al., 2001; Stuss et al., 2001).

For the 7 subjects who were scanned in both Experiments 1 and 2, two additional analyses were conducted to confirm that the TPJ-M was consistent across experiments. First, in all 7 subjects the TPJ-M defined in Experiment 1 overlapped strikingly with TPJ-M defined by the contrast of false belief (FB) versus false photograph (FP) stories in Experiment 2. Fig. 3a shows the overlap in a typical individual subject of the TPJ-M defined by these two tasks. Second, this overlap was confirmed with a functional ROI analysis. Voxels near the TPJ are more active during ToM than mechanical inference stories in these individual subjects in Experiment 1 ($P < 0.0001$, uncorrected) were probed for their response during Experiment 2. This independent ROI showed a much greater response to false belief than false photograph stories in Experiment 2 (mean FB PSC = 1.6, mean FP PSC = 0.7, *t* test $P < 0.02$; Fig. 3b). The reliability of the TPJ-M across experiments makes it unlikely that the results of Experiment 1 were the result of stimulus confounds or logical differences between conditions.

For the 14 subjects who saw all five conditions, the fMRI data were further analyzed within individually defined functional regions of interest (ROI) that included all voxels that met two criteria, i.e., they were significantly more active in at least half of the individual subjects during false belief than false photograph stories ($P < 0.0001$, uncorrected), and they fell within a sphere of 15-mm radius centered on the most significant voxel of clusters identified in the random effects group analysis ($P < 0.0001$, uncorrected) of the same contrast. Using these criteria, we identified ROIs in the TPJ-M bilaterally and right aSTS.

In the TPJ-M and the right aSTS, the BOLD signal change during desire stories was significantly greater than

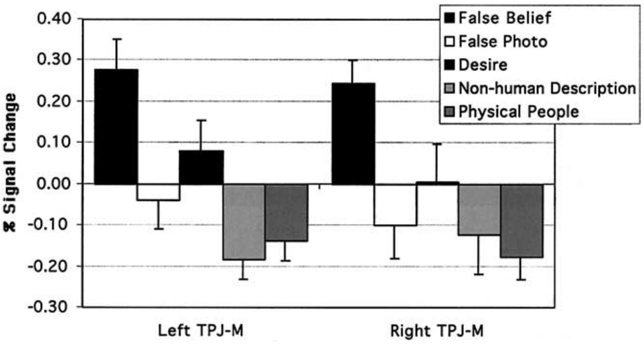


Fig. 4. Experiment 2. Average percent signal change in left and right TPJ-M, defined in individual subjects ($n = 14$) as voxels that respond significantly more to false belief (FB) than to false photo (FP) stories ($P <$ 0.0001, uncorrected. Response magnitude for these two conditions is illustrative only, since these data were used to determine the region of interest). In the TPJ-M bilaterally the BOLD response to physical people stories was significantly lower than to desire stories ($P < 0.05$), and not significantly different from nonhuman description stories ($P < 0.1$, repeated-analysis of variance). Response decreases are commonly observed in the TPJ vicinity during demanding nonsocial tasks (Shulman et al., 1997; Gusnard and Raichle, 2001).

during either physical people or nonhuman description stories (both paired samples t tests $P < 0.05$), which did not differ from each other (Fig. 4). The left and right TPJ-M did not differ. Thus, these regions are not involved in the detection of any person in verbal stories, but respond selectively to stories in which describe (or imply) characters' mental states. Did any regions show the predicted profile of a response to a person per se? At a lower threshold, a separate whole brain analysis ($P < 0.001$, uncorrected) of physical people > nonhuman descriptions revealed regions of frontal cortex [dorsal medial prefrontal (-3 57 39), and right lateral frontal cortex (39 15 54)].

Discussion

The results of Experiment 2 confirm that the TPJ-M shows an increased response to stimuli that invite ToM reasoning compared with logically similar nonsocial controls (false photograph stories). Second, the TPJ-M does not show an increased response to the mere presence of a person in the stimulus (physical people stories). The right and left TPJ-M responses to physical people stories did not differ, thus resolving the ambiguity of the apparently lateralized response to photographs of bodies in Experiment 1.

General discussion

In two experiments we found greater BOLD response in a region within the TPJ bilaterally (here called TPJ-M) while subjects read stories that describe or imply a character's goals and beliefs than during stories about nonhuman objects. This pattern is robust across subjects, tasks, and stimuli, and is not merely an effect of the difficulty or logical structure of false belief stories, since the TPJ-M did not respond to the more difficult and logically similar false photograph stories.

We asked whether the TPJ-M represents the simple presence of another person (possibly via detecting a human body and/or biological motion) or is involved specifically in ToM. We found that the TPJ-M was anatomically and functionally distinct from the nearby EBA (Downing et al., 2001), which responded preferentially to the visual appearance of human bodies, suggesting the presence of at least two distinct regions involved in social information processing.

A key innovation of this study over previous studies was the inclusion in Experiment 2 of physical people stories, which described the physical appearance of human bodies. Previous studies (Fletcher et al., 1995; Gallagher et al., 2000) have included "physical" stories describing acting people, which produced greater activation in the TPJ than a scrambled sentence control. Our data show that the TPJ-M response was no greater to stories that described other people in physical detail than that to stories describing the physical details of nonhuman objects—and was significantly lower than to stories that did invite a mental state interpretation (desire stories).

Could the TPJ-M activation reflect mental imagery of the biological motion or goal-directed action described in the false belief, human action, and desire stories? We think this is unlikely. Saxe, R., Xiao, D.K., Kovacs, G., Perrett, D., and Kanwisher, N. (unpublished data) found that the TPJ-M response to a movie of a walking person was much lower than its response to false belief stories. If the response of the TPJ-M to verbal stories was merely a consequence of subjects' imagining biological motion, we would predict the opposite. Also the TPJ-M was doubly dissociated from its neighbour, the pSTS-VA (visual analysis of action), which responded more to the movies than to verbal stories.

In all, our results show that a region of the TPJ[2] is involved in reasoning about other minds, not just in understanding stories involving people per se (Gallagher and Frith, 2003; p 80). But critically, neighbouring subregions of cortex have different functional profiles, highlighting the necessity of careful within-subject comparisons. The TPJ-M, identified here by responses to (false) belief stories,

---

[2] What is the relationship between the TPJ-M and attention? Selective attention leads to increases in regions of the TPJ during social perception tasks (e.g., Narumoto et al., 2001; Winston et al., 2002), and to decreases in regions of the TPJ during visual attention tasks (Shulman et al., 1997; Gusnard and Raichle, 2001; Jiang Y., Kanwisher, N., unpublished data). Downar et al. (2001) proposed "a role for the TPJ in detecting behaviourally relevant events in the sensory environment" (p. 1256) that is interfered with by demanding visual attention. One possibility is that the mental states of other people constitute a particular category of such "behaviourally relevant" stimuli. Alternatively, these results may reflect functionally and anatomically distinct subregions within the TPJ. Direct testing of the relationship between the TPJ-M and selective attention is an important avenue for future work.

may play a broad role in social and even moral cognition (Moll J et al., 2002; Greene and Haidt, 2003).

Acknowledgments

This work was funded by grants NEI 13455 and NIHM 66696. Our thanks especially to Yuhong Jiang for comments and conversation, and to Ben Balas, Robb Rutledge, Miles Shuman, and Amal Dorai for help with data collection and analysis.

Appendix

Experiment 1

Instructions: "Read each story silently to your self. Please make sure you understand what is happening; it is more important that you understand the story, than that you go as fast as possible. When you are done reading the story, press the button."

Theory of mind (ToM) sample story

A boy is making a paper mache project
for his art class. He spends hours
ripping newspaper into even strips.
Then he goes out to buy flour. His
mother comes home and throws all the
newspaper strips away.

Mechanical inference (MI) sample story

A pot of water was left on low heat
yesterday in case anybody wanted tea.
The pot stayed on the heat all night.
Nobody did drink tea, but this morning,
the water was gone.

Human action sample story

Jane is walking to work this
morning through a very industrial
area. In one place the crane is
taking up the whole sidewalk. To
get to her building, she has to
take a detour.

Experiment 2

Instructions: "Please read each story carefully. After each story, you will be given one fill-in-the-blanks question about the story. Underneath will be two words that could fill in the blank. Choose the correct word (to make the sentence true in the story) by pressing the left button to choose the left-hand word, and the right button to choose the right-hand word."

False belief (FB) sample story

John told Emily that he had a Porsche.
Actually, his car is a Ford. Emily
doesn't know anything about cars
though, so she believed John.
—
When Emily sees John's car she
thinks it is a
porsche ford

False photograph (FP) sample story

A photograph was taken of an apple hanging
on a tree branch. The film took half an hour to
develop. In the meantime, a strong
wind blew the apple to the ground.
—
The developed photograph shows the apple on the
ground branch

Desire sample story

For Susie's birthday, her parents decided
to have a picnic in the park. They wanted
ponies and games on the lawn. If it rained,
the children would have to play inside.
—
Susie's parents wanted to have her birthday
inside outside

Physical people sample story

Emily was always the tallest kid in her
class. In kindergarten she was already
over 4 feet tall. Now that she is in
college she is 6′4″. She is a head taller
than the others.
—
In kindergarten Emily was over
4 ft 6 ft
. . .tall

Nonhuman description sample story

Nine planets and their moons, plus various
lumps of debris called asteroids and
comets, make up the sun's solar system.
The earth is one of four rocky planets
in the inner solar system.
—
The solar system has
four nine
. . .planets.

References

Allison, T., Puce, A., et al., 2000. Social perception from visual cues: role of the STS region. Trends Cogn. Sci. 4, 267–278.

Bartsch, K., Wellman, H., 1995. Children Talk about the Mind. Oxford University Press, New York.

Brunet, E., Sarfati, Y., et al., 2000. A PET investigation of the attribution of intentions with a nonverbal task. Neuroimage 11, 157–166.

Carey, S., 1985. Conceptual Change in Childhood. MIT Press, Cambridge, MA.

Carruthers, P., Smith, P., 1996. Theories of Theories of Mind. Cambridge University Press, New York.

Castelli, F., Happe, F., et al., 2000. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. Neuroimage 12, 314–325.

Dennett, D., 1978. Beliefs about beliefs. Behav. Brain Sci. 1, 568–570.

Dennett, D., 1996. Kinds of Minds: Toward an Understanding of Consciousness. Basic Books New York, NY.

Downar, J., Crawley, A.P., et al., 2001. The effect of task relevance on the cortical response to changes in visual and auditory stimuli: an event-related fMRI study. Neuroimage 14, 1256–1267.

Downing, P.E., Jiang, Y., et al., 2001. A cortical area selective for visual processing of the human body. Science 293, 2470–2473.

Ferstl, E.C., von Cramon, D.Y., 2002. What does the frontomedian cortex contribute to language processing: coherence or theory of mind? Neuroimage 17, 1599–1612.

Fletcher, P.C., Happe, F., et al., 1995. Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. Cognition 57, 109–128.

Frith, C.D., Frith, U., 1999. Interacting minds—a biological basis. Science 286, 1692–1695.

Gallagher, H.L., Frith, C.D., 2003. Functional imaging of "theory of mind." Trends Cogn. Sci. 7, 77–83.

Gallagher, H.L., Happe, F., et al., 2000. Reading the mind in cartoons and stories: an fMRI study of "theory of mind" in verbal and nonverbal tasks. Neuropsychologia 38, 11–21.

Greene, J., Haidt, J., 2003. How (and where) does moral judgement work? Trends Cogn. Sci. 6, 517–523.

Grossman, E., Donnelly, M., et al., 2000. Brain areas involved in perception of biological motion. J. Cogn. Neurosci 12, 711–720.

Gusnard, D.A., Raichle, M.E., 2001. Searching for a baseline: functional imaging and the resting human brain. Nat. Rey. Neurosci. 2, 685–694.

Hoffman, E.A., Haxby, J.V., 2000. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. Nat. Neurosci. 3, 80–84.

Leslie, A., 1999. "Theory of mind" as a mechanism of selective attention, in: Gazzanigal, M. (Ed.), The New Cognitive Neurosciences. MIT Press, Cambridge, MA.

Malle, B.F., Moses, L.J., Baldwin, D.A. (Eds.), 2001. Intentions and Intentionality: foundations of Social Cognition. MIT Press, Cambridge, MA.

Moll, J., et al., 2002. The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. J. Neurosci 22, 2730–2736.

Narumoto, J., Okada, T., et al., 2001. Attention to emotion modulates fMRI activity in human right superior temporal sulcus. Brain Res. Cogn. Brain Res. 12, 225–231.

Rowe, A.D., Bullock, P.R., et al., 2001. "Theory of mind" impairments and their relationship to executive functioning following frontal lobe excisions. Brain 124 (Pt 3), 600–616.

Saxe, R., Xiao, D.K., et al., 2003. Distinct representations of bodies, actions and thoughts in posterior superior temporal sulcus, submitted.

Shulman, G.L., Corbetta, M., et al., 1997. Top-down modulation of early sensory cortex. Cereb Cortex 7, 193–206.

Stuss, D.T., Gallup Jr., G.G., et al., 2001. The frontal lobes are necessary for "theory of mind." Brain 124 (Pt 2), 279–286.

Tong, F., Nakayama, N., et al., 2000. Response properties of the human fusiform face area. Cogn. Neuropsychol. 17, 257–279.

Vogeley, K., Bussfeld, P., et al., 2001. Mind reading: neural mechanisms of theory of mind and self-perspective. Neuroimage 14 (1 Pt 1), 170–181.

Wellman, H.M., Gelman, S., 1992. Cognitive development: foundational theories of core domains. Annu. Rev. Psychol. 43, 337–375.

Wimmer, H., Perner, J., 1983. Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition 13, 103–128.

Winston, J.S., Strange, B.A., et al., 2002. Automatic and intentional brain responses during evaluation of trustworthiness of faces. Nat. Neurosci. 5, 277–283.

Zaitchik, D., 1990. When representations conflict with reality: the preschooler's problem with false beliefs and "false" photographs. Cognition 35, 41–68.