

Visualizing How One Brain Understands Another: A PET Study of Theory of Mind

Chadi Calarge, M.D.

Nancy C. Andreasen, M.D., Ph.D.

Daniel S. O'Leary, Ph.D.

Objective: Theory of mind (TOM), or "mentalizing," refers to the ability to attribute mental states to self and others. Inferring what people are thinking and feeling is an important aspect of human social interaction, and it is also an important aspect of both psychiatric diagnosis and treatment. The authors conducted a positron emission tomography (PET) study to examine the neural substrates of TOM, using a task that mimics real-life social interaction.

Method: Thirteen healthy volunteers underwent [^{15}O]H $_2\text{O}$ PET while performing an experimental task and a control task. During the experimental task they created a "story" about the mental state of a stranger whom they imagined encountering on a park bench. During the control task, they read aloud a story requiring no mental state attribution.

Results: The TOM task activated an extensive neural network that included the medial frontal cortex, the superior frontal cortex, the anterior and retrosplenial cingulate, and the anterior temporal pole; most of these activations were limited to the left hemisphere. In addition, the largest activation was in the contralateral right cerebellum, as well as the anterior vermis.

Conclusions: A language-based TOM task activated distributed brain regions that are important for representing mental states of the self and others, retrieving memory of personal experiences, and coordinating and monitoring the overall performance of the task. The activations in the medial frontal cortex replicate findings in previous TOM studies, while the activations in the cerebellum reinforce the growing evidence that the cerebellum performs cognitive functions in the human brain.

(*Am J Psychiatry* 2003; 160:1954–1964)

The ability to infer what other people are thinking and feeling is one of the most fundamental aspects of human social interaction. This capacity is also known as "mentalizing," having a "theory of mind," or possessing an "intentional stance" (1–7). This ability forms the basis of such diverse human activities as negotiating treaties, deciding to make a charitable donation, or selecting a day-care center for one's child. It is also the basis for many of the interactions that occur during the practice of psychiatry, such as conducting a diagnostic interview, assessing a person's motives or psychodynamics, performing psychotherapy, assessing the severity of a syndrome and the need for medications, or determining suicidal risk. Therefore, understanding how our brains and minds actually "mentalize" is a topic of great interest. The tools of functional neuroimaging now permit us to observe one person's mind as it attempts to infer what is occurring in another person's mind, thereby making a previously mysterious process accessible to scientific investigation.

Several recent studies have begun the process of exploring the neural basis of theory of mind (TOM): how networks and faculties within our brains and minds activate and interact when we perform mental activities that require us to understand the subjective reality of other people.

These studies have used a variety of imaging techniques, designs, and test materials.

In one of the first positron emission tomography (PET) studies on TOM, Goel et al. (8) showed activation of the left medial frontal gyrus and the left middle and superior temporal gyri when they compared normal subjects making inferences about the minds of others (from Columbus's era) to a baseline condition of making inferences about the physical world. Another early PET study compared subjects on comprehension of "TOM" stories, "physical" stories, and unlinked sentences, and it also revealed a selective activation of the left medial frontal lobe in the TOM condition (9). A more recent PET study (10) used a nonverbal task involving cartoons to compare attribution of intent with several control conditions, and it too found medial frontal activations, as well as others in the bilateral temporal lobes and cerebellum. In an fMRI study, Gallagher et al. (11) showed their subjects stories and cartoons requiring both TOM and non-TOM processing. They found activations in the medial prefrontal gyrus and in the temporoparietal junctions during the TOM task. Taken together, these studies have begun to identify some of the regions involved when human beings attribute mental states to others, and in particular they suggest that medial frontal brain regions play an important role.

Most of the studies to date have attempted to dissect components of TOM in designs that are relatively simple and therefore clean and elegant. Most have examined passive responses to stimuli rather than active engagement in the TOM task. However, mentalizing as it occurs in daily life is not a simple cognitive process, nor is it usually passive. Evidence from evolutionary biology suggests that a fully developed TOM ability is limited to human beings and requires multiple cognitive capacities and systems (1, 3, 12–16). It has been suggested that the capacities for language and episodic memory make some contributions to the highly developed human ability to mentalize (4, 10, 16–21).

Therefore, we chose to conduct a PET study of TOM using a task that requires subjects to actively place themselves in another person's place and to attribute mental states to that person. We asked them to imagine and describe the experience of a stranger met during a chance encounter on a park bench. While we measured regional cerebral blood flow (rCBF) using PET, healthy subjects verbally described what the stranger had experienced. We analyzed the subjects' utterances in order to assess their "mentalizing" capacities. The abilities examined during this task are similar to those used by psychiatrists when they take a patient's history and attempt to imagine the internal experience of another person. They are also similar to those used in many types of psychotherapy, including psychodynamic psychotherapy. Therefore, this study provides an opportunity to begin to understand the neural basis of the mental processes that are used in some aspects of the practice of psychiatry.

Method

Subjects

The subjects were 13 healthy right-handed volunteers (seven women and six men) recruited from the community. They were screened to rule out a current or past history of psychiatric illness by using a short version of the Comprehensive Assessment of Symptoms and History (22). They were also evaluated by means of a history and physical examination to rule out any current or past history of neurological or general medical illness. Their mean age was 26.5 years (SD=6.4), and their mean educational achievement was 14.6 years (SD=0.9). They had a mean full-scale IQ of 108 (SD=9). All gave written consent to a protocol approved by the University of Iowa Human Subjects Institutional Review Board.

Tasks

During the experimental task, referred to as the "TOM story," the subject was told: "Imagine you sat next to a woman (for men), or next to a man (for women) on a park bench and you realized she (he) was crying. Make up a story about what led up to her (his) crying." These instructions were presented to the subjects on a video monitor positioned 12–13 inches from their eyes. Subjects were given 30 seconds to read the instructions and plan their narrative before beginning their story. They were allowed to speak for approximately 100 seconds, and PET data were acquired during this time interval. The control/comparison task ("read story") was to read aloud a simple story requiring no attribution of mental states. This story was also presented on the video monitor. The subjects were required to read for 40 seconds. If they finished

reading the story before the time was up, they had to restart from the beginning. The experimental and control tasks both required the subjects to read instructions or other material and to speak aloud in continuous narrative speech, thereby isolating the TOM component in the experimental task. Subjects were audiotaped during the two conditions, and transcriptions were prepared from the tapes. Both conditions were timed so that the subject began speaking 10 seconds before the arrival of the bolus of [^{15}O]H $_2\text{O}$ in the brain. The two tasks were part of a larger eight-condition study of reading and language production. The order of the tasks in this study was counterbalanced. Results from the analysis of other conditions have not yet been reported.

Scoring

Each TOM story transcript was rated independently by two raters for the occurrence of mental state attributions to self or others. Every TOM story was divided into separate sentences or parts on the basis of what was being relayed. Each one of these parts is referred to as an utterance, which is defined as an idea that could be easily distinguished from the one that precedes or follows. Each utterance was scored as level 0 if no attribution of mental states was made, level I if one level of attribution of mental states to self or other was made, and level II if at least two levels of attributions of mental states to self or other were made (see Appendix 1 for examples of rating). Both raters had trained by reading the literature available as well as agreeing on guidelines to follow (17, 23–27). The interrater agreement was 76%. The percentage of mental state attributions was then determined by counting the total number of utterances containing attributions and dividing that number by the total number of utterances ($\times 100$). All utterances containing TOM statements used rather rhetorically (e.g., "...I guess" and "...you know") were not scored (28).

PET Imaging

Quantitative PET blood flow data were acquired on a GE 4096-plus whole-body scanner (General Electric Systems, Milwaukee) after a bolus injection of [^{15}O]H $_2\text{O}$ (29). To acquaint the subject with the imaging conditions and to ascertain stimulus timing, the time from injection to bolus arrival in the brain was individually measured by delivering a 15 mCi bolus during an initial scout injection. During the scout injection, the subject was asked to read a list of words presented on the monitor. All subsequent scans employed a 50 mCi [^{15}O]H $_2\text{O}$ intravenous bolus dose. Imaging began at the time of injection ($t=0$) and continued for 100 seconds in the form of 20 frames of 5 seconds each. Arterial blood was sampled from a catheter placed in the radial artery to obtain the input function needed for calculation of tissue perfusion (in milliliters per minute per 100 g of tissue).

On the basis of the time-activity curves over major cerebral arteries, the eight frames reflecting the 40 seconds of postbolus transit were summed and reconstructed into 2-mm voxels (128 \times 128 matrix) by using a Butterworth filter (order=6, cutoff frequency=0.35 Nyquist units) (30). By using this summed image and the measured arterial input function, the CBF was calculated on a pixel-by-pixel basis by means of the autoradiographic method (31). The CBF calculation was normalized by dividing it by the global CBF. To reduce anatomical variability, an 18-mm Hanning filter was applied. Imaging was repeated at approximately 15-minute intervals.

Magnetic Resonance Imaging

MR scans, used for anatomic localization of functional activity, were obtained for each subject with a standard T $_1$ -weighted three-dimensional spoiled gradient recall acquisition sequence on a 1.5-T GE Signa scanner (General Electric Systems, Milwaukee) (TE=5, TR=24, flip angle=40°, NEX=2, field of view=26, matrix=256 \times 192, 1.5-mm slice thickness).

TABLE 1. Measures of Performance on a Theory of Mind (TOM) Task and a Control Task of 13 Healthy Subjects in a PET Study of TOM

| Task and Measure | Mean | SD |
|--|------|------|
| TOM task^a | | |
| Number of utterances rated by level of mental state attributions to self or others | | |
| Level 0 (no attributions) | 6.3 | 3.7 |
| Level I (one level of attributions) | 8.3 | 2.8 |
| Level II (at least two levels of attributions) | 0.2 | 0.5 |
| Rate of mental state attributions per utterance | 0.59 | 0.22 |
| Number of utterances | 14.7 | 4.0 |
| Number of mental state attributions \times 100/ number of words | 2.7 | 0.9 |
| Rate of speech | | |
| Number of words | 257 | 69 |
| Number of words per minute | 163 | 26 |
| Control task^b | | |
| Rate of speech | | |
| Number of words | 114 | 19 |
| Number of words per minute | 172 | 29 |

^a Subjects were told to imagine sitting next to a stranger who was crying on a park bench and to make up a story about what led up to the person's crying. These instructions were presented to the subjects on a video monitor positioned 12–13 inches from their eyes. Subjects were given 30 seconds to read the instructions and plan their narrative before beginning their story. They were allowed to speak for approximately 100 seconds, during which PET data were acquired.

^b Subjects read aloud a simple story presented on a video monitor positioned 12–13 inches from their eyes. The subjects were required to read for 40 seconds. If they finished reading the story before the time was up, they had to restart from the beginning. PET data were acquired 10 seconds after the subjects began to read.

Image Analysis

The normalized quantitative PET blood flow images and MR images were analyzed by using the locally developed software package BRAINS (32–34). The outline of the brain was identified on the MR images by a combination of edge detection and manual tracing. MR scans were volume rendered; the anterior commissure-posterior commissure line was identified and used to realign the brains of all the subjects to a standard position to place each brain in standardized Talairach coordinate space (35). The PET image of each individual was then fit to that individual's MR scan by using a surface-fit algorithm (36). Subjects were checked for head movement with each injection, and images were individually refit as needed. The MR images of all the subjects were averaged, so that the data obtained with PET could be localized on coregistered MR images. The coregistered images were resampled and simultaneously visualized in all three orthogonal planes.

Statistical Analysis

Student's *t* test was used to analyze the differences in behavioral performance across groups and conditions.

Statistical analysis for a within-group comparison of the PET images was performed by using an adaptation (29) of the method of Worsley et al. (37). Images were resampled to 128 \times 128 \times 80 voxels by using the Talairach atlas bounding box. A within-subject subtraction of the findings for the experimental and control tasks was then performed, followed by across-subject averaging of the subtraction images and computation of voxel-by-voxel *t* tests of the changes in rCBF. Significant regions of activation were calculated on the *t*-map images by using a technique that corrects for the large number of voxel-by-voxel *t* tests, the lack of independence between voxels, and the resolution of the processed PET images (29, 37). Areas containing at least 50 voxels with a *t* value greater than 3.61 are reported in the tables, as well as the highest

t value (*t* max) and the total number of voxels in the region (equivalent to *p*=0.0005, uncorrected).

Results

Behavioral Data

Table 1 shows the behavioral data of the subjects, indicating their performance during the TOM task and the read-story task. The difference between conditions in the rate of word production per minute was not statistically significant (read-story task: mean=172, SD=29; TOM task: mean=163, SD=26) (*p*=0.40). Subjects produced a mean of 14.7 utterances (SD=4.0) and 257 words (SD=69) during the TOM task. Table 1 also shows the average number of mental state attributions at the different levels during the TOM task, as scored by the two raters. There were, on average, 6.3 level-0 utterances (SD=3.7), 8.3 (SD=2.8) level-I utterances, and 0.2 (SD=0.5) level-II utterance per story. Fifty-nine percent of the utterances contained attribution of mental states to self or others (i.e., mean=0.59, SD=0.22, of the total number of utterances contained level-I or level-II attributions).

Imaging Data

Table 2 shows the results of subtracting the read-story condition from the TOM-task condition. The table indicates the brain regions with a higher rCBF during one condition compared with the other, using region names based on inspection of the coregistered MR and PET images, as well as the *x*, *y*, and *z* coordinates from the Talairach atlas. Areas containing at least 50 voxels with a *t* value greater than 3.61 are reported in the tables, as well as the highest *t* value (*t* max) and the total number of voxels in the region. The positive *t* values represent the brain regions that showed greater blood flow during the TOM-task condition than during the read-story condition, while the negative *t* values show regions with greater blood flow during the read-story condition. The data are presented by using the Talairach atlas convention, with *x* referring to the left/right position with respect to a midsagittal plane, *y* referring to the rostral/caudate (anterior/posterior) position with respect to a verticofrontal plane defined by the anterior commissure, and *z* referring to the superior/inferior position with zero at the anterior-posterior commissure plane.

Figure 1, Figure 2, Figure 3, and Figure 4 show four sets of PET images that have been selected to illustrate the activations shown in Table 2. Visual display of results is shown in two ways. One presentation shows only the peaks, as defined by the volume measurement, superimposed on the composite average MR image from the 13 subjects. The other presentation, referred to as the "*t* map," shows the color-coded *t* values for all voxels in the image. The peak map and the *t* map provide complementary information. The former identifies areas of activation by using a strict definition based on a relatively arbitrary cutoff point, while the latter provides a more descriptive

TABLE 2. Brain Regions With Significant Increase in Activation During a Theory of Mind (TOM) Task and a Control Task in 13 Healthy Subjects in a PET Study of TOM

| Task and Brain Region | Brodmann's Area | t Max ^a | Talairach Coordinates | | | Number of Voxels |
|--|-----------------|--------------------|-----------------------|-----|-----|------------------|
| | | | x | y | z | |
| TOM task ^b | | | | | | |
| Medial frontal gyrus | 32/10 | 3.97 | 0 | 45 | 1 | 93 |
| Right anterior cingulate | 24/32 | 4.38 | 5 | 33 | 15 | 1,095 |
| Left superior frontal gyrus/paracingulate area | 6/9/32 | 5.99 | -20 | 10 | 50 | 4,292 |
| Left superior frontal gyrus | 8 | 4.31 | -15 | 31 | 34 | 361 |
| | 9 | 3.96 | -17 | 42 | 20 | 81 |
| Left inferior frontal gyrus | 47 | 3.90 | -34 | 15 | -8 | 66 |
| Left retrosplenial cingulate | 23 | 4.06 | -8 | -55 | 23 | 69 |
| Left temporal pole | 21 | 4.57 | -44 | 0 | -28 | 819 |
| Left angular gyrus | 39 | 4.11 | -48 | -64 | 25 | 73 |
| Left lateral cerebellum | | 4.50 | -46 | -72 | -34 | 375 |
| Cerebellar vermis anterior lobe | | 4.08 | -3 | -55 | -19 | 143 |
| Right cerebellum | | 7.79 | 29 | -80 | -24 | 10,360 |
| Cerebellar vermis anterior lobe | | 4.04 | 6 | -55 | -5 | 172 |
| Control task ^c | | | | | | |
| Left Heschl's gyrus | 41/42 | -3.94 | -44 | -26 | 6 | 62 |
| Left middle temporal gyrus | 37 | -4.00 | -46 | -58 | 4 | 96 |
| Right inferior frontal gyrus | 44 | -5.43 | 44 | 8 | 16 | 1,479 |
| | 44 | -5.19 | 42 | 9 | 26 | 366 |
| Right superior parietal lobe | 7 | -5.16 | 34 | -42 | 48 | 442 |
| | 7 | -3.97 | 24 | -72 | 40 | 170 |
| Right inferior temporal gyrus | 37 | -5.73 | 49 | -54 | -11 | 1,396 |
| Right calcarine fissure | 17 | -10.94 | 3 | -77 | 14 | 20,665 |

^a Highest absolute value of t in region.

^b Subjects were told to imagine sitting next to a stranger who was crying on a park bench and to make up a story about what led up to the person's crying. These instructions were presented to the subjects on a video monitor positioned 12–13 inches from their eyes. Subjects were given 30 seconds to read the instructions and plan their narrative before beginning their story. They were allowed to speak for approximately 100 seconds, during which PET data were acquired.

^c Subjects read aloud a simple story presented on a video monitor positioned 12–13 inches from their eyes. The subjects were required to read for 40 seconds. If they finished reading the story before the time was up, they had to restart from the beginning. PET data were acquired 10 seconds after the subjects began to read.

picture of the geography of the circuitry involved. Regions activated by the TOM task appear in shades of red/yellow, and regions that were used for the read-story task appear in shades of blue.

The TOM condition appears to activate an extensive network that is primarily in the left hemisphere. Examples of the components of this network are shown in Figure 1, Figure 2, Figure 3, and Figure 4. As has been noted in all previous functional imaging studies of TOM, the medial frontal cortex is activated. Multiple additional frontal activations are also observed in both the superior and inferior frontal regions; one of the superior frontal regions is on the boundary of the cingulate gyrus (i.e., the paracingulate), and activations are also present in both the anterior cingulate and the retrosplenial cingulate. Additional regions used during the TOM task include the angular gyrus of the parietal lobe and the anterior pole of the temporal lobe.

During the TOM task the largest activation occurs in the right cerebellum, however, and is composed of 10,360 voxels. This activation reflects the interaction of the right cerebellum with the contralateral left hemisphere cortical activations, as predicted by cross-hemispheric diaschisis. Three other smaller cerebellar activations are also present, including two in the anterior lobe of the vermis.

As shown in Table 2 and in Figure 1, Figure 2, Figure 3, and Figure 4, the control task activates regions that are en-

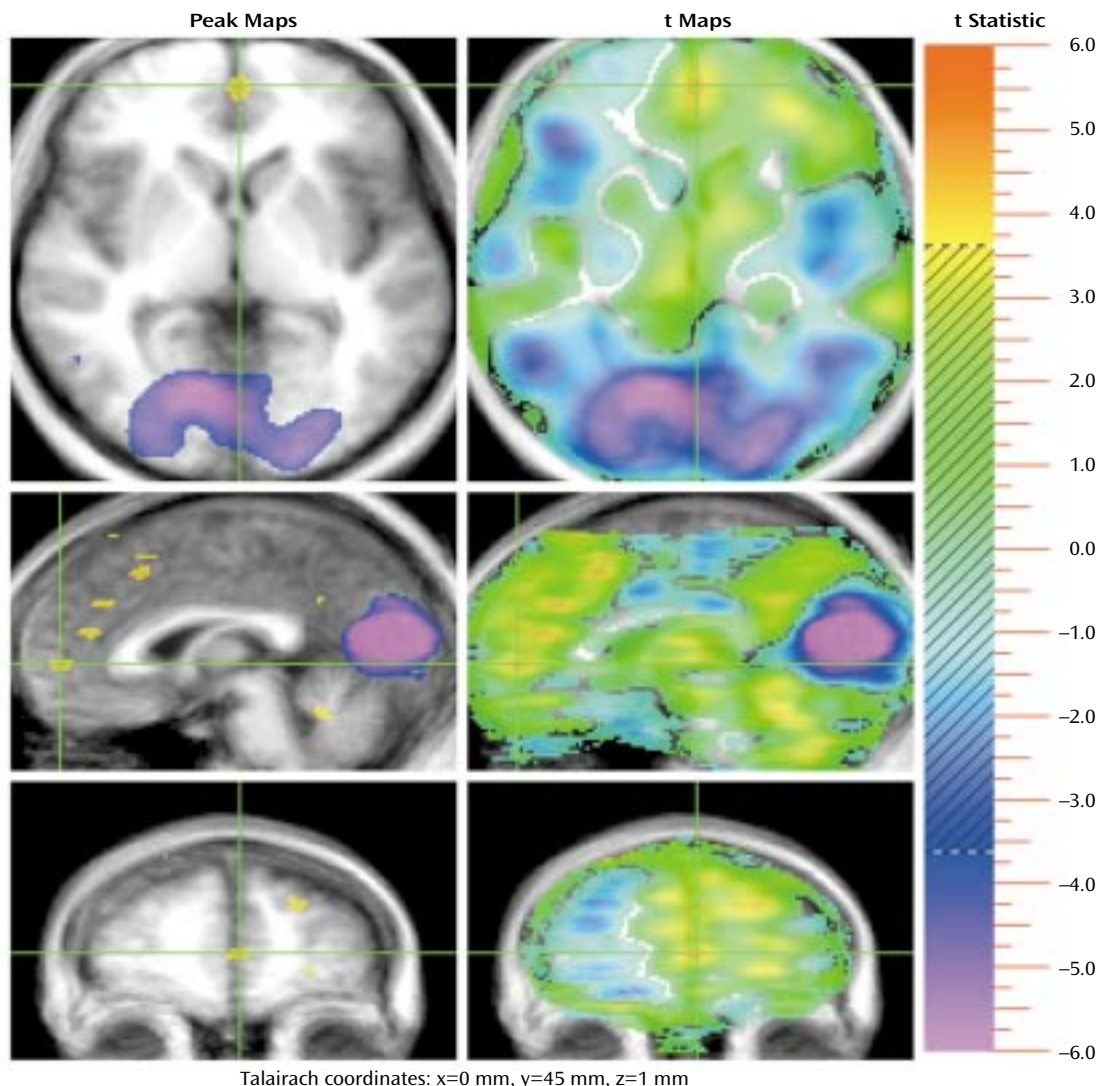
gaged by the process of reading a story. These activations occur primarily in the right hemisphere. As expected, the bilateral visual cortex is extensively activated, reflecting the need to visually scan the material being read aloud. Heschl's gyrus (auditory cortex) is also activated, as the subjects listen to themselves read. Other areas of activation due to reading include several inferior frontal, superior parietal, and inferior temporal regions.

Discussion

This study adds to the growing literature on TOM by examining a complex mentalizing task that requires the use of language and active imagining of another person's mental state. It replicates some earlier findings and adds some new ones. The TOM task investigated in this study activated a distributed group of brain regions that are engaged when individuals attempt to imagine and verbally describe the subjective state of another individual whom they have never met before. It is related to the state that psychiatrists think of as "the empathic mode."

The medial frontal cortex appears to be a crucial region for this TOM task. Like the four previous studies that have used PET or fMRI to investigate TOM (8–11), this study showed that the medial frontal region is activated when human beings attribute mental states (Figure 1). The function of this region is thought to be maintenance of a repre-

FIGURE 1. Sites in the Medial Frontal Lobe/Anterior Cingulate Showing Significant Increase in Activation During a Theory of Mind (TOM) Task^a Compared to a Control Task^a in 13 Healthy Subjects^{b,c}



^a In the TOM task, subjects were told to imagine sitting next to a stranger who was crying on a park bench and to make up a story about what led up to the person's crying. These instructions were presented to the subjects on a video monitor positioned 12–13 inches from their eyes. Subjects were given 30 seconds to read the instructions and plan their narrative before beginning their story. They were allowed to speak for approximately 100 seconds, during which PET data were acquired. In the control task, subjects read aloud a simple story presented on a video monitor positioned 12–13 inches from their eyes. The subjects were required to read for 40 seconds. If they finished reading the story before the time was up, they had to restart from the beginning. PET data were acquired 10 seconds after the subjects began to read.

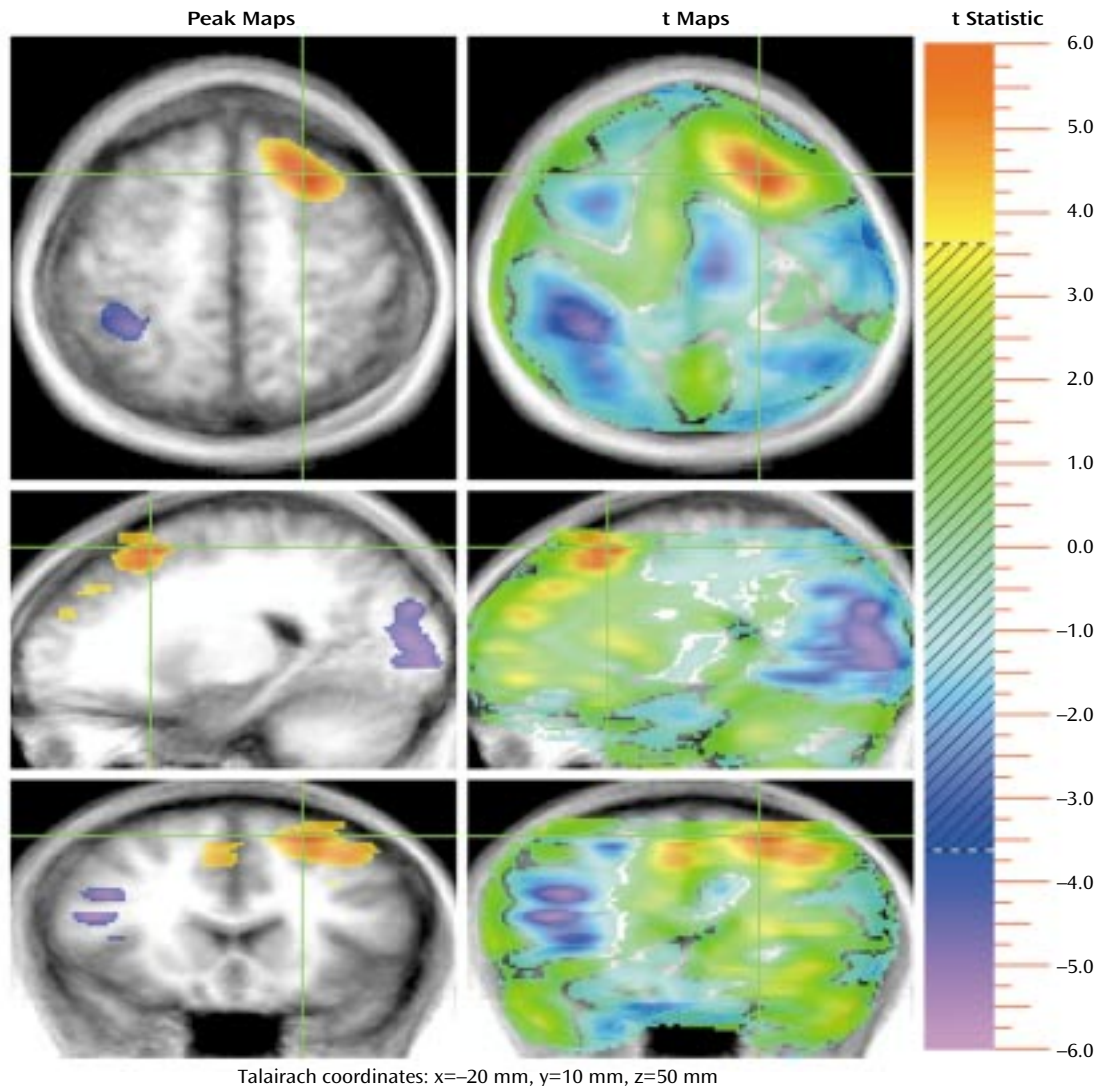
^b The left panel presents statistical maps of contiguous voxels exceeding a significance level of $t=3.61$ (equivalent to $p=0.0005$). The right panel presents statistical maps showing all t values. Both positive and negative peaks are represented. For each slice (axial, sagittal, and coronal planes), the green crosshairs show the location of the two other slices, and the intersection of the crosshairs indicates the pixel with the highest t value within the area marked by the crosshairs. Yellow and red colors indicate activations during the TOM task; blue and purple colors indicate activations during the control task. The hatched part of the color scale covers colors corresponding to data that did not reach statistical significance and, therefore, do not appear on the peak map (left).

^c Three discrete activations during the TOM task are shown in the medial frontal lobe/anterior cingulate and are seen most clearly in the sagittal plane. An activation in the cerebellar vermis during the TOM task is also seen in the sagittal plane. In the posterior brain, the large occipital activation during the control task is shown in both the transaxial and sagittal planes, reflecting the more prominent use of the visual cortex during this task.

sensation of the mental state of self (12). The previous four studies that also found medial frontal activation were all quite different in their designs and stimulus materials: pictures of various kinds of objects for which subjects must infer the objects' normal use and model another person's knowledge about the objects' use (8), inferential ver-

sus factual stories (9), cartoons versus stories that also contrast facts with jokes and deception (11), and cartoons that contrast intent with physical causality (10). Yet these different designs found medial frontal activation, predominately in the left hemisphere, during the TOM condition. Therefore, we can conclude that this region must be cen-

FIGURE 2. Sites in the Left Superior Frontal Gyrus/Paracingulate Area Showing Significant Increase in Activation During a Theory of Mind (TOM) Task Compared to a Control Task in 13 Healthy Subjects^{a,b}



^a See footnotes a and b in Figure 1.

^b The crosshairs are placed in a very large area of activation during the TOM task (4,292 voxels) in the left superior frontal gyrus and the paracingulate area, which is seen in all three planes. Activations during the control task are seen in the right superior parietal lobule (transaxial plane), the occipital cortex (sagittal plane), and the right inferior frontal lobe (coronal plane).

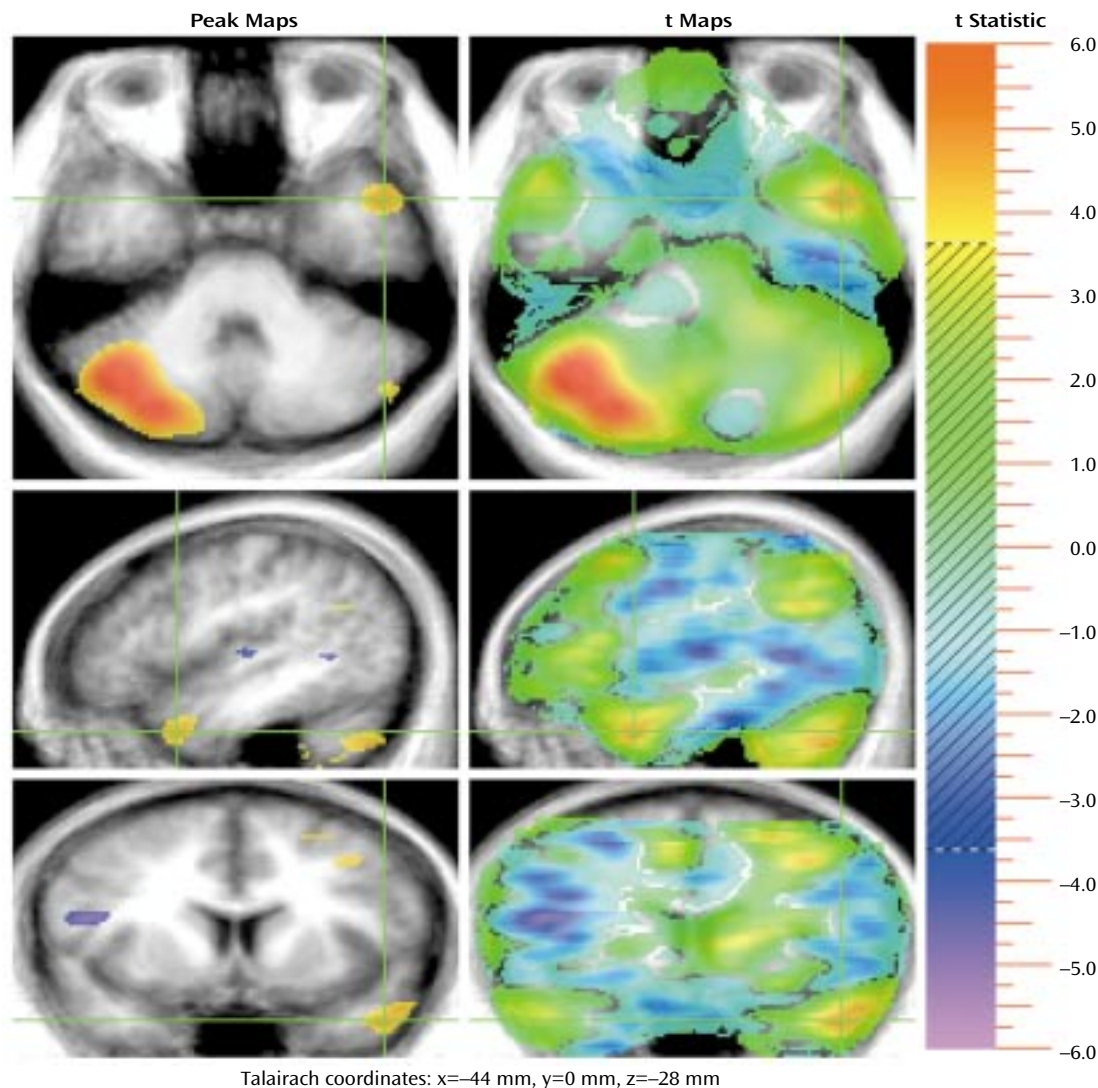
trally involved in the core process of mentalizing in a manner that is independent of the modality through which the TOM task is presented (e.g., visual versus verbal).

This study differs from its predecessors in two significant respects that may affect its results and illuminate other components of TOM. Unlike the previous studies, which required passive viewing of stimuli and some type of simple response (often pushing a button in response to a question), this study asked subjects to actively imagine the internal state of another person and to describe that internal state in extended verbal discourse. This task required using both language and memory to create a TOM discourse. As shown in the behavioral data, subjects made an attribution of mental states in 59% of their utterances, indicating that this task was appropriate for tapping into

TOM. There was no difference in the rates of words produced per minute between the experimental and control tasks. Therefore, we can conclude that the verbalization component of the TOM task was subtracted out through the control task. Furthermore, the classic "language production regions" (e.g., Broca's area, Wernicke's area, sensorimotor speech areas) were not seen as activated in the experimental task, adding to the evidence that this task successfully examined TOM.

Since making up a story about another person does require the creative and spontaneous use of language (which the control task did not), however, we cannot assume that language components were completely eliminated. Nor is such elimination necessarily desirable, given the extensive literature suggesting that language is a cru-

FIGURE 3. Sites in the Right Cerebellum Showing Significant Increase in Activation During a Theory of Mind (TOM) Task Compared to a Control Task in 13 Healthy Subjects^{a,b}



^a See footnotes a and b in Figure 1.

^b A large right cerebellar activation during the TOM task is seen most prominently in the transaxial plane. The crosshairs are placed in an area of activation during the TOM task in the left temporal pole, which is seen in all three planes.

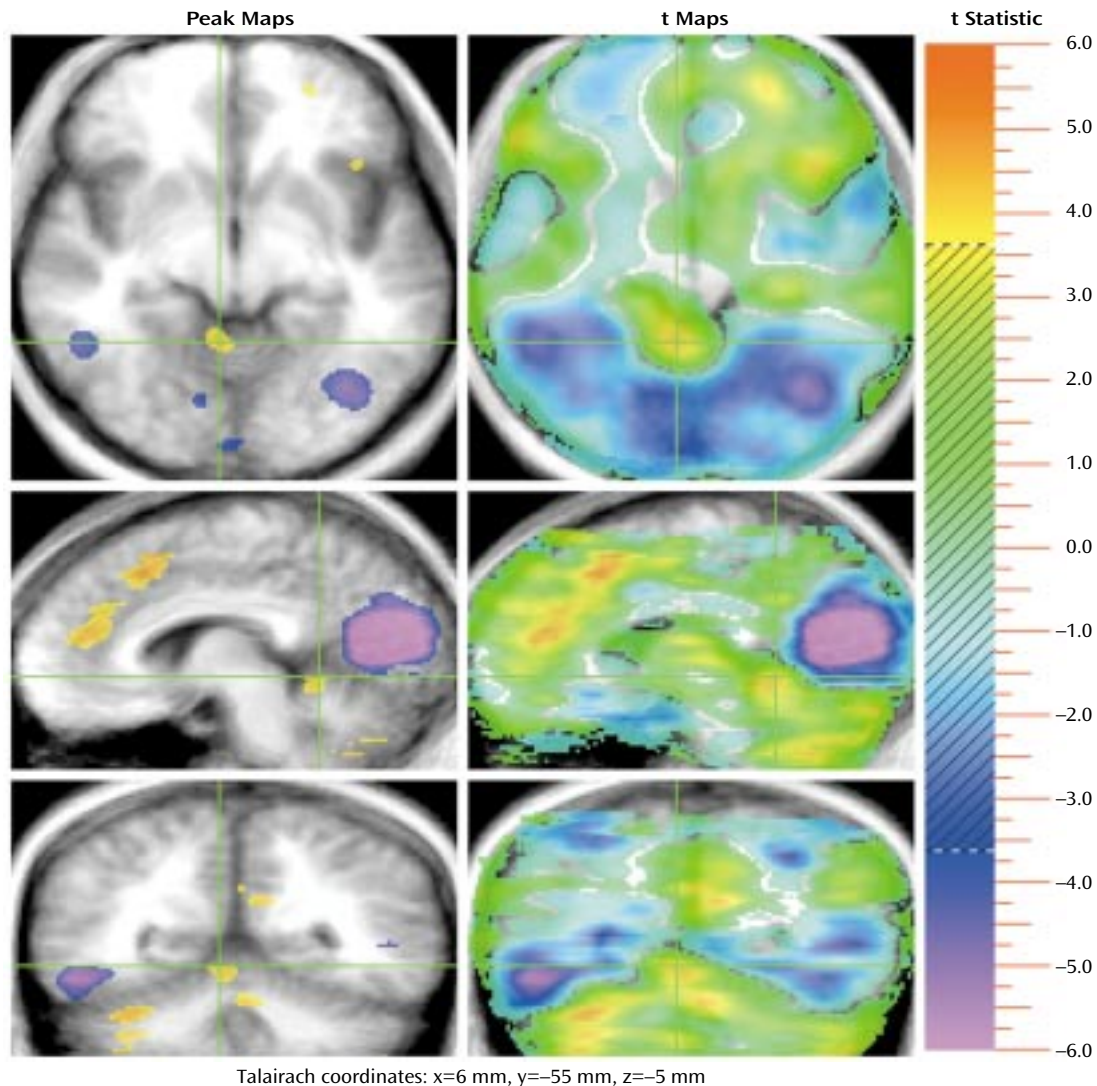
cial component of TOM (10, 17–19, 38). Likewise, it has been suggested that TOM recruits episodic memory (recollections of past personal experience) (16–21). TOM is a complex cognitive process, which is likely to require that subjects mentally reference their own past experiences as they attempt to imagine what the other person has experienced. Therefore, the distributed activations observed in this study must be interpreted in the light of the possible participation of these additional brain systems, which must to some extent participate in the “TOM system” when it is challenged by a complex empathic task.

The medial frontal activations observed in this study are part of a large cluster of frontal activations. One group occurs in the superior frontal regions, with three significant activations that include a more anterior activation (Talairach coordinates: $x=-17$, $y=42$, $z=20$), a more rostral one

(coordinates: $x=-15$, $y=31$, $z=34$), and finally the most posterior one (coordinates: $x=-20$, $y=10$, $z=50$) (Figure 2). The most posterior activation is very large (4,292 voxels) and divides into two separate activations if a higher significance threshold is used. At a threshold of 4.2, this activation separates into two: one centered in the superior frontal gyrus (coordinates: $x=-20$, $y=9$, $z=50$; 4.4 cc; t max=5.99) and one in the anterior cingulate (Brodmann's area 32) (coordinates: $x=-5$, $y=14$, $z=40$; 0.5 cc; t max=4.76). The function of these multiple superior frontal regions is not definitively known, but it is likely that they reflect various aspects of medial frontal functions. That is, they may contribute to the emotional, cognitive, and action components of representation of the self (12).

In addition, a second group of frontal activations occurs in several regions of the cingulate gyrus (Figure 1 and Fig-

FIGURE 4. Sites in the Cerebellar Vermis Anterior Lobe Showing Significant Increase in Activation During a Theory of Mind (TOM) Task Compared to a Control Task in 13 Healthy Subjects^{a,b}



^a See footnotes a and b in Figure 1.

^b Two activations occurred in the anterior lobe of the cerebellar vermis during the TOM task. These activations are clearly seen in the coronal plane; this plane shows a small portion of the large activation in the right lateral cerebellum. The transaxial and sagittal planes also show the vermis activation during the TOM task. In addition, the sagittal section shows portions of the activation in the paracingulate region during the TOM task. A large activation in the primary visual cortex (20,665 voxels) during the control task is shown in all three planes and also extends into the visual association cortex.

ure 4). The activations are bilateral, but the largest (coordinates: x=5, y=32, z=15; 1,095 voxels) is on the right (with some extension into the left hemisphere). The cingulate gyrus has been extensively studied in functional imaging studies, as well as in lesion and animal studies. A recent review proposed that the role of the anterior cingulate is to juxtapose emotions with focused problem solving, error recognition, self-control, and adaptive response to changing conditions (39). The cingulate gyrus is widely connected with diverse parts of the brain and appears to be central for coordinating and focusing attention on complex tasks. Several lesion, animal, and electrophysiological studies have identified an anterior “emotional” part of the

anterior cingulate and another more posterior “cognitive” region (40, 41). Thus, the multiple cingulate activations appear to reflect efforts to integrate the emotional and cognitive components of the TOM task.

A retrosplenial cingulate activation is also present. This region has been shown in several imaging studies to play a role in the expression of emotions, as well as in episodic memory (42–44). Working in conjunction with the other activated regions, the retrosplenial cortex may be used for focusing attention on an emotionally charged topic, retrieving personal experiences that will assist the empathic response, and attributing the retrieved information to self and others (4, 16, 20, 21). An additional group of activations

may also reflect the language and memory retrieval components of this particular TOM task. They include the activations in the angular gyrus and the anterior temporal pole (Figure 3). The angular gyrus is a parietal association region that has connections to the anterior temporal pole, as well as to prefrontal regions (19). These regions may be activated by the search for specific words, for story content, and for episodic memories. Both frontal and anterior temporal regions have been shown to be active in verbal memory tasks that require verbal recall of word lists, complex narratives, and episodic memory (42, 45, 46). Anterior temporal activations were also observed in all previous PET and fMRI TOM studies (8–11), suggesting that these regions may be a modality-independent part of the TOM system.

The largest activation during the TOM task occurs in the right cerebellum (10,360 voxels). This peak fills the lateral and medial aspects of the right cerebellum. There is a small mirror activation in the left lateral cerebellum. In addition, two activations occur in the anterior lobe of the cerebellar vermis (Figure 3 and Figure 4). These findings add to the rapidly growing evidence that the cerebellum is an important “cognitive organ” in the human brain. They indicate that the cerebellum plays a role in yet another complex “higher” nonmotor mental activity: performing a task that requires a theory of mind. Previous studies have shown that the cerebellum is activated in many other cognitive tasks such as memory for faces, verbal memory, episodic memory, error detection, attention, sensory activation and discrimination, and timing (47–60). Cerebellar activations were also seen in two earlier TOM studies (10, 11). The large right cerebellar activation reflects cross-hemispheric diaschisis: the multiple left hemisphere frontal, temporal, and parietal regions used in the TOM task are linked to the contralateral cerebellum, reflecting recognized anatomic pathways (47–58). The two activations in the vermis may be presumed to reflect a similar link to the midline “paralimbic” activations in the cingulate gyrus. Taken together, these large and multiple cerebellar activations indicate that the cerebellum is working in concert with the cerebral cortex to coordinate the process of imagining and describing the internal state of another person, interactively finding, checking, and monitoring the creation of the TOM story. It is of interest that a growing body of literature has suggested an abnormal cerebellar structure in autism, at both the cellular and anatomical levels (61–64). Impaired TOM, in fact, seems to be a hallmark of this neurodevelopmental illness (23).

Frith and his team (12) have made extensive contributions to the TOM literature, both experimentally and theoretically. They suggested that the brain may possess a “theory of mind system,” just as it possesses other systems that have become increasingly well-mapped, such as the language system, the facial recognition system, or the object recognition system. On the basis of their experimental work and studies of brain activity in higher primates, they proposed that the TOM system is composed of three ma-

ajor nodes: medial prefrontal, superior temporal sulcus, and inferior frontal. They proposed that the medial prefrontal node maintains representations of the mental state of the self, that the superior temporal sulcus detects the behavior of agents and analyzes the goals and outcomes of this behavior, and that the inferior frontal region maintains representations of actions and goals (12).

The present study, especially when examined in the context of other recent studies, adds another perspective to the nature of the TOM system in the human brain. Instead of attempting to eliminate the contributions of episodic memory and language, it includes them in the design, based on the recognition that most efforts to attribute mental states to other people depend heavily on these forms of mental activity. Two competing mechanisms have been proposed for our ability to mentalize. The simulation theory suggests that we draw from our repertoire of past experiences when we attempt to imagine the experiences of others, while the theory theory suggests that we formulate an abstract representation or set of rules (theory) and then apply that model/theory to recreate or imagine the experience of others (2, 26, 65, 66). Both of these mechanisms inevitably must conduct the simulation or formulate the theory using language, and both must also draw on personal past experience (episodic memory). Therefore, our approach suggests that both language and episodic memory should be considered as components of the TOM system.

The TOM network may consist of a core system and “supportive” components. Lesion studies, for example, have shown some brain areas to be essential for TOM processing (2, 67, 68). Medial frontal lesions, particularly right ventral lesions, impaired detection of deception in the study by Stuss et al. (2), for example, suggesting that this brain area is part of the core system. As for language, there is some evidence that developing normal language skills is also important for the acquisition of TOM (20). Impaired grammar in adulthood, however, does not seem to affect performance on mentalizing tasks (38). This finding suggests that although language centers might contribute significantly to normal TOM processing, some components (e.g., the phonological loop) might not be as essential, while others (e.g., semantic components) may be. In this study we have identified a TOM system that is widely distributed and that is made up of interactive nodes in frontal and anterior cingulate regions, associative and memory regions in the angular gyrus and anterior temporal lobe, and the cerebellum. Many of these regions have also been found to be active during mentalizing in other TOM imaging studies.

Regardless of how the TOM system is conceptualized, the replication of so many findings across multiple studies suggests that the tools of functional imaging can indeed be used to study complex mental activities. Using the mind to create scientific studies of how one human mind can un-

derstand another has become a realizable goal. Freud's project for a scientific psychology is now well under way.

APPENDIX 1. Examples of Rating of Utterances Based on Level of Occurrence of Attributions of Mental States to Self or Others in the Theory of Mind Task

Level 0: no attributions of mental states to self or others

She said her husband was killed in a car wreck a couple of years ago.

It was true her husband was having an affair.

She doesn't really say anything.

I went to the park today and sat down on a park bench next to a woman.

Level I: one level of attribution of mental states to self or others

She was crying because she missed him.

I tried to comfort her the best I could.

She makes me really sad, too, because I like cats a lot.

She didn't know what to do so she thought maybe she could walk around.

Level II: at least two levels of attributions of mental states to self or others

It is really sad about her cat, but it is nice in a way, too, to know that she felt so strongly about it.

He was crying because he told me the family doesn't know he absolutely hates his job.

He just didn't know what to do and couldn't convince them that he really wanted to stay at home.

Received May 15, 2002; revision received Dec. 3, 2002; accepted April 9, 2003. From University of Iowa Health Care, Department of Psychiatry, and Mental Health-Clinical Research Center, Department of Psychiatry, Roy J. and Lucille A. Carver College of Medicine, University of Iowa Hospitals and Clinics, Iowa City, Iowa; the Mental Illness and Neuroscience Discovery (MIND) Institute, Albuquerque, N.M.; and the Department of Psychiatry, University of New Mexico, Albuquerque, N.M. Address reprint requests to Dr. Andreasen, University of Iowa Health Care, Department of Psychiatry, 2880 JPP, 200 Hawkins Dr., Iowa City, IA 52242-1057.

Supported by NIMH grants MH-40856, MH-60990, MH-19113, and MHCRC43271.

The authors thank Alicia Bales and Eugene Zeien for assistance with data acquisition and data analysis.

References

- Frith C: Brain mechanisms for "having a theory of mind." *J Psychopharmacol* 1996; 10:9-15
- Stuss DT, Gallup GG Jr, Alexander MP: The frontal lobes are necessary for "theory of mind." *Brain* 2001; 124(part 2):279-286
- Heyes CM: Theory of mind in nonhuman primates. *Behav Brain Sci* 1998; 21:101-114
- Tancredi LR, Volkow ND: A theory of the mind/brain dichotomy with special reference to the contribution of positron emission tomography. *Perspect Biol Med* 1992; 35:549-571
- Premack D, Woodruff G: Does the chimpanzee have a theory of mind? *Behav Brain Sci* 1978; 1:515-526
- Dennett DC: Beliefs about beliefs. *Behav Brain Sci* 1978; 1:568-569
- Courtin C: The impact of sign language on the cognitive development of deaf children: the case of theory of mind. *J Deaf Studies and Deaf Education* 2000; 5:266-276
- Goel V, Grafman J, Sadato N, Hallett M: Modeling other minds. *Neuroreport* 1995; 6:1741-1746
- Fletcher PC, Happe F, Frith U, Baker SC, Dolan RJ, Frackowiak RS, Frith CD: Other minds in the brain: a functional imaging

study of "theory of mind" in story comprehension. *Cognition* 1995; 57:109-128

- Brunet E, Sarfati Y, Hardy-Bayle MC, Decety J: A PET investigation of the attribution of intentions with a nonverbal task. *Neuroimage* 2000; 11:157-166
- Gallagher HL, Happe F, Brunswick N, Fletcher PC, Frith U, Frith CD: Reading the mind in cartoons and stories: an fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia* 2000; 38:11-21
- Frith CD, Frith U: Interacting minds—a biological basis. *Science* 1999; 286:1692-1695
- Brothers L: Brain mechanisms of social cognition. *J Psychopharmacol* 1996; 10:2-8
- Bard KA: Imitation and mirror self-recognition may be developmental precursors to theory of mind in human and nonhuman primates. *Behav Brain Sci* 1998; 21:115
- Walker SF: Precursors to theories of mind in nonhuman brains. *Behav Brain Sci* 1998; 21:131-132
- Suddendorf T: Simpler for evolution: secondary representation in apes, children, and ancestors. *Behav Brain Sci* 1998; 21:131
- Bartsch K, Wellman HM: *Children Talk About the Mind*. New York, Oxford University Press, 1995
- Gray C, Russell P: Theory of mind in nonhuman primates: a question of language? *Behav Brain Sci* 1998; 21:121
- Mesulam MM: From sensation to cognition. *Brain* 1998; 121(part 6):1013-1052
- Nelson K: The psychological and social origins of autobiographical memory. *Psychol Med* 1993; 4:7-14
- Parker A: Primate cognitive neuroscience: what are the useful questions? *Behav Brain Sci* 1998; 21:128
- Andreasen NC, Flaum M, Arndt S: The Comprehensive Assessment of Symptoms and History (CASH): an instrument for assessing diagnosis and psychopathology. *Arch Gen Psychiatry* 1992; 49:615-623
- Baron-Cohen S: The autistic child's theory of mind: a case of specific developmental delay. *J Child Psychol Psychiatry* 1989; 30:285-297
- Marschark M, Green V, Hindmarsh G, Walker S: Understanding theory of mind in children who are deaf. *J Child Psychol Psychiatry* 2000; 41:1067-1073
- Shatz M, Wellman HM, Silber S: The acquisition of mental verbs: a systematic investigation of the first reference to mental state. *Cognition* 1983; 14:301-321
- Wellman HM: *The Child's Theory of Mind*. Cambridge, Mass, MIT Press, 1990
- Moore C, Furrow D, Chiasson L, Patriquin M: Developmental relationships between production and comprehension of mental terms. *First Language* 1994; 14:1-17
- Brown JR, Donelan-McCall N, Dunn J: Why talk about mental states? the significance of children's conversations with friends, siblings, and mothers. *Child Dev* 1996; 67:836-849
- Andreasen NC, O'Leary DS, Arndt S, Cizadlo T, Hurtig R, Rezaei K, Watkins GL, Boles Ponto LL, Hichwa RD: Short-term and long-term verbal memory: a positron emission tomography study. *Proc Natl Acad Sci USA* 1995; 92:5111-5115
- Hurtig RR, Hichwa RD, O'Leary DS, Boles Ponto LL, Narayana S, Watkins GL, Andreasen NC: Effects of timing and duration of cognitive activation in [¹⁵O]water PET studies. *J Cereb Blood Flow Metab* 1994; 14:423-430
- Raichle ME, Martin WR, Herscovitch P, Mintun MA, Markham J: Brain blood flow measured with intravenous H₂(¹⁵O), II: implementation and validation. *J Nucl Med* 1983; 24:790-798
- Andreasen NC, Cohen G, Harris G, Cizadlo T, Parkkinen J, Rezaei K, Swayze VW II: Image processing for the study of brain structure and function: problems and programs. *J Neuropsychiatry Clin Neurosci* 1992; 4:125-133

33. Andreasen NC, Cizadlo T, Harris G, Swayze V II, O'Leary DS, Cohen G, Ehrhardt J, Yuh WT: Voxel processing techniques for the antemortem study of neuroanatomy and neuropathology using magnetic resonance imaging. *J Neuropsychiatry Clin Neurosci* 1993; 5:121-130
34. Andreasen NC, Arndt S, Swayze V II, Cizadlo T, Flaum M, O'Leary D, Ehrhardt JC, Yuh WTC: Thalamic abnormalities in schizophrenia visualized through magnetic resonance image averaging. *Science* 1994; 266:294-298
35. Talairach J, Tournoux P: Co-Planar Stereotaxic Atlas of the Human Brain: Three-Dimensional Proportional System. New York, Thieme Medical, 1988
36. Cizadlo T, Andreasen NC, Zeien G, Rajaprabhakaran R, Harris G, O'Leary D, Swayze V, Arndt S, Hichwa R, Ehrhardt J, Yuh WTC: Image registration issues in the analysis of multiple-injection ¹⁵O H₂O PET studies: BRAINFIT. *Proceedings of the Society of Photo-Optical Instrumentation Engineers* 1994; 2168:423-430
37. Worsley KJ, Evans AC, Marrett S, Neelin P: A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab* 1992; 12:900-918
38. Varley R, Siegal M: Evidence for cognition without grammar from causal reasoning and "theory of mind" in an agrammatic aphasic patient. *Curr Biol* 2000; 10:723-726
39. Allman JM, Hakeem A, Erwin JM, Nimchinsky E, Hof P: The anterior cingulate cortex: the evolution of an interface between emotion and cognition. *Ann NY Acad Sci* 2001; 935:107-117
40. Devinsky O, Morrell MJ, Vogt BA: Contributions of anterior cingulate cortex to behaviour. *Brain* 1995; 118:279-230
41. Vogt BA, Nimchinsky EA, Vogt LJ, Hof PR: Human cingulate cortex: surface features, flat maps, and cytoarchitecture. *J Comp Neurol* 1995; 359:490-506
42. Andreasen NC, O'Leary DS, Cizadlo T, Arndt S, Rezai K, Watkins GL, Boles Ponto LL, Hichwa RD: Remembering the past: two facets of episodic memory explored with positron emission tomography. *Am J Psychiatry* 1995; 152:1576-1585
43. Maddock RJ: The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends Neurosci* 1999; 22:310-316
44. Vogt BA, Absher JR, Bush G: Human retrosplenial cortex: where is it and is it involved in emotion? *Trends Neurosci* 2000; 23: 195-197
45. Andreasen NC, O'Leary DS, Arndt S, Cizadlo T, Rezai K, Watkins GL, Boles Ponto LL, Hichwa RD: PET studies of memory, I: novel and practiced free recall of complex narratives. *Neuroimage* 1995; 2:284-295
46. Andreasen NC, O'Leary DS, Cizadlo T, Arndt S, Rezai K, Watkins GL, Boles Ponto LL, Hichwa RD: PET studies of memory, II: novel versus practiced free recall of word lists. *Neuroimage* 1995; 2:296-305
47. Akshoomoff NA, Courchesne E: A new role for the cerebellum in cognitive operations. *Behav Neurosci* 1992; 106:731-778
48. Allen G, Buxton RB, Wong EC, Courchesne E: Attentional activation of the cerebellum independent of motor involvement. *Science* 1997; 275:1940-1943
49. Andreasen NC, O'Leary DS, Paradiso S, Cizadlo T, Arndt S, Watkins GL, Ponto LL, Hichwa RD: The cerebellum plays a role in conscious episodic memory retrieval. *Hum Brain Mapp* 1999; 8:226-234
50. Fiez JA, Petersen SE, Cheney MK, Raichle ME: Impaired non-motor learning and error detection associated with cerebellar damage: a single case study. *Brain* 1992; 115(part 1):155-178
51. Gao JH, Parsons LM, Bower JM, Xiong J, Li J, Fox PT: Cerebellum implicated in sensory acquisition and discrimination rather than motor control. *Science* 1996; 272:545-547
52. Ito M: Movement and thought: identical control mechanisms by the cerebellum. *Trends Neurosci* 1993; 16:448-450
53. Keele SW, Ivry R: Does the cerebellum provide a common computation for diverse tasks? a timing hypothesis. *Ann NY Acad Sci* 1990; 608:179-207
54. Kim SG, Ugurbil K, Strick PL: Activation of a cerebellar output nucleus during cognitive processing. *Science* 1994; 265:949-951
55. Leiner HC, Leiner AL, Dow RS: Reappraising the cerebellum: what does the hindbrain contribute to the forebrain? *Behav Neurosci* 1989; 103:998-1008
56. Middleton FA, Strick PL: Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science* 1994; 266:458-461
57. Schmahmann JD: An emerging concept: the cerebellar contribution to higher function. *Arch Neurol* 1991; 48:1178-1187
58. Schmahmann JD, Pandya DN: Prefrontal cortex projections to the basilar pons in rhesus monkey: implications for the cerebellar contribution to higher function. *Neurosci Lett* 1995; 199: 175-178
59. Papathanassiou D, Etard O, Mellet E, Zago L, Mazoyer B, Tzourio-Mazoyer N: A common language network for comprehension and production: a contribution to the definition of language epicenters with PET. *Neuroimage* 2000; 11:347-357
60. Gabrieli JD, Poldrack RA, Desmond JE: The role of left prefrontal cortex in language and memory. *Proc Natl Acad Sci USA* 1998; 95:906-913
61. Carper RA, Courchesne E: Inverse correlation between frontal lobe and cerebellum sizes in children with autism. *Brain* 2000; 123(part 4):836-844
62. Hardan AY, Minshew NJ, Harenski K, Keshavan MS: Posterior fossa magnetic resonance imaging in autism. *J Am Acad Child Adolesc Psychiatry* 2001; 40:666-672
63. Marazziti D, Nardi I, Cassano GB: Autism, serotonin and cerebellum: might there be a connection? *CNS Spectrums* 1999; 3: 80-82
64. Muratori F, Cesari A, Casella C: Autism and cerebellum: an unusual finding with MRI. *Panminerva Med* 2001; 43:311-315
65. Adolphs R: Social cognition and the human brain. *Trends Cogn Sci* 1999; 3:469-479
66. Stich S, Nichols S: Folk psychology: simulation or tacit theory? *Mind & Language* 1992; 7:35-71
67. Siegal M, Varley R: Neural systems involved in "theory of mind." *Nat Rev Neurosci* 2002; 3:463-471
68. Winner E, Brownell H, Happe F, Blum A, Pincus D: Distinguishing lies from jokes: theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain Lang* 1998; 62:89-106