Neural Correlates of Detecting Pretense: Automatic Engagement of the Intentional Stance under Covert Conditions

Tim P. German, Jeffrey L. Niehaus, Meghan P. Roarty, Barry Giesbrecht, and Michael B. Miller

Abstract

■ Typically developing children begin to produce and understand pretend play between 18 and 24 months of age, and early pretense has been argued to be a candidate "core" capacity central to the deployment of representations of other peoples' mental states—"theory of mind." In a functional magnetic resonance imaging study, 16 healthy adult volunteers were imaged while watching short (5 sec) clips of actors who either performed simple everyday actions or pretended to perform a similar set of actions, under covert conditions (e.g., participants were not directed to attend to actors' mental states). There was increased activity in the medial prefrontal areas (Brodmann's areas [BA] 9/6/32, 9, and 10), inferior frontal gyrus bilaterally (BA 44, 47), temporo-parietal regions (BA 21 and 22), and parahippocampal areas, including the amygdala, when subjects viewed pretend actions as compared with real actions. This result suggests that at least some areas previously implicated in making explicit mental state judgments are also strongly activated in response to actions that call for mental state interpretation (e.g., pretense) even when there is no explicit instruction for "mind reading." This outcome is discussed in terms of accounts that propose "theory of mind" to be underwritten by automatic specialized mechanisms for the interpretation of the behavior of social agents. ■

INTRODUCTION

Social behavior rests on a foundational capacity, perhaps uniquely human, to understand the actions, reactions, and interactions of other social agents in terms of their mental states. This capacity is variously known as "theory of mind," commonsense psychology, belief–desire reasoning, mind reading, and as what Dennett (1987) termed "taking the intentional stance." Theory of mind appears to operate automatically, without effort, and in the typical case, develops rapidly over the first few years of life, without formal teaching, and largely independently of IQ (Leslie, 1994a). These facts suggest to many cognitive scientists that its acquisition and functioning is based on a reliably developing and perhaps modular cognitive mechanism (Scholl & Leslie, 1999; Leslie, 1994b, 2000b). Support for this view is found in the case of childhood autism, where there appears to be a specific cognitive deficit in mental state reasoning, which results in social, imaginative, and communicative deficits, alongside relatively spared abilities in other domains (Baron-Cohen, 1995; Leslie & Thaiss, 1992; Baron-Cohen, Leslie, & Frith, 1985).

Attention in cognitive neuroscience has more recently focused on identifying the brain areas that underwrite the capacity to deploy a theory of mind, utilizing methods of functional imaging including positron emission tomography (PET; Ruby & Decety, 2003; Gallagher, Jack, Roepstorff, & Frith, 2002; Brunet, Sarfarti, Hardy-Baylé, & Decety, 2000; Castelli, Happé, Frith, & Frith, 2000; Fletcher et al., 1995; Goel, Grafman, Sadato, & Hallet, 1995), functional magnetic resonance imaging (fMRI; Mason, Banfield, & Macrae, 2004; Ramnani & Miall, 2004; Saxe & Kanwisher, 2003; Ferstl & von Cramon, 2002; McCabe, Houser, Ryan, Smith, & Trouard, 2001; Vogeley et al., 2001; Gallagher et al., 2000; Baron-Cohen et al., 1999), and event-related potentials (ERP; Sabbagh & Taylor, 2000), as well as investigations based on patients with acquired or developmental brain lesions (Fine, Lumsden, & Blair, 2001; Lough, Gregory, & Hodges, 2001; Rowe, Bullock, Polkey, & Morris, 2001; Stuss, Gallup, & Alexander, 2001; Channon & Crawford, 2000; Stone, Baron-Cohen, & Knight, 1998).

This body of work, although deploying a range of tasks, populations, and methods, has converged on the identification of several brain regions that are thought to underlie theory of mind, either because they are commonly activated when subjects are required to engage in explicit reasoning involving the mental states of

University of California

other people (e.g., see Saxe, Carey, & Kanwisher, 2004; Gallagher & Frith, 2003, for reviews) or because damage to those areas results in deficits in social cognitive function (e.g., Rowe et al., 2001; Stuss et al., 2001; Stone et al., 1998). The areas most commonly implicated include areas of the medial prefrontal cortex (Brodmann's areas [BA] 9/32), areas of the posterior superior temporal gyrus at the junction with the parietal cortex, and anterior lateral temporal regions (e.g., temporal poles). Areas also implicated, although less consistently, include the orbito-frontal cortex (Baron-Cohen, Ring, Moriarty, & Schmitz, 1994) and the amygdala (Grèzes, Frith, & Passingham, 2004; Stone, Baron-Cohen, Calder, Keane, & Young, 2003; Fine et al., 2001; Baron-Cohen et al., 1999).

The activations seen in imaging studies of theory of mind (some of which are summarized in Table 1) indicate candidate brain areas that are involved in overt mental state judgment. In these studies, theory of mind activity is typically isolated by subtracting activity from control tasks designed to parallel mental state reasoning tasks in all aspects save the critical requirement to consider the mental states of others. Ideally, these tasks form a "minimal pair," differing only in mental state versus control content. For example, Fletcher et al. (1995) explicitly instructed subjects to solve story-based tasks by reasoning about the mental states of the protagonists and compared activations in that task to activations in control tasks where subjects were also explicitly directed to avoid considering mental states. In the study by Castelli et al. (2000), subjects were explicitly informed of three different types of animation they would see, including information that some would seem "as if [the triangles] were taking into account their reciprocal feelings and thoughts" (Castelli et al., 2000, p. 322), before being asked to view the animations.

Importantly, activations in some of the areas associated with theory of mind content are also seen in cases where the only difference between conditions is the

Table 1. Representative Medial Frontal Activations in Mental State Reasoning Tasks

| Task | Study | BA | Coordinates | | |
			x	y	z
Theory of mind story comprehension (vs. control stories)	Fletcher et al. (1995)	8	−12	42	40
Judging others knowledge about artifact function (vs. simple inference about function)	Goel et al. (1995)	9	−12	38	32
Judging mental states from eye region (vs. judging sex)	Baron-Cohen et al. (1999)	24	−8	36	0
		32	0	44	14
Theory of mind story and cartoon comprehension (vs. control stories/cartoons)	Gallagher et al. (2000)	10	−10	48	12
Explaining complex motion patterns of moving triangles (vs. random motion)	Castelli et al. (2000)	9	−4	60	32
Picture sequence completion (intention sequences vs. physical causality sequences)	Brunet et al. (2000)	8	4	56	44
Theory of mind story comprehension (self and other content) vs. control stories	Vogeley et al. (2001)	10	6	56	2
On-line cooperation with human (vs. computer opponent)	McCabe et al. (2001)	10	5	52	10
On-line strategic game with human (vs. computer opponent)	Gallagher et al. (2002)	9	−10	50	30
		10	8	54	12
Judging coherence in theory of mind sentence pairs (vs. language control task)	Ferstl and von Cramon (2002)	9	−19	49	30
Semantic judgments about persons vs. about objects	Mitchell et al. (2002)		0	54	21
			3	39	0
First- vs. third-person judgments	Ruby and Decety (2003)		0	20	70
Judging actions as performable by humans vs. by dogs	Mason et al. (2004)	8/9	10	48	34
Predicting action in person vs. computer	Ramnami and Miall (2004)		−8	56	24

need to take an explicit ''intentional stance'' (Gallagher et al., 2002). In this PET study, participants played a strategic game (paper, scissors, stone) either against what they thought was a human opponent or against what they thought was a computer following a range of known strategies. The human − computer subtraction was conducted to calculate the difference in activity in response to the same series of random trials inserted into the human or computer sequence, thus isolating the intentional stance (i.e., the requirement to reason, explicitly, about mental states) from associated features of mental state reasoning (e.g., detecting social agents, biological motion, mental state language, etc.). Gallagher et al. (2002) identified only the anterior paracingulate gyrus area as active when subjects thought they were playing a human versus a computer opponent, leading them to characterize this area (BA 9/32) as uniquely involved when participants take the intentional stance, ''even when there are no additional verbal or visual cues to assist mentalizing'' (2002, p. 819). This is in contrast to the other putative areas identified in off-line tasks (e.g., areas of the temporo-parietal cortex, the temporal poles) that are activated by ''signals important for mentalizing even when the subject is not adopting an intentional stance'' (ibid., p. 819).

In the current study, we address the question of whether areas associated with the deployment of mental state reasoning are activated in a task where no explicit judgments about mental states are made. Instead, we present adult subjects with stimuli in which we hypothesize that the intentional stance will be automatically engaged—cases where an actor performs an act of pretense. The detection and interpretation of pretense is one of the earliest documented theory of mind capacities emerging in typical development between 18 and 24 months (Leslie, 1987, 1994b, 2000a), and because children this age are too young to have much explicit knowledge about pretense nor require any kind of instruction to reason about mental states, it is assumed to be based on mechanisms that automatically create mental state representations for the child from the input of mother's behavior (German & Leslie, 2000, 2001; Leslie, 2000b). Thus, we predict that pretend acts will provoke the engagement of the intentional stance in the absence of overt instructions to consider the mental states of the actor.

Note that while a number of the prior studies have avoided explicitly instructing or cuing the participants to think about mental states (e.g., Mason et al., 2004; Saxe & Kanwisher, 2003; Brunet et al., 2000), these tasks do nevertheless typically involve an explicit judgment about human action. For example, Saxe and Kanwisher (2003) required participants to judge the content of a character's belief, while Mason et al. (2004) asked participants to decide if particular actions could be performed by humans or not. Our aim here is to assess whether brain areas associated with mental state reasoning can be activated where the task does not involve any explicit judgment about peoples' actions at all. Indeed, in the current task, participants are directed toward making judgments about entirely different aspects of the stimuli.

Interestingly, studies have been reported that report medial prefrontal cortex activations associated with tasks where, arguably, no explicit action or mental state judgments were required. Mitchell, Heatherton, and Macrae (2002) showed activations in medial prefrontal cortex in a task requiring individuals to judge whether certain adjectives applied to people versus objects. Although no explicit action judgment tasks were made, making this arguably a covert task, the task still required the explicit consideration of mental states, especially in the cases of certain adjectives (e.g., devious, anxious). Another candidate ''covert'' task was reported by Calder et al. (2002) who identified medial prefrontal activations associated with the processing of eye-direction information in a task where no explicit instructions to consider mental states were given. Finally, Singer et al. (2004) showed that areas of the anterior cingulate are activated during an ''empathy'' task in which female participants witnessed their partner receiving a painful stimulus, under no instructions to consider their beliefs or feelings.

In the current study, 16 healthy volunteers were imaged via fMRI as they watched short video clips of actors performing simple actions in short blocks (e.g., Figure 1). Half the actions involved the actors really performing the act in question (e.g., pouring tea into a cup, reaching a book from a shelf, etc.). The other half involved the actors pretending to perform such actions. Participants were initially asked to perform a cover task to ensure attention to the entire clips, which consisted in deciding whether the film clip had been prematurely edited. To allow for this, half the real and half the pretend clips were filmed in such a way that the action was interrupted by blue screen before the conclusion of the event. After each clip, the subjects were asked to respond via a button press as to whether the video clip was complete. To our knowledge, no other study has previously attempted to determine whether areas associated with mental state reasoning (such as medial prefrontal cortex) are activated in response to displays involving specific actions of social agents under covert conditions (i.e., conditions that call for no explicit mental state or action judgment).

RESULTS

Behavioral Data

All 16 subjects were appropriately engaged in the task (i.e., ''was the film complete''). The average correct response across subjects for the task was 88.31%, with a low of 67% and a high of 98%. It should be noted that

Figure 1. Example of actions viewed by subjects in the experiment. The figure shows the real action (putting a book on a shelf) and the corresponding pretend version. The dark line indicates the approximate point at which incomplete clips would end. Note that subjects would see only one example of each action being performed by a given actor.

errors were distributed among the clips across the scan and were not confined to the first clips in a block. Some clips were somewhat ambiguous, and thus errors were expected, and it was unlikely that the sequences of three clips of a given type were detectable by subjects.

Brain Activity Associated with Pretense

The critical comparison in this study was between those brain regions that were more active for viewing pretend actions than for viewing real actions in a covert condition. Table 2 and Figure 2 illustrate significant activations across the 16 subjects using a random effects model and a statistical threshold of $p < .005$ (uncorrected) and a voxel extent of 30 (see Methods). We chose to use a threshold that would reveal any regions that might be sensitive to the covertly pretense condition while still maintaining a statistical threshold and voxel extent that would survive a corrected probability criterion for clusters (Forman et al., 1995; Xiong, Gao, Lancaster, & Fox, 1995). Three regions in the medial prefrontal cortex were significantly active, including an anterior region in the superior frontal gyrus (BA 9), an area posterior to that encompassing the anterior cingulate and the medial frontal gyri (BA 32/6/9), and an area inferior to that in the medial frontal gyrus (BA 10). Several regions were also significantly active in the right and left ventrolateral prefrontal cortex, including the right and left inferior frontal gyrus or operculum (BA 47) and more superior regions in the right and left inferior frontal gyrus (BA 44). Two regions in the temporo-parietal region were active, the left middle temporal gyrus (BA 21) and the right superior temporal gyrus (BA 22). The right fusi-

form gyrus (BA 37), the left amygdala, and the right parahippocampal gyrus were also active. These regions have often been implicated in previous studies involving "theory of mind" tasks that have relied on explicit measures.

An activation in one of the critical regions mentioned above, the medial frontal gyrus (BA 10), represented the difference between two deactivations relative to baseline. Figure 3 illustrates an analysis of the percent signal change from baseline at the local maxima. In this case, the average signal intensity across the block for both the pretend and real conditions was less than the signal intensity during the baseline state (i.e., viewing the blue screen). In contrast, the average signal intensity in the anterior cingulate for both the pretend and real conditions was greater than the baseline state. Yet, in both cases, the signal intensity associated with the pretend condition was significantly greater than the signal intensity associated with the real condition. Although little is known about the difference between two activations versus the difference between two deactivations, it has been suggested that in the case of two deactivations, the condition closer to the baseline state may represent mental operations that overlap with the default processing mode of the brain (Gusnard, Akbudak, Shulman, & Raichle, 2001). Several researchers have linked various cognitive processes to this overlap, including person knowledge (Mitchell et al., 2002), self-referencing (Kelley et al., 2002), and coherence processing (Ferstl & von Cramon, 2002). It should be noted, however, that other regions in the medial prefrontal cortex that are associated with pretense in our study do not reflect the difference between two deactivations.

Table 2. Regions Based on Peak Activations of Increased Activity Associated with Viewing Pretend Actions Compared with Viewing Real Actions in the Covert Condition

	BA	x	y	z	Voxels	z Score
Anterior cingulate/ medial frontal gyri	32/6/9	6	31	32	102	3.87
Superior frontal gyrus	9	−9	57	25	140	3.58
Medial frontal gyrus	10	−3	52	−3	51	3.87
R inferior frontal gyrus	47	33	20	−11	335	4.26
R inferior frontal gyrus	44	56	18	10	54	3.63
L inferior frontal gyrus	47	−39	14	−13	238	3.65
L inferior frontal gyrus	44	−48	10	16	31	3.18
L middle frontal gyrus	10	−30	53	6	39	3.44
L middle frontal gyrus	11	−45	40	−12	32	3.36
L middle temporal gyrus	21	−50	−52	3	63	3.56
R superior temporal gyrus	22	59	−58	14	56	3.66
R inferior temporal gyrus	19	48	−70	−2	130	3.56
R fusiform gyrus	37	42	−53	−12	33	3.35
L amygdala		−18	−4	−12	33	3.35
R parahippocampal gyrus		21	−10	−22	36	3.71
R thalamus		21	−11	12	131	3.92
Caudate		−3	9	9	30	3.09
Cerebellum		12	−60	−40	195	3.83

R = right; L = left; BA = Brodmann's area; x, y, and z are in Talairach coordinates; all activations that survived a statistical threshold of $p < .005$ (uncorrected) and a voxel extent of 30.

The opposite contrast, comparing brain regions that were more active for viewing real actions than for viewing pretend actions in the covert condition, produced just one significant activation in the group analysis (using a random effects model and a statistical threshold of $p < .005$, uncorrected, and a voxel extent of 30, as illustrated in Figure 2). That activation was in the right precentral gyrus (BA 6).

Brain Activity Associated with Pretend Actions with Real Objects in the Covert Condition

We were concerned that activations associated with viewing pretend actions may only appear if the pretend action included an object that was missing from the scene. So half of the pretend actions included the object used in the real action, while the other half of the

pretend actions did not include this object. As illustrated in Figure 4, analyzing only those pretend actions that included a real object in comparison with the real actions revealed significant activations in the group analysis ($p < .05$, uncorrected, and a voxel extent of 10) in all the same regions implicated when collapsing across conditions with objects and those without with the exception of the activations in the medial temporal lobe (left amygdala and right parahippocampus).

DISCUSSION

Areas of the brain typically associated with mental state reasoning under explicit and overt instructions include the (1) medial prefrontal cortex, (2) areas of the posterior superior temporal gyrus, and (3) the temporal poles (Gallagher & Frith, 2003). In the current study, a covert task was employed to determine whether theory of mind areas are responsive to theory of mind content in the absence of any instruction to reason about the mental states of social agents or any requirement to make a judgment about the actions of social agents in the scene. The results showed increased activations in areas of medial prefrontal and inferior prefrontal cortex and in areas of the temporo-parietal junction in response to pretend versus real actions, consistent with previous imaging studies. There was no evidence in this study of any anterior lateral temporal lobe (temporal pole) activations. In addition, there were further notable activations in bilateral medial temporal lobe regions (left amygdala and right parahippocampus). We discuss this latter activation pattern first, before moving to discuss the areas more typically seen in theory of mind neuroimaging studies.

The role of the amygdala in mental state reasoning has been a matter of some debate. Damage to the amygdala in various primates can cause disturbance of social behavior (Kling, 1972), and humans with bilateral amygdala damage have problems in emotion recognition from facial expressions and social judgments (Adolphs, Tranel, & Damasio, 1998), as well as problems in explicit theory of mind tasks involving judging the mental states of other people (Stone et al., 2003; Fine et al., 2001). Further, the amygdala may be a critical structure that is damaged in autism (Abell et al., 1999; Aylward et al., 1999), resulting in social cognition failures forming the core of the disorder (Baron-Cohen et al., 1999).

However, other theorists suggest that the amygdala may play a limited role in social cognition, citing data from mature and neonatal rhesus monkeys in which the amygdala was selectively and completely lesioned (Amaral et al., 2003). Despite some social and emotional disruptions in these animals in the evaluation of threats in the environment, the mature animals demonstrated normal to increased levels of social interaction with conspecifics and neonatal lesioned animals engaged in
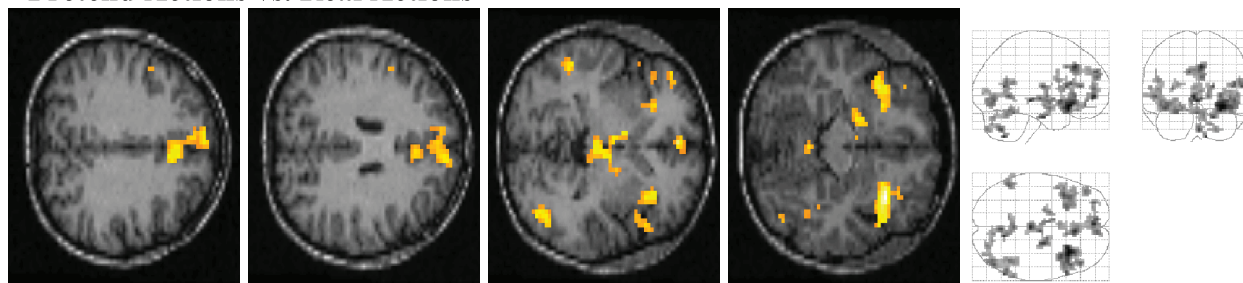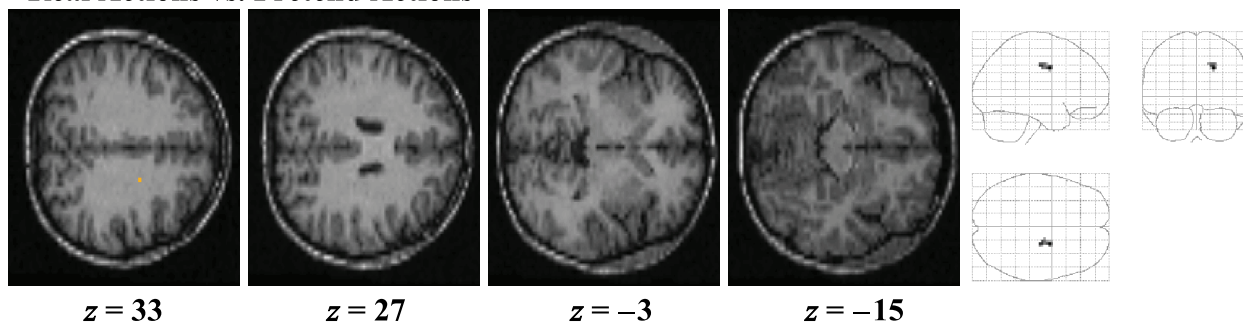
Figure 2. Four axial sections for two separate contrasts show significant activations that exceeded the statistical threshold. Functional images are superimposed on one subject's high-resolution anatomical image with the left side of the image corresponding to the left side of the brain. Adjacent to the axial sections are the glass brain representations revealing all the significant activations for that particular contrast. Viewing pretend actions relative to viewing real actions significantly increased activations in the medial prefrontal cortex, ventrolateral prefrontal cortex, medial temporal lobe, and in the temporo-parietal junction, whereas viewing real actions relative to pretend actions only produced a single activation in the right precentral gyrus.

typical social behaviors (e.g., facial expressions, grooming, and play) during development.

Increased activity in the amygdala has only rarely been reported in neuroimaging studies of mental state reasoning (Gallagher & Frith, 2003), although one possible reason might be the relative difficulty in demonstrating
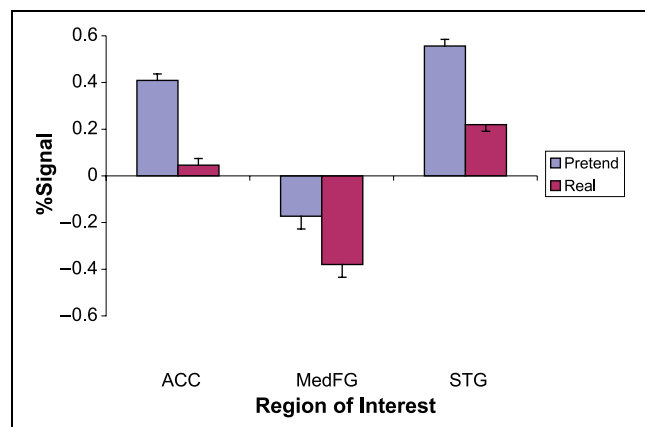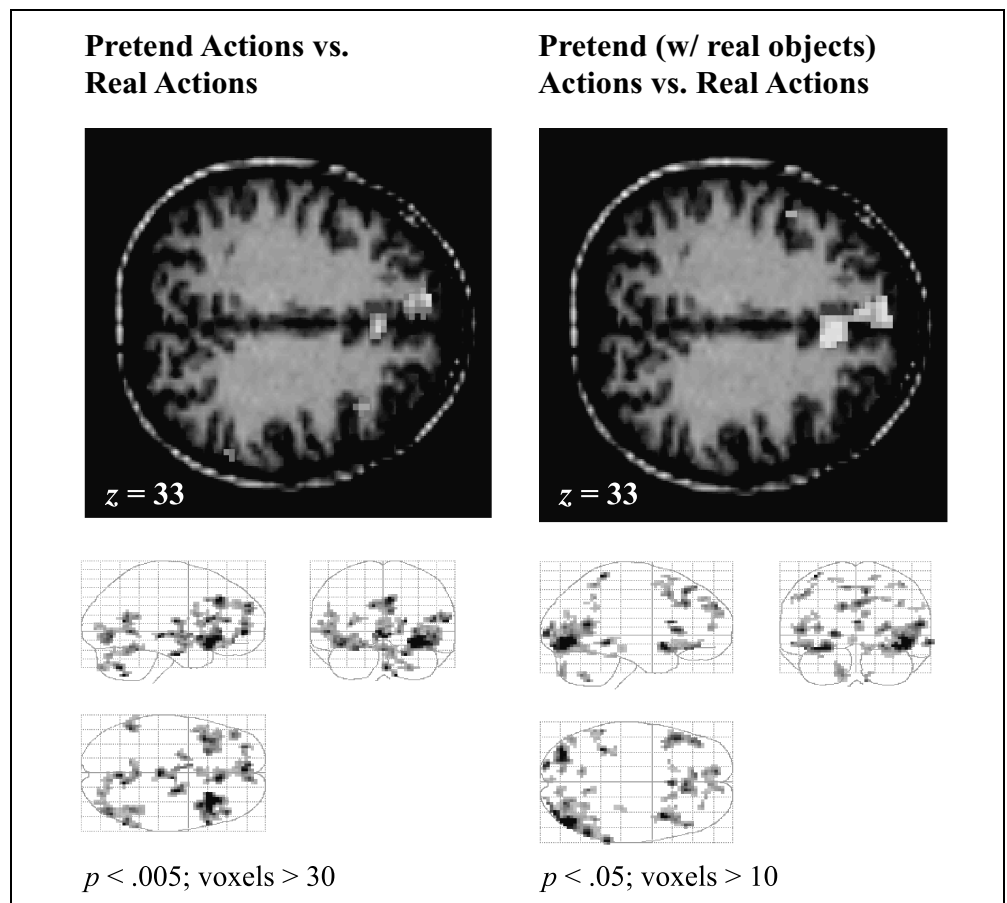

Figure 3. Peak signal change relative to the resting baseline at the local maxima of three key regions (ACC = anterior cingulate; MeFG = medial frontal gyrus; STG = superior temporal gyrus) for the pretend and real conditions.

activations in medial temporal lobe structures, as has been discussed in the case of hippocampus activations in memory encoding and retrieval (Schacter & Wagner, 1999). Note that two neuroimaging studies that have reported amygdala activation both involve the explicit calculation of affective mental states from the eye region of faces (Wicker, Perrett, Baron-Cohen, & Decety, 2003; Baron-Cohen et al., 1999), rendering uncertain the role the amygdala plays beyond mediating responses to the emotional content of the stimuli. This interpretation is also offered to explain amygdala activation demonstrated in a recent study looking at brain regions sensitive to calculating the deceptive intent of an actor (Grèzes et al., 2004). Note also that the amygdala and parahippocampal gyrus activity in the current study did not depend on processing information from facial expressions, nor on processing any other emotion information; the actors in both real and pretend video clips in the current study maintained neutral expressions throughout the actions. Thus, the current results provide qualified support for the possible role of the amygdala in supporting mental state calculations (Brothers, 1990).

The current study also identified areas of the brain more typically associated with theory of mind content including the medial prefrontal cortex (BA 32/6/9, 9, and 10), inferior frontal cortex (BA 44 and 47), and regions

Figure 4. Pretend actions with real objects compared to real actions (middle panels; using a more liberal threshold) produced similar activations as pretend actions with both real objects and missing objects.

Pretend Actions vs. Real Actions

Pretend (w/ real objects) Actions vs. Real Actions

z = 33

z = 33

p < .005; voxels > 30

p < .05; voxels > 10

of the temporo-parietal junction (BA 21, 22). However, most theory of mind neuroimaging studies employ explicit theory of mind tasks, where judgments about the actions of social agents are confounded with instructions to think about mental states in executing the task, be it comprehension of verbal stories (Vogeley et al., 2001; Gallagher et al., 2000; Fletcher et al., 1995), understanding mental state-based humorous cartoons (Gallagher et al., 2000), completing picture sequences based on agents' intentions (Brunet et al., 2000), judging the likely knowledge of another person (Goel et al., 1995), or judging the intentions of or explaining the motion patterns of triangles engaging in complex mental state-based interactions (Castelli et al., 2000).

In one study, the explicit instruction to take the intentional stance was isolated from content difference in the signal between the experimental (theory of mind) and control conditions (Gallagher et al., 2002). In this study, subjects played a strategic game (paper, scissors, stone) against either a human opponent (theory of mind condition) or against a computer following one of several known strategies (control condition) for several trials, before being confronted with a series of the exact same sequence of trials in each condition. The only difference between conditions, therefore, was that the participants in the theory of mind condition were in-

structed to think about the beliefs of their opponent, while those in the computer control condition were not; all the properties of the signal being processed were otherwise identical. This study found increased activity in the paracingulate gyrus (BA 32/9), as reported for other explicit theory of mind tasks, but interestingly, there were no activations seen in other areas typically associated with theory of mind tasks (e.g., temporo-parietal areas or the temporal poles).

Gallagher et al. (2002) argued that this circumscribed region of medial prefrontal cortex is critically involved in explicit mental state reasoning—taking the intentional stance—even in the absence of other behavioral cues associated with social agents, such as mental state language (Fletcher et al., 1995), visual depictions of humans (Gallagher et al., 2000; Baron-Cohen et al., 1999), or cues to social agency such as biological motion (Castelli et al., 2000), which have been shown in other studies to activate areas of superior temporal cortex (Grossman & Blake, 2001; see Allison, Puce, & McCarthy, 2000, for a review). According to this analysis, although superior temporal and temporal parietal areas might be involved in some aspects of mental state reasoning (such as the detection and interpretation of goal directed actions), it is in medial prefrontal cortex regions where representations of beliefs and other

knowledge-based mental states (e.g., pretends, expectation) are calculated (Gallagher & Frith, 2003). This picture is supported by evidence suggesting that lesions to frontal areas seem to cause disruptions in overt mental state reasoning (e.g., Rowe et al., 2001; Lough et al., 2001; Stuss et al., 2001; Channon & Crawford, 2000; Stone et al., 1998).

The current study demonstrates that it is not necessary to explicitly instruct participants to take the intentional stance, as did Gallagher et al. (2002), or indeed make any kind of overt judgment about mental states (e.g., Saxe & Kanwisher, 2003; Brunet et al., 2000) to activate these medial frontal areas; they can also be engaged by mental state content in situations where the behavior of a social agent is such that a mental state interpretation is called for—a phenomenon we dub the automatic engagement of the intentional stance. Partial support for this idea is provided by a previous study that demonstrated medial prefrontal cortex activations associated with a covert task and theory of mind related content (Calder et al., 2002). In that study, increased regional cerebral blood flow was observed in medial prefrontal regions (i.e., BA 8/9, coordinates 2, 42, and 36) in response to increasing proportions of horizontally averted gaze. The authors attributed this result to the automatic engagement of processing of the possible targets of the pictured person's mental states, in cases where the goal or focus of the actor was not obvious (e.g., the face was not looking at the subject herself; see also Singer et al., 2004; Mitchell et al., 2002, for cases of theory of mind areas active in tasks with no explicit judgments about others' mental states). Note also that the medial prefrontal and temporo-parietal areas are among those identified as having high levels of resting-state metabolic activity (Gusnard et al., 2001), and this is consistent with the notion that spontaneous ongoing automatic mental state calculation may occur in these regions.

Although theory of mind is sometimes characterized as a "complex higher cognitive function and a ubiquitous element of social reasoning and problem solving" (Gallagher et al., 2002, p. 814), there is a long tradition in cognitive developmental research that has assumed that theory of mind knowledge is initially acquired by cognitive mechanisms that operate automatically over the behavior of social agents to calculate representations of the underlying mental states (see, e.g., Leslie, 1987, 1994a, 2000a, 2000b; see also German & Leslie, 2000, 2001). A motivating consideration for this view is the fact that very young children are capable of interpreting even nonliteral behavior as such from early in development—understanding pretense emerges sometime between 18 and 24 months in the typical case—long before any explicit knowledge about pretense has been acquired and long before success is achieved on the basic belief–desire reasoning problems that are argued to mark mature theory of mind skills (e.g., the

false-belief task, Wellman, Cross, & Watson, 2001; see also Bloom & German, 2000).

In a pretend situation, such as when mother talks into a banana as if it were a telephone, the child is not confused nor does she learn that bananas are an additional class of object that can be labeled by the word "telephone." Instead, she is capable of calculating mother's underlying mental state (Leslie, 1987), most likely aided by particular signals of pretense that the mother provides, which differ from those involved in real action (Lillard & Witherington, 2004). Because 18-month-old children cannot plausibly know very much about mental states explicitly (see, e.g., German & Leslie, 2001; Lillard, 1993, for discussion), Leslie (1987) proposed a computational model in which the behavior of social agents is taken as input to a specialized mechanism that automatically calculates from that behavioral description a mental state representation that makes explicit information about the event, including the agent (mother) and a mental state (e.g., pretense) that is taken to a proposition ("is a telephone") in respect to some object anchored in the real world (of the banana; Leslie, 1987, 1994a, 2000b). The proposition is "decoupled" from typical input–output relations so as to avoid confusion with "primary" representations of the real world (see Leslie, 1994b, for further discussion).

This mechanism, the "theory of mind mechanism," is proposed to underwrite the capacity for the later development of theory of mind knowledge. Early theory of mind abilities can be plausibly thought of as comprising mechanisms that allow for mental state information to be attended to, even in very young children without much explicit knowledge about mental states, and even though mental states cannot be directly seen, heard, or felt and without the benefit of any formal instructions to do so. Being able to attend to mental states is a prerequisite for learning about them (German & Leslie, 2000; Leslie, 2000a). It is our contention that the current study provides support for the continued operation of reliably developing mechanisms for the detection and representations of mental states—systems that are engaged automatically in response to certain streams of perceived behavior.

One possible alternative characterization of the results reported here might be in terms of whether the pretend films cause activations that are in response to "novelty." Arguably, participants might have seen pretend actions as more novel than real actions. There are two ways to read this alternative. One interpretation is that the pattern of brain responses identified here might be simply a result of some general reaction to novel stimuli, or stimuli that signal a novel response type are required. It has been reported previously that prefrontal ERP responses (e.g., the P3a, occurring 300–400 msec post-stimulus) are seen when participants receive cues, indicating they should change rules in the Wisconsin Card Sorting Task (e.g., Barceló, Periáòez, & Knight, 2002)

and lesions to the lateral prefrontal cortex result in an attenuated novelty response (Knight, 1984). Hippocampal regions are also implicated, given that patients with posterior hippocampal lesions show attenuated ERP responses to novelty (Knight, 1996). Functional imaging data also suggest that a test–retest-reliable lateral temporal signal in response to novel auditory stimuli can occur in both bilateral superior temporal and bilateral inferior frontal gyri (Kiehl & Liddle, 2003).

Although the current study cannot rule out that activations observed to pretense were partially caused by novelty, there are reasons to suspect that the pattern of responses is not entirely explicable that way. First, novelty has been associated with lateral rather than medial activations in frontal cortex (Kiehl & Liddle, 2003; Kirino, Belger, Goldman-Rakic, & McCarthy, 2000; Knight, 1984). Second, Ferstl and von Cramon (2002) showed that in a task requiring participants to judge the coherence of sentences, it was coherent and not incoherent sentences that resulted in medial frontal activations. Plausibly, incoherent sentences would be more likely to be construed as novel than coherent sentences.

Alternatively, perhaps, the mental state regions are engaged not because of novelty in general, but rather because of novelty within the domain of human action. Perhaps, a range of atypical human actions, including but not limited to pretense, result in the automatic engagement of mental state calculations? It is exactly according to this sense of novelty that the current predictions were made. Pretense was selected as a scenario likely to be an important behavioral cue to the calculation of mental states requiring representations decoupled from reality, given its important early emergence in infancy as a marker of mental state representation. It is not claimed here that the activations seen in this study should be thought of as caused if and only if pretense is detected, but rather that pretend actions should be one case where the intentional stance is reliably and automatically engaged, with no need for any prior instruction to attend to mental states.

It follows from this perspective that other classes of human action might also result in the engagement of mental state reasoning areas. Mitchell et al. (2002) showed theory of mind-type activations in response to judgments made about people as opposed to objects, and Singer et al. (2004) showed that anterior cingulate activations were associated with empathy toward a close friend's pain; in both cases, there was no explicit direction to consider the mental states of others. We speculate that other cases where activations might be expected will involve cases where the goal that the agent pursues is atypical (indeed, pretense might be thought of as a special case within this class, involving a "knowing" violation of a typical goal). Other kinds of actions that violate the actors' goals (e.g., where someone fails to gain an object because they are mistaken about the location of that object) might also result in theory of mind areas to be activated automatically (i.e., with no instruction to calculate the mental state content).

A final issue concerns the extent to which activations in mental state reasoning tasks such as that reported in the current study, and indeed more broadly, should be thought of as resulting from processes that have to do with the deployment of representations of mental states themselves (e.g., conceptual competence) or as resulting from processes (perhaps that have a more domain general, penetrable character) associated with performing calculations over those special representations (e.g., maintenance of representations in working memory, inhibition of alternative mental state contents, selection among those alternatives, etc.).

Distinguishing theory of mind-specific activations from associated task performance activations requires the adoption of "minimal pair" designs. For example, Saxe and Kanwisher (2003) showed that activations in temporo-parietal areas are specific to belief reasoning by controlling for task structure using a nonmental analog (e.g., the "false-photo" task, Zaitchik, 1990). This approach will be important in deciding whether other "theory of mind" areas are activated because they perform (relatively) domain general processing resources shared by belief reasoning and other tasks with similar computational structure (see also Saxe et al., 2004; Ferstl & Von Cramon, 2002).

Whereas the cognitive developmental literature has very clearly articulated the competence–performance distinction in thinking about developing performance on the false-belief task (see e.g., Bloom & German, 2000, for a review), the possible separable contributions to the activation patterns seen in theory of mind neuroimaging studies of (1) neural areas representing mental state concepts themselves (e.g., mechanisms producing mental state representations) and (2) neural areas responsible for maintaining, manipulating, and selecting between candidate representations with different content has not yet been clearly delineated (see Leslie et al., in press, for specific models of belief–desire reasoning that articulate this distinction).

Conclusions

In summary, the current study shows that a set of neural mechanisms for the interpretation of the actions of social agents are automatically engaged when a viewer is confronted by instances of social behavior that require a decoupled mental state representation to be calculated—in this case, an act of pretense. These mechanisms appear to comprise circuits in temporo-parietal regions, possibly associated with the detection of goal directed social action (e.g., Frith & Frith, 1999), and areas of the prefrontal cortex that have been assumed to reflect the explicit adoption of the "intentional

stance,'' including the calculation of representations decoupled from reality (e.g., Gallagher et al., 2002). It is our view that research on further cases where epistemic mental state calculations might be engaged by the actions of social agents without explicit experimental instruction will be an important addition to research efforts based on tasks involving explicit and overt tasks of social reasoning. Via this dual route, the functional and neural organization of this fundamental social capacity may be more rapidly elucidated.

METHODS

Subjects

Sixteen subjects between the ages of 18 and 29 were recruited from the Dartmouth College community (8 men and 8 women). No subject reported any abnormal neurological history and all had normal or corrected-to-normal visual acuity. Subjects were paid for their participation. All subjects gave informed consent in accordance with the guidelines set by the Committee for the Protection of Human Subjects at Dartmouth College and by the Human Subjects Committee at the University of California, Santa Barbara, CA.

Materials and Design

We created a series of video clips of actors performing simple acts (e.g., putting a hat on a hook, pouring a cup of tea). There were 12 different actions (see Appendix), each one performed by one of four different actors (2 men and 2 women). Each clip was filmed (on a Canon ZR 45 DV Camcorder), for each actor, both as a real action (the actor really performed the act) and as a pretend action (the actor pretended to perform the action). For half of the pretend actions, one or more objects for the act was missing (e.g., no teapot), while for the other half, all objects were present, and the pretense consisted of miming the act with the objects instead of really performing it. Actors were instructed not to make overt pretense ''displays'' to the camera (e.g., significant facial affect, overt exaggeration to the camera), because as indicated, we were interested in isolating a minimal difference in content signal between pretense and real acts as far as possible.

Four sets of clips were then created from this raw material by editing the actions, such that half the time, the clip ended before the act was finished (incomplete clips), and the rest of the time, the act finished before the clip ended. Thus, there were four kinds of clips in the study (PC = pretend complete; PI = pretend incomplete; RC = real complete, and RI = real incomplete). Each clip was approximately 6 sec long, and these were edited in iMovie (Apple Computer, Cupertino, CA) to create trials that comprised a short section of blue screen (~1 sec), followed by a video clip of one of the four types, followed by a further short section of blue screen (1 sec) before a white response screen (2 sec), reminding the participant to respond and indicating the relevant options on the appropriate side (e.g., left = complete, right = incomplete). Each trial thus lasted 10 sec. The clips varied slightly in length around a mean length of 6 sec, with time added or subtracted to the blue screen, such that the trial length was constant.

These clips were arranged into blocks of three clips for a total of 30 sec per block. Each block contained just one of the four kinds of clips. Selection was constrained, such that each action performed by each actor appeared in only one possible film type, and therefore subjects would not see the same actor performing the same clip as both pretend and real or as both complete and incomplete. The 16 blocks of clips, along with 4 rest blocks consisting entirely of blue screen, were then arranged into a random sequence and edited together before being exported to DVD (iDVD, Apple). The total length of each DVD film was 10 min and 10 sec (e.g., 20 blocks of 30 sec each plus 10 sec of blue screen at the outset of each film).

Behavioral Procedure

Subjects were instructed that they would be seeing a sequence of short video clips depicting people performing simple everyday actions. They were told to watch the clips to the end, because in some cases, the clips have been edited such that they would finish before the action was complete, but in other cases, the clips would finish only after the action was complete. They were asked to make responses with button presses on a box held in the left hand in response to complete clips and in the right hand in responses to incomplete clips, only after they saw the white response screen after each clip. Responses were recorded to determine whether participants responded appropriately.

Functional Magnetic Resonance Imaging Parameters

Imaging was performed at Dartmouth College on a 1.5-T whole-body scanner (General Electric Medical Systems, Signa, Milwaukee, WI) with a standard head coil. Visual stimuli were presented via a standard DVD player and projected to a screen positioned at the head end of the bore by an Epson (model ELP-7000) LCD projector. Subjects viewed the screen through a mirror mounted on top of the head coil. A fiber-optic, light-sensitive key-press that interfaced with the PsyScope Button Box (New Micros, Dallas, TX) was used to record subjects' responses. Cushions were used to minimize head movement.

Anatomical images were acquired using a high-resolution 3-D spoiled gradient recovery sequence

(124 sagittal slices, TE = 6 msec, TR = 2500 msec, flip angle = 25°, voxel size = 1 × 1 × 1.2 mm). Each session included two functional runs. Functional images were collected using a gradient spin-echo, echo-planar sequence sensitive to BOLD contrast (T2*, 240 repetitions, TR = 2500 msec, TE = 35 msec, flip angle = 90°, 3.75 × 3.75 mm in-plane resolution, 25 axial slices at 4.5-mm slice thickness and 1-mm skip between slices allowed for imaging of the whole brain). The first four scans of each functional run were dropped to allow for T1 saturation to stabilize. Signal loss was observed in small portions of the orbito-frontal cortex due to the sinus cavities.

Functional Magnetic Resonance Imaging Data Analyses

fMRI data were analyzed using Statistical Parametric Mapping software (SPM2, Wellcome Department of Cognitive Neurology, London, UK; Friston et al., 1995). Functional data were first corrected for differences in slice acquisition timing by resampling all slices in time to match the first slice, then realigned across the two runs to correct for head movement, coregistered with each participant's anatomical data, and then transformed into a standard anatomical space based on the ICBM 152 brain template (Montreal Neurological Institute), which approximates Talairach and Tournoux's (1988) atlas space. Images were resampled into 3-mm cubic voxels and then spatially smoothed using an 8-mm full width half maximum isotropic Gaussian kernel.

For each participant, a general linear model incorporating task effects (modeled as a box-car function convolved with the canonical hemodynamic response function) and a linear trend was used to compute t contrast images (weighted parameter estimates) for each comparison at each voxel. These individual contrast images were then submitted to a second-level, random-effects analysis to create group images. The resulting statistical maps were thresholded at $p < .005$ (uncorrected). To reduce the rate of false positives, a spatial extent threshold of 30 contiguous voxels was also applied (Poline, Worsley, Evans, & Friston, 1997; Forman et al., 1995; Xiong et al., 1995; Friston, Worsley, Frackowiak, Mazziotta, & Evans, 1994).

A peak signal change analysis was performed at the local maxima of the key regions of interest. This analysis entailed computing the percent signal intensity during the pretend, real, and resting-baseline epochs on a subject-by-subject basis, starting at 7.5 sec after the start of the epoch (i.e., allowing the hemodynamic response to peak) and extending to the end of the epoch. Peak signal relative to the resting baseline was then derived in the pretend and real conditions by subtracting the resting baseline activity.

APPENDIX

Actions for covert task
Throw paper ball into trash
Opening a jar
Putting on shoe
Getting a cup from a shelf
Eat some yogurt with a spoon
Driving a screw into wall
Blowing out a candle
Crack an egg into a pan
Peeling a potato
Lighting a match
Unscrewing a light bulb
Slice a potato

REFERENCES

Abell, F., Krams, M., Ashburner, J., Passingham, R., Friston, K., Frackowiak, R., Happé, F., Frith, C., & Frith, U. (1999). The neuroanatomy of autism: A voxel-based whole brain analysis of structural scans. *NeuroReport, 10,* 1647–1651.

Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature, 393,* 470–474.

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

Amaral, D. G., Bauman, M. D., Capitanio, J. P., Lavenex, P., Mason, W. A., Mauldin-Jourdain, M. L., & Mendoza, S. P. (2003). The amygdala: Is it an essential component of the neural network for social cognition? *Neuropsychologia, 41,* 517–522.

Aylward, E. H., Minshew, N. J., Goldstein, G., Honeycutt, N. A., Augustine, A. M., Yates, K. O., Barta, P. E., & Pearlson, G. D. (1999). MRI volumes of amygdala and hippocampus in non-mentally retarded autistic adolescents and adults. *Neurology, 53,* 2145–2150.

Barceló, F., Periáñez, J. A., & Knight, R. T. (2002). Think

differently: A brain orienting response to task novelty. *NeuroReport, 13,* 1887–1892.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind.* Cambridge: MIT Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a theory of mind? *Cognition, 21,* 37–46.

Baron-Cohen, S., Ring, H., Moriarty, J., & Schmitz, B. (1994). Recognition of mental state terms: Clinical findings in children with autism and a functional neuroimaging study of normal adults. *British Journal of Psychiatry, 165,* 640–649.

Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., & Williams, S. C. (1999). Social intelligence in the normal and autistic brain: An fMRI study. *European Journal of Neuroscience, 11,* 1891–1898.

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77,* B25–B31.

Brothers, L. (1990). The social brain: A project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience, 1,* 27–51.

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

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

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

Channon, S., & Crawford, S. (2000). The effects of anterior lesions on performance on a story comprehension test: Left anterior impairment on a theory of mind-type task. *Neuropsychologia, 38,* 1006–1017.

Dennett, D. C. (1987). *The intentional stance.* Cambridge: MIT Press.

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.

Fine, C., Lumsden, J., & Blair, R. J. R. (2001). Dissociation between "theory of mind" and executive functions in a patient with early left amygdala damage. *Brain, 124,* 287–298.

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

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine, 33,* 636–647.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. B., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2,* 189–210.

Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., & Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping, 1,* 214–220.

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 in Cognitive Sciences, 7,* 77–83.

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

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

German, T. P., & Leslie, A. M. (2000). Attending to and learning about mental states. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 229–252). Hove: Psychology Press.

German, T. P., & Leslie, A. M. (2001). Children's inferences from knowing to pretending and believing. *British Journal of Developmental Psychology, 19,* 59–83.

Goel, V., Grafman, J., Sadato, N., & Hallett, M. (1995). Modelling other minds. *NeuroReport, 6,* 1741–1746.

Grèzes, J., Frith, C. D., & Passingham, R. (2004). Brain mechanisms for inferring deceit in the actions of others. *Journal of Neuroscience, 24,* 5500–5505.

Grossman, E. D., & Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. *Vision Research, 41,* 1475–1482.

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences, U.S.A., 98,* 4259–4264.

Kelley, W. M., Macrea, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience, 14,* 785–794.

Kiehl, K. A., & Liddle, P. F. (2003). Reproducibility of the hemodynamic response to auditory oddball stimuli: A six-week test–retest study. *Human Brain Mapping, 18,* 42–52.

Kirino, E., Belger, A., Goldman-Rakic, P., & McCarthy, G. (2000). Prefrontal activation evoked by infrequent target and novel stimuli in a visual detection task: An event related functional magnetic resonance imaging study. *The Journal of Neuroscience, 20,* 6612–6618.

Kling, A. (1972). Effects of amygdalectomy on socio-affective behavior in non-human primates. In B. E. Eleftheriou (Ed.), *Neurobiology of the amygdala* (pp. 511–536). New York: Plenum.

Knight, R. T. (1984). Decreased response to novel stimuli after prefrontal lesions in man. *Electroencephalography and Clinical Neurophysiology: Evoked Potentials, 59,* 9–20.

Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection. *Nature, 383,* 256–259.

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review, 94,* 412–426.

Leslie, A. M. (1994a). Pretending and believing: Issues in the theory of ToMM. *Cognition, 50,* 211–238.

Leslie, A. M. (1994b). ToMM, ToBy, and agency: Core architecture and domain specificity. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119–148). New York: Cambridge University Press.

Leslie, A. M. (2000a). How to acquire a "representational theory of mind." In D. Sperber (Ed.), *Metarepresentations: A multidisciplinary perspective* (pp. 197–223). Oxford, UK: Oxford University Press.

Leslie, A. M. (2000b). "Theory of mind" as a mechanism of selective attention. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 1235–1247). Cambridge: MIT Press.

Leslie, A. M., German, T. P., & Pollizi, P. (in press). Belief-desire reasoning as a process of selection. *Cognitive Psychology*.

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition, 43,* 225–251.

Lillard, A. S. (1993). Young children's conceptualization of pretense: Action or mental representational state? *Child Development, 64,* 372–386.

Lillard, A. S., & Witherington, D. (2004). Mothers' behavior modifications during pretense snacks and their possible signal value for toddlers. *Developmental Psychology, 40,* 95–113.

Lough, S., Gregory, C., & Hodges, J. R. (2001). Dissociation of social cognition and executive function in frontal variant frontotemporal dementia. *Neurocase, 7,* 123–130.

Mason, M. F., Banfield, J. F., & Macrea, C. N. (2004). Thinking about actions: The neural substrates of person knowledge. *Cerebral Cortex, 14,* 209–214.

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

Mitchell., J. P., Heatherton, T. F., & Macrea, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences, U.S.A., 99,* 15238–15243.

Poline, J. B., Worsley, K. J., Evans, A. C., & Friston, K. J. (1997). Combining spatial extent and peak intensity to test for activations in functional imaging. *Neuroimage, 5,* 83–96.

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

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,* 600–616.

Ruby, P., & Decety, J. (2003). What you believe versus what you think they believe: A neuroimaging study of conceptual perspective taking. *European Journal of Neuroscience, 17,* 2475–2480.

Sabbagh, M. A., & Taylor, M. (2000). Neural correlates of the theory-of-mind reasoning: An event-related potential study. *Psychological Science, 11,* 46–50.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology, 55,* 87–124.

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

Schacter, D. L., & Wagner, A. D. (1999). Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus, 9,* 7–24.

Scholl, B. J., & Leslie, A. M. (1999). Modularity, development and "theory of mind." *Mind and Language, 14,* 131–153.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves affective but not sensory components of pain. *Science, 303,* 1157–1162.

Stone, V. E., Baron-Cohen, S., Calder, A., Keane, J., & Young, A. (2003). Acquired theory of mind impairments in individuals with bilateral amygdala lesions. *Neuropsychologia, 41,* 209–220.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10,* 640–656.

Stuss, D. T., Gallop, G. G., Jr., & Alexander, M. P. (2001). The frontal lobes are necessary for "theory of mind." *Brain, 124,* 279–286.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the brain.* New York: Thieme.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, F., Happé, F., Falkai, P., Maier, W., Shah, N. J., Fink, G. R., & Zilles, K. (2001). Mind reading: Neural mechanisms of theory of mind and self-perspective. *Neuroimage, 14,* 170–181.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72,* 655–684.

Wicker, B., Perret, D. I., Baron-Cohen, S., & Decety, J. (2003). Being the target of another's emotion: A PET study. *Neuropsychologia, 41,* 139–146.

Xiong, J., Gao, J., Lancaster, J. L., & Fox, P. T. (1995). Clustered pixels analysis for functional MRI activation studies of the human brain. *Human Brain Mapping, 3,* 287–301.

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