THE ROYAL SOCIETY

Published online 4 February 2003

Development and neurophysiology of mentalizing

Uta Frith[1] and Christopher D. Frith[2*]

[1]*Institute of Cognitive Neuroscience, and* [2]*Wellcome Department of Imaging Neuroscience, University College London, Queen Square, London WC1N 3AR, UK*

The mentalizing (theory of mind) system of the brain is probably in operation from *ca*. 18 months of age, allowing implicit attribution of intentions and other mental states. Between the ages of 4 and 6 years explicit mentalizing becomes possible, and from this age children are able to explain the misleading reasons that have given rise to a false belief. Neuroimaging studies of mentalizing have so far only been carried out in adults. They reveal a system with three components consistently activated during both implicit and explicit mentalizing tasks: medial prefrontal cortex (MPFC), temporal poles and posterior superior temporal sulcus (STS). The functions of these components can be elucidated, to some extent, from their role in other tasks used in neuroimaging studies. Thus, the MPFC region is probably the basis of the decoupling mechanism that distinguishes mental state representations from physical state representations; the STS region is probably the basis of the detection of agency, and the temporal poles might be involved in access to social knowledge in the form of scripts. The activation of these components in concert appears to be critical to mentalizing.

Keywords: mentalizing; theory of mind; medial prefrontal cortex; anterior cingulated cortex; temporal poles; superior temporal sulcus

1. DEVELOPMENT OF MENTALIZING

In 1978 a paper by Premack and Woodruff appeared with the provocative title 'Does the chimpanzee have a "theory of mind"?' (Premack & Woodruff 1978). The phrase 'theory of mind' was not to be taken literally of course, and certainly it did not imply the possession of an explicit philosophical theory about the contents of the mind. Instead, it crystallized the question of whether the mind of the chimpanzee works like the human mind, in that it makes the implicit assumption that the behaviour of others is determined by their desires, attitudes and beliefs. These are not states of the world, but states of the mind. Over the years, alternatives for the term 'theory of mind', such as 'ToM', 'mentalizing' and 'intentional stance', have also come into use. We will mainly use the term 'mentalizing'.

Premack and Woodruff in their seminal paper reported studies that tested the possibility that chimpanzees are implicitly aware that different individuals can have different thoughts and use this ability to predict their behaviour. One of the more striking outcomes of this social insight would be the ability to deceive others and to understand deception. The results of the experiments were equivocal and subsequent studies have remained tantalizing (Byrne & Whiten 1988; Heyes 1998; Povinelli & Bering 2002). While some studies reported an incipient but not very robust theory of mind in the chimpanzee and other great apes, the verdict fell the other way for monkeys: they do not show any evidence of the ability to attribute mental states (Cheney & Seyfarth 1990).

In contrast to the uncertainty about mentalizing in other species, the development of a fluent mentalizing ability, with far-reaching consequences for social insight, is undoubtedly a human accomplishment. How does this ability develop? When do children first show evidence of mentalizing? Evidence might come from explicit mental state language ('I *think* my brother is *pretending* to be a ghost'), but mentalizing might also be implicit in behaviour (far from being frightened, the child removes the sheet to reveal her brother underneath). In his commentary on Premack and Woodruff's paper, Dennett (1978) proposed a stringent test for the presence of theory of mind, the prediction of another person's behaviour on the basis of this person's false belief. A true belief would not do, as in this case it would be impossible to decide unequivocally whether the other person behaves in accordance with reality or in accordance with his or her own belief about reality. So, if the child runs towards the curtain when another person is hiding there, this may be because the other person is indeed there, or because the child believes the other person to be there. A new experimental paradigm was needed, and this was created by Heinz Wimmer and Josef Perner (1983). This paradigm opened the door to a new era in the study of social cognition. It goes like this: Maxi has some chocolate and puts it into a blue cupboard. Maxi goes out. Now his mother comes in and moves the chocolate to a green cupboard. Maxi comes back to get his chocolate. Where will Maxi look for the chocolate? The answer is of course: Maxi will look in the blue cupboard, because this is where he falsely believes the chocolate to be. Control questions checked that the child understood the sequence of events: where

*Author for correspondence (cfrith@fil.ion.ucl.ac.uk).

One contribution of 15 to a Theme Issue 'Decoding, imitating and influencing the actions of others: the mechanisms of social interaction'.

is the chocolate really? Do you remember where Maxi put the chocolate in the beginning?

A series of subsequent studies established that children of *ca*. 4 years of age, but no younger, begin to understand this scenario and can verbally explain it when asked. At age 5 years over 90%, and at age 6 years all children, could understand the task (Baron-Cohen *et al.* 1985; Perner *et al.* 1987). Other researchers used variants of this task with essentially similar results. Studies were also carried out in other cultures indicating the universality of this clear developmental phenomenon (Avis & Harris 1991).

(a) *From age 5 years*

Perner & Wimmer (1985) devised a more difficult task that required the attribution of a belief about another person's belief, a so-called second-order task. Here, Mary believes that John believes that something is the case. Children from the age of *ca*. 5 or 6 years effortlessly understand this task (Sullivan *et al.* 1994). Even more complex scenarios are used in suspense stories with detectives and spies where people carry around secrets and resort to bluff, and double bluff if necessary. These plots are popular from late childhood onwards and do not seem to require much mental effort. Of course to know about the full range of mentalizing situations and to use this knowledge to predict other people's behaviour, experience is necessary. There are many shades of social insight and social competence in adults. The successful Machiavellian individual probably has to practice for many years, and benefits from the study of suitable handbooks. Niccolo Machiavelli's (1469–1527) treatise on political acumen in *The Prince* is still unsurpassed.

(b) *From age 3 years*

But what happens before the age of five? Do young children not act as if they knew that other people had thoughts and that thoughts are different from physical states? Of course they do. A number of experimental paradigms suitable for younger ages have been invented to demonstrate this. Three-year-olds certainly know the difference between physical and mental entities. For instance, Wellman & Estes (1986) told children that one character had a biscuit and another was thinking about a biscuit. Children had no trouble saying which biscuit could be touched.

From 3 years of age or earlier children use words which refer to mental states, 'I *thought* it was an alligator. Now I *know* it's a crocodile', is an example quoted by Shatz *et al.* (1983) from a 3-year-old. Examples of mental state words in use by many 2-year-olds are *want, wish* and *pretend*.

The false-belief scenario with Maxi and the chocolate, which at first glance is quite complicated, has been transformed into a little play that can be watched by young children aged 3 years. In this way, Clements & Perner (1994, 2001) were able to show that when Maxi comes back to look for his chocolate, 3-year-olds reliably look first at the door near the blue cupboard, where he initially put the chocolate rather than the door near the green cupboard. Nevertheless, when asked the test question, the same children point towards the green cupboard, and give the wrong answer.

Three-year-olds also have an incipient understanding of the difference between knowing, thinking and guessing. Masangkay *et al.* (1974) and Flavell *et al.* (1981) showed that children aged 3 years, but not younger ones, could tell that if there were different pictures on each side of a card, the person sitting opposite would see a different picture when the card was held up. Hogrefe *et al.* (1986) showed that 3-year-olds realize that only the person who has looked inside a box knows what is inside it, but not another person, who did not look inside. However, such understanding is evident even earlier in the right communicative context. In the context of requesting an object, 2-year-olds show themselves to be sensitive to the knowledge state of a parent. They actively direct their mother's attention to the location of an object, if, unbeknown to her, the object had been moved (O'Neill 1996). Four-year-olds are less dependent on this context and can give reasons why seeing leads to knowing, and not seeing to not knowing (e.g. O'Neill & Gopnik 1991; Povinelli & deBlois 1992). Remarkably, when tested in implicit form, infants from as young as 18 months of age appear to have a practical understanding of this logic (Poulin-Dubois *et al.* 2003). The infants in this study were surprised, and looked longer, if a woman pointed to the wrong place after she had observed where another person hid an object. By contrast, they were not surprised and, did not look longer, when she had been unable to observe the hiding place.

(c) *From 18 months of age*

The age of 18 months (or thereabouts) is, in many respects, a developmental watershed, which marks the end of infancy. Thus, beginning at around this time, language learning takes off rapidly. This may be because from that time onwards word learning is facilitated by the ability to track a speaker's intention when he or she utters a word (Baldwin & Moses 1996; Bloom 2000). The child knows when the mother is naming an object for the benefit of the child rather than saying words that have nothing to do with the object the child is holding at the time. Without making this distinction the child would learn accidental sound and object associations. In fact such errors are rare. This age is also significant for the onset of pretend play. As Leslie (1987) cogently argued, the understanding of pretence is an unequivocal manifestation of the ability to mentalize. Leslie's well-known example is the mother playfully picking up a banana and pretending to telephone. The child laughs and does not get confused about the property of telephones and bananas. To prevent such confusion the child must have the ability to represent the attitude the mother takes to the banana. This has to be different from the representation of the banana's real life use. A possible cognitive mechanism suggested by Leslie was termed 'decoupling'. This term vividly conveys the need to keep separate representations of real events from representations of thoughts that no longer need to refer to such events.

The examples of pretend play and rapid language acquisition involve the joint attention of two people. Mother and child jointly attend to the object being named or to the object that is the target of pretence. When is joint attention first documented? The answer depends on whether strict or lenient criteria are used. The minimum requirement for joint attention is that both infant and adult look together at a third object. But this may be acci-

dental or contrived. A more stringent requirement is that one person's attention towards the object is deliberately drawn there by the other person, starting with a direct gaze. From approximately 12 months of age infants tend to look automatically at a target that an adult is looking at (Butterworth & Jarrett 1991). However, this achievement is not as impressive as it seems, as this only happens when the target is already within the infant's point of view. It is not until approximately 18 months that the infant reliably turns towards a goal that an adult is pointing to or gazing at, when this goal is not already in the line of vision (Butterworth 1991; Caron *et al.* 1997). Using the most stringent criterion one might therefore date joint attention from 18 months, even though joint looking and gaze following can be observed much earlier. Strictly defined joint attention indicates an implicit awareness of the fact that different people can pay attention to different things at the same time, and of the fact that their attention can be 'directed' to coincide with one's own interests. The development of joint attention between 14 and 24 months has been shown to have an orderly progression by Carpenter *et al.* (1998) and to be correlated with other significant developments in social competence. Reliable imitation of intentional actions performed by others, regardless of whether these actions reach their goal, also emerges at approximately 18 months, as demonstrated in a classic study by Meltzoff (1995).

At this stage, infants also seem to respond to a novel toy by taking into account their mother's emotional expression: they will not approach it if she signals fear (Repacholi 1998). Children at this age understand eye gaze as a communicative tool. They know that a person cannot see through an obstacle and they try to remove the hands if their mother covers her eyes when they want to show her a picture (Lempers *et al.* 1977).

At earlier ages, examples of mentalizing have rarely been reported, and this may indicate that the index behaviour is less robust at younger ages. One highly interesting study by Onishi & Baillargeon (2002) suggests that appropriate methods using length of looking time, can reveal an implicit form of false-belief understanding in children aged 15.5 months.

(d) *From 12 months of age*

There are some achievements from the age of 12 months (or thereabouts) that may well be vital milestones on the road to the development of mentalizing and suggest a dawning awareness of mental states such as intentions and desires. Perhaps the most impressive achievement is that from the age of 1 year onwards, infants can respond to an object as an intentional agent, purely on the basis of its interactive behaviour with another person (Johnson 2003).

Some of the most important tools for communication outside language come from looking and pointing gestures. They allow even infants to predict the action of agents. Woodward *et al.* (2001) showed that from 12 months of age but not before, there is a primitive understanding that gaze involves a relation between a person and the object of her gaze.

From approximately 12 months of age infants use information about an adult's gaze direction and positive emotional expression to predict that the adult will reach for the object (Phillips *et al.* 2002). This indicates an early ability to appreciate that a person may have different goals and that these goals may have different meanings. Sodian & Thoermer (2003) demonstrated that infants expect agents to grasp the object that they look at, rather than another object that is also present. However, if a pointing gesture was used as a cue for grasping instead of gaze, infants were less surprised if the agent grasped the other object instead.

(e) *From nine months of age*

Gergely *et al.* (1995; see also Csibra 2003) obtained evidence from an ingenious experiment concerning infants' ability to reason about goals. They call it the principle of rationality: infants aged between 9 and 12 months expect agents to approach a goal in the most economic way. They are surprised if an agent does not do so, but jumps instead over an invisible hurdle. This demonstrates that they can separately represent goals of agents and the means used to reach the goal. The ability to represent goals and the ability to reason 'rationally' are likely to be an important prerequisite of the ability to represent intentions.

(f) *From six months of age*

Infants at about this age are surprised if an object moves on its own, but not if a person does (Spelke *et al.* 1995). This suggests that they can distinguish animate agents by the fact that they are self-propelled. By this definition a self-propelled agent need not be a biological creature, but can be a mechanical toy or even a car. The importance of agents is not that they are biological entities but that they may move unpredictably 'of their own will'. The representation of the action of agents is likely to be an essential requirement for the representation of the intention of agents.

Woodward (1998) showed that infants expected a human hand to reach towards the same goal objects when its location had been changed rather than for a different object that would have been easy to reach. By contrast, the infants did not show this differentiated expectation when no human hand was used, but instead a mechanical rod. The distinction between biological and mechanical movement is probably another prerequisite for the understanding of intentions. As we shall see in § 4b, in adults, specific regions in the STS of the brain are active in response to these different types of movement. That the difference is detected at such an early age suggests that these regions mature early and that learning must be ultra-fast.

(g) *From three months of age*

The range of behaviours that can be observed in the early months of life is quite limited, and this limits the sources of evidence. However, it is clear that infants only a few weeks old smile more and vocalize more towards people than towards objects (Legerstee 1992). This could well suggest an innate preference for social stimuli.

Not only eye movements but also other forms of biological motion seem to have a privileged status in attracting infants' attention at an extremely early age. They track objects with self-propelled movement (Crichton & Lange Kuettner 1999). They also show more interest in the kinematic patterns of point-light displays of

a person walking than of random movement (Bertenthal *et al.* 1984).

The ability to react reflexively to movement of gaze as a priming cue for one's own eye movement is likely to be innate as it can already be observed at three months of age (Hood *et al.* 1998). This is different from the voluntary following of the general direction of an adult's gaze, which is not accomplished until *ca.* 12–18 months of age. The same observable action, gaze following, is guided by different mechanisms and thus can mean very different things. It is unlikely that the early gaze reflex evident at age three months rests on the same neural substrate as the type of sophisticated gaze following seen at age 18 months that implies the ability to mentalize.

2. CONCLUSIONS

Evidence of mentalizing becomes abundant only from *ca.* 18 months of age. Accomplishments at, and just before, 12 months of age are nevertheless astounding in their own right. They suggest that the infant can represent separately agents, goals and means of getting to the goal. Representing the visible goals of agents, however, is not the same as representing the invisible intentions in agents. It is unclear whether, and how, this early ability relates to the later understanding of intentions. Intentions, after all, can result in actions that may be thwarted or never fulfilled. So far, clear evidence for understanding intentions is only available from 18-month-olds, at the same time as they begin to understand other mental states.

One remarkable fact about the studies reviewed is that they suggest universal developmental stages, applicable to all children, notwithstanding individual differences in the speed of development. For this reason it is possible to identify abnormal development in those children who appear to have a faulty mentalizing mechanism. This is suggested to be the case in autism (Baron-Cohen *et al.* 1985).

Perhaps it is difficult to find evidence for the intentional stance in the first year of life because there are limits set by the experience that is available to young infants, but there are also limits set by the state of maturation of the brain. The presence of developmental abnormalities in brain function that affect mentalizing would not be readily discovered at this young age. Would enhanced experience at this stage be helpful? Possibly, but even if experience is available, innate mechanisms may not be mature enough to take advantage of it. Cognitive mechanisms may go through a number of developmental stages, and this could well be the case for the mentalizing system.

Tentatively, we can conclude that an implicit version of the intentional stance emerges first, concerned with desires, goals and intentions. This is usually dated at around 18 months. At 18–24 months there is a convergence of several important developmental milestones, including a true understanding of joint attention, deliberate imitation and the ability to track a speaker's intention while learning words. There is also evidence for the ability to understand knowing and seeing at an implicit level, and possibly even an implicit understanding of false belief.

In summary, we can probably assume that the understanding of many mental states (wanting, intending, knowing, pretending and believing) is already available in implicit form to 2-year-olds and governs their behaviour as well as their understanding of other people's behaviour. We would therefore expect that if functional brain imaging were done in children aged 2 years (for example, while watching an agent performing actions that do not reach their goal versus a robot performing mechanical goal-directed actions), the mentalizing system of the brain (see § 4) would already be in operation. Conversely, in children with autism, the presumed fault in this system should show up at this age too.

We can also conclude that another major leap in the development of mentalizing occurs between the ages of 4 and 6 years. It is only from 6 years of age onwards that we can safely attribute to a normally developing child a full and explicit awareness of mental states and their role in the explanation and prediction of other people's behaviour. What explains this significant change? Different theories are currently debated. One assumes that the change is extraneous to mentalizing but has to do with the executive components of false-belief tasks (e.g. Russell 1996). Another theory postulates that only the older child can apply the full ability to simulate another person's mental states, moving freely from their own to another's perspective (Harris 1991). A third proposal is that the child behaves like a theorist who, from time to time, is compelled by the facts to change his concepts about the physical and social world (Gopnik & Wellman 1994). While all these theories might help to explain changes in task performance, an even more parsimonious theory is that the mentalizing mechanism itself makes another leap in development at *ca.* 4 years of age. If it were possible to make visible the mentalizing system in the brain during implicit watching of a false-belief scenario before and after the observed changes in explicit task performance, this question might be answered.

(a) *What role for early components of social cognition in mentalizing?*

While other primitive neural mechanisms may facilitate social learning, we do not know whether they contribute directly to the social insight that is facilitated by the intentional stance. It is possible that strong connections between the brain regions that subserve these mechanisms, strengthened through learning, eventually give rise to the ability to mentalize. It is also possible that an additional neural mechanism is needed for the development of this ability, which is, after all, of late origin in terms of evolution.

We can only speculate about the role of early-appearing components of social cognition in mentalizing. There are three such functions, which might be particularly relevant. First, there is the preference for social stimuli. Evidence from behavioural and electrophysiological studies suggests that even newborn infants are responsive to human faces and preferentially orient towards stimuli that resemble faces. In adults the fusiform gyrus and STS are thought to subserve this function (Chao *et al.* 1999; Allison *et al.* 2000). In newborn babies, however, these cortical areas are not yet mature, and subcortical regions are probably involved (Johnson & Morton 1991), Second, an agency detection mechanism might be the basis of the sensitivity of three-month-olds to biological motion and eye movement. This mechanism in adults is thought to be sub-

served by the STS. Third, there may be a mechanism that enables an understanding of the meaning of actions, a differentiation of the goals of actions and the means to reach them. Mirror neurons, situated in the ventral part of the lateral premotor cortex, might be involved in such a mechanism (Rizzolatti *et al.* 2002).

Might these potentially innate components (a preference for conspecifics, a predisposition to detect agency and a predisposition to understand actions), contribute to the development of mentalizing? They might be necessary prerequisites. However, by themselves they are not sufficient for the development of mentalizing. This follows from the assumption that they are shared with a great many other species, most of which do not possess a trace of mentalizing ability. As we shall see in the review of neuroimaging studies (§ 4), the neural components of the mentalizing system comprise some of the putative prerequisites that developmental studies have demonstrated. However, the mentalizing system comprises additional components whose function in development is as yet unknown. We speculate that only when all these components are connected together in the brain are both necessary and sufficient conditions for mentalizing present. One of the reasons that we cannot make more precise links from the detailed and ingenious behavioural studies with infants and young children to neuroimaging studies with adults is our lack of knowledge of the developing human brain either in terms of structure or function.

The role of learning and experience in the development of mentalizing still needs to be investigated. Different individuals have different experiences and this is likely to be reflected in their mentalizing competence. So far, studies have rarely focused on individual differences, and thus our knowledge is currently very limited. Wellman *et al.* (2000) report that the first achievement of explicit false-belief understanding can vary from between 2 years six months to 6 years. Some evidence exists that the presence of older siblings facilitates the understanding of false beliefs (after age 4 years) (Ruffman *et al.* 1998), and it is widely believed that girls achieve the developmental milestones of mentalizing somewhat earlier than boys. While cross-cultural studies do not suggest marked differences in early achievements, it is obvious that cultural differences could play a large, if not dominant, role in the development of the content of an adult theory of mind (Lillard 1998).

3. NEUROIMAGING STUDIES OF MENTALIZING

Neuroimaging provides another kind of evidence about the nature and components of the ability to mentalize. All the studies carried out so far have concerned adults rather than children. Most studies have been modelled on the story of Maxi and the chocolate. For example, while being scanned the volunteer reads a series of very short stories in which the behaviour of the protagonist is determined by his or her false belief about the situation. An example is the 'burglar story' from a set of stories testing mentalizing ability (Happé 1994).

'A burglar who has just robbed a shop is making his getaway. As he is running home, a policeman on his beat sees him drop his glove. He doesn't know the man is a burglar, he just wants to tell him he dropped his glove. But when the policeman shouts out to the burglar, "Hey, you! Stop!" The burglar turns round, sees the policeman, and gives himself up. He puts his hands up and admits that he did the break-in at the local shop'.

Subsequently the volunteer is asked to explain the burglar's behaviour. An appropriate answer would be that the burglar falsely believes that the policeman knows he has just robbed the shop. Reading and understanding such stories engages many processes in addition to mentalizing and so control stories, matched for difficulty, are necessary. Such stories also involve people, but the critical events are explained in terms of physical causality.

'A burglar is about to break in to a jewellers' shop. He skilfully picks the lock on the shop door. Carefully he crawls under the electronic detector beam. If he breaks this beam it will set off the alarm. Quietly he opens the door of the storeroom and sees the gems glittering. As he reaches out, however, he steps on something soft. He hears a screech and something small and furry runs out past him towards the shop door. Immediately the alarm sounds'.

In this example, the appropriate answer to the question, 'Why did the alarm go off?' would be because some animal had triggered it.

4. A NEURAL SYSTEM FOR MENTALIZING

In the first study to use such stories (Fletcher *et al.* 1995) a comparison of mentalizing with physical stories revealed activity in the MPFC, posterior cingulate and right posterior STS. In comparison with a low-level baseline of unlinked sentences, activity was also seen in the temporal poles, bilaterally. The MPFC seemed to be particularly linked to mentalizing since it was the only area that was not also activated by the physical stories. Two subsequent fMRI studies used the same stories and obtained very similar results (Gallagher *et al.* 2000; Vogeley *et al.* 2001). Activity was seen in the MPFC, temporal poles and STS when reading mentalizing stories compared with physical stories, although in Vogeley *et al.* the STS activity was most marked in a novel condition in which the volunteer imagined herself as the protagonist in a mentalizing story.

Two studies have presented mentalizing scenarios using drawings rather than words. Brunet *et al.* (2000) presented cartoon strips in which the sequence could only be understood in terms of the goals and intentions of the protagonist. Gallagher *et al.* (2000) used cartoons without captions in which the jokes involved false beliefs. Again in both these studies activity was observed in the MPFC, temporal poles and STS.

Goel *et al.* (1995) used a very different task to engage mentalizing. Volunteers were shown objects and had to indicate whether or not Christopher Columbus would have known what each object was used for. Such a decision involves inferring something about the knowledge and beliefs of someone who lived 500 years ago. In comparison to various control tasks activity was again seen in MPFC, temporal pole and STS.

Berthoz *et al.* (2002) have reported a study of social norm transgression that also involved mentalizing. Volunteers read short vignettes in which social transgressions occurred. These could be accidental or deliberate. An example of an accidental transgression is as follows;

'Joanna is invited for a Japanese dinner at her friend's house. She has a bite of the first course, chokes and spits out the food while she is choking'. Volunteers were asked to try and imagine how the character in the story would feel. In comparison with matched stories in which no transgression occurred, both deliberate and accidental transgressions elicited activity in the MPFC, temporal poles and STS. Activity was also seen in areas responding to aversive emotional expressions such as anger.

An implicit mentalizing task which activated all three of these areas was based on the observation by Heider & Simmel (1944) that people will attribute intentions and desires to moving geometric shapes if these movements are of sufficient complexity. Castelli *et al.* (2000), using positron emission tomography, presented an animated sequence in which two triangles interacted with each other. The more the observers attributed mental states to the triangles the greater the activity in the MPFC, temporal pole and STS. Schultz *et al.* (2003) used a similar task and, using fMRI, observed activations in the same regions. In both these studies where mentalizing was elicited by the movements of abstract shapes, the activity in the temporal pole extended into the amygdala and activity was also seen in the fusiform gyrus.

All these studies, except possibly those using passive viewing of animations, have involved explicit mentalizing since the subjects were asked to describe the mental states of other people or make decisions based on the mental states of other people. In addition, in all these studies, mentalizing was elicited by the material presented. The approach is analogous to studies in which the colour area in the visual system is identified by comparing the activity elicited by stimuli with and without colour (Zeki *et al.* 1991). An alternative approach is to keep the stimulus material constant and change the attitude of the volunteer. For example, the same visual stimulus is presented, but the volunteer is required to attend to colour in one condition and to motion in another (Corbetta 1993). Two studies have used this approach to identify brain areas associated with mentalizing. McCabe *et al.* (2001) scanned volunteers while they played an economic game with another person. In this game mutual cooperation between players increases the amount of money that can be won. In the comparison task the volunteers believed they were playing with a computer that used fixed rules. Gallagher *et al.* (2002) scanned volunteers while they played the game 'Stone–Paper–Scissors'. This is a competitive game in which success depends upon predicting what the other player will do next. In this study the comparison condition was also created by telling the volunteers that they were playing against a computer. In fact, the sequence of the opponent's moves was the same in both conditions.

In these studies the volunteers were not explicitly instructed to mentalize while performing their task. However, an intensive debriefing of the volunteers in the study of 'Stone–Paper–Scissors' confirmed that they had engaged in mentalizing while playing against a person. They described guessing and second guessing their opponent's responses and felt that they could understand and 'go along with' what their opponent was doing. Playing against a computer felt distinctly different. The volunteers considered that the computer was in principle very predictable, but the rules it used might be difficult to detect. They also felt that the computer might be too fast for them to keep up with.

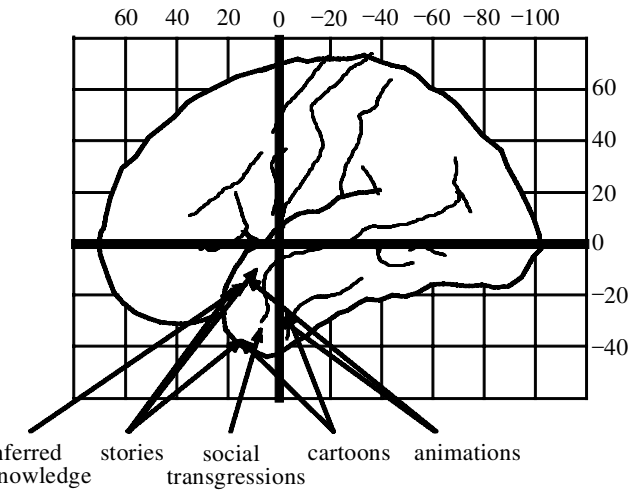

Figure 1. Lateral view of the brain shown in the Talairach coordinate system. Peak activations in the temporal pole are shown for five different tasks used in 10 studies of mentalizing. Where activation was bilateral the two sides have been combined. Inferred knowledge: Goel *et al.* (1995); stories: Fletcher *et al.* (1995), Gallagher *et al.* (2000), Vogeley *et al.* (2001), Ferstl & von Cramon (2002); social transgressions: Berthoz *et al.* (2002); cartoons: Brunet *et al.* (2000), Gallagher *et al.* (2000); animations: Castelli *et al.* (2000), Schultz *et al.* (2003).

Both studies revealed activity in the MPFC when the volunteers believed that they were interacting with another person. But this was the only area that was more active in this condition than in the condition where they believed they were playing against a computer. This dissociation between the MPFC and the other regions suggests that the posterior regions are more concerned with the nature of the sensory signals that elicit mentalizing, whereas MPFC activity reflects the attitude taken towards those signals. In order to explore the precise role of the various areas in the mentalizing network we shall now consider studies that activate some or all of these areas, but which were not explicitly designed to engage mentalizing.

In this review we have restricted ourselves to those imaging studies that declare mentalizing as an experimental variable and that have used appropriate controls and statistical analysis. Furthermore, throughout our review we have relied on those studies that report their results in Tailarach space. Without such standardized indicators of the location of changes of activity in the critical conditions, a comparison with other studies is not possible.

(a) *Temporal pole*

Five different mentalizing tasks as used in 10 studies have elicited activity in the temporal poles bilaterally, with somewhat greater effects on the left (figure 1). This region of the anterior temporal lobe is a site for the potential convergence of all sensory modalities and also limbic inputs (Moran *et al.* 1987). As shown in figure 2, this region is frequently activated in studies of language and semantics, although in these cases the activity is restricted to the left

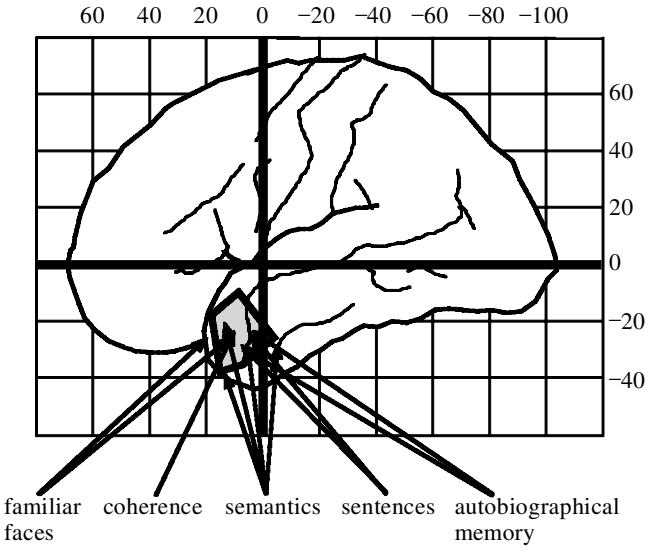

Figure 2. Lateral view of the brain shown in the Talairach coordinate system. The shaded area shows the region activated by the studies of mentalizing shown in detail in figure 1. Peak activations are shown for 11 studies of other processes that activate adjacent regions of the temporal poles. Familiar faces and voices: Nakamura *et al.* (2000, 2001); coherence: Maguire *et al.* (1999); semantics: Vandenberge *et al.* (1996), Noppeney & Price (2002*a*,*b*); sentences: Bottini *et al.* (1994), Vandenberghe *et al.* (2002); autobiographical memory: Fink *et al.* (1996), Maguire & Mummery (1999), Maguire *et al.* (2000).

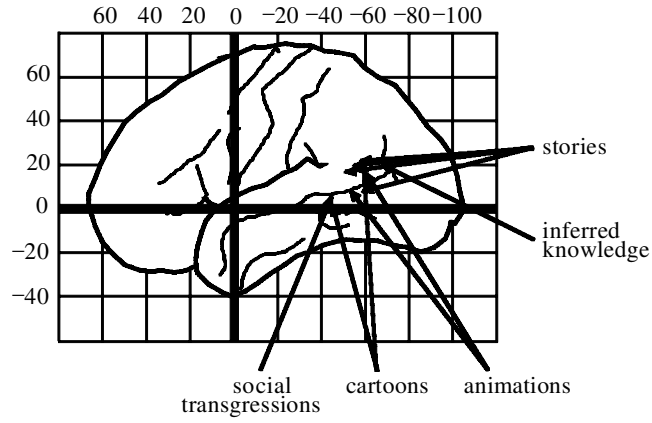

Figure 3. Lateral view of the brain shown in the Talairach coordinate system. Peak activations in the posterior STS are shown for 10 studies of mentalizing. Where activation was bilateral the two sides have been combined. Stories: Fletcher *et al.* (1995), Gallagher *et al.* (2000), Vogeley *et al.* (2001), Ferstl & von Cramon (2002); inferred knowledge: Goel *et al.* (1995); animations: Castelli *et al.* (2000), Schultz *et al.* (2003); cartoons: Brunet *et al.* (2000), Gallagher *et al.* (2000); social transgressions: Berthoz *et al.* (2002).

temporal pole. In particular, this region is activated when sentences are compared with unrelated word strings (Bottini *et al.* 1994; Vandenberghe *et al.* 2002), when narratives are compared with nonsense (Mazoyer *et al.* 1993) or with unrelated sentence strings (Fletcher *et al.* 1995), and when highly coherent narratives are compared with less coherent narratives (Maguire *et al.* 1999). The same region is also activated when volunteers make semantic decisions (e.g. Which is more similar to cow? Horse or bear? (Vandenberghe *et al.* 1996; see also Noppeney & Price 2002*a*,*b*)) In addition, this area is activated during memory retrieval. This is particularly the case during retrieval from autobiographical memory (Fink *et al.* 1996; Maguire & Mummery 1999; Maguire *et al.* 2000), during the incidental retrieval of emotional context in single-word recognition (Maratos *et al.* 2001) and during the recognition of familiar faces, scenes and voices (Nakamura *et al.* 2000, 2001).

We tentatively conclude that this region is concerned with generating, on the basis of past experience, a wider semantic and emotional context for the material currently being processed. This function would aid the interpretation of stories and pictures whether or not they involve mentalizing. One component of the wider semantic context is sometimes referred to as a 'script' (Schank & Abelson 1977). Scripts are built up through experience and record the particular goals and activities that take place in a particular setting at a particular time. A much used example is the 'restaurant script' which leads us to expect that we will first get the menu, then order, taste the wine, and so on. Identifying which script is most appropriate to a situation will be of considerable help in predicting what people are going to do. The temporal poles, especially on the left, may well be concerned with the retrieval of scripts. Patients with semantic dementia show atrophy in the anterior temporal lobes, especially on the left (Chan *et al.* 2001). As this atrophy progresses, these patients lose knowledge of all but the simplest and most concrete scripts (Funnell 2001).

Scripts provide a useful framework within which mentalizing can be applied. Events rarely conform exactly to the established script and mentalizing is needed to understand the deviations.

(b) *Posterior STS*

Mentalizing tasks elicit activity in the posterior STS (temporo-parietal junction extending towards the angular gyrus) bilaterally with somewhat greater effect on the right (see figure 3). The same 10 studies as shown in figure 1 are represented in this diagram. Figure 4 shows activations of this region by 19 other studies, mostly concerned with living agents and biological motion. The posterior STS is also a multimodal convergence zone with connections to the limbic system (Barnes & Pandya 1992). It is well known that this region is activated when volunteers observe biological motion (see Allison et al. 2000; Puce & Perrett 2003). Activation is seen during presentation of moving bodies and parts of bodies (Grezes *et al.* 1998; Puce *et al.* 1998; Campbell *et al.* 2001), while hearing speech and seeing speaking mouths (Calvert *et al.* 2000), and during presentations of action reduced to moving points of light (Bonda *et al.* 1996; Grossman *et al.* 2000; Grezes *et al.* 2001). The location of the maximum response to biological motion is *ca.* 10 mm superior and anterior to V5, which responds to visual motion in general (Zeki *et al.* 1991). However, this region of the STS is also activated by static images of faces and animals (e.g. Chao *et al.* 1999) especially when attending to eye gaze (Wicker *et al.* 1998; Hoffman & Haxby 2000), by names of animals (e.g. Chao *et al.* 1999), and by making semantic decisions

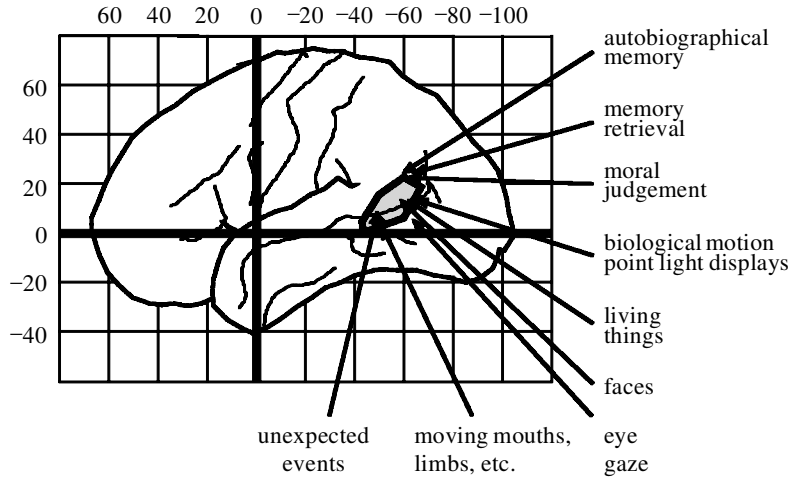

Figure 4. Lateral view of the brain shown in the Talairach coordinate system. The shaded area shows the region activated by studies of mentalizing shown in detail in figure 3. Peak activations are shown for 19 studies of other processes that activate adjacent regions of the STS. Autobiographical memory: Vandenberge *et al.* (1996), Maguire & Mummery (1999), Maguire *et al.* (2000); memory retrieval: Maratos *et al.* (2001), Lee *et al.* (2002); moral judgement: Greene *et al.* (2001); biological motion, point light displays: Bonda *et al.* (1996), Grossman *et al.* (2000), Grèzes *et al.* (2001); living things: Price *et al.* (1997), Chao *et al.* (1999); static faces: Chao *et al.* (1999); eye gaze: Wicker *et al.* (1998), Hoffman & Haxby (2000); biological motion, mouths, eyes, hands: Puce *et al.* (1998), Grezes *et al.* (1998), Calvert *et al.* (2000), Campbell *et al.* (2001); unexpected events: Downar *et al.* (2000), Corbetta *et al.* (2000).

about living things (e.g. Price *et al.* 1997). These observations suggest that this region is activated when observing the behaviour of living things and also when retrieving information about the behaviour of living things. An adjacent area closer to the angular gyrus is also activated by retrieval from semantic memory (e.g. Vandenberghe *et al.* 1996; Maratos *et al.* 2001; Lee *et al.* 2002) and from autobiographical memory (e.g. Maguire & Mummery 1999; Maguire *et al.* 2000). Whether this activity is specific to retrieval of memories about living things is not yet known.

An interesting set of parallel observations have been made about the area of fusiform gyrus that was activated in the two studies that elicited mentalizing using animations (Castelli *et al.* 2000; Schultz *et al.* 2003). This is also an area that seems to be concerned with knowledge about living things such as faces and animals (Chao *et al.* 1999). Presumably the knowledge in this region in the ventral stream primarily concerns the appearance of living things, their form and colour, rather than their patterns of behaviour. For example, this region is more active than the STS when volunteers make decisions about the identity of faces (Hoffman & Haxby 2000).

There is, however, another kind of event that elicits activity in the STS and does not specifically involve living things. An unexpected change of stimulation in any modality elicits activity in the same location as biological motion (Corbetta *et al.* 2000; Downar *et al.* 2000). Furthermore, learning to follow complex but predictable patterns of movement activates this region (Maquet *et al.* 2003). These results suggest that this region is not specifically concerned with the behaviour of living things, but with complex behaviour whatever its source. Nevertheless, we suggest that sudden changes of stimulation and complex patterns of movement are far more likely to be associated with living things than with mechanical or physical systems.

Knowledge about complex behaviour and, in particular, the ability to predict the next move in a sequence of behaviour is extremely valuable in any social interaction and could underlie some of the precursors of mentalizing, like gaze following and joint attention. Indeed it is known that activity in the STS increases when volunteers are asked to attend to gaze direction (Hoffman & Haxby 2000). The mentalizing system goes one step further and uses the observed patterns of behaviour to perceive the mental states that underlie this behaviour.

(c) *MPFC*

All 12 mentalizing tasks available to this review have elicited activity in the MPFC, with the interactive game-playing tasks (McCabe *et al.* 2001; Gallagher *et al.* 2002) activating this region only (see figure 5). The medial prefrontal region activated by these studies is the most anterior part of the paracingulate cortex, where it lies anterior to the genu of the corpus callosum and the ACC proper. The MPFC has direct connections to the temporal pole and to the STS (Bachevalier *et al.* 1997). The paracingulate cortex (BA 32) is often considered to be part of the ACC that incorporates the cytoarchitectonically defined Brodmann areas 24, 25 and 33. The ACC is an ancient structure that has been broadly defined by Broca as belonging to the limbic lobe (Bush *et al.* 2000).

However, the existence of an unusual type of projection neuron (spindle cell) found in the sub-areas of the ACC 24a, 24b and 24c in the human, and some other higher primates (pongids and hominids) but not monkeys, is evidence that the ACC has undergone changes in recent evolution (Nimchinsky *et al.* 1999). Furthermore, in humans these cells are not present at birth, but first appear at approximately four months of age (Allman *et al.* 2001). However, BA 32 has been described as cytoarchitectonically a cingulo-frontal transition area (Devinsky *et al.* 1995) and therefore anatomically (and speculatively

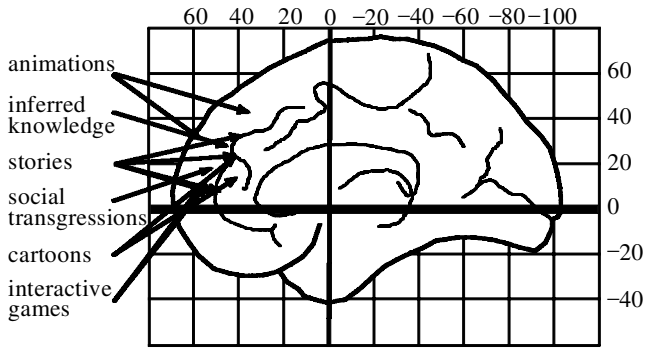

Figure 5. Medial view of the brain shown in the Talairach coordinate system. Peak activations are shown for the 10 studies of mentalizing shown in figures 1 and 3. In addition, two studies using a sixth mentalizing task (interactive games) are included which only activated the MPFC. Animations: Castelli *et al.* (2000), Schultz *et al.* (2003); inferred knowledge: Goel *et al.* (1995); stories: Fletcher *et al.* (1995), Gallagher *et al.* (2000), Vogeley *et al.* (2001), Ferstl & von Cramon (2002); social transgressions: Berthoz *et al.* (2002); cartoons: Brunet *et al.* (2000), Gallagher *et al.* (2000); interactive games: McCabe *et al.* (2001), Gallagher *et al.* (2002).

functionally) distinct from the ACC proper. It remains to be seen whether the recent evolutionary changes observed in the ACC are relevant to the more anterior region of medial frontal lobe where activations associated with mentalizing are observed. Recent anatomical changes in this region would be consistent with the observation that mentalizing has never been observed in monkeys (Cheney & Seyfarth 1990) and can only be found in a most rudimentary form in great apes (Byrne & Whiten 1988; Povinelli & Preuss 1995; Heyes 1998).

Evidence from anatomy and from functional studies shows that the ACC can be divided into distinct areas with different functions, as indicated in figure 6. In terms of the nomenclature of Picard & Strick (1996) the mentalizing region overlaps with, but is mostly anterior to, the rCZa. In terms of the functional nomenclature of Bush *et al.* (2000) the mentalizing region overlaps with the emotional division of the ACC.

(i) *Executive processes*

One plausible characterization of mentalizing tasks is that they involve complex problem solving of the type required by executive tasks, but this idea is not supported by imaging studies. Many kinds of executive tasks are known to activate the ACC. Duncan & Owen (2000) have performed a careful meta-analysis of such tasks showing that increasing the difficulty in a wide range of tasks activates the same region of the ACC whatever the nature of the task. However, all but one of the 26 peak activations that they list lie posterior to the mentalizing region, being centred instead in the rCZp. The mean coordinates derived from the meta-analysis of Stroop-like tasks from Barch *et al.* (2001) also lie in this division of the ACC (see figure 6). Independent confirmation of this distinction between executive tasks and theory of mind tasks comes from studies of patients with lesions. Patients can be found who perform executive tasks very badly, while still performing mentalizing tasks well (Varley *et al.* 2001) and vice versa (Fine *et al.* 2001). Rowe *et al.* (2001) observed that patients with frontal-lobe lesions performed badly on mentalizing tasks and executive tasks. However, within this group poor performance on one type of task was not related to poor performance on the other type of task.

(ii) *Representing emotion*

The recent meta-analysis of Phan *et al.* (2002) shows that tasks involving emotion can elicit activity in most regions of the ACC including the mentalizing region. What is the difference between the kinds of emotional task that activate the different divisions of the ACC? Lane (2000) highlights an important distinction between having an emotional experience and attending to an emotional experience. Lane *et al.* (1998) studied the effects of having an emotional experience by comparing responses to emotional experiences (happiness, sadness and disgust) with those to neutral experiences. The contrast revealed activity in the ACC, but in a posterior part at the border of the rostral cingulate zone and the cCZ. In another study, Lane *et al.* (1997) investigated the effects of attending to an emotion. Volunteers were shown emotionally arousing scenes. In one condition they indicated whether the scene was indoors or outdoors, while in the other condition they indicated the emotion aroused in them by the picture. When volunteers attended to their emotional experience, activity was seen in the mentalizing region, just anterior to the rCZa. The same distinction was observed by Gusnard *et al.* (2001) in a replication of Lane *et al.* (1997). Volunteers were shown pleasant, unpleasant or neutral scenes and were asked to indicate either their emotional response or whether the scenes were indoors or outdoors. Emotionally laden scenes elicited activity in the posterior ACC (cCZ at the border with supplementary motor area) whatever the task, while attention to emotion increased activity in the mentalizing region.

Petrovic & Ingvar (2002) have pointed out that a very similar distinction can be found in the study of pain. As stimuli become increasingly noxious, increases in activity are seen in the cCZ. However, the *perception* of pain does not relate directly to the nature of the stimulus, but can be altered by cognitive manipulations such as hypnotic suggestion, distraction or placebo analgesia. Variations in the perception of pain are related to activity in the rCZa overlapping with the mentalizing region. These studies of emotion and pain suggest that first-order representations of these states are located in the cCZ where correlates of arousal and stress are also observed (Critchley *et al.* 2000). Second-order representations of these states, available for attention and report, are located in the rCZa. We call these representations second order because they do not reflect the physical nature of the stimulus, but the mental attitude to that stimulus. To use the terminology of Leslie (1994), these representations are *decoupled* from the physical world and are no longer subject to normal input–output relations.

This formulation is consistent with our earlier suggestion (Frith & Frith 1999) that the mentalizing region of the MPFC is engaged when we attend to our own mental states as well as the mental states of others. Other situations where attention to mental states of the self activates this region include attention to the irrelevant thoughts that occur during scanning (McGuire *et al.* 1996) and atten-

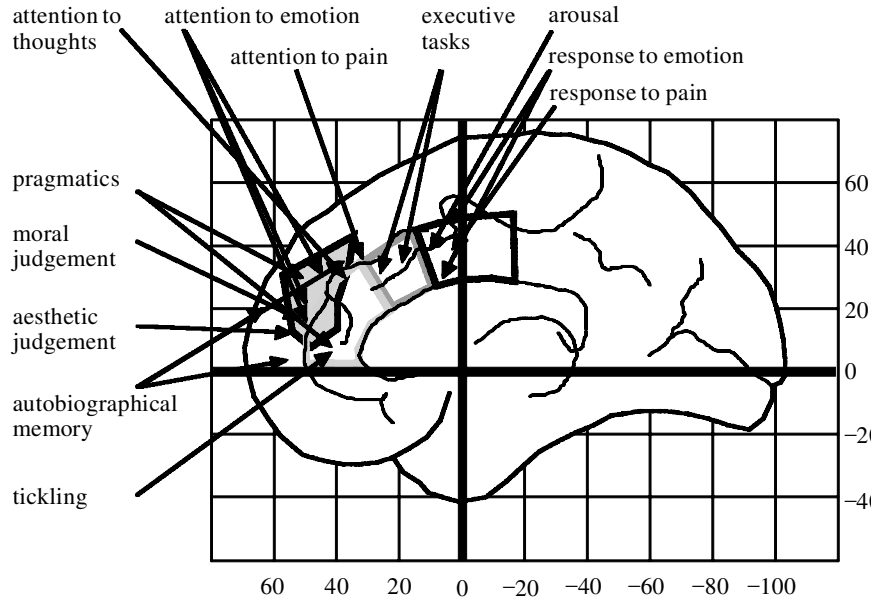

Figure 6. Medial view of the brain shown in the Talairach coordinate system. The shaded area shows the region activated by studies of mentalizing shown in detail in figure 5. Peak activations are shown for 19 studies of other processes that activate the ACC and adjacent MPFC. Approximate divisions of the ACC are shown. From right to left: cCZ, rCZ, rCZa. Arousal: Critchley *et al.* (2000, 2001); response to emotion: Lane *et al.* (1998), Gusnard *et al.* (2001); response to pain: Rainville *et al.* (1999), Petrovic & Ingvar (2002); executive tasks: Duncan & Owen (2000), Barch *et al.* (2001); attention to emotion: Lane *et al.* (1997), Gusnard *et al.* (2001); attention to pain: Rainville *et al.* (1999), Petrovic & Ingvar, (2002); attention to thoughts: McGuire *et al.* (1996); pragmatics: Bottini *et al.* (1994), Ferstl & von Cramon (2002); moral judgement: Greene *et al.* (2001); aesthetic judgement: Zysset *et al.* (2002); autobiographical memory: Maguire & Mummery (1999), Maguire *et al.* (2000); tickling: Blakemore *et al.* (1998).

tion to being tickled (Blakemore *et al.* 1998). We would also include two other tasks as examples of attending to the emotional states of the self although this was not necessarily the interpretation given by the authors. Zysset *et al.* (2002) observed activation in the mentalizing area when volunteers evaluated things (for example, answering the question, 'Do you like Leipzig?'). Greene *et al.* (2001) observed activation in the same area when volunteers considered moral dilemmas. We suggest that one component in the answering of such questions involves attending to the emotion aroused by topic (Does the thought of Leipzig make me happy or sad? How distressed would I feel if I had to take this particular course of action?). It is notable that the moral dilemma study of Greene *et al.* also activated the STS component of the mentalizing system, and it may be argued that this task is also a mentalizing task.

(iii) *Autobiographical memory*

A rather different task that can require representations of the self is autobiographical memory. Tulving (1985) has suggested that there is a form of autobiographical or episodic memory in which we perform 'mental time travel' and relive our past experiences (autonoetic memory). This would be a case of representing a past mental state clearly decoupled from current reality. In a series of studies Eleanor Maguire and her colleagues (Maguire & Mummery 1999; Maguire *et al.* 2000, 2001) have shown that retrieval from autobiographical memory reliably activates the mentalizing region of the MPFC (in addition to medial temporal lobe structures and the STS). Autobiographical memory tasks can often be solved simply on the basis of a feeling of familiarity rather than truly reliving the event and it is usually difficult to relate these different processes to specific brain regions. However, Maguire *et al.* (2001) also studied patient Jon who has considerable memory problems associated with early and severe damage to his hippocampi. Jon spontaneously makes the distinction between past events that he can clearly remember happening and others that he knows a lot about, but does not recall the event occurring. Memories of events where he clearly remembered them happening were associated with greater activity in the MPFC. This effect was independent of his ratings of emotional intensity and valence for the events.

(iv) *Pragmatics*

The studies we have discussed so far are consistent with our suggestion that the area of MPFC activated in mentalizing tasks is concerned with the representation of the mental states of the self and others decoupled from reality. There is one last set of studies that at first glance cannot be so easily incorporated into this scheme. Ferstl & von Cramon (2002) have shown that a certain kind of language task activates the same region of the MPFC as a simple mentalizing task. In both cases volunteers heard pairs of sentences. Examples of these sentence pairs include (i) 'Mary's exam was about to begin. Her palms were sweaty'; and (ii) 'The lights have been on since last night. The car doesn't start'. In the mentalizing task volunteers had to think about the motivations and feelings of the people in the sentences of type (i). In the language condition they had to decide whether there was a logical connection between the two sentences of type (ii). In comparison to a control task both conditions elicited activations in the mentalizing region of the MPFC.

Interpreting the two unlinked sentences in these

examples depends upon an aspect of language processing often referred to as pragmatics. In many real-life cases the understanding of an utterance cannot be based solely on the meanings of the individual words (semantics) or upon the grammar by which they are connected (syntax). It has been proposed that a successful understanding of an utterance depends upon perceiving the intention of the speaker (Grice 1957). The idea that the purpose of utterances is for the listener to recognize the intention of the speaker has been elaborated by Sperber & Wilson (1995) in their theory of relevance. If this analysis is correct then pragmatics, the understanding of utterances, depends upon mentalizing whether or not this is required by the task instruction. This would apply also to the type (ii) sentences used by Ferstl & van Cramon (2002) where logical connections had to be found. For instance the example above may evoke the idea that 'someone (stupidly or maliciously) left the lights on'.

The need for mentalizing is particularly clear in non-literal figures of speech such as metaphor and irony. Sperber and Wilson analyse the example in which a mother says to her daughter, 'Your room is a pigsty'. How is the daughter to understand this? Her room is not literally a pigsty, but it shares with pigsties the characteristic of being very messy and untidy. But why didn't the mother simply say, 'Your room is very messy and untidy'? This utterance would accurately describe the state of the room. The value of the metaphor in this example is that it not only conveys the state of the room, but also, as the mother intends, her displeasure at this state. We would therefore expect that metaphors, in comparison to literal statements, would activate the mentalizing area and this expectation has been confirmed in the study of Bottini *et al.* (1994). Irony (e.g. 'Peter is well read. He's even heard of Shakespeare') is an even more extreme example than metaphor since the listener has to recognize that the speaker intends to convey a meaning opposite to the literal content of the words (i.e. Peter is not at all well read). In such cases the meaning is decoupled from the words. We are not aware of any imaging study, but we would predict that the understanding of sarcasm or irony would activate the mentalizing network.

One aspect of pragmatics that has received little attention to date is the initiation of communication by calling someone's name or by gazing at them intently. These are sometimes referred to as 'ostensive' signals. Such stimuli normally signal the intention to communicate and therefore 'guarantee relevance' in Sperber and Wilson's terminology. The effects of such ostensive signals were examined in a recent neuroimaging study (Kampe *et al.* 2003). Subjects were asked to respond to a rare target, while they viewed a series of faces with direct or averted gaze (versus scrambled faces) or listened to voices calling either the subject's own name or another name (versus scrambled voices). The results showed that independent of modality the initiation of communication activated two components of the mentalizing system, the MPFC and temporal poles. This study is consistent with other neuroimaging studies of pragmatics in demonstrating that the relationship between communicative and mentalizing functions is remarkably close.

5. CONCLUSIONS

We conclude, from the facts available to date, that the region of the MPFC associated with mentalizing tasks is activated whenever people are attending to certain states of the self or others. These states, which are usually referred to as mental states, must be decoupled from reality. To understand the response to pain, whether it is my pain or someone else's pain, I must represent, not the noxiousness of the stimulus, but how I or the other person perceive the pain. Likewise, it is not the unpleasantness of the picture that determines our emotional response to it, but the unpleasantness we feel. Such decoupled representations are also needed for mentalizing. What determines our behaviour is not the state of the world, but our beliefs about the state of the world. Activity in the MPFC is connected with the creation of these decoupled representations of beliefs about the world. In the case of false beliefs there is a discrepancy between the belief and the actual state of the world. However, we are not claiming that activity in the MPFC signals these discrepancies. This would be equivalent to error detection. We are claiming that the MPFC is equally active when true beliefs are involved. This is because beliefs may or may not map onto the actual state of the world. This would also be true for other mental states such as wishes, intentions and pretence. Activity in the MPFC signals that these representations are decoupled from the real world to which they may or may not correspond. Thus, the role of this particular region of the MPFC would be analogous to that of the more posterior region (rCZp) where neuronal activity signals the existence of response conflict or multiple response possibilities rather than errors (Petit *et al.* 1998; Botvinick *et al.* 1999).

Mentalizing is not only about representing our own thoughts, feelings and beliefs as distinct from reality. It is also about representing the mental states of other people. Clearly, other components of the mentalizing system need to supply the content of these thoughts, feelings and beliefs and their relation to people's actions. This knowledge is supplied partly from our knowledge of the world based on past experience applied to the current situation and partly from our observations and predictions about people's current behaviour (STS). Both types of knowledge help to understand the content of mental states and their relation to actions, and may be accessible via temporal poles and the STS. By identifying the roles of the regions in this way it should be possible to link the various precursors of mentalizing that emerge during the first 4 years of life to specific components of the brain's mature mentalizing system. This will have to await the development of suitable methods for using fMRI techniques to study infants and young children.

The authors are grateful to Sarah Blakemore, Paul Fletcher, Josef Perner and Beate Sodian for their comments on this paper. This work was supported by MRC grant no. G961 7036 to U.F. C.D.F. is supported by the Wellcome Trust.

REFERENCES

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

Allman, J. M., Hakeem, A., Erwin, J. M., Nimchinsky, E. & Hof, P. 2001 The anterior cingulate cortex—the evolution of an interface between emotion and cognition. *Unity Know. Converg. Nat. Hum. Sci.* **935**, 107–117.

Avis, J. & Harris, P. L. 1991 Belief–desire reasoning among Baka children—evidence for a universal conception of mind. *Child Dev.* **62**, 460–467.

Bachevalier, J., Meunier, M., Lu, M. X. & Ungerleider, L. G. 1997 Thalamic and temporal cortex input to medial prefrontal cortex in rhesus monkeys. *Exp. Brain Res.* **115**, 430–444.

Baldwin, D. A. & Moses, L. J. 1996 The ontogeny of social information gathering. *Child Dev.* **67**, 1915–1939.

Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J. & Snyder, A. 2001 Anterior cingulate cortex and response conflict: effects of response modality and processing domain. *Cereb. Cortex* **11**, 837–848.

Barnes, C. L. & Pandya, D. N. 1992 Efferent cortical connections of multimodal cortex of the superior temporal sulcus in the rhesus monkey. *J. Comp. Neurol.* **318**, 222–244.

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

Bertenthal, B. I., Proffitt, D. R. & Cutting, J. E. 1984 Infant sensitivity to figural coherence in biomechanical motions. *J. Exp. Child Psychol.* **37**, 213–230.

Berthoz, S., Armony, J. L., Blair, R. J. R. & Dolan, R. J. 2002 An fMRI study of intentional and unintentional (embarrassing) violations of social norms. *Brain* **125**, 1696–1708.

Blakemore, J.-S., Wolpert, D. M. & Frith, C. D. 1998 Central cancellation of self-produced tickle sensation. *Nature Neurosci.* **1**, 635–639.

Bloom, P. 2000 *How children learn the meanings of words*. Cambridge, MA: MIT Press.

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

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., Frackowiak, R. S. J. & Frith, C. D. 1994 The role of the right hemisphere in the interpretation of figurative aspects of language: a positron emission tomography activation study. *Brain* **117**, 1241–1253.

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S. & Cohen, J. D. 1999 Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* **402**, 179–181.

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

Bush, G., Luu, P. & Posner, M. I. 2000 Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci.* **4**, 215–222.

Butterworth, G. 1991 The ontogeny and phylogeny of joint visual attention. In *Natural theories of mind: evolution, development and simulation of everyday mindreading* (ed. A. Whiten), pp. 223–232. Cambridge, MA: Blackwell.

Butterworth, G. & Jarrett, N. 1991 What minds have in common is space—spatial mechanisms serving joint visual-attention in infancy. *Br. J. Dev. Psychol.* **9**, 55–72.

Byrne, R. W. & Whiten, A. 1988 *Machiavellian intelligence*. Oxford: Clarendon Press.

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

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., Brammer, M. J. & David, A. S. 2001 Cortical substrates for the perception of face actions: an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). *Cogn. Brain Res.* **12**, 233–243.

Caron, A. J., Caron, R., Roberts, J. & Brooks, R. 1997 Infant sensitivity to deviations in dynamic facial–vocal displays: the role of eye regard. *Devl Psychol.* **33**, 802–813.

Carpenter, M., Nagell, K. & Tomasello, M. 1998 Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* **63**, 176.

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

Chan, D., Fox, N. C., Scahill, R. I., Crum, W. R., Whitwell, J. L., Leschziner, G., Rossor, A. M., Stevens, J. M., Cipolotti, L. & Rossor, M. N. 2001 Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. *Ann. Neurol.* **49**, 433–442.

Chao, L. L., Haxby, J. V. & Martin, A. 1999 Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neurosci.* **2**, 913–919.

Cheney, D. L. & Seyfarth, R. M. 1990 *How monkeys see the world: inside the mind of another species*. University of Chicago Press.

Clements, W. A. & Perner, J. 1994 Implicit understanding of belief. *Cogn. Dev.* **9**, 377–395.

Clements, W. A. & Perner, J. 2001 When actions really do speak louder than words, but only explicitly; young children's understanding of false belief in action. *Br. J. Devl Psychol.* **19**, 413–432.

Corbetta, M. 1993 Positron emission tomography as a tool to study human vision and attention. *Proc. Natl Acad. Sci. USA* **90**, 10 901–10 903.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P. & Shulman, G. L. 2000 Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neurosci.* **3**, 292–297.

Crichton, M. T. & Lange Kuettner, C. 1999 Animacy and propulsion in infancy: tracking, waving and reaching to self-propelled and induced moving objects. *Devl Sci.* **2**, 318–324.

Critchley, H. D., Corfield, D. R., Chandler, M. P., Mathias, C. J. & Dolan, R. J. 2000 Cerebral correlates of autonomic cardiovascular arousal: a functional neuroimaging investigation in humans. *J. Physiol.* (*Lond.*) **523**, 259–270.

Critchley, H. D., Mathias, C. T. & Dolan, R. J. 2001 Neuroanatomical basis for first- and second-order representations of bodily states. *Nature Neurosci.* **4**, 207–212.

Csibra, G. 2003 Teleological and referential understanding of action in infancy. *Phil. Trans. R. Soc. Lond.* B **358**, 447–458. (DOI 10.1098/rstb.2002.1235.)

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

Devinsky, O., Morrell, M. J. & Vogt, B. A. 1995 Contributions of anterior cingulate cortex to behavior. *Brain* **118**, 279–306.

Downar, J., Crawley, A. P., Mikulis, D. J. & Davis, K. D. 2000 A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neurosci.* **3**, 277–283.

Duncan, J. & Owen, A. M. 2000 Dissociative methods in the study of frontal lobe function. In *Control of cognitive processes*, vol. XVIII (ed. S. Monsell & J. Driver), pp. 567–576. Cambridge, MA: 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.

Fink, G. R., Markowitsch, H. J., Reinkemeier, M., Bruckbauer, T., Kessler, J. & Heiss, W. D. 1996 Cerebral rep-

resentation of one's own past: neural networks involved in autobiographical memory. *J. Neurosci.* **16**, 4275–4282.
Flavell, J. H., Everett, B. A., Croft, K. & Flavell, E. R. 1981 Young children's knowledge about visual-perception—further evidence for the level 1–level 2 distinction. *Devl Psychol.* **17**, 99–103.
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* **44**, 283–296; **57**, 109–128.
Frith, C. D. & Frith, U. 1999 Interacting minds—a biological basis. *Science* **286**, 1692–1695.
Funnell, E. 2001 Evidence for scripts in semantic dementia. Implications for theories of semantic memory. *Cogn. Neuropsychol.* **18**, 323–341.
Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U. & Frith, C. D. 2000 Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia* **38**, 11–21.
Gallagher, H. L., Jack, A. I., Roepstorff, A. & Frith, C. D. 2002 Imaging the intentional stance. *NeuroImage* **16**, 814–821.
Gergely, G., Knadasdy, Z., Csibra, G. & Biro, S. 1995 Taking the international stance at 12 months of age. *Cognition* **56**, 165–193.
Goel, V., Grafman, J. N. S. & Hallett, M. 1995 Modelling other minds. *NeuroReport* **6**, 1741–1746.
Gopnik, A. & Wellman, H. M. 1994 The theory theory. In *Mapping the mind: domain specificity in cognition and culture* (ed. L. A. Hirschfeld & S. A. Gelman), pp. 257–293. New York: Cambridge University Press.
Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M. & Cohen, J. D. 2001 An fMRI investigation of emotional engagement in moral judgment. *Science* **293**, 2105–2108.
Grezes, J., Costes, N. & Decety, J. 1998 Top-down effect of strategy on the percption of human biological motion: a PET investigation. *Cogn. Neuropsychol.* **15**, 553–582.
Grezes, J., Fonlupt, P., Bertenthal, B., Delon-Martin, C., Segebarth, C. & Decety, J. 2001 Does perception of biological motion rely on specific brain regions? *NeuroImage* **13**, 775–785.
Grice, H. P. 1957 Meaning. *Phil. Rev.* **66**, 377–388.
Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G. & Blake, R. 2000 Brain areas involved in perception of biological motion. *J. Cogn. Neurosci.* **12**, 711–720.
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. *Proc. Natl Acad. Sci. USA* **98**, 4259–4264.
Happé, F. G. E. 1994 An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped and normal children and adults. *J. Autism Devl Disord.* **24**, 129–154.
Harris, P. L. 1991 The work of the imagination. In *Natural theories of mind: evolution, development and simulation of everyday mindreading* (ed. A. Whiten), pp. 283–304. Cambridge, MA: Blackwell.
Heider, F. & Simmel, M. 1944 An experimental study of apparent behaviour. *Am. J. Psychol.* **57**, 243–259.
Heyes, C. M. 1998 Theory of mind in nonhuman primates. *Behav. Brain Sci.* **21**, 101–134.
Hoffman, E. A. & Haxby, J. V. 2000 Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neurosci.* **3**, 80–84.
Hogrefe, G. J., Wimmer, H. & Perner, J. 1986 Ignorance versus false belief—a developmental lag in attribution of epistemic states. *Child Dev.* **57**, 567–582.
Hood, B. M., Willen, J. D. & Driver, J. 1998 Adult's eyes trigger shifts of visual attention in human infants. *Psychol. Sci.* **9**, 131–134.
Johnson, M. H. & Morton, J. 1991 *Biology and cognitive development: the case of face recognition*. Oxford: Blackwell.
Johnson, S. C. 2003 Detecting agents. *Phil. Trans. R. Soc. Lond.* B **358**, 549–559. (DOI 10.1098/rstb.2002.1237.)
Kampe, K., Frith, C. D. & Frith, U. 2003 'Hey John': signals conveying communicative intention towards the self activate brain regions associated with mentalizing regardless of modality. *J. Neurosci.* (In the press.)
Lane, R. D. 2000 Neural correlates of conscious emotional experience. In *Cognitive neuroscience of emotion* (ed. R. D. Lane & L. Nadel), pp. 345–370. London: Oxford University Press.
Lane, R. D., Fink, G. R., Chua, P. M. & Dolan, R. J. 1997 Neural activation during selective attention to subjective emotional responses. *NeuroReport* **8**, 3969–3972.
Lane, R. D., Reiman, E. M., Axelrod, B., Yun, L. S., Holmes, A. & Schwartz, G. E. 1998 Neural correlates of levels of emotional awareness: evidence of an interaction between emotion and attention in the anterior cingulate cortex. *J. Cogn. Neurosci.* **10**, 525–535.
Lee, A. C. H., Robbins, T. W., Graham, K. S. & Owen, A. M. 2002 'Pray or prey?' Dissociation of semantic memory retrieval from episodic memory processes using positron emission tomography and a novel homophone task. *Neuroimage* **16**, 724–735.
Legerstee, M. 1992 A review of the animate/inanimate distinction in infancy. *Early Dev. Parent.* **1**, 59–67.
Lempers, J. D., Flavell, E. R. & Flavell, J. H. 1977 The development in very young children of tacit knowledge concerning visual perception. *Genet. Psychol. Monogr.* **95**, 3–53.
Leslie, A. M. 1987 Pretence and representation: the origins of 'theory of mind'. *Psychol. Rev.* **94**, 412–426.
Leslie, A. M. 1994 Pretending and believing: issues in the theory of mind TOMM. *Cognition* **50**, 211–238.
Lillard, A. 1998 Ethnopsychologies: cultural variations in theories of mind. *Psychol. Bull.* **123**, 3–32.
McCabe, K., Houser, D., Ryan, L., Smith, V. & Trouard, T. 2001 A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA* **98**, 11 832–11 835.
McGuire, P. K., Paulesu, E., Frackowiak, R. S. J. & Frith, C. D. 1996 Brain activity during stimulus independent thought. *NeuroReport* **7**, 2095–2099.
Maguire, E. A. & Mummery, C. J. 1999 Differential modulation of a common memory retrieval network revealed by positron emission tomography. *Hippocampus* **9**, 54–61.
Maguire, E. A., Frith, C. D. & Morris, R. G. M. 1999 The functional neuroanatomy of comprehension and memory: the importance of prior knowledge. *Brain* **122**, 1839–1850.
Maguire, E. A., Mummery, C. J. & Buchel, C. 2000 Patterns of hippocampal–cortical interaction dissociate temporal lobe memory subsystems. *Hippocampus* **10**, 475–482.
Maguire, E. A., Vargha-Khadem, F. & Mishkin, M. 2001 The effects of bilateral hippocampal damage on fMRI regional activations and interactions during memory retrieval. *Brain* **124**, 1156–1170.
Maquet, P., Schwartz, S., Passingham, R. & Frith, C. D. 2003 Sleep-related consolidation of a visuo-motor skill: brain mechanisms as assessed by fMRI. *J. Neurosci.* (In the press.)
Maratos, E. J., Dolan, R. J., Morris, J. S., Henson, R. N. A. & Rugg, M. D. 2001 Neural activity associated with episodic memory for emotional context. *Neuropsychologia* **39**, 910–920.
Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E. & Flavell, T. H. 1974 The early

development of inferences about the visual percepts of others. *Child Dev* **45**, 357–366.
Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L. & Mehler, J. 1993 The cortical representation of speech. *J. Cogn. Neurosci.* **5**, 467–479.
Meltzoff, A. N. 1995 Understanding the intentions of others—re-enactment of intended acts by 18-month-old children. *Devl Psychol.* **31**, 838–850.
Moran, M. A., Mufson, E. J. & Mesulam, M. M. 1987 Neural inputs into the temporopolar cortex of the rhesus monkey. *J. Comp. Neurol.* **256**, 88–103.
Nakamura, K. (and 10 others) 2000 Functional delineation of the human occipito-temporal areas related to face and scene processing—a PET study. *Brain* **123**, 1903–1912.
Nakamura, K. (and 10 others) 2001 Neural substrates for recognition of familiar voices: a PET study. *Neuropsychologia* **39**, 1047–1054.
Nimchinsky, E. A., Gilissen, E., Allman, J. M., Perl, D. P., Erwin, J. M. & Hof, P. R. 1999 A neuronal morphologic type unique to humans and great apes. *Proc. Natl Acad. Sci. USA* **96**, 5268–5273.
Noppeney, U. & Price, C. J. 2002*a* A PET study of stimulus- and task-induced semantic processing. *NeuroImage* **15**, 927–935.
Noppeney, U. & Price, C. J. 2002*b* Retrieval of visual, auditory, and abstract semantics. *NeuroImage* **15**, 917–926.
O'Neill, D. 1996 Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Dev.* **67**, 659–677.
O'Neill, D. & Gopnik, A. 1991 Young children's ability to identify the sources of their beliefs. *Devl Psychol.* **27**, 390–397.
Onishi, K. H. & Baillargeon, R. 2002 15-month-old infants' understanding of false belief. *Int. Conf. Infant Studies, Toronto, Canada.*
Perner, J. & Wimmer, H. 1985 'John thinks that Mary thinks that...': attribution of second-order beliefs by 5- to 10-year-old children. *J. Exp. Child Psychol.* **39**, 437–471.
Perner, J., Leekam, S. R. & Wimmer, H. 1987 2-year-olds difficulty with false belief—the case for a conceptual deficit. *Br. J. Devl Psychol.* **5**, 125–137.
Petit, L., Courtney, S. M., Ungerleider, L. G. & Haxby, J. V. 1998 Sustained activity in the medial wall during working memory delays. *J. Neurosci.* **18**, 9429–9437.
Petrovic, P. & Ingvar, M. 2002 Imaging cognitive modulation of pain processing. *Pain* **95**, 1–5.
Phan, K. L., Wager, T., Taylor, S. F. & Liberzon, I. 2002 Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage* **16**, 331–348.
Phillips, A. T., Wellman, H. M. & Spelke, E. S. 2002 Infants' ability, to connect gaze and emotional expression to intentional action. *Cognition* **85**, 53–78.
Picard, N. & Strick, P. L. 1996 Motor areas of the medial wall: a review of their location and functional activation. *Cereb. Cortex* **6**, 342–353.
Poulin-Dubois, D., Tilden, J., Sodian, B., Metz, U. & Schöppner, B. 2003 Implicit understanding of the seeing–knowing relation in 14- to 24-month-old children. (Submitted.)
Povinelli, D. J. & Bering, J. M. 2002 The mentality of apes revisited. *Curr. Direct. Psychol. Sci.* **11**, 115–119.
Povinelli, D. J. & deBlois, S. 1992 Young children's (*Homo sapiens*) understanding of knowledge formation in themselves and others. *J. Comp. Psychol.* **106**, 228–238.
Povinelli, D. J. & Preuss, T. M. 1995 Theory of mind: evolutionary history of a cognitive specialization. *Trends Neurosci.* **18**, 418–424.
Premack, D. & Woodruff, G. 1978 Does the chimpanzee have a theory of mind? *Behav. Brain. Sci.* **1**, 515–526.
Price, C. J., Moore, C. J., Humphreys, G. W. & Wise, R. J. S. 1997 Segregating semantic from phonological processes during reading. *J. Cogn. Neurosci.* **9**, 727–733.
Puce, A. & Perrett, D. 2003 Electrophysiology and brain imaging of biological motion. *Phil. Trans. R. Soc. Lond.* B **358**, 435–445. (DOI 10.1098/rstb.2002.1221.)
Puce, A., Allison, T., Bentin, S., Gore, J. C. & McCarthy, G. 1998 Temporal cortex activation in humans viewing eye and mouth movements. *J. Neurosci.* **18**, 2188–2199.
Rainville, P., Hofbauer, R. K., Paus, T., Duncan, G. H., Bushnell, M. C. & Price, D. D. 1999 Cerebral mechanisms of hypnotic induction and suggestion. *J. Cogn. Neurosci.* **11**, 110–125.
Repacholi, B. M. 1998 Infants' use of attentional cues to identify the referent of another person's emotional expression. *Devl Psychol.* **34**, 1017–1025.
Rizzolatti, G., Fogassi, L. & Gallese, V. 2002 Motor and cogntiive functions of the ventral premotor cortex. *Curr. Opin. Neurobiol.* **12**, 149–154.
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.
Ruffman, T., Perner, J., Naito, M., Parkin, L. & Clements, W. A. 1998 Older (but not younger) siblings facilitate false belief understanding. *Devl Psychol.* **34**, 161–174.
Russell, J. 1996 *Agency: its role in mental development*. Oxford: Erlbaum.
Schank, R. C. & Abelson, R. P. 1977 *Scripts, plans, goals and understanding: an inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.
Schultz, R. T., Grelotti, D. J., Klin, A., Kleinman, J., Van der Gaag, C., Marois, R. & Skudlarski, P. 2003 The role of the fusiform face area in social cognition: implications for the pathobiology of autism. *Phil. Trans. R. Soc. Lond.* B **358**, 415–424. (DOI 10.1098/rstb.2002.1208.)
Shatz, M., Wellman, H. M. & Silber, S. 1983 The acquisition of mental verbs—a systematic investigation of the 1st reference to mental state. *Cognition* **14**, 301–321.
Sodian, B. & Thoermer, C. 2003 Infants' understanding of looking, pointing and reaching as cues to goal-directed action. (Submitted.)
Spelke, E. S., Phillips, A. & Woodward, A. L. 1995 Infants' knowledge of object motion and human action. In *Causal cognition: a multidisciplinary debate. Symp. Fyssen Foundation* (ed. D. Sperber & D. Premack), pp. 44–78. New York: Clarendon Press/Oxford University Press.
Sperber, D. & Wilson, D. 1995 *Relevance: communication and cognition*. Oxford: Blackwell Scientific.
Sullivan, K., Zaitchik, D. & Tagerflusberg, H. 1994 Preschoolers can attribute 2nd-order beliefs. *Devl Psychol.* **30**, 395–402.
Tulving, E. 1985 Memory and consciousness. *Can. Psychol.* **26**, 1–12.
Vandenberghe, R., Price, C., Wise, R., Josephs, O. & Frackowiak, R. S. J. 1996 Functional anatomy of a common semantic system for words and pictures. *Nature* **383**, 254–256.
Vandenberghe, R., Nobre, A. C. & Price, C. J. 2002 The response of left temporal cortex to sentences. *J. Cogn. Neurosci.* **14**, 550–560.
Varley, R., Siegal, M. & Want, S. C. 2001 Severe impairment in grammar does not preclude theory of mind. *Neurocase* **7**, 489–493.
Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happe, 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. & Estes, D. 1986 Early understanding of mental entities—a re-examination of childhood realism. *Child Dev.* **57**, 910–923.

Wellman, H. M., Phillips, A. T. & Rodriguez, T. 2000 Young children's understanding of perception, desire, and emotion. *Child Dev.* **71**, 895–912.

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

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

Woodward, A. L. 1998 Infants selectively encode the goal object of an actor's reach. *Cognition* **69**, 1–34.

Woodward, A. L., Sommerville, J. A. & Guajardo, J. J. 2001 How infants make sense of intentional action. In *Intentions and intentionality: foundations of social cognition* (ed. B. F. Malle & L. J. Moses), pp. 149–169. Cambridge, MA: MIT Press.

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C. & Frackowiak, R. S. 1991 A direct demonstration of functional specialization in human visual cortex. *J. Neurosci.* **11**, 641–649.

Zysset, S., Huber, O., Ferstl, E. & von Cramon, D. Y. 2002 The anterior frontomedian cortex and evaluative judgment: an fMRI study. *NeuroImage* **15**, 983–991.

GLOSSARY

ACC: anterior cingulate cortex
cCZ: caudal cingulate zone
fMRI: functional magnetic resonance imaging
rCZp: rostral cingulate zone, posterior part
rCZa: rostral cingulate zone, anterior part
MPFC: medial prefrontal cortex
STS: superior temporal sulcus