

Neuroimaging of empathy

Tom F. D. Farrow

Academic Clinical Psychiatry, University of Sheffield

12.1 Introduction

Empathy, the ability or process 'to identify with and understand another's situation, feelings and motives' would initially appear an unlikely candidate for neuroimaging research. Being aware of, and interpreting, other's behaviour on an emotional level is likely to be recently evolved and hence a 'high-level' cognitive process. Such complex brain processes are generally considered as unlikely to have a dedicated brain region serving them, or to be easy to isolate for examination.

This chapter will describe how empathy has been dissected into a set of component cognitive processes, how brain imaging researchers have designed experiments to examine various combinations of these components, and what these findings may tell us about empathy's neurophysiological basis.

12.2 A neuroimaging primer

It may be useful to begin by summarizing the field of neuroimaging, and highlighting which aspects may be of relevance. Structural neuroimaging concerns the physical size and integrity of brain tissue, and in as much as there may be a relationship between size and function, if we could identify brain regions which were part of an empathy system or circuit, then investigating their size or integrity may be informative (presuming that we can objectively measure subjects' behavioural empathic levels). Functional neuroimaging utilizes surrogate markers (normally regional blood flow) to infer which parts of the brain are 'active' whilst a specific task or mental process is undertaken. It is assumed that increases in blood flow are associated with increased functional activity, though such activity may represent the firing of either excitatory or inhibitory neurones. Positron emission tomography (PET; a type of functional neuroimaging) utilizes a fast decaying

radioactive dye (often oxygen-15) injected into the body to track the brain's regional blood supply. By repetitively performing a single task over a relatively long time (3–5 min) it is possible to see where the blood is directed and hence, by inference, where the brain is activated. Functional magnetic resonance imaging (fMRI) is a non-invasive neuroimaging technique that relies on the differing properties of oxy-haemoglobin and deoxy-haemoglobin (oxygenated and deoxygenated blood) in a magnetic field. This allows mapping of the distribution of oxygenated blood (and by inference neuronal activity) in response to a particular task. fMRI scans are designed such that subjects perform contrasting tasks, the demands of which are matched as far as possible, so as only to differ by the specific cognitive process of interest. By subtracting the 'baseline' task from the 'active' task, many of the background and other processes unrelated to the cognitive process of interest (e.g. perceiving the noise of the scanner) are excluded. Tasks in an fMRI scanner are typically performed in an alternating fashion with each block lasting 10–20 s, or as an 'event-related' design where neural responses to single tasks (each typically <3 s duration) are recorded and summated (c.f. PET imaging). In fMRI, brain activations to complex psychological paradigms such as empathy (as opposed to, say, visual cortex activation when viewing a flashing chequerboard pattern) are rarely strong enough for reliable and meaningful areas to be identified in single subjects. It is therefore more common for reported activations to be group-averaged. The error inherent in the reported neuroanatomical location of activations is therefore heavily influenced by inter-subject variability in brain anatomy (commonly quoted to be in the order of 8–12 mm). Areas of activation or structural change are reported by two methods – neuroanatomical name (i.e. brain region) and Brodmann's Area (BA; a functional and cytoarchitectonic parcellation of brain grey matter, similar but not identical to the brain's gyral folds). For the non-specialist reader, areas relevant to this chapter are shown in Figure 12.1.

Functional neuroimaging techniques such as PET and fMRI also allow examination of brain changes within a group over time, such as those changes induced by psychotherapeutic or pharmacological intervention and accompanying symptom resolution. Ultimately, successful functional neuroimaging of complex psychological brain processes is almost entirely reliant on the ability of the task performed in the scanner to engage the cognitive process of interest.

12.3 Component cognitive processes relevant to empathy

For the purposes of designing functional neuroimaging paradigms to tap empathy, proposed component cognitive processes need to be elucidated. Therefore, any

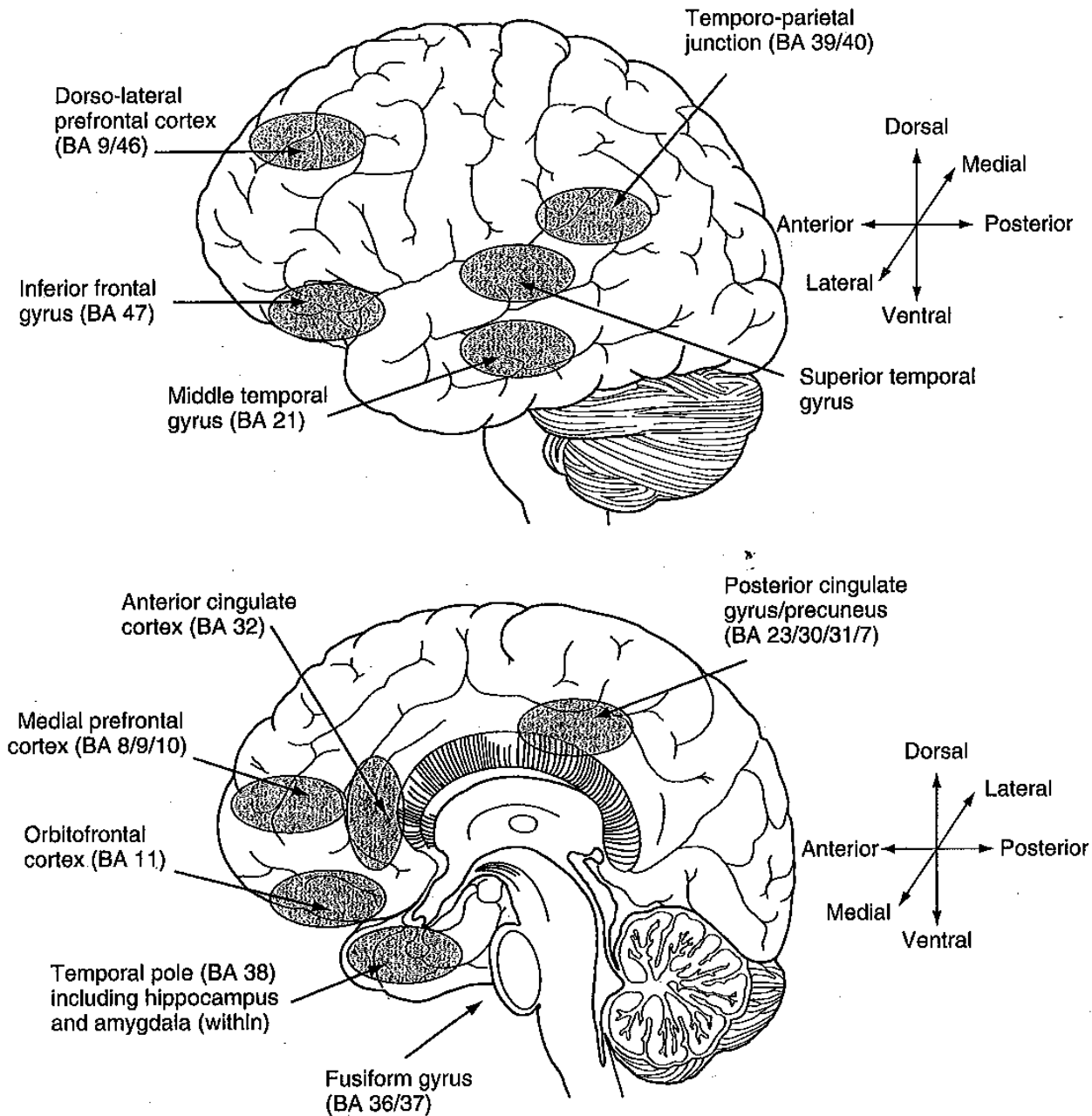


Figure 12.1. View of the brain from the left showing lateral (top picture) and midline (bottom picture) surfaces. Areas relevant to this chapter are identified. These are identified for the non-specialist reader and do not represent the shape or extent of each region

neuroimaging study which has examined any of these component cognitive processes (in isolation or in groups) will also be of interest in understanding empathy's neurophysiological basis. These component processes and areas of interest include:

1. Attending to socially relevant stimuli including facial (expression) and body (posture, movements) perception. This perceptual ability is central to social

cognition – ‘the mental processes underlying social perception and social judgements’. Empathy is an aspect of social cognition and dependent on our interaction with, and perception of, the people around us. When based on deliberation and intent (and thereby requiring motivational and attentional systems), empathic judgements may also revolve around such issues as the morality of the situation in question.

2. Theory of Mind (ToM) – ‘the attribution of independent mental states to self and others in order to explain and predict others’ behaviour’. This capacity is often described as seeing the world from someone else’s point of view or perspective. A common task in ToM-based paradigms is the ‘false belief task’. This is often classified according to whether the ToM attribution is first order (‘A thinks X’) or second order (‘A thinks B thinks X’).
3. Self-awareness or the ability to self-reflect is required to understand and project emotions onto others.
4. Mirror neurones. These specialist cells are activated in association with, for example, moving one’s hand but also when watching another’s hand movements. The discovery that certain regions of the brain activate in response not only to performing a task, but also to watching another person perform that task, suggests that we may interpret others’ actions (and emotions) by simulating them in our own brains (see Chapter 23 on the perception-action model and Chapter 24 on simulation theory).
5. Emotion processing and emotional and affective intuition. Empathy can also be a more automatic, subconscious response and involve little if any deliberate reasoning. This automatic cognitive response, which is often associated with a measurable autonomic (bodily, visceral) component, may involve separate brain structures to those involved in more ‘conscious’ empathic cognitions.

Any attempt to use neuroimaging to investigate empathy as a whole should take the following points into consideration:

1. There is almost certainly no ‘empathy centre’, but rather integrated activity in groups or networks of interconnected brain regions.
2. Empathy is not a unitary concept and can probably be elicited by a number of methods (e.g. with or without a face-processing component). A ‘definitive’ empathy paradigm therefore does not exist. This is likely to lead to different brain activations to empathy paradigms, dependent on which component processes are evoked.
3. An MRI scanner is a hostile and ‘ecologically invalid’ environment. Even if paradigms are capable of evoking ‘active empathy’, the real-world transferability of this is debatable. The subtle distinctions between empathy, sympathy and ‘detached concern’, and the possibility that the cognitive and affective

components of empathy are dissociable mean that results should be treated with caution and interpreted in the context of the cognitive paradigm that was used to evoke them.

Results of functional neuroimaging of empathy paradigms can be interpreted in the light of hypothesized regions (see Section 12.4), but due to the preliminary nature of research in this field, post hoc interpretation of non-hypothesized regions may also shed light on understanding empathy's neurophysiological basis.

12.4 Neuroimaging studies of component parts of empathy

12.4.1 Socially relevant stimuli

Perception of socially relevant stimuli has been mainly localized to the temporal lobes (Adolphs, 2001). Areas within the posterior superior temporal sulcus [Brodmann's Area (BA) 39] are activated in response to biologically and socially salient visual motion stimuli (Allison *et al.*, 2000). However, face perception, as possibly the most important flag of social relevance, has been reported to be divided between the fusiform gyrus (BA 36/37; invariant features, e.g. identity) and the posterior superior temporal gyrus (BA 39/40; expression and gaze; Haxby *et al.*, 2000).

Empathy may involve little if any deliberate reasoning (Greene & Haidt, 2002) or may be a conscious cognitive choice. The decision as whether to empathize (when empathy is based upon deliberation and intent) may be heavily dependent on the perceived morality (reason, emotion and affective intuition) of the situation. Greene and colleagues in an fMRI study (Greene *et al.*, 2001) used moral dilemmas as probes, and identified three regions: medial frontal gyrus (BA 9/10), posterior cingulate gyrus (BA 31) and bilateral angular gyrus (BA 39; in the temporo-parietal junction area) as differentiating moral from non-moral judgments. Another study of the neural basis of moral emotions (Moll *et al.*, 2002) replicated (amongst other regions) medial prefrontal cortex (BA 10/46 and 9) activation, further associating this region with social behaviours.

12.4.2 Theory of Mind and social cognition

Functional imaging studies of healthy volunteers performing Theory of Mind (ToM) tasks (Fletcher *et al.*, 1995; Gallagher *et al.*, 2000; Goel *et al.*, 1995) have all implicated a specific region of left prefrontal cortex (BA 8/9/10). Goel and colleagues in a PET study had subjects deciding whether Christopher Columbus (the fifteenth century European explorer) could infer the function of an object from a picture (i.e. seeing the world from Christopher Columbus' perspective). Fletcher and colleagues in another PET study had subjects reading, and answering

questions about, short stories, half of which required ToM to correctly respond. Gallagher and colleagues in an fMRI study repeated the paradigm used by Fletcher and colleagues and additionally had subjects viewing cartoons which required ToM to interpret. The consistency of left prefrontal cortex activation in response to these radically different paradigms and across imaging modalities is striking. While dorso-medial prefrontal cortex (DMPFC) is associated with tasks emphasizing self reference and ToM tasks (Fletcher *et al.*, 1995; Gallagher & Frith, 2003; Goel *et al.*, 1995), dorso-lateral prefrontal cortex (DLPFC) is associated with working memory and executive functions. Other areas activated by these paradigms included posterior cingulate cortex [BA 31; associated with internal (visceral) state monitoring and self behaviour evaluation], precuneus (BA 7; associated with episodic memory retrieval and mental imagery), left temporal lobe (BA 21, 38) and temporo-parietal junction/posterior superior temporal gyrus/angular gyrus (BA 39/40). A PET study repetition of Fletcher and colleagues' 1995 study (Happé *et al.*, 1996) in patients with Asperger's syndrome (who show ToM deficits) found no activity in the previously reported left prefrontal region with, rather, an adjacent, more ventral area of medial prefrontal cortex (BA 9/10) being activated.

ToM has been further subdivided into 'hot' and 'cold' cognition or reasoning (Goel & Dolan, 2003) differentiating the inference of others' epistemic states (beliefs, knowledge, focus of attention) from that of others' affective states (emotions, preferences, beneficent or hostile intentions). Goel and Dolan (2003) reported that a reciprocal prefrontal activation pattern exists as a function of emotional saliency. Specifically, whereas 'cold' reasoning activated DLPFC and suppressed ventro-medial prefrontal cortex (VMPFC), the reverse occurred in response to 'hot' cognition. It is noteworthy however in this study, that whilst the DLPFC activations were left lateralized, the VMPFC activations were midline or bilateral. A review of functional neuroimaging of ToM (Gallagher & Frith, 2003) identified a set of three areas which are repeatedly activated by, and associated with, ToM paradigms; namely, the anterior paracingulate cortex (a division of medial prefrontal cortex approximately corresponding to BA 9/32), the superior temporal sulcus (STS; BA 39/40) and the temporal poles bilaterally (BA 38). Gallagher and Frith (2003) argue that the anterior paracingulate cortex is involved in 'mentalizing' or is 'the location of the cognitive mechanism underpinning the ability to represent mental states "decoupled" from reality', whereas the STS and temporal poles are involved in abilities that aid mentalizing. While the STS may be involved in the perception of intentional behaviour [biological motion or stimuli which signal the actions and intentions of another individual (Allison *et al.*, 2000)], the temporal poles are involved in the retrieval from memory of personal

(semantic and episodic) experiences. Other possibly relevant cognitive paradigms which have activated the STS include being the target of another's emotions (Wicker *et al.*, 2003a). The temporo-parietal junction (BA 39/40; an area of posterior STS) has been reported to be specific to reasoning about the content of others' mental states (as opposed to another's simple physical presence; Saxe & Kanwisher, 2003). Other regions which may be more related to social tasks in general rather than being prerequisites for ToM (but are sometimes activated in response to ToM paradigms) include the amygdala and orbitofrontal cortex. The amygdala (particularly left) is most frequently associated with fear perception and conditioning, and memory modulation central to emotional processing, whilst the orbitofrontal cortex is associated with reward and decision-making in the context of emotional situations. Lesions of the orbitofrontal cortex may lead to ToM deficits (Stone *et al.*, 1998).

Some lesional data on the role of the amygdala (Fine *et al.*, 2001) and frontal cortex (Rowe *et al.*, 2001) in ToM, and research on whether ToM and general executive functions are interdependent are slightly contradictory however to the previously reported functional neuroimaging findings. Fine and colleagues (2001) report a case study of a 32-year-old patient with a congenital or early-onset left amygdala lesion. Despite being severely impaired in their ability to represent mental states, the patient was unimpaired on a wide range of general executive functioning tests. This appears to contradict the interpretive assumptions of many ToM neuroimaging studies which suggest that ToM is mediated by general executive functioning. Rowe and colleagues (2001) studied 31 patients with unilateral frontal lobe lesions following neurosurgery. All patients exhibited significant impairment on both first- and second-order ToM false belief tests, and on a range of executive functioning tests. However again, as in the study by Fine and colleagues (2001), the executive functioning deficits appeared to be independent of the ToM impairments. This may highlight a possible distinction between what is essential for ToM attributions (as revealed by lesional studies) and brain regions involved in ToM attributions in an 'optimal system' (as revealed by neuroimaging in healthy subjects). A further lesional study examining five patients with bilateral orbitofrontal damage and five patients with left DLPFC damage (Stone *et al.*, 1998) showed that orbitofrontal lesion patients performed similarly to those with Asperger's syndrome (i.e. performing poorly on second-order ToM tasks), but that the DLPFC patients showed no such deficits. As well as a 'false belief' task the authors had a 'faux pas' task, again on which patients with bilateral orbitofrontal damage performed poorly. Faux pas tasks include an element of empathy as they require both an understanding of a false or mistaken belief and an empathic inference of the effect that it has on another person.

12.4.3 Self-awareness

Empathy may also be reliant on self-awareness and the ability to consciously reflect on one's sense of self, as this ability guides our social interactions. An fMRI study (Johnson *et al.*, 2002) had 11 volunteers consciously reflecting on their own traits, abilities and attitudes (e.g. 'I'd rather be alone', 'I have a quick temper') contrasted with factual knowledge judgements (e.g. 'You need water to live', 'Ten seconds is more than a minute'). All 11 subjects individually (as well as in a group activation analysis) activated anterior medial prefrontal cortex (BA 9/10; right lateralized in 5, left lateralized in 3 and midline in 3) and posterior cingulate cortex (BA 23/30/31). The posterior cingulate cortex is often associated with retrieval of episodic autobiographical memories, but in this context its role in the evaluation of emotional salience of stimuli or mediating an interaction between memory retrieval and emotion (Maddock, 1999) may be more relevant. A posterior cingulate/precuneus (BA 7/31) activation was reported in another study to dissociate between interacting with a human or computer partner and 'inferring the intentions of real social partners with whom they are directly interacting and whose behaviour has consequences for their material well-being' (Rilling *et al.*, 2004).

12.4.4 Mirror neurones

Some groups of neurones are called 'mirroring neurones' or 'mirror neurones' because they are activated both by performing and observing an intentional action (Gallese & Goldman, 1998; see also Schulkin, 2000). Neurophysiological evidence would further suggest that these mirror neurones may be active even when crucial parts of the actions are obscured and can only be inferred (Umiltà *et al.*, 2001). Mirror neurones are pertinent to empathy, in the debate as to whether empathy is underpinned by 'simulation theory' or 'theory-theory' (Preston & de Waal, 2002). While simulation theory states that our ability to recognize and reason about other people's states of mind is an example of experience projection, i.e. we know others' minds by simulation, theory-theory states that we employ a theory to make attributions of mental states of others, i.e. our understanding of mind is a framework or a theory analogous to scientific theories. Simulation theory may be particularly relevant in understanding situations which are not easily encoded into language (e.g. emotionally salient ones), thereby suggesting a role for conscious experience in social cognition. Mirror neurones would support simulation theory rather than theory-theory as a mechanism for individuals to detect and interpret conspecifics' mental states and to facilitate a more general mind-reading ability. Mirror neurones are relevant to empathy because they would enable an organism to detect certain mental states of observed conspecifics, possibly a precursor to, or component of, a more general mind-reading ability.

12.4.5 General emotional processing

Individual neuroimaging studies of general emotion have activated a very wide range of areas, probably partly due to heterogeneity in task design, imaging methods and analysis. Meta-analytic reviews (e.g. Phan *et al.*, 2002) have sought to report consistent inter-study findings. The role of medial prefrontal cortex (MPFC; BA 9) in general emotional processing and the anterior cingulate cortex in imagery and emotional tasks requiring cognitive demand are probably the two most robust and reproducible finds with relevance to empathy. A more specific role for the MPFC in internally attended emotional states and the cognitive aspects (e.g. identification and appraisal) of emotional processing is now beginning to emerge. As emotion and social cognition appear to depend on some of the same brain regions, this may explain why social cognition and empathy deficits often coexist in disorders such as autism, schizophrenia, depression and post-traumatic stress disorder (PTSD; Grady & Keightley, 2002).

12.5 Studies of empathy itself

It should initially be re-iterated that a definitive empathy paradigm almost certainly does not exist, due partly to the wide-ranging definition and partly to the fact that different stimuli may evoke empathic responses, but that these may not be identical neurophysiologically. To date there have been few studies which have attempted to neuroimage empathy as a unitary concept, and all have approached the 'challenge' in individual ways. At this early stage of research in the 'field' it is therefore preferable to look for overlaps in findings rather than a consensus.

12.5.1 Emotional mimicry studies

In a hypothesis encompassing a 'seamless integration among perception, socially relevant mimicry [the "chameleon effect"], emotional experience and empathy', one study investigated the existence of a human mirroring system for affective facial expressions, and how this underpins empathy (Leslie *et al.*, 2004). Subjects underwent fMRI whilst passively viewing, actively imitating or independently generating facial expressions. Results suggested that conscious imitation of facial expression is dissociable from unconscious mimicry (hypothesized to underlie empathy) when passively viewing faces and that this latter ability is localized to right ventral premotor cortex [BA 6 (immediately posterior to Broca's area)], thereby providing evidence for a motor theory of empathy.

A very similar paradigm to that used by Leslie and colleagues was used by a different group (Carr *et al.*, 2003), but was approached from a different perspective and thus interpreted in a different light. The only substantive difference between the two studies was that Carr and colleagues used still pictures of faces (or parts

thereof) whereas Leslie and colleagues used short (2-s) video clips. Carr and colleagues' paper postulates a fronto-temporal circuit connected via the insula to the limbic system, thereby characterizing empathy as a process by which action representation modulates emotional activity. The structural neuroanatomical basis and flow of information of the 'action representation' part of this model is explicitly hypothesized as follows:

... the circuit of frontoparietal networks interacting with the superior temporal cortex is critical for action representation. This ... circuit is composed of inferior frontal and posterior parietal neurons that discharge during the execution and also the observation of an action (mirror neurons), and of superior temporal neurons that discharge only during the observation of an action. ... [T]his circuit is critical for imitation and ... within this circuit, information processing ... flow[s] as follows. (i) The superior temporal cortex codes an early visual description of the action and sends this information to posterior parietal mirror neurons ... (ii) The posterior parietal cortex codes the precise kinesthetic aspect of the movement and sends this information to inferior frontal mirror neurons ... (iii) The inferior frontal cortex codes the goal of the action ... (iv) Efferent copies of motor plans are sent from parietal and frontal mirror areas back to the superior temporal cortex, such that a matching mechanism between the visual description of the observed action and the predicted sensory consequences of the planned imitative action can occur. (v) Once the visual description of the observed action and the predicted sensory consequences of the planned imitative action are matched, imitation can be initiated.

A largely overlapping network including the premotor face area, dorsal inferior frontal gyrus (BA 44/45), superior temporal sulcus, insula and amygdala was activated by both observing and imitating emotional facial expressions, with the fronto-temporal network being significantly more activated by imitation than observation.

12.5.2 Disgust

Another study investigating whether there is a common neural basis to understanding and experiencing an emotion (Wicker *et al.*, 2003b) used fMRI to examine subjects' neural response to odorant-induced disgust and viewing faces expressing disgust. Both tasks activated right anterior insula and inferior frontal gyrus. In a similar way therefore to motor mirror neurones, observing an emotion appears to activate a neural representation of experiencing an emotion, thereby providing a mechanism for understanding others' behaviour.

12.5.3 Pain

In a further fMRI study of an understanding/experiencing neural overlap (Singer *et al.*, 2004), pain (self experienced and observing a loved one receiving) was

examined. While bilateral anterior insula and rostral anterior cingulate cortex were common to both tasks, additional areas (sensorimotor cortex, somatosensory cortex and caudal anterior cingulate cortex) were activated exclusively in the experiencing pain condition. The authors conclude that 'only that part of the pain network associated with its affective qualities, but not its sensory qualities, mediates empathy'. Interestingly, the anterior insula and anterior cingulate cortex activations correlated with subjects' individual empathy scores. These findings again suggest that our ability to empathize has evolved from a system for representing our internal bodily states and subjective feeling states. Another study used perceiving other's pain to investigate a central component of empathy, the 'interpersonal sharing of affect' (Jackson *et al.*, 2005). Perceiving and rating other's pain was associated with activation of regions including anterior cingulate cortex (ACC) and insula. Furthermore the change in ACC activity (between viewing painful and non-painful scenarios) was positively correlated with subjects' ratings of others' pain, a possible surrogate measure of the extent of interpersonal empathic engagement. The specific area of ACC activated was the rostro-dorsal region (sometimes referred to as the 'cognitive cingulate' and associated with error monitoring and selecting among competing responses) as opposed to the caudo-ventral anterior cingulate, more associated with autonomic responses. This may be relevant when distinguishing whether an affective (bodily sensation) empathic engagement has occurred.

12.5.4 Empathy and sympathy

In contrast to empathy, sympathy might be defined as 'The affinity, association or relationship between persons wherein whatever affects one similarly affects the other', possibly distinguishing it as an 'intellectual understanding' of another person compared with empathy's 'emotional knowing'. These two concepts have a large overlap however, and differentiating them for the purposes of neuroimaging is extremely challenging. A PET study of a combination of sympathy and empathy, with concomitant skin conductance response (SCR; i.e. affective, bodily arousal) recording (Decety & Chaminade, 2003) had subjects watching individuals recounting sad or neutral stories with congruent or incongruent affect (neutral, happy or sad facial expression). Decety and colleagues reported right inferior parietal lobule (BA 39/40) activation to be associated with 'shared representations' or 'concern for others ... [simulating] the affective experiences of others ... the self [taking] the perspective of others', while general emotional content activated left inferior frontal gyrus and bilateral temporal poles. VMPFC (BA 32) activation and the largest SCR response were associated in this study with social conflict arising from the mismatched story content and facial expression condition. This

latter finding may have resonance with the idea that the strength of empathic response is to some degree modulated by the difference in affect between the 'empathizer' and the subject of their empathy.

12.5.5 Empathy and forgiveness

In an fMRI study of empathy and forgivability (Farrow *et al.*, 2001), healthy subjects were required to make judgements from the perspective of another person as to what would be the most likely explanation for their affective state, or which of two crimes was more forgivable. Forgivability comprises multiple cognitive components, one of which may be the ability to empathize with others, including an aggressor (Denton & Martin, 1998). Empathic judgements activated left medial frontal gyrus (BA 9), left inferior frontal gyrus (BA 47), posterior cingulate/precuneus (BA 31/7) and left middle temporal gyrus (BA 21) whilst forgivability judgements activated left medial frontal gyrus (BA 9/10) and posterior cingulate cortex (BA 31). Both paradigms also activated orbitofrontal cortex (BA 11), possibly due to this region's role in evaluating the relative merits of two options. The most striking difference between the activations to empathic and forgivability judgements was the left middle temporal gyrus activation. While suggesting that attempting to understand others is physiologically distinct from determining the forgivability of their actions, the difference may also be due to the fact that the forgivability scenarios were based around unknown individuals, while the empathic scenarios were based on personal acquaintances. This difference in the paradigm design may highlight an important distinction of whether empathy is differentially applied to known or unknown (simulated) individuals and whether these are associated with different brain substrates.

These empathy and forgivability judgement paradigms have been repeated in patients with PTSD (Farrow *et al.*, 2005) and schizophrenia (Lee *et al.*, 2003). Both studies involved scanning patients on two occasions, 3 to 5 months apart, during which time PTSD patients received a course of cognitive behavioural therapy (CBT) and patients with schizophrenia received 'treatment as usual' (anti-psychotic medication). The latter (Lee *et al.*, 2003) study also conducted the empathy and forgivability paradigms on a new group of 14 healthy subjects on two occasions. Whereas the patients with PTSD and schizophrenia showed increased activation in task-relevant regions (e.g. left medial prefrontal cortex, posterior cingulate gyrus) at the second scan, healthy subjects revealed a *reduced* amount. These early data illustrate the potential for mapping the brain's response to treatment interventions in a systematic manner, in this case using an empathy paradigm, a central component of which (social cognition) is known to be abnormal in both patients with PTSD and schizophrenia (Grady & Keightley, 2002).

12.6 Other relevant neuroimaging studies

Two further neuroimaging studies which may inform our understanding of the neurophysiology of empathy examined social exclusion (Eisenberger *et al.*, 2003) and aggressive behaviour (Pietrini *et al.*, 2000). Social exclusion involves distress at being 'left out' and a lack of the soothing feeling of being in the presence of others. This could be reframed (though the authors of the article in question do not) as a feeling that others are not empathizing with you. In line with results from neuroimaging studies of pain, ACC was activated, and correlated with levels of self-reported distress. ACC activity and its role in conflict monitoring were modulated by right ventral prefrontal cortex, a brain region associated with inhibition of negative affect (Hariri *et al.*, 2000).

The presence of empathy acts as a mitigator of aggressive behaviour (Björkqvist & Österman, 2000), possibly particularly physical aggression. In a PET study of imaginal aggressive behaviour (Pietrini *et al.*, 2000), healthy subjects had associated *decreased* activity (functional deactivation) in VMPFC and orbitofrontal cortex. The possible interpretation of these results as inhibition of empathy towards someone that you are about to hurt is compelling. As further supporting evidence, patients with personality disorders and a history of aggressive behaviour are reported to have reduced ventral prefrontal cortex glucose utilization (Goyer *et al.*, 1994) and grey-matter volumes (Raine *et al.*, 2000).

Finally, lesional data have suggested that right ventro-medial cortex is most consistently associated with empathic deficits (Shamay-Tsoory *et al.*, 2003), but that other neuropsychological functions (such as ToM and cognitive flexibility) and related deficits (e.g. understanding sarcasm; Shamay *et al.*, 2002) are so intimately related that disentangling them is extremely complex.

12.7 Conclusions

Ten years of neuroimaging of the postulated component parts of empathy and five years of increasingly sophisticated 'full' empathy paradigms have begun to provide a consensus as to which brain regions form a core network and which regions may be specific to the subtleties of the individual cognitive probes used. The 'empathic experience' may ultimately be shown to be irreducible, but until we are able to convincingly elicit 'true' empathy in the scanner, the investigation of component parts is a good foundation on which to base further work. Medial prefrontal cortex, posterior cingulate and various temporal lobe regions (superior temporal sulcus/temporo-parietal junction/temporal pole) have all been so frequently and consistently reported as to be considered as 'core'. The roles of (mainly subregions such as) ACC, orbitofrontal cortex, amygdala, insula and precuneus are associated

with specific components or aspects which differentiate between the various interpretations of how to design a paradigm to tap empathy. The connection of superior temporal regions and inferior frontal cortexes to the limbic system via the insula [a critical relay from action representation to emotion (Carr *et al.*, 2003)] may prove to be an important pathway in the neuropathology of many neuropsychiatric disorders with dysempathy as a central feature. Furthermore, studies of ToM in Asperger's syndrome (Happé *et al.*, 1996) and empathy and forgiveness in PTSD (Farrow *et al.*, 2005) and schizophrenia (Lee *et al.*, 2003) would suggest that alterations to the location and extent of activations are present in certain disorders. As stated previously, detailed interpretation of brain activations to empathy paradigms is only meaningful in the context of exactly which task was performed and cannot presently be considered generalizable to empathy as a whole. The concurrent recording of the skin-conductance response and brain activation has recently begun (e.g. Decety & Chaminade, 2003) and may significantly enhance our understanding of how the affective and cognitive components of empathy interact.

REFERENCES

- Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiology*, *11*, 231–239.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, *4*, 267–278.
- Björkqvist, K., & Österman, K. (2000). Social intelligence – empathy = aggression? *Aggression and Violent Behavior*, *5*, 191–200.
- Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 5497–5502.
- Decety, J., & Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia*, *41*, 127–138.
- Denton, R. T., & Martin M. T. (1998). Defining forgiveness: an empirical exploration of process and role. *The American Journal of Family Therapy*, *26*, 281–292.
- Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science*, *302*, 290–292.
- Farrow, T. F. D., Hunter, M. D., Wilkinson, I. D., *et al.* (2005). Quantifiable change in functional brain response to empathic and forgiveness judgments with resolution of posttraumatic stress disorder. *Psychiatry Research – Neuroimaging*, *140*(1), 45–53.
- Farrow, T. F. D., Zheng, Y., Wilkinson, I. D., *et al.* (2001). Investigating the functional anatomy of empathy and forgiveness. *NeuroReport*, *12*, 2433–2438.

- 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., *et al.* (1995). Other minds in the brain: a functional imaging study of 'theory of mind' in story comprehension. *Cognition*, *57*, 109–128.
- 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., *et al.* (2000). Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia*, *38*, 11–21.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*, 493–501.
- Goel, V., & Dolan, R. J. (2003). Reciprocal neural response within lateral and ventral medial prefrontal cortex during hot and cold reasoning. *NeuroImage*, *20*, 2314–2321.
- Goel, V., Grafman, J., Sadato, N., & Hallet, M. (1995). Modeling other minds. *NeuroReport*, *6*, 1741–1746.
- Goyer, P. F., Andreason, P. J., Semple, W. E., *et al.* (1994). Positron-emission tomography and personality disorders. *Neuropsychopharmacology*, *10*, 21–28.
- Grady, C. L., & Keightley, M. L. (2002). Studies of altered social cognition in neuropsychiatric disorders using functional neuroimaging. *Canadian Journal of Psychiatry*, *47*, 327–336.
- Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*, 517–523.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgement. *Science*, *293*, 2105–2108.
- Happé, F., Ehlers, S., Fletcher, P., *et al.* (1996). 'Theory of mind' in the brain. Evidence from a PET scan study of Asperger Syndrome. *NeuroReport*, *8*, 197–201.
- Hariri, A. R., Bookheimer, S. Y., & Mazziotta, J. C. (2000). Modulating emotional responses: effects of a neocortical network on the limbic system. *NeuroReport*, *11*(1), 43–48.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233.
- Jackson, P. L., Meltzoff, A. N., & Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage*, *24*, 771–779.
- Johnson, S. C., Baxter, L. C., Wilder, L. S., *et al.* (2002). Neural correlates of self reflection. *Brain*, *125*, 1808–1814.
- Lee, K. H., Egleston, P. N., Brown, W. H., *et al.* (2003). Treatment induced changes in brain activation in empathy in people with schizophrenia. *NeuroImage*, *19*, 117.
- Leslie, K. R., Johnson-Frey, S. H., & Grafton, S. T. (2004). Functional imaging of face and hand imitation: towards a motor theory of empathy. *NeuroImage*, *21*, 601–607.
- Maddock, R. J. (1999). The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends in Neurosciences*, *22*, 310–316.
- Moll, J., de Oliveira-Souza, R., Eslinger, P. J., *et al.* (2002). The neural correlates of moral sensitivity: a functional magnetic resonance investigation of basic and moral emotions. *The Journal of Neuroscience*, *22*, 2730–2736.

- 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.
- Pietrini, P., Guazzelli, M., Basso, G., Jaffe, K., & Grafman, J. (2000). Neural correlates of imaginal aggressive behavior assessed by positron emission tomography in healthy subjects. *American Journal of Psychiatry*, *157*, 1772–1781.
- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*, 1–72.
- Raine, A., Lencz, T., Bihrlé, S., LaCasse, L., & Colletti, P. (2000). Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. *Archives of General Psychiatry*, *57*, 119–127.
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage*, *22*, 1694–1703.
- Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). 'Theory of mind' impairments and their relationship to executive functioning following frontal lobe excisions. *Brain*, *124*, 600–616.
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in 'theory of mind'. *NeuroImage*, *19*, 1835–1842.
- Schulkin, J. (2000). Theory of mind and mirroring neurons. *Trends in Cognitive Sciences*, *4*, 252–256.
- Shamay, S. G., Tomer, R., & Aharon-Peretz, J. (2002). Deficits in understanding sarcasm in patients with prefrontal lesion is related to impaired empathic ability. *Brain and Cognition*, *48*, 558–563.
- Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., & Aharon-Peretz, J. (2003). Characterization of empathy deficits following prefrontal brain damage: the role of the right ventromedial prefrontal cortex. *Journal of Cognitive Neuroscience*, *15*, 324–337.
- Singer, T., Seymour, B., O'Doherty, J., *et al.* (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*, 1157–1162.
- Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience*, *10*, 640–656.
- Umiltà, M. A., Kohler, E., Gallese, V., *et al.* (2001). I know what you are doing: a neurophysiological study. *Neuron*, *31*, 155–165.
- Wicker, B., Keysers, C., Plailly, J., *et al.* (2003b). Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron*, *40*, 655–664.
- Wicker, B., Perrett, D. I., Baron-Cohen, S., & Decety, J. (2003a). Being the target of another's emotion: a PET study. *Neuropsychologia*, *41*, 139–146.

2007

Empathy in Mental Illness

Tom F. D. Farrow

University of Sheffield

and

Peter W. R. Woodruff

University of Sheffield

 **CAMBRIDGE**
UNIVERSITY PRESS