Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations

R.J.R. Blair

Mood and Anxiety Program, National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services, 15K North Drive, Room 206, MSC 2670, Bethesda, MD 20892-2670, USA

Received 1 February 2005
Available online 12 September 2005

Abstract

Empathy is a lay term that is becoming increasingly viewed as a unitary function within the field of cognitive neuroscience. In this paper, a selective review of the empathy literature is provided. It is argued from this literature that empathy is not a unitary system but rather a loose collection of partially dissociable neurocognitive systems. In particular, three main divisions can be made: cognitive empathy (or Theory of Mind), motor empathy, and emotional empathy. The two main psychiatric disorders associated with empathic dysfunction are considered: autism and psychopathy. It is argued that individuals with autism show difficulties with cognitive and motor empathy but less clear difficulties with respect to emotional empathy. In contrast, individuals with psychopathy show clear difficulties with a specific form of emotional empathy but no indications of impairment with cognitive and motor empathy.

Published by Elsevier Inc.

Keywords: Empathy; Psychopath; Autism

1. Introduction

The aim of this paper was to selectively review and consider the form of the separable functions that are grouped together in lay, and frequently scientific, parlance within the term “empathy.”
First, an attempt will be made to specify the dissociable neurocognitive systems that are grouped together within this term. In short, an attempt will be made to define empathy. Second, the two psychiatric conditions most associated with empathic dysfunction will be considered: autism and psychopathy. The aim here will be to consider the degree to which individuals with these conditions show selective deficits in the separable functions/neurocognitive systems that are grouped together as empathy.

2. Empathy

2.1. Defining empathy

Perhaps the most common definition of empathy is “an affective response more appropriate to someone else’s situation than to one’s own” (Hoffman, 1987; p. 48); i.e., empathy is an emotional reaction in an observer to the affective state of another individual. To some, empathy is considered to be a unitary process. Thus, Preston and de Waal (2002) have argued that “empathy [is] a super-ordinate category that includes all sub-classes of phenomena that share the same mechanism. This includes emotional contagion, sympathy, cognitive empathy, helping behavior, etc.” (p. 4). However, that position will not be supported here. Instead, it will be argued that the term “empathy” subsumes a variety of dissociable neurocognitive processes. Three main divisions, each reliant on at least partially dissociable neural systems, will be identified: cognitive, motor, and emotional empathy.

The term cognitive empathy has been used where the individual represents the internal mental state of another individual. Cognitive empathy is effectively Theory of Mind. Motor empathy occurs when the individual mirrors the motor responses of the observed actor. More recently this notion of “motor” empathy has been incorporated within the new Perception–Action model of emotional empathy (Preston & de Waal, 2002). There are at least two main forms of emotional empathy. The first and main form is a response to the emotional displays of others; their facial and vocal expressions and body movements. A second form is a response to other emotional stimuli, for example, a response to a phrase such as “Adam just lost his house.”

2.2. Theory of Mind (“cognitive empathy”)

Theory of Mind refers to the ability to represent the mental states of others, i.e., their thoughts, desires, beliefs, intentions, and knowledge (Frith, 1989; Leslie, 1987; Premack & Woodruff, 1978). Theory of Mind allows the attribution of mental states to self and others to explain and predict behavior. A series of neuroimaging studies have examined the neural systems engaged during the representation of the mental states of others relative to comparison conditions not requiring such representation. This work indicates the importance of medial prefrontal cortex (especially anterior paracingulate cortex), the temporal-parietal junction, and the temporal poles for the representation of the mental states of others (Brunet, Sarfati, Hardy-Bayle, & Decety, 2000; Castelli, Happe, Frith, & Frith, 2000; Fletcher et al., 1995; Gallagher et al., 2000; Goel, Grafman, Sadato, & Hallett, 1995; Vogeley et al., 2001); see, for a review, Frith (2001). However, it should be noted, confirmatory neuropsychological work to demonstrate the crucial role of these regions remains in
its infancy. Moreover, some early studies have not been supportive, at least for a role for anterior paracingulate cortex in Theory of Mind (Bird, Castelli, Malik, Frith, & Husain, 2004).

In addition to being considered a form of empathy in its own right, the ability to represent the mental states of others has been considered to be necessary for emotional empathy to occur (Batson, Fultz, & Schoenrade, 1987; Feshbach, 1987). Within these positions, representations of the internal mental state of another are assumed to act as stimuli for the activation of the affective, empathic response (Batson et al., 1987). Feshbach (1978, 1987), for example, viewed empathy to be a function of three processes: first, the cognitive ability to discriminate affective cues in others; Second, the more mature cognitive skills entailed in assuming the perspective and role of another person; and finally, emotional responsiveness (i.e., the ability to experience emotions); (Feshbach, 1978, 1987). According to Feshbach (1987), “empathy is conceived to be the outcome of cognitive and affective processes that operate conjointly” (p. 273). In short, individuals who cannot represent the mental states of would should not be able to empathically respond to others.

2.3. Motor empathy and the perception–action model

Motor empathy is defined as the tendency to automatically mimic and synchronize facial expressions, vocalizations, postures, and movements with those of another person (Hatfield, Cacioppo, & Rapson, 1994). It has long been considered a “primitive form of sympathy” (Smith, 1966; originally published 1759; Spencer, 1870). More recently, a neurocognitive account of motor empathy has been developed (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003; Decety & Jackson, 2004). This account relies heavily on the recent discovery of mirror neurons. Mirror neurons are neurons which show activity during the execution and also the observation of an action (Rizzolatti, Fogassi, & Gallese, 2001). While human neuroimaging work cannot indicate commonalities in the activation of individual neurons when executing and observing an action, such work has shown that the neural circuit involved in action execution overlaps with that activated when actions are observed (Blakemore & Decety, 2001).

Within the account, the perception of another individual’s state activates the observer’s corresponding representations, which in turn activate somatic and autonomic responses. At the anatomical level (see Carr et al., 2003), the suggestion is that the superior temporal cortex (STC) codes an early visual description of the action and sends this information to posterior parietal mirror neurons which code the precise kinesthetic aspect of the movement and then send this information to inferior frontal (BA 44/45) mirror neurons. The inferior frontal mirror neurons are thought to code the goal of the action. Connections from superior temporal, parietal, and inferior frontal cortices to the insula are thought to allow this representation information to generate emotional responses through limbic areas (Carr et al., 2003); see Fig. 1.

In short, empathy necessarily relies on the activation of neurons that code an action description (STC) and mirror neurons (parietal and inferior frontal cortex). Thus, lesions to any of these areas should significantly and globally disrupt empathy. In line with this position, two recent studies studies investigating which cortical regions, when damaged, most effected expression recognition stressed the importance of the inferior parietal cortex, a region that is proximal to STC (Adolphs, Damasio, Cooper, & Damasio, 2000; Adolphs, Damasio, Tranel, & Damasio, 1996). Lesions of frontal cortex also disrupt expression recognition although, it should be noted, the disruption has been associated with damage to orbital frontal cortex rather than inferior frontal (BA 44/45).
cortex (Blair & Cipolotti, 2000; Hornak, Rolls, & Wade, 1996; Hornak et al., 2003). Interestingly, lesions of medial regions of orbital frontal cortex, rather than lesions of inferior frontal cortex, have also been associated with deficits in affective empathy as measured by self-report (Shamay-Tsoory, Tomer, Goldsher, Berger, & Aharon-Peretz, 2004).

Neuroimaging data examining the neural response to the facial expressions of others provide some support for this model. Facial expressions are frequently associated with activity in superior temporal sulcus (Halgren, Raij, Marinkovic, Jousmaki, & Hari, 2000; Kesler/West et al., 2001; Phillips et al., 1997; Pizzagalli et al., 2002; Streit et al., 1999) as are face stimuli more generally (Haxby, Hoffman, & Gobbini, 2000). In addition, three neuroimaging studies have observed inferior frontal (BA 44/45) cortex activity to emotional expressions (George et al., 1993; Gorno-Tempini et al., 2001; Nakamura et al., 1999) although, it should be noted, many other studies have not. More direct evidence was potentially provided by Carr et al. (2003). In this study, participants just viewed facial stimuli or viewed facial stimuli while they were asked to “internally generate the target emotion.” They reported that imitation and viewing of emotion “activated a largely similar network of brain areas.” However, there are several difficulties with this study that make interpretation difficult. First, the extent of similarity of the networks is difficult to quantify; conjunction analysis was not used and of the 36 peak activations identified, only 12 were significantly activated in both conditions (17 were only significantly active during imitation while 7 were significantly active only during observation). Second, both conditions involved the viewing of face stimuli. For this reason alone, a degree of overlap of neural response would be expected. Moreover, a second neuroimaging study indicated a dissociation between the neural systems involved in face imitation and those involved in the response to a viewed expression (Leslie, Johnson-Frey, & Grafton, 2004).

2.4. Emotional empathy

While there are indications that facial expressions are processed from visual cortex via temporal cortex and onto limbic areas (the cortical pathway: retinogeniculostriate–extrastriate–fusiform)
there have also been claims of a subcortical pathway (retinocollicular–pulvinar–amygdalar); (Adolphs, 2002; de Gelder, Vroomen, Pourtois, & Weiskrantz, 1999; Morris, Ohman, & Dolan, 1999; Pizzagalli, Regard, & Lehmann, 1999). These two routes for expression processing mirror those previously suggested to be involved in aversive conditioning (Armony, Servan-Schreiber, Romanski, Cohen, & LeDoux, 1997; LeDoux, 2000). Thus, information on conditioned stimuli during auditory fear conditioning can be mediated by projections to the amygdala from either the auditory thalamus or auditory cortex (Campeau & Davis, 1995; LeDoux, Sakaguchi, & Reis, 1984; Romanski & LeDoux, 1992a, 1992b). The subcortical route is thought to provide coarse stimulus processing while the cortical route is thought to allow more precise stimulus encoding and allow discrimination learning (Armony et al., 1997; LeDoux, 2000).

Early data supporting the existence of the subcortical pathway suggested, surprisingly, that it could, like the cortical pathway, provide sufficiently fine-grained stimulus encoding to allow discrimination learning (Morris et al., 1999). However, later work has not supported this ( Vuilleumier, Armony, Driver, & Dolan, 2003). Vuilleumier et al. (2003) observed that the pulvinar and superior colliculus responded to low frequency but not high frequency facial expressions; i.e., coarse-grained but not fine-grained information. Additional support for the suggestion of a subcortical pathway has been provided by a work with G.Y., a patient with a long-standing right-sided hemianopia following occipital lobe damage at the age of 8 years (de Gelder et al., 1999). This “blindsight” patient showed some ability to discriminate between different facial expressions in his blind hemifield. Moreover, G.Y. demonstrated differential amygdala responses to fearful versus happy expressions when these were presented to both the blind and seeing hemi- fields. However, striate and fusiform activity only occurred in response to stimuli presented to the seeing hemifield (Morris, DeGelder, Weiskrantz, & Dolan, 2001).

It is argued that facial expressions of emotion have specific communicatory functions, that they impart specific information to the observer (Blair, 2003a). From this view, empathy, at least to facial and vocal emotional expressions, is the “translation” of the communication by the observer. It is argued that because of the different implications of these communicatory signals they are translated in several separable systems (Blair, 2003a). These will be considered in turn below.

It is argued that fearfulness, sadness, and happiness are reinforcers that modulate the probability that a particular behavior will be performed in the future (Blair, 2003a). Indeed, fearful faces have been seen as aversive unconditioned stimuli that rapidly convey information to others that a novel stimulus is aversive and should be avoided (Mineka & Cook, 1993). Similarly, it has been argued that sad facial expressions also act as aversive unconditioned stimuli discouraging actions that caused the display of sadness in another individual and motivating reparatory behaviors (Blair, 1995). Happy expressions, in contrast, are appetitive unconditioned stimuli which increase the probability of actions to which they appear causally related (Matthews & Wells, 1999).

The amygdala has been implicated in aversive and appetitive conditioning including instrumental learning (Everitt, Cardinal, Hall, Parkinson, & Robbins, 2000; Killcross, Robbins, & Everitt, 1997; LeDoux, 2000). In line with the suggested role of fearful, sad, and happy expressions as reinforcers, neuroimaging studies, with a few exceptions (Kesler/West et al., 2001), have generally found that fearful, sad, and happy expressions all modulate amygdala activity (Baird et al., 1999; Blair, Morris, Frith, Perrett, & Dolan, 1999; Breiter et al., 1996; Drevets, Lowry, Gautier, Perrett, & Kupfer, 2000; Morris et al., 1996; Phillips et al., 1997, 1998; Schneider, Gur, & Muenz, 1994), though it should be noted that happy expressions have been reported to both
increase and decrease amygdala activity (Breiter et al., 1996; Morris et al., 1996). The neuropsychological literature supports the neuroimaging literature regarding the importance of the amygdala in the processing of fearful expressions. There have been occasional suggestions that amygdala damage leads to general expression recognition impairment but these have been based on reports on patients whose lesions extend considerably beyond the amygdala (Rapcsak et al., 2000). Instead, amygdala lesions have been consistently associated with impairment in the recognition of fearful expressions (Adolphs, Tranel, Damasio, & Damasio, 1994; Adolphs et al., 1999; Calder, Young, Rowland, & Perrett, 1996; Schmolck & Squire, 2001). Impairment in the processing of sad expressions is not uncommonly found in patients with amygdala lesions (Adolphs & Tranel, 2004; Adolphs et al., 1999; Schmolck & Squire, 2001). Indeed, a review of patient performance across studies reported that approximately 50% of patients with amygdala damage present with impairment for the recognition of sad expressions (Fine & Blair, 2000). Amygdala lesions rarely result in impairment in the recognition of happy expressions (Adolphs et al., 1999; Fine & Blair, 2000). However, this may reflect the ease with which happy expressions are recognized (Ekman & Friesen, 1976).

Disgusted expressions are also reinforcers but reinforcers that most frequently provide valence information about foods (Rozin, Haidt, & McCauley, 1993). Disgusted expressions are particularly important for the rapid transmission of taste aversions; the observer is warned not to approach the food that the emoter is displaying the disgust reaction to. Several imaging studies have shown the importance of the insula in its response to primary disgust stimuli (i.e., aversive tastes/odors) (Small et al., 2003; Zald, Lee, Fluegel, & Pardo, 1998). Moreover, insula lesions have been found to block the acquisition and expression of taste aversion learning (Cubero, Thiele, & Bernstein, 1999). In other words, the insula allows the representation of the aversive taste that can then be associated with the sensory qualities of the novel food. If disgusted expressions are important for taste aversion learning then they too should recruit the insula. Neuroimaging work shows that disgusted expressions do indeed engage the insula in neuroimaging work (Phillips et al., 1997, 1998; Sprengelmeyer, Rausch, Eysel, & Przuntek, 1998) while neuropsychological work shows that patients with damage to the insula present with selective impairment for the recognition of disgusted expressions (Calder, Keane, Manes, Antoun, & Young, 2000; Sprengelmeyer et al., 1996). Moreover, the regions of anterior insula activated by primary disgust stimuli are also activated by disgusted expressions (Wicker et al., 2003). In other words, the insula allows the representation of the aversive taste whether this be a primary disgust stimulus (i.e., a taste or odor) or a communication of aversive taste through the disgusted expression of another. This can then be associated with the sensory qualities of the novel food.

Angry expressions are known to curtail the behavior of others in situations where social rules or expectations have been violated (Averill, 1982). It has been argued that displays of anger or embarrassment do not act as unconditioned stimuli for aversive conditioning or instrumental learning but rather as important signals to modulate current behavioral responding, particularly in situations involving hierarchy interactions (Blair, 2003a; Blair & Cipolotti, 2000; Keltner & Anderson, 2000). They appear to serve to inform the observer to stop the current behavioral action rather than to convey any information as to whether that action should be initiated in the future. Angry expressions, in short, trigger response reversal (Blair & Cipolotti, 2000; Blair et al., 1999), a function that is orbital and particularly ventrolateral frontal cortex is involved in (Cools, Clark, Owen, & Robbins, 2002; Dias, Robbins, & Roberts, 1996; O'Doherty,
Kringelbach, Rolls, Hornak, & Andrews, 2001). The neural response to angry expressions involves ventrolateral frontal cortex (Blair et al., 1999; Kesler/West et al., 2001; Sprengelmeyer et al., 1998) particularly when angry expressions cue reversal (Kringelbach & Rolls, 2003).

2.5. Summary: Multiple forms of “empathy”

Fig. 2 synthesizes the current position on the multiple forms of empathy described above: Theory of Mind, motor empathy, and emotional empathy. This position assumes that these three forms of empathy do share a degree of anatomical overlap—all appear to rely on the integrity of neurons in superior temporal regions. However, the position assumes that beyond superior temporal cortex, all can operate independent of one another. Motor and emotional empathy are likely to be relatively automatic and, with respect to facial expressions, may occur in parallel. Both may help to filter an appropriate verbal response if the individual is asked to name the expression even if, as it is argued here, their primary computational results are different. In the case of motor empathy, a congruent motor response is primed. In the case of emotional empathy, the unconditioned stimulus (the expression) can be paired with the object in the environment that elicited the expression in the displayer.

The degree to which representing the mental states of others (Theory of Mind) can occur automatically is uncertain. Studies examining the neural correlates of Theory of Mind have asked participants to represent the mental states of story/cartoon protagonists (Brunet et al.,...
2000; Castelli et al., 2000; Fletcher et al., 1995; Gallagher et al., 2000; Goel et al., 1995; Vogeley et al., 2001). However, it is interesting to note that a study examining judgments of embarrassment/inappropriateness of intentional/unintentional conventional norm violations did report activity in regions associated with Theory of Mind (anterior paracingulate cortex, the temporal-parietal junction, and inferior temporal cortex). Indeed, these areas were all significantly more active when processing an intentional as opposed to an unintentional norm violation (Berthoz, Armony, Blair, & Dolan, 2002). This suggests that calculations of the mental states of others may occur even during paradigms where the task demands do not stipulate that this should occur. Of course, while it is relatively easy to specify the conditions that automatically activate motor empathy (a moving conspecific) or emotional empathy (a facial or vocal expression), it is unclear at the present time what the stimulus conditions that would automatically elicit Theory of Mind might be.

It is important to note that there are other forms of empathy that have not been considered above. We can have empathic reactions to another individual’s verbal statements; e.g., “Adam just lost his house.” It could be argued that when hearing about such a situation, we need to simulate the state (imagine that we have lost our own house) and then process the corresponding emotional state. However, given that the loss of a house is an aversive [conditioned] stimulus (as it has to be even within a simulationist position) we do not need to simulate the state to achieve an emotional response. The representation of the conditioned stimulus “loss of the house” (or a generalization from a proximal conditioned stimulus) will be sufficient to generate the emotional response. The system then only needs to calculate that the origin of this emotional response is the aversive stimulus of “loss of the house,” related to John, to feel empathy for John. For this form of empathy to occur, it is necessary for the individual to have previously associated “loss of the house” or a similar stimulus with an aversive state.

This conditioning approach is worth bearing in mind when interpreting recent studies examining the empathic response to observed pain in others. Work has shown that painful stimuli consistently activate the secondary somatosensory cortex, insular regions, and the anterior cingulate cortex (Davis, 2000; Peyron, Laurent, & Garcia-Larrea, 2000). Secondary somatosensory cortex is likely involved in the representation of the somatosensory stimulus while activation within the insula and anterior cingulate represents the pain itself. Interestingly, the sight of another individual in pain or the thought of another individual in pain activates both anterior cingulate and the insula (Jackson, Meltzoff, & Decety, 2005; Morrison, Lloyd, di Pellegrino, & Roberts, 2004; Singer et al., 2004); i.e., pain conditioned stimuli (the sight of a hand receiving a painful stimulus) can elicit a pain conditioned response.

In short, empathy is not a unitary capacity mediated by a unitary system but a variety of different functions that are mediated by partially separable systems. There may be overlap, Theory of Mind, motor empathy, and emotional empathy (with the exception of that seen to stimuli associated with pain and perhaps touch; see immediately above) all recruit superior temporal cortex, but there is considerable specificity also (see Fig. 2). This means that when we consider whether a particular psychiatric population is associated with empathic dysfunction, for this assertion to be meaningful, we need to be able to stipulate the form of the functional impairment shown by patients within that population. In the next sections of this paper, this will be attempted for the two psychiatric disorders most often associated with empathic dysfunction: autism and psychopathy.
3. Autism

Autism is a severe developmental disorder described by the American Psychiatric Association’s diagnostic and statistical manual (DSM-IV) as “the presence of markedly abnormal or impaired development in social interaction and communication and a markedly restricted repertoire of activities and interests” (American Psychiatric Association, 1994, p. 66). The main criteria for the diagnosis in DSM-IV can be summarized as qualitative impairment in social communication, and restricted and repetitive patterns of behavior and interests. These criteria must be evident before 3 years of age. As long as autism has been recognized, the idea has existed that people with autism are impaired in their ability to enter into emotional relationships. Thus, Kanner, the psychiatrist who originally described the disorder in 1943, wrote “these children have come into the world with an innate inability to form the usual, biologically provided affective contact with other people, just as other children come into the world with innate physical or intellectual handicaps” (Kanner, 1943, p. 250).

With respect to empathy dysfunction in autism, Baron-Cohen and colleagues originally demonstrated Theory of Mind impairment in children with autism (Baron-Cohen, Leslie, & Frith, 1985). This impairment has been consistently replicated and reviews of this literature can be found elsewhere (Baron-Cohen, 1995; Hill & Frith, 2003). Neuroimaging work on individuals with Asperger syndrome/high functioning individuals with autism has examined neural responses during task performance for measures involving the representation of mental states [verbal theory of mind scenarios Happé et al. (1996) or animated shapes using interactivity with implied intentions (Castelli, Frith, Happé, & Frith, 2002)]. These studies have reported reduced activation in the three brain regions critical to mentalizing in normal individuals (medial prefrontal cortex, temporal-parietal junction, and the temporal poles); (Castelli et al., 2002; Happé et al., 1996); Happé et al. (1996) only found reduced medial prefrontal activation.

There are strong reasons to believe that motor empathy is likely to be impaired in autism. Certainly, there is considerable evidence that people with autism show deficits in imitation (see, for reviews of the literature, Smith & Bryson, 1994; Williams, Whiten, & Singh, 2004). However, as yet, no studies have investigated the neural responses of individuals with autism during imitation.

With respect to emotional empathy, there have been many suggestions that autism is due to an innate impairment in the ability to perceive and respond to the affective expressions of others, and that this deficit leads to their profound difficulties in social interaction (Hobson, 1993). Many have reported that children with autism have difficulty recognizing the emotional expressions of others (Bormann-Kischkel, Vilsmeier, & Baude, 1995; Hobson, 1986; Howard et al., 2000) with a recent claim suggesting that this is specific for fearful expressions (Howard et al., 2000). However, the above only applies to studies where the groups have not been matched on mental age. When they are matched for verbal mental age, children with autism have usually been found to be unimpaired in facial affect recognition (Adolphs, Sears, & Piven, 2001; Ozonoff, Pennington, & Rogers, 1990; Prior, Dahlstrom, & Squires, 1990). In addition, several studies have found the emotion processing impairment to be pronounced only when the emotion is a complex ‘cognitive’ emotion requiring the representation of the mental states of others, such as surprise or embarrassment (Baron-Cohen, Spitz, & Cross, 1993; Bormann-Kischkel et al., 1995; Capps, Yirmiya, & Sigman, 1992). Previous work has implicated regions associated with Theory of Mind
(medial prefrontal cortex, temporal-parietal junction, and the temporal poles) in the mediation of embarrassment (Berthoz et al., 2002).

Individuals with autism do show some general face processing impairments. For example, they show reduced recognition and memory of faces (Blair, Frith, Smith, Abell, & Cipolotti, 2002; Boucher & Lewis, 1992; Klin et al., 1999) and abnormal eye scan paths when viewing faces (Klin, Jones, Schultz, Volkmar, & Cohen, 2002). However, individuals with autism are not prosopagnosic and, as noted above, when matched for verbal mental age, are usually unimpaired in facial affect recognition (Adolphs et al., 2001; Ozonoff et al., 1990; Prior et al., 1990). Neuroimaging studies have indicated that three specific areas are critically involved in face processing: the lateral occipital gyri, bilateral regions in the lateral fusiform gyrus, and the posterior superior temporal sulcus (Haxby, Hoffman, & Gobbini, 2002; Kanwisher, McDermott, & Chun, 1997; Kanwisher, Stanley, & Harris, 2000). However, despite the overlap between Theory of Mind, motor empathy, and face recognition in the recruitment of superior temporal regions, and indications that this region shows reduced activation during Theory of Mind tasks in individuals with autism (Castelli et al., 2002), neuroimaging studies of face processing typically report reduced activity in fusiform rather than superior temporal cortex. This reduced face-related fusiform activity has been seen in individuals with autism viewing the neutral and emotional expressions of strangers (Critchley et al., 2000; Hall, Szechtman, & Nahmias, 2003; Hubl et al., 2003; Pierce, Muller, Ambrose, Allen, & Courchesne, 2001; Schultz et al., 2000); though, for contradictory findings, see (Hadjikhani et al., 2004).

In summary, individuals with autism are clearly impaired with respect to both Theory of Mind and motor empathy (given the impairment in imitation). This probably reflects a feature of the pathophysiology associated with autism, perhaps related to the functioning of the superior temporal cortex. However, it is unclear whether there is any evidence of impairment for emotional empathy. Certainly, the current evidence is weak. However, if appropriate evidence does emerge and given indications of reduced superior temporal cortical activity in individuals with autism (Castelli et al., 2002), the impairment is likely to be general, reflecting a more general representational difficulty with faces, rather than specific for a particular expression-specific type of emotional learning.

4. Psychopathy

The classification of psychopathy was introduced by Hare (1980, 1991) and has proved to be a useful predictor of future risk (Hare, 1991). Individuals with psychopathy are marked by pronounced emotional (considerably reduced empathy and guilt) and behavioral disturbance (criminal activity and, frequently, violence); (Frick, O’Brien, Wootton, & McBurnett, 1994; Hare, 1980, 1991; Harpur, Hare, & Hakstian, 1989). Psychopathy is a developmental disorder, usually appearing in early childhood (certainly by eight years of age) and continuing throughout the lifespan (Harpur & Hare, 1994). Individuals with psychopathy represent a subset, perhaps 25% (Hart & Hare, 1996), of individuals who meet criteria for the psychiatric classifications of Conduct Disorder and Antisocial Personality Disorder.

Psychopathy can be considered one of the prototypical disorders associated with empathic dysfunction. Reference to empathic dysfunction is part of the diagnostic criteria of psychopathy.
The very ability to inflict serious harm to others repeatedly can be, and is (Hare, 1991), an indicator of a profound disturbance in an appropriate “empathic” response to the suffering of another.

5. Psychopathy and Theory of Mind

There are no indications of Theory of Mind impairment in individuals with psychopathy. Three out of four studies assessing the ability of individuals with psychopathy on Theory of Mind measures have reported no impairment (Blair et al., 1996; Richell et al., 2003; Widom, 1978). Only one study has reported impairment and this used a rating scale that is not a typical measure of Theory of Mind (Widom, 1976).

In addition to the above work on individuals with psychopathy, it is important to note that even in the broader spectrum of antisocial individuals, there is little data suggesting any link between Theory of Mind impairment and antisocial behavior. Hughes, Dunn, and White (1998) did find some indication of Theory of Mind impairment in their “hard-to-manage” preschoolers relative to the comparison group. However, Happé and Frith found no impairment in their children with Emotional and Behavioral difficulties (Happé & Frith, 1996). Similarly, a study of school bullies found no indications of Theory of Mind impairment (Sutton, Smith, & Swettenham, 1999). In addition, Sutton and colleagues also found no relationship between Theory of Mind performance on the advanced Eyes task (Baron-Cohen, Wheelwright, & Joliffe, 1997) and “Disruptive Behavior Disorder” symptoms in children aged 11–13 years (Sutton, Reeves, & Keogh, 2000).

In summary, the profound empathic dysfunction reported in the clinical description of psychopathy (Hare, 1991) does not involve Theory of Mind impairment. Individuals with psychopathy are unimpaired on measures of Theory of Mind. Indeed, there are no indications that any populations who show heightened levels of antisocial behavior are associated with Theory of Mind impairment.

6. Psychopathy and emotional empathy

As noted above, the clinical description of psychopathy includes reference to a lack of empathy. This description has been substantiated empirically. Individuals with psychopathy show reduced vicarious conditioning; i.e., reduced autonomic responses to stimuli associated with the distress of another individual (Aniskiewicz, 1979; House & Milligan, 1976); though, for a failure to replicate, see (Sutker, 1970). In addition, both adults with psychopathy and children with psychopathic tendencies show reduced autonomic responses to the sad expressions of others (Blair, 1999; Blair, Jones, Clark, & Smith, 1997). Several studies have examined the ability of individuals with psychopathy to recognize the facial or vocal emotional expressions of others (Blair, Budhani, Colledge, & Scott, 2005; Blair, Colledge, Murray, & Mitchell, 2001; Blair et al., in press-b; Blair, Mitchell, et al., 2002; Kosson, Suchy, Mayer, & Libby, 2002; Stevens, Charman, & Blair, 2001). While children with psychopathic tendencies have shown impairment in the recognition of sad expressions (Blair et al., 2001; Stevens et al., 2001), this has not been found in adults with psychopathy (with one exception; Dolan, personal communication), probably because of the ease...
with which this expression can be recognized. In all studies (with one exception: Kosson et al., 2002), children with psychopathic tendencies and adults with psychopathy have shown impairment in the recognition of fearful expressions. Kosson et al. (2002) reported some difficulty with the recognition of disgusted expressions (but only when the participants were responding with their left rather than right hand). Blair et al. (in press) also found some impairment in the adults with psychopathy for the recognition of disgusted expressions, however, this deficit was not present if the effect of IQ was co-varied out.

In short, the empathic dysfunction shown by individuals with psychopathy appears relatively selective. Individuals with psychopathy are impaired when processing fearful, sad (in adulthood, if responsiveness is indexed by skin conductance responses (SCRs), in childhood whether by SCR or recognition score), and possibly disgusted expressions. No study has yet reported that individuals with psychopathy show impairment for the processing of angry, happy or surprised expressions.

Earlier, it was suggested that facial expressions have specific communicatory functions that are “translated” by specific neurocognitive systems (Blair, 2003a). One of these systems was described as responsive to the aversive and appetitive unconditioned stimuli of fearful, sad, and happy expressions and consequently modulating the probability that any stimulus associated with these expressions will be avoided or approached in the future. The suggestion is that individuals with psychopathy have dysfunction primarily in this system. It is possible that they also have dysfunction in the second system responsive to disgusted expressions (Kosson et al., 2002), however, in the absence of additional data that possibility will not be considered here.

As noted above, the primary neural system responsible for orchestrating an emotional response to fearful, sad, and happy expressions is the amygdala (Baird et al., 1999; Blair et al., 1999; Breiter et al., 1996; Drevets et al., 2000; Morris et al., 1996; Phillips et al., 1997, 1998; Schneider et al., 1994). This suggests that amygdala dysfunction in psychopathy. There is considerable data in line with this suggestion (Blair, 2001, 2002, 2003b; Blair & Frith, 2000; Patrick, 1994). Thus, individuals with psychopathy show reduced amygdaloid volume relative to comparison individuals (Tiihonen et al., 2000) and reduced amygdala activation during emotional memory (Kiehl et al., 2001) and aversive conditioning tasks (Viet et al., 2002). Human and animal neuropsychological work has informed us that the effects of amygdala lesions include impairment in: (1) aversive conditioning (Bechara, Damasio, Damasio, & Lee, 1999; Bechara et al., 1995; LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998); (2) the augmentation of the startle reflex to visual threat primes (Angrilli et al., 1996; Funayama, Grillon, Davis, & Phelps, 2001); and (3) passive avoidance learning (Ambrogi Lorenzini, Baldi, Bucherelli, Sacchetti, & Tassoni, 1999). Individuals with psychopathy are also impaired in the above tasks (Blair et al., 2001; Flor, Birbaumer, Hermann, Ziegler, & Patrick, 2002; Herpertz et al., 2001; Levenston, Patrick, Bradley, & Lang, 2000; Lykken, 1957; Newman & Kosson, 1986).

While individuals with psychopathy are impaired in the processing of fearful and sad expressions, they show no impairment for the processing of happy expressions (Blair et al., 2001, 2002, 2005, in press-b; Kosson et al., 2002; Stevens et al., 2001). While this is consistent with the neuropsychological literature documenting the consequences of amygdala lesions (Adolphs et al., 1994, 1999; Calder et al., 1996; Schmolck & Squire, 2001), it is less consistent with the neuroimaging literature (see above) which suggests a role for the amygdala in the processing of happy expressions. Of course, the absence of impairment for happy expressions in individuals with
psychopathy might reflect the ease with which they are recognized (i.e., an intact amygdala is not necessary for naming happy expressions).

The fact that individuals with psychopathy show no impairment for angry expressions is also of interest. Neurological patients following lesions of orbital and ventral frontal cortex show general difficulties with processing expressions which may be particularly marked for angry expressions (Blair & Cipolotti, 2000; Hornak et al., 1996, 2003). There have been claims that such patients (sometimes described as presenting with “acquired sociopathy”), who also show heightened levels of aggression, can be considered a neurological model for psychopathy (Damasio, 1994). However, the aggression of patients with “acquired sociopathy” show heightened levels of reactive aggression (aggression in response to threat or frustration) whether the lesion occurs in child- or adult-hood (Anderson, Bechara, Damasio, Tranel, & Damasio, 1999; Blair & Cipolotti, 2000; Grafman, Schwab, Warden, Pridgen, & Brown, 1996). Such patients never show heightened levels of goal directed, instrumental aggression. Individuals with psychopathy, in contrast, do show heightened levels of instrumental and reactive aggression (Cornell et al., 1996; Williamson, Hare, & Wong, 1987). This difference in forms of aggression is one of the several data incompatible with suggestions that “acquired sociopathy” is a neurological model for psychopathy (Blair, 2001). However, there are developmental disorders that show heightened risks for reactive but not instrumental aggression: childhood Bipolar Disorder or Intermittent Explosive Disorder (Coccaro, 1998; Leibenluft, Blair, Charney, & Pine, 2003). Patients with both disorders show general difficulties with processing expressions but their difficulty is particularly marked for angry expressions (Best, Williams, & Coccaro, 2002; McClure, Pope, Hoberman, Pine, & Leibenluft, 2003). In addition, patients with these disorders show problems with response reversal, a function reliant on orbital/ventrolateral prefrontal cortex (Best et al., 2002; Gorrindo, Blair, Budhani, Pine, & Leibenluft, in press). In other words, “acquired sociopathy” may be a neurological model for these developmental disorders that show a rather different “empathy impairment” from individuals with psychopathy. Their impairment appears to be linked to an inability to regulate their emotional responses (on the basis of social cue and other information) such that they are more predisposed to reactive aggressive outbursts (Blair, 2004).

In summary, individuals with psychopathy show no Theory of Mind impairment. With respect to motor empathy, no definitive conclusions can be drawn. To my knowledge, no studies have formally assessed motor empathy in psychopathy (e.g., by assessing imitation in individuals with psychopathy). However, it is worth noting that there are no indications of the general expression recognition impairment in individuals with psychopathy as would be predicted to exist if there was motor empathy impairment in individuals with this disorder. In short, it is likely that motor empathy, like Theory of Mind, is unimpaired in individuals with this disorder.

Individuals with psychopathy do show a selective emotional empathy dysfunction; they are impaired in the processing of fearful, sad, and possibly disgusted expressions. Indeed, this impairment is likely to be at the heart of the disorder (Blair, 1995). Individuals who are indifferent to the fear and sadness of others are individuals who are difficult to socialize through effective socialization practices as empathy induction. Empathy induction involves the socializer focusing the attention of the transgressor on the distress of the victim (and presumably heightens the salience of the aversive stimulus of the victim’s distress). While the greater use of empathy induction and other positive forms of parenting reduce the probability of antisocial behavior in healthy children, they have no significant effect on the probability of antisocial behavior in children with psycho-
pathic tendencies (Wootton, Frick, Shelton, & Silverthorn, 1997). In other words, if we could find means to increase the empathic reaction of children with psychopathic tendencies, we might be able to considerably improve the prognosis of this disorder. The investigation of such means is one of the main foci of our research at the moment.

7. General conclusions

Empathy is a lay term that has been co-opted into scientific parlance. As a lay term, it can be used to describe many different neurocognitive functions. Three have been described in this paper: Theory of Mind, motor empathy, and emotional empathy. These three involve at least partially dissociable neurocognitive architectures although, a common region to all three appears to be the superior temporal cortex.

There are two psychiatric conditions that have been most associated with empathic dysfunction: autism and psychopathy. Interestingly, the empathic dysfunction shown by these two disorders doubly dissociates. Individuals with autism show impairment with Theory of Mind and motor empathy (as indexed by their difficulties with imitation). However, there are fewer indications that individuals with autism show difficulties with emotional empathy and, if they do, their impairment is certainly not of the selective form seen in individuals with psychopathy. Individuals with psychopathy show no difficulties with Theory of Mind and probably motor empathy. However, they show marked and selective difficulties with emotional empathy; their impairment is for the processing of sad, fearful, and possibly disgusted expressions. This impairment appears related to amygdala dysfunction and is likely to underpin the difficulty in their socialization.

Two other psychiatric disorders also show a form of empathy dysfunction; childhood Bipolar Disorder and Intermittent Explosive Disorder. These disorders have not been assessed with respect to Theory of Mind and motor empathy. However, they do show general expression recognition difficulties that are particularly marked for angry expressions. This impairment leads to difficulties in regulating emotional responses with respect to others, particularly in hierarchy situations, and increases the risk for reactive aggressive outbursts (Blair, 2004). The empathy impairment and its implications appears similar to that seen in individuals with “acquired sociopathy”; neurological patients presenting with social difficulties following lesions of orbital/ventrolateral prefrontal cortex.

In short, empathy is a general term for a collection of specific neurocognitive functions. These functions may be disrupted selectively. Such selective disruptions are at the heart of specific psychiatric disorders: autism, psychopathy, childhood bipolar disorder, and intermittent explosive disorder. The continued elucidation of these forms of empathy dysfunction will one day make likely the successful treatment of these disorders.

References


