The Neural Basis of Empathy

Boris C. Bernhardt and Tania Singer

Department of Social Neuroscience, Max-Planck Institute of Human Cognitive and Brain Sciences, Stephanstraße 1a, 04109 Leipzig, Germany; email: bernhardt@cbs.mpg.de, singer@cbs.mpg.de

Keywords
social neuroscience, insula, cingulate cortex, fMRI, emotion

Abstract
Empathy—the ability to share the feelings of others—is fundamental to our emotional and social lives. Previous human imaging studies focusing on empathy for others’ pain have consistently shown activations in regions also involved in the direct pain experience, particularly anterior insula and anterior and midcingulate cortex. These findings suggest that empathy is, in part, based on shared representations for firsthand and vicarious experiences of affective states. Empathic responses are not static but can be modulated by person characteristics, such as degree of alexithymia. It has also been shown that contextual appraisal, including perceived fairness or group membership of others, may modulate empathic neuronal activations. Empathy often involves coactivations in further networks associated with social cognition, depending on the specific situation and information available in the environment. Empathy-related insular and cingulate activity may reflect domain-general computations representing and predicting feeling states in self and others, likely guiding adaptive homeostatic responses and goal-directed behavior in dynamic social contexts.
INTRODUCTION

Empathy is a crucial component of human emotional experience and social interaction. The ability to share the affective states of our closest ones and complete strangers allows us to predict and understand their feelings, motivations, and actions. Extending previous work from philosophy and behavioral psychology (Batson 2009, de Vignemont & Singer 2006, Eisenberg 2000, Hoffman 2000), advances in social neuroscience have provided important new insights into the brain basis of empathy.

In this review, we outline the main results of brain imaging studies that have investigated the neural underpinnings of human empathy. Using mainly functional magnetic resonance imaging (fMRI), the majority of studies suggest that observing affective states in others activates brain networks also involved in the firsthand experience of these states, confirming the notion that empathy is, in part, based on shared networks (de Vignemont & Singer 2006, Keysers & Gazzola 2007, Preston & de Waal 2002). In particular, anterior insula (AI) and dorsal-anterior/anterior-midcingulate cortex (dACC/aMCC) play central roles in vicarious responses in the domain of disgust, pleasant or unpleasant tastes, physical and emotional pain, and other social emotions such as embarrassment or admiration (for recent meta-analyses, see Fan et al. 2011, Kurth et al. 2010, Lamm et al. 2011; for anatomical orientation, see Figure 1a). On the basis of their structural and functional patterns of connectivity, and their involvement in other functional processes at the interface of sensory, affective, and cognitive domains, regions such as AI and dACC/aMCC may generally contribute to the generation of subjective experiences and adaptive responses to actual and predicted states in the self and others. These general processes may then subsume empathy as a special case. We also highlight evidence that additional networks involved in social cognition can be flexibly corecruited during empathic understanding, depending on the particular situation and information available in the environment. Moreover, we summarize studies that have identified multiple factors that modulate or even counteract empathy. For example, initial evidence suggests that empathic responses can be counteracted by opposing motivational systems, such as the desire for revenge or Schadenfreude, closely related to activation in brain areas implicated in reward-processing. Finally, we outline future research avenues.

DEFINING EMPATHY

Despite a long tradition of philosophy and behavioral psychology research (Batson 2009, Eisenberg 2000, Eisenberg & Fabes 1990, Hoffman 2000, Wispe 1986), empathy has no universally accepted definition, and the different phenomena it subsumes remain debatable (Batson 2009, Blair 2005, de Vignemont & Singer 2006, Preston & de Waal 2002). However, previous conceptual work on empathy has greatly facilitated the design and interpretation of empirical studies that assess empathic traits through self-report measures and empathic states through controlled observational
Empathy occurs when the observation or imagination of affective states in another induces shared states in the observer (de Vignemont & Singer 2006, Singer & Lamm 2009). This state is also associated with knowledge that the target is the source of the affective state in the self. This reading of empathy necessarily involves components of affective sharing, self-awareness, and self-other distinction (for other more general notions of empathy see Baron-Cohen & Wheelwright 2004, Blair & Mislach 2011). Based mostly on results from empathy experiments, especially those investigating the brain processes underlying the empathic experience, may deepen our understanding of this phenomenon at the interface of social interactions and internal feeling states and ultimately promise to disentangle the conceptual web surrounding empathy.

A relatively specific notion claims that empathy occurs when the observation or imagination of affective states in another induces shared states in the observer (de Vignemont & Singer 2006, Singer & Lamm 2009). This state is also associated with knowledge that the target is the source of the affective state in the self. This reading of empathy necessarily involves components of affective sharing, self-awareness, and self-other distinction (for other more general notions of empathy see Baron-Cohen & Wheelwright 2004, Blair & Mislach 2011). Therefore, empathy differs from basic sharing-only phenomena such as emotional contagion and mimicry. Indeed, neither contagion nor mimicry requires a distinction about whether the origin of the affective experience is within the observer or was triggered by another person. Emotional contagion and empathy, for example when watching a friend in distress, can lead to personal distress, a self-centered and aversive response in the observer (Eisenberg & Fabes 1990). In contrast, during empathic concern, sympathy, or compassion, vicarious responses involve a feeling of concern for the other’s suffering that induces a motivation to alleviate the suffering, but not necessarily any sharing of feelings. Whereas empathizing with a sad person may result in a feeling of sadness in the self, sympathy and compassion often result in a feeling of loving or caring for that person and a motivation to relieve their suffering (Baumeister & Vohs 2007, Klimecki & Singer 2012, Singer & Steinbeis 2009). This motivation may then be transformed into prosocial behavior (Batson et al. 2007). In our own understanding, emotional contagion underlies affect sharing; this can be followed by other-oriented feelings such as compassion, sympathy, and empathic concern, which may further promote prosocial behavior; conversely, contagion and empathy may also induce aversive distress responses that can lead to withdrawal behavior motivated by the desire to protect oneself from negative emotions (Klimecki & Singer 2012).

**Empathy in the Brain**

In their seminal article on empathy in 2002, Preston & de Waal suggested that the observation and imagination of others in a given emotional state automatically activates a corresponding representation in the observer, along with its associated autonomic and somatic responses (Preston & de Waal 2002). This hypothesis was inspired by accounts that suggested a close link between action and perception through common coding schemes (Prinz 1984, 2005). Moreover, the discovery of mirror neurons, a class of neurons in monkey premotor and parietal cortices activated during execution and observation of actions, provided a neural mechanism for shared representations in the domain of action understanding (Gallese et al. 2004, Keysers & Gazzola 2007, Rizzolatti et al. 2001). Subsequent studies, based predominantly on fMRI, have investigated empathic brain responses for a variety of states including pain (Morrison et al. 2004, Singer et al. 2004), disgust (Benuzzi et al. 2008, Jabbi et al. 2007, Wicker et al. 2003), fear (de Gelder et al. 2004), anxiety (Prehn-Kristensen et al. 2009), anger (de Greck et al. 2012), sadness (Harrison et al. 2006), neutral touch (Blakemore et al. 2005, Ebisch et al. 2008, Keysers et al. 2004), pleasant affect (Jabbi et al. 2007), reward (Mobs et al. 2009), and higher-order emotions such as social exclusion (Masten et al. 2011) and embarrassment (Krach et al. 2011). Based mostly on results from empathy for pain, these studies showed that empathic responses recruit, to some extent, brain areas similar to those engaged during the corresponding first-person state. In the following section, we first highlight findings from studies on empathy for pain and then summarize evidence of empathic responses for other emotions and sensations.
Empathy for Pain

Empathy for pain has been studied frequently, owing to the robustness of pain in inducing empathy. The firsthand pain experience is generally aversive; moreover, it motivates behavioral responses to reduce the noxious stimulation (Price 2000) and can induce forms of warning communication to conspecifics (Craig 2004). Furthermore, observing others in pain can motivate helping behavior (Hein et al. 2010) and is often experienced as unpleasant and even painful for the observers themselves. Last, the neural circuits involved in pain are relatively well understood (Apkarian et al. 2005, Bushnell et al. 1999, Craig 2003, Peyron et al. 2000, Rainville 2002; Duerden & Albanese 2012).

Firsthand pain experience consistently activates networks in premotor and prefrontal, primary and secondary somatosensory cortices (S1 and S2), dACC/aMCC, and insula, along with thalamic and brain stem regions such as the periaqueductal gray (PAG) (Apkarian et al. 2005, Bushnell et al. 1999, Derbyshire 2000, Peyron et al. 2000, Duerden & Albanese 2012). Activations, albeit less consistently, have also been shown in the amygdala and cerebellum (Apkarian et al. 2005, Duerden & Albanese 2012). Most of these regions receive parallel input from multiple nociceptive pathways (Apkarian et al. 2005). Somatosensory regions and adjacent posterior insula are thought to encode the more sensory-discriminative components of pain. A case study showed that a patient with a large lesion in the postcentral gyrus and parietal operculum, comprising S1 and S2, lost discriminative aspects of pain perception, without overt loss of pain affect (Ploner et al. 1999). S2 responses have been shown to correlate with objective stimulus intensity, but not with affective ratings (Maihofner et al. 2006). Other studies have shown a contralateral bias for pain processing in subregions of S2 and posterior insula, suggesting a representation in these areas of the sensory-discriminative attributes, such as the stimulus location, of painful stimuli relative to body side in these areas (Bingel et al. 2003). Conversely, regions such as AI and dACC/aMCC are thought to encode more affective-motivational dimensions of pain (Price 2000). ACC and insula responses vary not simply as a function of noxious input but rather as a function of subjectively felt pain intensity (Kong et al. 2008). In ACC, activity correlates positively with ratings of pain unpleasantness (Rainville et al. 1997) but does not correlate much with stimulus intensity (Peyron et al. 2000). Moreover, insula and ACC responses to painful stimuli can be influenced by the emotional context, suggesting interaction effects within the affective domain (Phillips et al. 2003).

To investigate brain responses during empathy for pain, Singer and colleagues studied females who were accompanied by their romantic partners (Singer et al. 2004). In one condition, the female, lying in the scanner, received a painful shock via an electrode attached to her hand. In the other condition, the male partner who was seated next to the MRI scanner and whose hand could be seen by the female via a mirror received the shock. In both conditions, abstract visual cues indicated to the female who would receive painful stimulation. The authors observed activity in AI, dACC, brain stem, and cerebellum when females received the shock directly and, most importantly, when they vicariously felt their partners’ pain. The presentation of facial expressions of others in pain (Botvinick et al. 2005, Lamm et al. 2007a, Saarela et al. 2007), or of body parts receiving painful stimulation (Jackson et al. 2005, Lamm et al. 2007b), has elicited similar findings. The consistency of activations in parts of the pain networks elicited by firsthand experience as well as during vicariously felt pain has thus been taken to support the hypothesis that empathy involves shared representations.

More specifically, employing statistical conjunction analysis, several studies quantified the extent of shared activations in first-person pain and empathy. Comparing average activation patterns in these two conditions within a group of subjects, overlapping regions were located in insular and cingulate regions (Jackson...
et al. 2006, Morrison et al. 2004, Singer et al. 2004). To extend these findings, Morrison & Downing (2007) studied fMRI signals of individual subjects in native anatomical space, minimizing confounds introduced by image preprocessing. They observed activation overlaps in 6 of 11 subjects in aMCC, at the transition between otherwise nonoverlapping regions activated by directly and vicariously felt pain. Although these findings also indicate divergent activations underlying firsthand pain and empathy, they further support a role of shared representations in empathy.

However, voxel-wise conjunctions do not necessarily indicate shared representations on the neuronal level. A typical voxel in an fMRI experiment has a resolution of around 3 mm per side, and its signal relates to the activity of thousands of neurons within possibly different neuronal populations. Future studies employing fMRI adaptation (Grill-Spector & Malach 2001, Henson & Rugg 2003) or multivariate pattern analysis may more selectively probe commonalities in activations of specific neuronal populations (Norman et al. 2006). Indeed, in a recent multivoxel pattern analysis, bilateral AI regions exhibited a similar spatial distribution of cortical fMRI activity when seeing another person’s hand in pain compared to firsthand pain, providing relatively strong evidence for similar neuronal populations involved in both conditions (Corradi-Dell’Acqua et al. 2011).

Empathic responses to others’ pain in somatosensory regions have been less consistently reported. Using transcranial magnetic stimulation (TMS), Avenanti and colleagues (2005) demonstrated that watching a video of a needle pricking a specific hand muscle reduces motor excitability of the equivalent muscle in the observer, similar to the freezing response that would occur if pain was directly administered. This reduction in motor excitability correlated with pain-intensity ratings, but not with those of pain unpleasantness. Importantly, no effect was seen when participants watched a cotton bud touching the same muscle or when the needle prick was applied to a different part of the hand, a foot, or a tomato. Thus, although not directly showing activations in somatosensory cortices, this study suggested that attentively watching pain applied to the other’s body parts interferes with somatosensory processing.

A recent meta-analysis by Lamm and colleagues (2011) on 32 fMRI studies of empathy for pain confirmed that observing pain in others most robustly activated AI, extending into the inferior frontal gyrus (IFG) and dACC/aMCC (Figure 1a). Moreover, by classifying previous experiments into those employing abstract visual cues to signal pain in others (cue-based paradigms, Figure 1b) and those showing pictures of body parts receiving pain (picture-based paradigms, Figure 1c) (Lamm et al. 2011), the study yielded further quantitative insights on the role of somatosensory regions in empathy. Indeed, during cue-based designs, activations in S1 and S2 contralateral to the stimulated hand were observed only in self-related but not in vicarious experiences of pain. Conversely, picture-based designs induced activity in both S1 and S2 during the other-related condition. However, similar activity was also elicited to a large extent by nonpainful control pictures and did not seem to be lateralized to a specific hemisphere. These results thus suggest that somatosensory activation sometimes observed in picture-based empathy for pain paradigms may rather be due to unspecific activation based on the perception of touch and movement of body parts and not due to empathy for pain itself (for similar arguments, see Keysers et al. 2010).

Directly comparing activation patterns of both design types, the meta-analysis of Lamm and colleagues also revealed an important divergence in terms of distributed network coactivations (Lamm et al. 2011). Indeed, cue-based studies preferentially activated regions such as ventral medial prefrontal cortex (PFC), superior temporal cortex (STC), and posterior regions such as the temporo-parietal junction (TPJ) and precuneus/posterior cingulate cortex (PCU/PCC) (Figure 1d). These areas are generally thought to play a role in processes related to Theory of Mind or mentalizing.

IFG: Inferior frontal gyrus
PFC: prefrontal cortex
TPJ: temporo-parietal junction
Theory of mind (mentalizing): ability to infer and represent beliefs and desires
Adapted from Jackson et al. (2005), Lamm et al. (2011), Lamm et al. (2007b), with permission.

Empathy for Other Emotions and Sensations

Studies based on vicarious responses to affective states other than pain, such as social exclusion (Masten et al. 2011), disgust (Jabbi et al. 2008, Wicker et al. 2003), anxiety (Prehn-Kristensen et al. 2009), and taste (Jabbi et al. 2007), have reproduced a central role of AI and ACC/MCC during empathy. Using disgusting odorants Wicker and colleagues observed that subregions of AI and ACC were activated during direct inhalation and when viewing the disgusted faces of people inhaling the probes (Wicker et al. 2003). In a follow-up experiment, the

Figure 1

Meta-analytic findings on empathy for pain. (a) A meta-analysis of 32 previous empathy-for-pain studies revealed consistent activations in anterior insula (AI) extending into the inferior frontal gyrus (IFG) and anterior midcingulate cortex (ACC/MCC) (Lamm et al. 2011). This meta-analysis also classified studies into different experimental paradigms. (b) In cue-based paradigms, pain in others is signaled via abstract cues. In the example stimuli, colored arrows indicate whether the other or the self will receive a nonpainful sensation or a painful shock. This paradigm type does not explicitly provide depictions of painful situations, and thus may more likely rely on internally generated processes and exclude effects of emotion contagion. (c) In picture-based paradigms, pictures or videos that depict limbs of target persons in painful situations are shown to the observer. In the example stimuli, one image indicates pain in the other, whereas the other image does not (Jackson et al. 2005, Lamm et al. 2007b). In addition to eliciting empathy, this paradigm form may also elicit sensorimotor processes. (d) Higher activations during cue-based than during picture-based studies were found in so-called mentalizing or Theory of Mind networks, including temporo-parietal junction (TPJ), ventromedial prefrontal cortex (vmPFC), middle/superior temporal gyrus (M/STG), precuneus and posterior cingulate cortex (PCC) (Lamm et al. 2011). (e) Higher activations during picture-based than during cue-based paradigms were found in so-called mirror-neuron networks, such as the inferior-parietal cortex (IPC) and IFG, as well as in AI and dorsomedial and dorsolateral prefrontal cortex (dIPFC) (Lamm et al. 2011). Adapted from Jackson et al. (2005), Lamm et al. (2011), Lamm et al. (2007b), with permission.
authors confirmed common AI activation during the observation and imagination of disgust in others that overlapped with activations when subjects tasted bitter liquids themselves (Jabbi et al. 2008). Similar to findings in the domain of pain, common AI activations were accompanied by differential coactivation across these various conditions. Indeed, while AI activation showed increased functional connectivity only with IFG regions during the observation of disgust, the direct experience and imagination of disgust were related to more extended network coactivations (Jabbi et al. 2008).

Insular and adjacent frontal-opercular regions are also activated when subjects witnessed positive affective states. As in the perception of disgusted facial expression, AI activity was reported when subjects observed pleased facial expressions in others (Jabbi et al. 2007). Moreover, a recent study that induced compassion and admiration reported activation in AI, dACC, and hypothalamus. Interestingly, AI responses had a faster onset when witnessing physical pain compared with social pain or admiration for positive attributes of others (Immordino-Yang et al. 2009). Using meta-analysis, Fan and colleagues (2011) summarized empathic responses across various domains. Although the included studies were mostly based on pain, AI and ACC activation could also be confirmed when subjects observed fear, happiness, disgust, or anxiety in others.

Preliminary evidence also indicates that AI and ACC/MCC may not necessarily be involved in the vicarious sharing of all states. Studies based on the observation of neutral touch reported shared activations in somatosensory cortices, but not in limbic structures (Blakemore et al. 2005, Ebisch et al. 2008, Keysers et al. 2004). Moreover, a recent study that measured subjects who observed socially desirable others being rewarded demonstrated activations in the ventral striatum, a region involved in reward processing (Mobbs et al. 2009). The perceived similarity between the target and observer correlated with increased activity in ventral ACC, possibly mediating an effect of self-relevance in vicarious reward.

INSULA AND ACC: CONNECTIVITY AND FUNCTIONS

Relatively consistent activations of AI and ACC/MCC in empathy suggest an important role of these two regions in vicariously sharing many emotions and sensations. However, joint insular and cingulate activations in vicarious emotions do not imply that these regions are empathy regions per se. Instead, these regions are known to participate in a multitude of sensory, affective, cognitive, and motivational processes (see the 2010 Special Issue on Insula in Brain Structure & Function). In Craig’s influential model, insular cortex plays a major role in representing and integrating internal and emotional feeling states; ACC, in turn, forms the motivational and action-related counterpart (Craig 2002, 2009). The diverse functional involvement of these regions also suggests that empathy might be a special case of general computational processes related to representing and predicting affective states in the self and others and of guiding adaptive homeostatic and behavioral responses (Singer et al. 2009, Singer & Lamm 2009). In the following section, we describe evidence for the functional implications of these regions based on their patterns of connectivity, their roles across multiple domains, and their frequent coactivation.

Insula

Connectivity and functional data support that the insula plays an important integrative role in sensation, affect, and cognition. Buried within the Sylvian fissure at the interface of frontal, temporal, and parietal lobes (Figure 2a,b; Ture et al. 1999), the insula is cytoarchitectonically defined by a rostrocaudal transition from agranular AI to granular PI (Gallay et al. 2011, Mesulam & Mufson 1982a). Tract-tracing experiments in nonhuman primates suggest that AI is densely connected with prefrontal regions, such as orbitofrontal cortex (OFC) and dorsolateral PFC, temporo-limbic regions, such as temporal poles, parahippocampal
cortices, amygdala, and cingulate cortex, and subcortical targets in the thalamus, basal ganglia, and brain stem (Amaral & Price 1984, Augustine 1996, Fudge et al. 2005, Mesulam & Mufson 1982b, Mufson & Mesulam 1982). Different segments of the insula are highly interconnected themselves, allowing a bidirectional flow of information between anterior and midposterior segments (Craig 2009). Patterns of connectivity in animals have recently been reproduced in humans using diffusion tractography (Cerliani et al. 2012, Nanetti et al. 2009) and resting-state fMRI signal correlations (Cauda et al. 2011, Deen et al. 2012). These patterns indicate a central role, especially of AI, in integrating interoceptive and affective information (Craig 2009, Critchley et al. 2004, Kurth et al. 2010). According to Craig’s model, information of the body’s physiological state is mapped to more posterior insular segments and subsequently rerepresented in the AI, where it may become consciously accessible, enabling a subjective affective experience and global feeling state (Craig 2002, 2009).
Cingulate Cortex

In the limbic system, the cingulate cortex represents the motivational-premotor counterpart of the rather sensory-predictive insula (Craig 2009) and has long been acknowledged as a hub region in affective, cognitive, and motor control phenomena (Paus 2001). Encircling the corpus callosum ventrally, dorsally, and posteriorly, it comprises at least four different cytoarchitectonic subregions, namely ACC, MCC, PCC, and retrosplenial cortex (Vogt et al. 2005). These subregions differ in terms of connectivity, as indicated by animal tract-tracing data as well as diffusion tractography and fMRI signal correlations in humans (Beckmann et al. 2009, Margulies et al. 2007, Vogt & Pandya 1987). Within subregions, connection patterns may also vary significantly. Indeed, whereas rostral ACC densely connects to lateral and orbital PFC and temporo-limbic regions (Pandya et al. 1981, Vogt & Pandya 1987), caudal divisions around dACC/aMCC show a relative increase of functional connections to sensorimotor regions (Margulies et al. 2007). This region receives direct projections from ascending pain pathways (Dum et al. 2009) and is interconnected with the insula (Mesulam & Mufson 1982b, Mufson & Mesulam 1982, Vogt & Pandya 1987), amygdala (Morecraft et al. 2007), ventral striatum (Kunishio & Haber 1994), and PAG (Hardy & Leichnetz 1981). In addition to the direct experience of pain and empathy, other negative affective states and anticipation thereof activate dACC/aMCC (Buchel et al. 1998, Ploghaus et al. 1999, Porro et al. 2002). Moreover, dACC/aMCC involvement in various attentionally or cognitively demanding tasks indicates that this region may implement general monitoring and control processes across multiple domains (Paus 2001). A recent meta-analysis of 939 studies showed overlapping activations in dACC/aMCC during negative affect, pain, and cognitive control (Shackman et al. 2011). The authors suggested that this region synthesizes information about unlearned and learned reinforcers. This may then bias adaptive responding in motor centers responsible for expressing affect and executing goal-directed behavior and ultimately guide behavior in uncertain, potentially aversive environments.

Interceptive Network Interactions

AI and ACC/MCC share a close functional relationship within various flexibly recruited and distributed networks (Craig 2009, Devinsky et al. 1995, Medford & Critchley 2010, Seeley et al. 2007, Sridharan et al. 2008, Taylor et al. 2009). In their seminal article, Devinsky and colleagues acknowledged that AI and ACC/MCC, together with limbic and subcortical regions such as OFC, amygdala, PAG, and ventral striatum, form a coherent network that assesses the motivational content of internal and external stimuli to regulate context-dependent behaviors (Devinsky et al. 1995). Integrating evidence across multiple domains of joint AI and ACC/MCC activation, Medford & Critchley (2010) recently suggested that while the AI forms an input region of a system that is based on self-awareness, these global emotional feeling states are ultimately rerepresented in ACC to control, select, and prepare appropriate responses. Indeed, a close functional relationship between AI and ACC/MCC was recently shown using resting-state fMRI connectivity analysis (Taylor et al. 2009). The authors suggested that this link may enable an integration of interoceptive information with salience. Seeley et al. (2007) also suggested a role in saliency processing in their study observing a correlation between the degree of functional coupling of AI and ACC together with limbic cortical and subcortical regions and anxiety ratings outside the scanner. Using Granger causality analysis of fMRI signals, the study showed that these salience networks may switch between otherwise relatively anticorrelated executive task-activated networks such as dorsolateral PFC and posterior parietal cortex and default-mode networks such as ventromedial PFC and PCU (Sridharan et al. 2008), which are generally more active during stimulus-independent thought, self-projection,

Recently, Singer, Critchley, and Preuschoff provided a framework that related findings of AI activations during empathy and affective states with neuroeconomic reports of a role of AI in uncertainty processing (Singer et al. 2009). Indeed, these studies have also shown AI activations during the processing of risk, risk prediction error, and uncertainty in decision making (Grinband et al. 2006, Huettel et al. 2006, Kuhnen & Knutson 2005, Paulus et al. 2003, Preuschoff et al. 2008). According to this model, AI integrates modality-specific information from multiple feeling states and uncertainty information with individual risk preferences and contextual information. These computations are thought to contribute to the generation of current and predictive feeling states and may ultimately facilitate error-based learning in the affective domain as a prerequisite for successful decision making under uncertain conditions. These representations in AI enable the formation of affective predictions related to the self but also related to predictions of other people’s feeling states. Finally, insula computations can be fed to valuation regions such as the OFC and ventral striatum and also to ACC for response selection and control. Strong interconnections of AI and ACC, and their hub-like position in multiple functional networks, also make them ideally suited to integrate interoceptive information with contextual input into global feeling states, ultimately allowing for modulation of decisions and action responses (Singer et al. 2009, Singer & Lamm 2009). Empathizing with others may thus relate to the involvement of AI and ACC in generating forward models of feeling states for others that, together with certainty computations, may enable one to predict and understand the social and affective behavior of others.

AI and ACC contain a distinctive class of spindle-shaped cells, the Von Economo neurons (Allman et al. 2010, Craig 2009, Von Economo 1926). Their large size and relatively simple dendritic morphology make them suitable for rapid communication between AI and ACC, allowing a fast integration of global affective states, motivation, control, and behavior in dynamic situational contexts (Allman et al. 2010). Comparative histological assessments suggest that these cells are numerous in adult humans; fewer are found in infants, great apes, elephants, and whales (Allman et al. 2010). Also based on the observation that these cells may be selectively destroyed in frontotemporal dementia, a neurodegenerative disorder associated with deficient empathy and socio-emotional functioning (Seeley et al. 2006), some investigators have suggested that Von Economo neurons play a role in empathy, social awareness, and self-control (Allman et al. 2010, Craig 2009).

**MODULATION OF EMPATHY**

Together with their frequent activation across various situations, patterns of structural connectivity of AI and ACC suggest that these two regions may integrate information from a range of different domains to allow the flexible selection of adaptive responses. Indeed, in the domain of empathy, ample data have shown how vicarious responses in AI and subsequent overt behavior can be modulated by various factors, such as those related to individual traits and situational contexts (Figure 3).

**Person Characteristics**

Individual differences in person characteristics likely affect empathic responses. To measure empathic traits, several relatively easy, reliable, and reproducible self-report questionnaires have been developed, including the Interpersonal Reactivity Index (IRI) (Davis 1983) and the Balanced Emotional Empathy Scale (BEES) (Mehrabian 1997). In empathy for pain, such scales have been correlated with empathic responses. Although results have been mixed (Lamm et al. 2011), some studies have shown a modulation of empathic responses by
empathic traits (Jabbi et al. 2007; Saarela et al. 2007; Singer et al. 2004, 2006). For example, Singer and colleagues (2004, 2006) reported that scores on the IRI empathic concern subscale and BESs correlated with AI and dACC activity during vicariously felt pain delivered by abstract cues. In another study, researchers displayed faces in pain and found similar correlations in AI and IFG (Saarela et al. 2007). Extending these findings to the domain of taste, Keysers’s group observed correlations between fronto-insular activations and several IRI subscales when subjects witnessed disgusted and pleased facial expressions (Jabbi et al. 2007). In conclusion, these experiments suggest that empathy traits may, under some conditions, modulate empathic brain responses. The meta-analysis by Lamm and colleagues (2011) has also summarized data for a modulation of empathic responses by trait measures and observed that, compared with trait measures of empathy, more robust brain-behavior correlations can be depicted when using state measures of felt empathy or unpleasantness in a trial-by-trial fashion or state questionnaires of felt empathic concern in the given situation.

Several studies have shown that empathic responses may be affected by alexithymia. Although this trait is present in ~10% of the general population (Salminen et al. 1999), elevated proportions are found in patients with autism spectrum disorder (ASD), a neurodevelopmental condition associated with communicative and social deficits (Hill et al. 2004). In a study by Silani and colleagues (2008), alexithymia and empathy self-report scores in controls and individuals with high-functioning ASD were found to be correlated with AI activity when subjects had to interocept on their feelings. A similar modulation was also observed during a cue-based empathy for pain design, in which empathy-related brain responses in AI were strongly modulated by the degree of alexithymia in controls and individuals with ASD (Bird et al. 2010). Indeed, the greater the participants’ deficits in understanding their own emotions, regardless of whether they were control subjects or patients, the less activation they showed in AI while empathizing with people present in the same room undergoing painful experiences (Figure 3a). These results thus confirm the hypothesis that representations in AI underlie

**Figure 3**

Modulation of empathy. Several cue-based studies have shown a modulation of empathy-related brain responses in fronto-insular regions, such as anterior insula (AI; left and center panels), associated in some studies with an antagonistic response in the nucleus accumbens of the ventral striatum (NAcc; right panels). (a) Modulation of empathy by personal traits, such as the degree of alexithymia (Bird et al. 2010). Activity in an independent region of interest in the left AI during empathy for pain in others was shown to correlate with alexithymia scores [Toronto Alexithymia Scale (TAS), a self-report measure of alexithymia] in controls (green) and individuals with autism spectrum disorder (ASD). The stronger the participants’ deficit was in understanding their own emotions, the lower empathy-related activation in left AI was when witnessing another person suffering pain. This effect was seen in patients and controls alike. (b, c) Modulation of empathy by perceived fairness (Singer et al. 2006). (b) If a target person behaved unfairly in an economic game prior to scanning, men (blue) but not women (purple) showed reduced bilateral AI activity when the unfair but not fair person is in pain. (c) The decrease in activity in AI was paralleled in men but not women by an increase in activation in left NAcc when unfair others receive pain relative to fair others. (Lower panel) The degree of activity in left NAcc was correlated to the subjectively expressed desire for revenge in men. (d, e) Modulation of empathy by perceived ingroup or outgroup membership (Hein et al. 2010). (d) Male soccer fans showed reduced left AI activity when fans favoring an opposing team received pain relative to painful stimulation of fans favoring the same team. (Upper right panel) Attitude toward outgroup member furthermore correlates with AI activity, with reduced AI activity for outgroup members that subjects did not view positively. (Lower panel) The stronger the participants’ AI responses were to ingroup pain relative to outgroup pain, the more they helped the ingroup member relative to the outgroup member in a subsequent session, in which subjects could choose to receive pain in order to reduce the other’s pain. (e) Increased NAcc activity in male soccer fans when disliked soccer fans of the opposing team received painful stimulation. (Lower panel) Increased NAcc activity when an outgroup member was in pain predicts help in delivering painful stimulation to an outgroup member in subsequent helping sessions. For the exact region of interest specifications, please see the original articles. Adapted from Bird et al. (2010), Hein et al. (2010), Singer et al. (2006) with permission.
representations of our own feeling states, which in turn form the basis for understanding the feelings of others. Thus, understanding your own feeling states may be a prerequisite to engage in vicarious simulation for a better understanding of other people’s feeling states. Finally, in addition to providing insight into the modulation of empathic responses in healthy subjects, these studies promote a more differentiated picture of social deficits observed in ASD.

**Contextual Appraisal**

Attribution of specific traits to the target may influence empathic responses of the observer. In an early study focusing on modulation of empathy-related brain responses in the domain of pain in AI, Singer and colleagues (2006) asked participants to engage in a game with confederates, who followed either fair or unfair strategies, prior to scanning. During subsequent scanning, participants watched cues indicating pain in the confederate. Measuring empathic responses, the authors reproduced well-known patterns of brain activity in the insula and dorsal ACC when fair confederates received painful shocks. However, in males these activations were reduced when unfair players received shocks (Figure 3b); instead, males but not females showed increased activation in the nucleus accumbens of the ventral striatum (NAcc) (Figure 3c, upper panel), which was further correlated with the desire for revenge (Figure 3c, lower panel). These findings thus provided initial evidence for a link between fairness behavior and the degree of affective sharing. Moreover, activation in valence and reward-prediction areas such as NAcc or OFC (O’Doherty et al. 2004, Rolls 2004, Schultz 2000) may provide a neurobiological mechanism that helps reinforce punishment of defectors in social situations (Fehr 2008).

These results were extended in a study focusing on the modulation of empathy-related brain responses in AI as a function of perceived group membership (Hein et al. 2010). As in the case of fairness, participants showed stronger empathy-related brain responses in AI toward ingroup compared with outgroup members (Figure 3d, left panel); this result was also reflected by a correlation between AI activity and their impression of the outgroup members, with stronger AI activation for positively valued outgroup members and reduced AI activations when outgroup members were seen negatively (Figure 3d, upper right panel). Moreover, participants who valued outgroup members more negatively showed increased NAcc activation when observing them in pain, compared to outgroup members who were valued more favorably (Figure 3e, upper panel).

Last, the stronger the participants’ AI responses were to ingroup pain relative to outgroup pain, the more they helped the ingroup member relative to the outgroup member in a subsequent session, in which subjects could choose to receive pain in order to reduce the other’s pain (Figure 3d, lower right panel). Conversely, increased NAcc activity when observing outgroup members in pain predicted the absence of subsequent helping (Figure 3e, lower panel). Thus, these two studies suggest that activation in reward-related areas such as NAcc in the context of empathy-for-pain paradigms may reflect an antagonistic motivation to empathy, possibly related to feelings of revenge and Schadenfreude, counteracting empathic resonance in regions such as AI. Whenever NAcc activation is high while witnessing the suffering of another person, activation in AI is low. This activation pattern may ultimately be associated with a lack of engagement in prosocial behavior. Therefore, the social evaluation of the suffering person may modulate the balance of the two systems, which in turn motivate either egoistic or altruistic behavior.

Whereas the above-mentioned studies utilized implicit contextual information, several experimenters provided the participants with explicit information that was hypothesized to influence the evaluation of the observed situation and, in turn, empathy (Decety et al. 2009a; Lamm et al. 2007a,b). Lamm and colleagues, for example, showed a series of needle injections and also informed the participants that
injections were administered to either normal, pain-sensitive hands or anesthetized hands undergoing biopsy (Lamm et al. 2007b). In another study from the same laboratory, participants watched faces of patients undergoing a special painful sound-based treatment, and they were provided with additional information about the treatment success or failure (Lamm et al. 2007a). As in the case of implicit information in the previous studies (Hein et al. 2010, Singer et al. 2006), these explicit contextual factors modulated hemodynamic responses in regions relevant to empathy, such as AI and ACC.

The above results strongly indicate that individual person characteristics and contextual appraisal modulate empathic brain responses. Moreover, there is also some, albeit less consistent, evidence for a modulation through other factors, including attention (Gu & Han 2007, but also see Avenanti et al. 2006 and Gu et al. 2010), neuropeptides such as oxytocin (Riem et al. 2011, but also see Singer et al. 2008), and perspective taking (Cheng et al. 2010, Lamm et al. 2007a). Studying the effects of perspective taking, Cheng and colleagues (2010) recently observed that adopting the perspective of the self and a loved one in pain, compared with that of a stranger in pain, leads to increased pain ratings (Cheng et al. 2010). Moreover, self- and loved-one perspectives induce AI and ACC activations, whereas adopting the stranger perspective results in superior frontal gyrus and TPJ activity. These findings suggest that imagining oneself, or a loved one, in pain will trigger responses in empathy networks. Conversely, projecting a stranger into the same situation may recruit regions also involved in self-other distinction. These network activations may thus contribute to reevaluating the affective meaning for the self of a given situation and may ultimately be a form of emotion-regulation strategy. Thereby, it can allow the subject to adapt responses to current situational demands, for instance, by increasing empathy for loved ones or to reduce distress when seeing strangers in pain.

CONCLUSIONS AND OPEN QUESTIONS

This review pulls together the growing evidence for the existence of shared representations activated by the firsthand experience and the vicarious experience of sensations or emotions experienced by another person. Recent meta-analyses on empathy fMRI studies revealed a special role of AI and dACC/aMCC in empathy for many, though not all, feelings and sensations. Representation in these interoceptive regions may be involved in integrating current and predictive information of feeling states in the self and others. These representations may ultimately enable adaptive responses to the social and affective behavior of others by promoting forms of other-oriented prosocial helping or self-oriented withdrawal behavior to counteract distress (Klimecki & Singer 2012), possibly via the ACC/MCC.

Questioning early automaticity assumptions of empathic brain responses (Preston & de Waal 2002), several recent studies focusing on the modulation of these empathy-related brain responses revealed that activation in these regions is not fixed but can be dynamically modulated by several factors related to situational context or person characteristics. Furthermore, depending on the particular situation and information available in the environment, empathic responses may furthermore coengage other neural networks relevant for social cognition such as those observed in mentalizing or action observation. Future research should further explore the complex interaction between different routes of social cognition in producing an empathic understanding of other peoples’ mental lives.

Additional open questions remain, such as the identification of the computational processes carried out by AI and dACC/aMCC. In this regard, the role of Von Economo neurons in mediating interoceptive signal exchange may be of interest. Moreover, lesion studies and animal recordings, together with methodological advancements in large-scale network modeling, meta-analysis, pattern classification, or...
repetition suppression study design may help to elucidate further the exact role of different regions involved in empathic networks (Bullmore & Sporns 2009, Caruana et al. 2011, Kurth et al. 2010, Shamay-Tsoory et al. 2009, Yarkoni et al. 2011).

Only initial evidence suggests affective plasticity and long-term alterations of brain systems involved in empathy and related positive affect such as compassion. For example, two recent cross-sectional studies by Lutz and colleagues have shown increased functional activation in insula and mentalizing networks in compassion meditation experts compared with novice meditators (Lutz et al. 2008, 2009). In a recent longitudinal study, Leiberg and colleagues (2011) furthermore observed that even one-day compassion training may enhance prosocial helping behavior, and preliminary evidence suggests that this may relate to altered functional activations in brain areas associated with positive affect, reward such as the mid-insula, medial OFC, putamen, and ventral tegmental area (Klimecki et al. 2011). Training compassionate responses may therefore increase the resiliency to aversive events, possibly by upregulating networks associated with positive affect, reward, and attachment (Klimecki et al. 2011).

Insights into plastic alterations in networks underlying empathy and social cognition in general can also be gained from investigating clinical and subclinical groups associated with deficient empathic ability, such as individuals with conduct disorder, ASD, and alexithymia (Bird et al. 2010, Silani et al. 2008). These studies may provide further insights into abnormal predispositions to unempathic responding, how neural plasticity can go awry, and the influence of deficient empathy on aggressive, impulsive, and selfish behavior (Birbaumer et al. 2005, Boccardi et al. 2010, Decety et al. 2009b, Kiehl et al. 2001, Sterzer et al. 2007, Tiitinen et al. 2008).

Last, despite abundant previous research carried out by developmental behavioral psychologists on the ontogeny of empathy in childhood (Eisenberg 2000, Knafo et al. 2008), the field of social neuroscience has only just started to address important developmental brain changes related to our ability to empathize and the relationship of these changes to moral reasoning and prosocial behavior (Decety & Michalska 2010, Singer 2006).

In the past few years, we have begun to understand better the neural basis of empathy and related states such as compassion. Identifying crucial subcomponents and brain network interactions involved in empathy sheds important light on the generation of this multifaceted experience at the heart of human emotional and social behavior. Ultimately, such insights may guide the development of strategies for circumventing aversive behavior and burnout syndromes in caregivers and physicians (Halifax 2010, Hojat et al. 2009, Klimecki & Singer 2012) and may lead to advances in nourishing socio-affective competences in children and in adults suffering from conduct disorders and ASD.

### SUMMARY POINTS

1. Neuroimaging studies on empathy for pain consistently revealed activations in AI and dACC/aMCC when directly experiencing pain as well as when empathizing with the pain of others, suggesting that empathy depends, in part, on shared representations.

2. The vicarious experience of affective states other than pain, such as social exclusion, disgust, anxiety, and taste, also activates AI and dACC/aMCC; nevertheless, initial evidence suggests that these regions may not necessarily be involved in the vicarious sharing of all states.
3. Depending on the situational context and information available in the environment, empathic responses may involve a corecruitment of so-called mirror-neuron networks and regions involved in theory of mind or mentalizing.

4. Empathic brain responses are not fixed but may be modulated by person characteristics such as degree of alexithymia or contextual appraisal such as perceived fairness of another person or group membership.

5. Being crucial hubs in human interoceptive cortex, AI and dACC/aMCC may perform domain-general computations that represent and predict feeling states and guide responses to the emotional experience of the self and others.

DISCLOSURE STATEMENT
The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS
We thank our colleagues Jonathan Smallwood, Haakon Engen, Cade McCall, and Joshua Grant for their insightful comments on earlier versions of the manuscript. We thank Alfred Anwander, Claus Lamm, and Grit Hein for kindly providing some of the data presented in the figures. We thank Sandra Zurborg for carefully proofreading the final version of this article. We thank Stefan Liebig for providing the illustrations in Figure 1. We apologize to all the investigators whose research could not be appropriately cited owing to space limitations. This work was funded by the European Research Council under the European Community’s Seventh Framework Program (FP7/2007-2013)/ERC Grant agreement no. 205557 [EMPATHICBRAIN].

LITERATURE CITED


Jackson PL, Brunet E, Meltzoff AN, Decety J. 2006. Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia* 44:752–61


Contents

The Neural Basis of Empathy
Boris C. Bernhardt and Tania Singer ........................................ 1

Cellular Pathways of Hereditary Spastic Paraplegia
Craig Blackstone ................................................................. 25

Functional Consequences of Mutations in Postsynaptic Scaffolding Proteins and Relevance to Psychiatric Disorders
Jonathan T. Ting, João Peça, and Guoping Feng .......................... 49

The Attention System of the Human Brain: 20 Years After
Steven E. Petersen and Michael I. Posner ................................ 73

Primary Visual Cortex: Awareness and Blindsight
David A. Leopold ............................................................... 91

Evolution of Synapse Complexity and Diversity
Richard D. Emes and Seth G. N. Grant ................................. 111

Social Control of the Brain
Russell D. Fernald ............................................................... 133

Under Pressure: Cellular and Molecular Responses During Glaucoma, a Common Neurodegeneration with Axonopathy
Robert W. Nickells, Gareth R. Howell, Ileana Sato, and Simon W. M. John ............... 153

Early Events in Axon/Dendrite Polarization
Pei-lin Cheng and Mu-ming Poo ........................................ 181

Mechanisms of Gamma Oscillations
György Buzsáki and Xiao-Jing Wang ................................... 203

The Restless Engram: Consolidations Never End
Yadin Dudai ................................................................. 227

The Physiology of the Axon Initial Segment
Kevin J. Bender and Laurence O. Trussell ................................ 249

Attractor Dynamics of Spatially Correlated Neural Activity in the Limbic System
James J. Knierim and Kechen Zhang ................................. 267

Neural Basis of Reinforcement Learning and Decision Making
Daeyeol Lee, Hyojung Seo, and Min Whan Jung ...................... 287
Critical-Period Plasticity in the Visual Cortex

Christiaan N. Levelt and Mark Hülbener ................................................................. 309

What Is the Brain-Cancer Connection?

Lei Cao and Matthew J. During ................................................................. 331

The Role of Organizers in Patterning the Nervous System

Clemens Kiecker and Andrew Lumsden ............................................................ 347

The Complement System: An Unexpected Role in Synaptic Pruning During Development and Disease

Alexander H. Stephan, Ben A. Barres, and Beth Stevens ...................................... 369

Brain Plasticity Through the Life Span: Learning to Learn and Action Video Games

Daphne Bavelier, C. Shawn Green, Alexandre Pouget, and Paul Schrater .................. 391

The Pathophysiology of Fragile X (and What It Teaches Us about Synapses)

Asha L. Bhakar, Gül Dölen, and Mark F. Bear ...................................................... 417

Central and Peripheral Circadian Clocks in Mammals

Jennifer A. Mohawk, Carla B. Green, and Joseph S. Takahashi .............................. 445

Decision-Related Activity in Sensory Neurons: Correlations Among Neurons and with Behavior

Hendrikje Nienborg, Marlene R. Cohen, and Bruce G. Cumming ............................ 463

Compressed Sensing, Sparsity, and Dimensionality in Neuronal Information Processing and Data Analysis

Surya Ganguli and Haim Sompolinsky .............................................................. 485

The Auditory Hair Cell Ribbon Synapse: From Assembly to Function

Saaid Safieddine, Aziz El-Amraoui, and Christine Petit ....................................... 509

Multiple Functions of Endocannabinoid Signaling in the Brain

István Katona and Tamás F. Freund ..................................................................... 529

Circuits for Skilled Reaching and Grasping

Bror Alstermark and Tadashi Iba ....................................................................... 559

Indexes

Cumulative Index of Contributing Authors, Volumes 26–35 ................................. 579

Cumulative Index of Chapter Titles, Volumes 26–35 ............................................. 583

Errata

An online log of corrections to *Annual Review of Neuroscience* articles may be found at http://neuro.annualreviews.org/
New From Annual Reviews:

**Annual Review of Statistics and Its Application**

Volume 1 • Online January 2014 • http://statistics.annualreviews.org

Editor: **Stephen E. Fienberg**, Carnegie Mellon University
Associate Editors: **Nancy Reid**, University of Toronto
**Stephen M. Stigler**, University of Chicago

The *Annual Review of Statistics and Its Application* aims to inform statisticians and quantitative methodologists, as well as all scientists and users of statistics about major methodological advances and the computational tools that allow for their implementation. It will include developments in the field of statistics, including theoretical statistical underpinnings of new methodology, as well as developments in specific application domains such as biostatistics and bioinformatics, economics, machine learning, psychology, sociology, and aspects of the physical sciences.

Complimentary online access to the first volume will be available until January 2015.

**Table of Contents:**

- A Systematic Statistical Approach to Evaluating Evidence from Observational Studies, David Madigan, Paul E. Stang, Jesse A. Berlin, Martijn Schuemie, J. Marc Overhage, Marc A. Suchard, Bill Dumouchel, Abraham G. Hartzema, Patrick B. Ryan
- The Role of Statistics in the Discovery of a Higgs Boson, David A. van Dyk
- Brain Imaging Analysis, F. DuBois Bowman
- Statistics and Climate, Peter Guttorp
- Climate Simulators and Climate Projections, Jonathan Rougier, Michael Goldstein
- Probabilistic Forecasting, Tilmann Gneiting, Matthias Katzfuss
- Bayesian Computational Tools, Christian P. Robert
- Bayesian Computation Via Markov Chain Monte Carlo, Radu V. Craiu, Jeffrey S. Rosenthal
- Build, Compute, Critique, Repeat: Data Analysis with Latent Variable Models, David M. Blei
- Structured Regularizers for High-Dimensional Problems: Statistical and Computational Issues, Martin J. Wainwright
- High-Dimensional Statistics with a View Toward Applications in Biology, Peter Bühlmann, Markus Kalisch, Lukas Meier
- Next-Generation Statistical Genetics: Modeling, Penalization, and Optimization in High-Dimensional Data, Kenneth Lange, Jeannette C. Papp, Janet S. Sinsheimer, Eric M. Sobel
- Breaking Bad: Two Decades of Life-Course Data Analysis in Criminology, Developmental Psychology, and Beyond, Elena A. Erosheva, Ross L. Matsueda, Donatello Telesca
- Event History Analysis, Niels Keiding
- Statistical Evaluation of Forensic DNA Profile Evidence, Christopher D. Steele, David J. Balding
- Using League Table Rankings in Public Policy Formation: Statistical Issues, Harvey Goldstein
- Statistical Ecology, Ruth King
- Estimating the Number of Species in Microbial Diversity Studies, John Bunge, Amy Willis, Fiona Walsh
- Dynamic Treatment Regimes, Bibhas Chakraborty, Susan A. Murphy
- Statistics and Related Topics in Single-Molecule Biophysics, Hong Qian, S.C. Kou
- Statistics and Quantitative Risk Management for Banking and Insurance, Paul Embrechts, Marius Hofert

Access this and all other Annual Reviews journals via your institution at www.annualreviews.org.