The phenomenon of empathy entails the ability to share the affective experiences of others. In recent years social neuroscience made considerable progress in revealing the mechanisms that enable a person to feel what another is feeling. The present review provides an in-depth and critical discussion of these findings. Consistent evidence shows that sharing the emotions of others is associated with activation in neural structures that are also active during the first-hand experience of that emotion. Part of the neural activation shared between self- and other-related experiences seems to be rather automatically activated. However, recent studies also show that empathy is a highly flexible phenomenon, and that vicarious responses are malleable with respect to a number of factors—such as contextual appraisal, the interpersonal relationship between empathizer and other, or the perspective adopted during observation of the other. Future investigations are needed to provide more detailed insights into these factors and their neural underpinnings. Questions such as whether individual differences in empathy can be explained by stable personality traits, whether we can train ourselves to be more empathic, and how empathy relates to prosocial behavior are of utmost relevance for both science and society.

Key words: empathy; social neuroscience; pain; fMRI; anterior insula (AI); anterior cingulate cortex (ACC); prosocial behavior; empathic concern, altruism; emotion contagion

Introduction

Being able to understand our conspecifics’ mental and affective states is a cornerstone of our lives as “social animals.” It enables us not only to communicate and interact with each other in effective and pleasant ways, but also to predict the actions, intentions, and feelings of others. How ordinary the ability to empathize with others appears to us only becomes evident when things go wrong, as when we are misunderstood by someone else and by consequence our feelings get hurt. But even in those cases our immediate affective reaction enables the other person to become aware of the misunderstanding and the emotional consequences of his or her actions. This ability to share others’ feelings ultimately results in a better understanding of the present and future mental states and actions of the people around us and possibly promotes prosocial behavior.

In recent years, the field of social neuroscience has begun to shed light on the neural underpinnings of the phenomenon of empathy. The aim of this review chapter is to give an overview of this research, to discuss shortcomings, and to provide recommendations for future research. Since we mostly focus on studies in the domain of empathy for pain, evidence concerning other emotions is only partially incorporated (see Decety & Jackson 2004; Decety & Lamm 2006; Hein & Singer 2008; Singer & Leiberg 2009; Singer 2006 for more exhaustive coverage of related issues). The present chapter is divided into three parts. First, we provide definitions of empathy and related terms such as compassion, emotion contagion, and personal distress. This is intended to illustrate that sharing the feelings of others, defined here as...
“empathy,” is only one part of a large spectrum of a person’s possible vicarious responses toward others. In the second part, current evidence from social neuroscience studies of empathy will be selectively reviewed. This review will focus on a discussion of the shared representations account of empathy, which is currently the dominant neuroscientifically motivated approach to understanding the mechanisms underlying empathy. In addition, we will reflect on the role of bottom-up and top-down influences in generating and modulating empathic responses and discuss the different motivational consequences carried by the vicarious responses of empathy versus personal distress. In the final part, we propose research questions and domains that should be given special attention by future empathy researchers.

**Definition of Terms: Empathy and Its Sisters**

Despite the word’s linguistic roots in ancient Greek—from empatheia (passion), which is composed of “en” (in) and “pathos” (feeling)—the scientific scrutiny of empathy has a relatively short history that can be dated back to its use in philosophical aesthetics. From there the English term originated as a direct translation of the German Einfühlung (“feeling into” something), a term that was originally proposed as a tool for analyzing works of art and nature, but later developed into a more general mechanism for recognizing each other as “minded creatures” (Stüber 2008). After this initial philosophically motivated period of scrutiny, most of the research on empathy was performed by developmental and social psychologists (e.g., Batson 1991, 2009; Hoffman 2000; Eisenberg 2000; Eisenberg & Strayer 1987). Naturally, social psychologists showed particular interest in this inherently social phenomenon (e.g., Batson 1991; Davis 1994). A great deal of social psychological research has been devoted to the question which perceptual, affective, and cognitive mechanisms enable us to “put ourselves into someone else’s shoes.” Surprisingly, it took quite some time for the field of neuroscience and in particular of functional neuroimaging to dare to make contributions to this challenging pursuit (Carr et al. 2003; Keysers et al. 2004; Morrison et al. 2004; Singer et al. 2004; Wicker et al. 2003). This might be attributed to the complexities inherent in this multidimensional psychological phenomenon as well as to the methodological challenges of bringing such an idiosyncratic and context-dependent phenomenon into a scientific environment that requires well-controlled and reproducible experiments.

At a basic phenomenological level, empathy denotes an affective response to the directly perceived, imagined, or inferred feeling state of another being (for an excellent overview of the various “things called empathy,” see Batson 2009). In our own understanding, empathy occurs when an observer perceives or imagines someone else’s (i.e., the target’s) affect and this triggers a response such that the observer partially feels what the target is feeling. De Vignemont and Singer (2006, p. 435), for example, define empathy as follows: We “empathize” with others when we have (1) an affective state (2) which is isomorphic to another person’s affective state, (3) which was elicited by observing or imagining another person’s affective state, and (4) when we know that the other person’s affective state is the source of our own affective state. However, there are almost as many definitions of empathy as there are researchers in the field. It is therefore essential to briefly review and define the relevant key concepts and components generally associated with the broad concept of empathy, such as mimicry, emotional contagion, sympathy, and compassion. Although these concepts each refer to a different phenomenon, they usually occur in concert. In most cases, mimicry or emotional contagion precedes empathy, which precedes sympathy and compassion, which in turn may precede prosocial behavior.
The Big and Little Sisters of Empathy

First, we need to distinguish between mimicry and emotional contagion, which can both contribute substantially to an empathic response. Mimicry is defined as the tendency to automatically synchronize affective expressions, vocalizations, postures, and movements with those of another person (Hatfield et al. 1994). Our understanding of its role as a low-level mechanism contributing to empathy derives from a multitude of studies using facial electromyography. These studies demonstrate that when an observer perceives another person’s affective facial expressions, such as a smile or a frown, corresponding affective expressions result in the observer (see Dimberg & Oehman 1996, for a review). Relying on the facial feedback hypothesis, according to which one appraises one’s own emotions by perceiving their bodily concomitants, Sonnby-Borgström (2002) proposed that mimicry enables one to automatically share and understand another’s emotions. Her proposal also receives support from studies showing a (notably, weak) correlation between the strength of the mimicry response and trait measures of empathy.

However, facial responses are not only triggered when we observe others, but also when we are exposed to negatively or positively valenced visual stimuli without social relevance (Cacioppo et al. 2000, for a review). In addition, several investigations demonstrate the influence of top-down processes on mimicry, such as those associated with the relationship between empathizer and target (Lakin & Chartrand 2003), the affective state of the observer (Moody et al. 2007; Niedenthal et al. 2001), or the perspective from which pain in others is witnessed (Lamm et al. 2008). These observations cast doubt on the assumption that mimicry represents some sort of automatic or “hard-wired” motor resonance with another person’s affective display. Furthermore, mimicry seems to serve a social function in increasing rapport and fondness between self and other, raising the question whether this function evolved for communicative rather than for epistemological reasons (Chartrand & Bargh 1999; van Baaren et al. 2004).

Emotional contagion is another process that is related to but distinct from empathy. It denotes the tendency to “catch” other people’s emotions and has alternately been labeled “primitive empathy” (Hatfield et al. 2009). For example, babies start crying when they hear other babies crying—long before they develop a sense of a self separate from others. Recently, initial evidence for involuntary pupillary contagion was found in an fMRI study (Harrison et al. 2006). Participants were presented with photos of sad faces with various pupil sizes. Their own pupil size was significantly smaller when they viewed sad faces with small as compared to larger pupils, and the Edinger–Westphal nucleus in the brainstem, which controls pupil size, was specifically engaged by this contagious effect. Activation in this subcortical structure provides evidence that pupillary contagion occurs outside of awareness and may represent a precursor of empathy. This study also demonstrates the strong overlap between mimicry and emotional contagion (e.g., Hatfield et al. 1993). Note, however, that there are cases in which mimicry occurs without an emotional component and other cases in which emotions are automatically elicited by observing others’ emotional states without the involvement of motor mimicry.

On a conceptual level, neither emotional contagion nor mimicry cannot account for the full-blown experience of empathy. In our understanding empathy crucially depends upon self-awareness and self/other distinction; in other words, on our ability to distinguish between whether the source of our affective experience lies within ourselves or was triggered by the other (de Vignemont & Singer 2006; Decety & Jackson 2004; Decety & Lamm 2006). Without this ability, witnessing someone else’s emotions could, for example, result, purely, in personal distress and a self-centered response in the observer. We therefore regard mimicry and emotional contagion as important, yet distinct
and neither necessary nor sufficient processes for the experience of empathy.

On the level of vicariously felt responses, we have to distinguish between empathy, sympathy, empathic concern, and compassion. In all four cases, affective changes are induced in the observer in response to the perceived or imagined affective state of another person. However, while empathy involves feelings that are isomorphic to those of the other person, sympathy, empathic concern, and compassion do not necessarily involve shared feelings. For example, empathizing with a person feeling sad will result in a feeling of sadness in the self, whereas sympathizing with, being empathically concerned, or feeling compassion for a sad person will result in either pity or compassionate love for the person, but not sadness. Also, when an observer notices that someone is jealous of him, he will most likely not start feeling jealous himself—though he might show sympathy or compassion for the jealous person (for similar arguments, see de Vignemont & Singer 2006). Note also that the terms empathic concern, sympathy, and compassion have sometimes been treated as synonymous (cf. Batson 2009) and that the most widely used psychometric measures of empathic concern (see below) involve self-report of compassionate, sympathetic, or tender feelings. Therefore, the crucial distinction between the term empathy and those like sympathy, empathic concern, and compassion is that empathy denotes that the observer’s emotions reflect affective sharing (“feeling with” the other person) while compassion, sympathy, empathic concern denote that the observer’s emotions are inherently other oriented (“feeling for” the other person).

Further, folk psychological accounts usually relate the occurrence of empathy to prosocial and altruistic, other-oriented motivations (i.e., a motivation with the goal to increase the other person’s well-being or welfare). Empathy, though, does not necessarily carry such motivations, and real-life examples of how empathy can “go awry” (from a prosocial point of view) abound. For example, a torturer may use empathy in order to sense how to increase his victim’s suffering; in competitive environments (from sports to business operations to warfare), successful tactics take into account the negative affective effects that an action will have on the opponent; and experiencing too much empathy can lead to an aversive distress response and selfish instead of other-oriented behavior (see below). In general, however, empathy is conceived to be a first necessary step in a chain that begins with affect sharing, followed by understanding the other person’s feelings, which then motivates other-related concern and finally engagement in helping behavior. Empathy and prosocial behavior are thus closely linked on a conceptual level. Notably, while consistent evidence for the link between “feeling for” (empathic concern) and prosocial behavior exists (e.g., Batson 1991; Eisenberg 2000; Eisenberg et al. 1989), a clearcut empirical demonstration of a link between empathy and prosocial behavior is still missing.

Empathy Research in the Context of Social Neuroscience: The Shared Network Hypothesis

Various scholars have proposed that we come to understand the actions, sensations, and emotions of others by the activation of neural representations corresponding to those states. Inspired by earlier perception—action models in the domain of action understanding (Prinz 2005), Preston and de Waal (2002) proposed a neuroscientific model of empathy, one which suggests that observing or imagining another person in a particular emotional state automatically activates a representation of that state in the observer, along with its associated autonomic and somatic responses. Other authors have also suggested that shared neural representations play a general role in understanding other people’s mental states. They claim that shared representations provide us with a simulation of their corresponding sensorimotor, affective, or mental states (Gallese 2003a;
Goldman 2006). Such accounts hold that the capacity to project ourselves imaginatively into another person’s perspective by simulating their mental activity using our own mental apparatus lies at the root of our mature mind-reading abilities, and the reasoning of these accounts has been extended to the domains of actions and feelings: To understand what another person is doing we simulate his movements using our own motor program; to understand what another person is feeling, we simulate his feelings using our own affective programs (see also Keysers & Gazzola 2006). Indeed, this so-called shared representations account of social interaction and intersubjectivity has become the dominant explanation of the hemodynamic activation patterns observed in recent fMRI studies of empathy.

The majority of social neuroscience studies on empathy used the observation of pain in others as a model paradigm to evoke empathic responses (de Vignemont & Singer 2006; Decety & Lamm 2006; Singer & Leiberg 2009; see, e.g., Jabbi et al. 2007 or Wicker et al. 2003, for examples using other emotional states). One common finding of these investigations is that vicariously experiencing pain activates part of the neural network that is also activated when we are in pain ourselves. For example, Singer and colleagues (2004) recruited couples and measured hemodynamic responses triggered by painful stimulation of the female partner via an electrode attached to her right hand. In another condition the same painful stimulation was applied to the male partner who was seated next to the MRI scanner and whose hand could be seen via a mirror system by the female partner lying in the scanner. Differently colored flashes of light on a screen pointed to either the male or the female partner’s hand, indicating which of them would receive painful stimulation. This procedure enabled the measurement of pain-related brain activation when pain was applied either to the scanned subject (firsthand experience of pain) or to her partner (empathy for pain). The results suggest that parts of the so-called pain matrix (Derbyshire 2000), which consists of the brain areas involved in the processing of pain, were activated when participants experienced pain themselves as well as when they saw a signal indicating that their loved one would experience pain. These areas—in particular, bilateral anterior insula (AI), the dorsal anterior cingulate cortex (ACC), brain stem, and the cerebellum—are involved in the processing of the affective component of pain; in other words, they encode how unpleasant or aversive the subjectively felt pain is. Thus, both the firsthand experience of pain and the knowledge that a beloved partner is experiencing pain activates the same affective brain circuits—suggesting that our own neural response reflects our partner’s negative affect.

This initial finding of shared neural activation between self and other has been replicated and extended using a variety of paradigms and methods. While initially evidence was restricted to the affective component of pain—as indicated by activation restricted to areas involved in coding the affective—motivational aspects of the feeling of pain—there is now evidence that areas associated with somatosensory processing can also be activated when we witness another person’s pain, in particular, when our attention is directed to the somatosensory aspects of the pain experience (e.g., Bufalari et al. 2007; Cheng et al. 2008; Lamm & Decety 2008; Lamm et al. 2007b). Bufalari and colleagues (2007), for example, demonstrated that the amplitude of an event-related potential component known to be generated in primary somatosensory cortex (P45) is modulated by seeing a needle piercing another person’s hand. In a similar vein, a recent fMRI study showed that (contralateral) right primary somatosensory cortex was activated when participants witnessed another person’s left hand being pierced (Lamm & Decety 2008; Lamm et al. 2007b). This activation overlapped with individually determined somatosensory representations of touch of the hand that had been determined in a separate localizer run. The latter finding is an important contribution to the fMRI literature, which has
so far demonstrated shared somatosensory activations only in secondary somatosensory cortex (e.g., Jackson et al. 2006a; Singer et al. 2006) or in parts of primary somatosensory cortex that are not somatotopically specific for the affected body part (Lamm et al. 2007b). Furthermore, a transcranial magnetic stimulation (TMS) study by Avenanti et al. (2006) demonstrated that motor-evoked potentials are modulated when participants observe a needle deeply penetrate the target's hand, but not when they observe a pinprick. This suggests that it is the saliency of the somatosensory quality of pain that determines whether somatosensory areas will or will not be involved in empathy for pain.

To investigate the areas involved in affective sharing during pain empathy more precisely, detailed analyses of activation clusters in the cingulate and insular cortices have recently been performed (Decety & Lamm 2009; Jackson et al. 2006b; Morrison & Downing 2007). These analyses suggest that there is reliable overlap when activation in these areas during firsthand and vicarious experience of pain are compared, but they also show that the majority of voxels in the insula and the cingulate cortex do not overlap. In particular, a recent meta-analysis compared published localizations for the experience of pain to those reported for empathy for pain. The results suggest a more posterior–midinsular activation pattern for the firsthand experience of pain (Decety & Lamm 2009). While this might have been expected for the hemisphere contralateral to the stimulated body part, it is surprising for the ipsilateral (right) hemisphere. The same meta-analysis also suggests overlapping, yet largely distinct activation patterns in medial and anterior cingulate cortex—a finding that is in line with detailed single-subject analyses in native anatomical space (Morrison & Downing 2007).

What is it, then, that is encoded or processed in shared activation areas such as the anterior insula or the medial and anterior cingulate cortex? It has been suggested that these regions represent a crucial part of the human interoceptive cortex (Craig 2003) and subserve neural representations of internal bodily states such as information about temperature, lust, hunger, bodily arousal states, and information from the gut (Craig 2002, 2003; Critchley 2005; Critchley et al. 2004; Damasio 1994). Based on anatomical observations in nonhuman species, Craig (2002, 2003) developed a detailed anatomical model suggesting that an image of the body's internal state is first mapped to the brain by afferents that provide input to thalamic nuclei, sensorimotor cortices, and posterior dorsal insula. Direct activation of both the AI and the ACC may correspond to the simultaneous generation of both a feeling and an affective motivation with its concomitant autonomic effects. The very same structures (AI and ACC) that play a crucial role in representing one's own subjective feeling states also seem to be crucial in processing vicarious feelings. Based on this proposal, Singer and colleagues (2004) extended an interoceptive model of emotions to the domain of empathy and suggested that cortical re-representations in AI of bodily states may have a dual function. First, they may allow us to form subjective representations of our own feelings. These representations not only allow us to understand our own feelings when emotional stimuli are present, but also to predict the bodily effects of anticipated emotional stimuli to our bodies. Second, they may serve as the visceral correlate of a prospective simulation of how something may feel for others. This may then help us to understand the emotional significance of a particular stimulus and its likely consequences. In accordance with this view, it is noteworthy that the anticipation of pain has been found to activate more anterior insular regions, whereas the actual experience of pain recruits more posterior insular regions—confirming the postulated role of more posterior insular regions in modality-specific, primary representations of pain and more anterior regions in the secondary representations of the anticipatory negative affect related to pain (Ploghaus et al. 1999). In line with this observation, in pain empathy studies activity in posterior insular cortices was observed
only when participants were experiencing pain themselves (Lamm et al. 2007a; Singer 2006; Singer et al. 2004), whereas activity in AI was observed when participants were experiencing pain themselves and when vicariously feeling someone else’s pain. In a similar vein, instructions to imagine pain from a first-person perspective or to specifically assess the somatosensory consequences of vicariously perceived pain revealed more posterior activation clusters in the insular cortex (Jackson et al. 2006a; Lamm et al. 2007a,b).

The important role of anterior insular cortex for empathy has also been demonstrated in other domains. AI is known to be crucially involved in processing sensations and emotions such as taste or disgust. Similar neural responses were elicited when participants viewed movies of disgusted faces and when they smelled disgusting odors themselves (Wicker et al. 2003) as well as when they viewed videos showing people sampling pleasant or unpleasant tastes and when they sampled the different tastes themselves (Jabbi et al. 2007). Interestingly, the latter study also revealed that brain responses in anterior insular cortex were positively correlated with individual differences in empathy measured by empathy questionnaires.

The observation of similar neural activations during the firsthand vs. the vicarious experience of various sensations and emotions (e.g., disgust, taste, pain) raises the question whether these activations can indeed be interpreted as shared representations. Shared activations are certainly a good indicator of shared representations. However, apart from the fact that we do not know how observations on one level (the psychological/representational level) are mapped to those on another (the neural one), none of the currently available human neuroimaging methods directly measures activity of single neurons or neural networks (however, see Hutchison et al. 1999). Therefore, two fMRI activation maps with overlapping clusters might still result from differing neural activity. One way to resolve this problem is the use of repetition-priming or fMRI-adaptation designs (Grill-Spector & Malach 2001; Henson & Rugg 2003). Such an approach has already been used in the domain of action observation (Dimstein et al. 2007) and mentalizing (Jenkins et al. 2008), and similar studies are now underway in the domain of empathy research. These studies will show whether activation overlaps in the observer result from the activation of the same neural resources.

Finally, a word of caution is required as to whether the observed involvement of AI or ACC is specifically related to empathy or to emotional contagion (or both). One indication that shared activations during empathy for pain are not simply related to emotional contagion is provided by studies of Singer et al. (2004, 2006). Here, arbitrary visual cues rather than explicit affective displays or other potentially contagious stimuli such as depictions of painful events were used to assess empathic responses. In addition, a recent fMRI study by Lamm et al. (2009) detected substantial responses in AI and ACC when participants were presented with visual stimuli depicting situations that were clearly nonpainful for them (touch of another person’s hand by a Q-tip), but known to be painful for the target. Together, these studies render it less plausible that their neural responses were directly triggered by the perception of overt (and emotionally contagious) correlates of the other person’s emotional state.

Nevertheless, additional studies are required to investigate how the brain actually differentiates between self- and other-related feelings and to dissociate the respective contributions of the two related phenomena.

In sum, while substantial empirical evidence suggests that shared neural activations are at the root of sharing feelings, sensations, and actions of others, additional research is required to clarify whether these activations are actually shared on the neural level, to what extent we share both the somatosensory antecedents or only the affective consequences of another’s affective state, what constitutes the functional significance of these shared activations, and how we can distinguish between emotional
contagion vs. empathy on the neural level. More sophisticated experimental designs, analytical approaches, as well as closer collaborations between philosophers of mind, neurophysiologists, introspection-based contemplatives, cognitive scientists, and social neuroscientists are required to yield more detailed answers to these questions.

**Empathy—Bottom-Up vs. Top-Down Processes**

An important aspect of most neuroscientifically motivated models of empathy is that the activation of shared representations in the observer is initiated mostly automatically and without conscious awareness. For example, in most studies of empathy in the domains of pain, touch, and disgust, participants are not informed that the goal of the study is to investigate empathy-related neural responses. Rather, they are instructed to passively watch a scene or movie depicting a person expressing an emotion or being touched (Blakemore et al. 2005; Keysers et al. 2004; Singer et al. 2004, 2006; Wicker et al. 2003). Nevertheless, this situation alone is sufficient to engage brain networks representing the firsthand experience of affect or touch. Some authors have therefore suggested that we automatically share other people's feelings, a hypothesis in line with earlier perception–action models of motor behavior and imitation and with their extension to the domain of empathy (Gallese 2003b; Preston & de Waal 2002). Preston and de Waal's model (2002), for example, stresses the importance of automatic and perceptually driven processes such as emotional contagion and mimicry. It should be noted though that the term automatic in this case refers to a process that does not require conscious and effortful processing but can be inhibited or controlled (cf. Bargh 1994). In addition, attention to the target's affective state is required for triggering the postulated cascade of events starting with emotional contagion and, ultimately, resulting in a full-blown empathic experience. The crucial role that is attributed to attention, inhibitory, and other executive control processes also documents that empathy is not a purely sensory-driven process in which affective states are induced in the observer solely by means of bottom-up processes. On the contrary, it has long been argued that contextual appraisal, cognitive processes, and top-down control are important constituents of human empathy. For example, in the eighteenth century, the philosopher and economist Adam Smith proposed that imagination enables us to project ourselves into the place of other persons, experiencing sensations that are generally similar to, although typically weaker than, those of the other person (Smith 1759/1976). Most contemporary neuroscientific models of empathy endorse this view and stress the importance of top-down control and contextual appraisal for either the generation of an empathic response or for modulating an existing one induced by the above-mentioned bottom-up processes (deVigenemont & Singer 2006; Decety & Lamm 2006; Hein & Singer 2008; Singer 2006).

Decety and Lamm (2006), for example, proposed a model in which bottom-up (i.e., direct matching between perception and action) and top-down (i.e., regulation, contextual appraisal, and control) information processes are fundamentally intertwined in the generation and modulation of empathy. In this model, bottom-up processes account for direct emotion sharing which is automatically activated (unless inhibited) by perceptual input. On the other end, executive functions implemented in the prefrontal and cingulate cortex serve to regulate both cognition and emotion through selective attention and self-regulation. This meta-cognitive level is continuously updated by bottom-up information, and in return controls the lower level by providing top-down feedback. Thus, top-down regulation, through executive functions, modulates lower levels and adds flexibility, making the individual less dependent on external cues. The meta-cognitive feedback loop also plays a crucial role in taking into account one's own mental competence in order to react (or not) to the affective states of others. This
model should be supplemented by top-down processes that are not classically associated with executive function and its associated neural structures, in particular those in the medial and dorsolateral prefrontal cortex. Mental imagery, for example, has been shown to result in shared representations in both the motor and the sensory-affective domain without considerable prefrontal activation (e.g., Decety & Grezes 2006; Ogino et al. 2007). In addition, contextual appraisal and target evaluation in empathy-for-pain paradigms predominantly activated areas such as the orbitofrontal cortex or the ventral striatum (e.g., Lamm et al. 2007b; Singer et al. 2006), in other words, structures involved in affective evaluation, reward, and punishment.

Several recent findings provide strong support for top-down influences on empathy. These findings document the flexibility of the human mind in responding to others and show that empathy is not an all-or-none phenomenon. Here, we want to focus on the role of attention, contextual appraisal, perspective-taking, and the relationship between empathizer and target as salient examples of how top-down control affects empathic responses (see Hein & Singer 2008 for a more exhaustive account).

Two recent fMRI studies and an event-related potential study demonstrate that the way we attend to the emotions of others substantially modulates our empathic responses to them. In one study participants were exposed to pictures of painful situations—such as cutting one’s finger or getting one’s hand caught in the door. When evaluating the painful consequences of those situations, participants showed activation in large parts of the pain matrix—a now typical response given the evidence reviewed above. However, when participants were instructed to count the number of hands shown in a picture—an experimental manipulation that was devised to distract their attention from the inflicted pain—no activation in insular or cingulate cortices was observed (Gu & Han 2007). In a follow-up electroencephalographic study the same group demonstrated that it is mainly a late event-related potential component that differs between the attention and no-attention conditions, while an early component is not affected by the focus of attentional set (Fan & Han 2008). The early response might correspond to the postulated automatic activation of shared representations, representations which might then be interpreted by later neural processes either inhibiting or augmenting the behavioral relevance of the initial affective resonance. Surprisingly, however, only the early response correlated with subjective reports of the amount of perceived pain in the other person as well as the unpleasantness experienced by the observer.

Recent results also support the assumption that the contextual appraisal of a situation rather than its sensory input alone determines the empathizer’s neural and behavioral response. In one fMRI study, participants viewed pictures of needle injections and of tissue biopsies performed on human hands—with the latter consisting of the insertion of a biopsy needle into an anesthetized hand (Lamm et al. 2007b). Hence, the actual painful consequences of two basically identical visual stimuli had to be taken into account in order to show appropriate vicarious responses. The neural structures supporting this appraisal process included the medial dorsal and orbitofrontal cortex (OFC) and the right temporoparietal junction (rTPJ). Involvement of the OFC was associated with reevaluating the valence of the seemingly aversive, yet actually neutral biopsy stimulus. Activation of the rTPJ and dorsal medial prefrontal cortex, however, was linked to self/other distinction and self-awareness. These capacities enabled the observers to distinguish between a probably automatically triggered, self-centered response to the aversive stimulus and the knowledge that such a response was actually not appropriate given the contextual information about the biopsy’s affective consequences. Notably, a recent meta-analysis of neuroimaging studies of attention, theory of mind, the experience of agency, and empathy demonstrated
that all these phenomena resulted in largely overlapping activations in the rTPJ (Decety & Lamm 2007). This finding also raises awareness that, in many cases, so-called high-level phenomena can be explained more parsimoniously by underlying low-level processes—such as the orienting of attention to external cues, or, as recently proposed, the switching between externally and internally oriented modes of information processing (Corbetta et al. 2008).

Perspective taking is another prominent example of how top-down processes shape empathic responses and social understanding (Avenanti et al. 2006; Jackson et al. 2006a; Lamm et al. 2007a, 2008; Ruby & Decety 2004; Stotland 1969). There is long-standing interest in how placing of oneself into another’s shoes affects one’s vicarious response to the other, and whether this differentially affects altruistic and prosocial behavior (Batson et al. 1997; Batson et al. 1987; Underwood & Moore 1982). Social psychology studies suggest that adopting a so-called imagine-other perspective (focusing on the feelings and thoughts of the other) promotes empathic concern and an other-oriented (altruistic) motivation while explicitly imagining oneself (“imagine-self”) to be in the target’s distressful situation results in heightened personal distress and an egoistic motivation to reduce that distress (via withdrawal or an aversive response).

Recent fMRI and TMS studies investigated the neural correlates of this phenomenon (Avenanti et al. 2006; Jackson et al. 2005; Lamm et al. 2007a). In the study by Lamm et al. (2007a), participants watched videos of patients undergoing painful auditory treatment either by imagining that they were in the patient’s place (imagine-self) or by focusing on the patient’s feelings and affective expressions (imagine-other). The results indicated higher personal distress and less empathic concern during the imagine-self perspective, which was associated with stronger hemodynamic responses in brain regions coding the motivational—affective dimensions of pain, such as the bilateral medial and anterior insula, the amygdala, as well as various structures involved in action control. Imagining oneself in a potentially harmful situation might therefore trigger a stronger aversive response than imagining someone else in the same situation. In addition, it might engage systems that help the observer to evaluate the affective meaning of similar situations, rendering these representations a basis for future projections into another person’s situation. Avenanti et al. (2006) used similar perspective-taking manipulations to assess their effects on motor inhibition. In line with fMRI findings, TMS results suggest that perspective-taking effects do not operate on the level of primary sensorimotor representations. Together with the absence of activation differences in visual areas in the fMRI study by Lamm et al. (2007a), this suggests that information about the other person “enters” the neural system in a similar way for both perspectives but is then processed and appraised differently (but see also Avenanti et al. 2008).

As anyone knows from personal experience, our attitudes toward others vary greatly and therefore can affect empathic responses to them. Lanzetta’s group provided the first investigations of the psychophysiological correlates of how our relationship to someone else affects our response to their emotions (Englis et al. 1982; Lanzetta & Englis 1989). Facial EMG and other measures of autonomic nervous system activity demonstrated that being in a competitive gaming relationship results in counterempathetic responses—in other words, joyful affect when the competitor loses and negative affect when he or she wins. In a similar vein, a recent fMRI study (Singer et al. 2006) assessed the modulation of empathic brain responses to another person’s pain as a function of the perceived fairness of the other person and the gender of the observer. As in previous empathy studies, empathy-related activation in ACC and AI was observed for both genders when a fair, liked player was in pain. However, men, but not women, showed an absence of such empathy-related activity when seeing an
unfair and disliked player in pain. Instead, men showed increased activation in areas associated with reward (nucleus accumbens), with activation in these areas correlating positively with their desire for revenge (as assessed by self-report after MRI scanning). These results suggest that, at least in men, a desire for revenge won over empathic motivation when they were confronted with someone experiencing pain who they believed deserved to be punished.

These results also point to gender as another important variable in the investigation of empathetic responses. In fact, there has been a longstanding debate about whether women actually possess more empathy, as expressed by higher scores in various self-report measures, or whether the different questionnaire scores can be explained by social desirability and demand effects (Davis 1994; Eisenberg & Lennon 1983). While social neuroscience should be able to make a valid contribution to this question soon (see Han et al. 2008 for preliminary evidence), the same amount of caution as in behavioral studies has to be applied with respect to experimental demand and social desirability effects. It will be a major challenge to disentangle neural responses that can be attributed to internal (“true”) or external (“socially desirable,” demand characteristics) empathetic motivations. In addition, gender-specific physiological differences such as head size, skull thickness, but also individual differences in hormonal state at the time of investigation have to be considered as potential confounds.

To summarize, there is considerable evidence that empathy is substantially modulated by top-down processes such as attention or the contextual appraisal of a situation. Notably, there seem to be at least two different ways in which top-down processes can affect the empathic response. One way is to either inhibit or amplify representations that have been activated via sensory channels and mechanisms associated with perception-action coupling. The second way is to generate empathic responses by means of imagination or anticipation of the other’s state in the absence of any bottom-up stimulation.

As for the former, based on available evidence, we speculate that cortical areas associated with executive function, contextual appraisal, and attention play a major role. This would include cortical structures such as the dorsolateral prefrontal cortex, medial and anterior cingulate cortex, right ventral premotor cortex, inferior and superior parietal cortex, and orbitofrontal cortex (e.g., Aron et al. 2004; Corbetta et al. 2008; Duncan & Owen 2000; Ridderinkhof et al. 2004; Rolls 2004). These areas interact with structures encoding the bottom-up driven affective responses, such as the anterior insula, amygdala, and possibly also parts of the ventral striatum. The latter way in which top-down processes affect empathy through mental imagery and top-down generation of feelings becomes particularly important when minimal sensory information about the other is available, requiring the use of context information, affective memory, and self-to-other projection to infer the affective condition of the other person. Depending on the situation and the information available to the empathizer, these processes might be implemented in distributed networks that are recruited during the imagery of action, sensation, and affect. In line with this reasoning, a recent fMRI study showed that the first-person imagery of pain activates structures also involved in empathy for pain as well as in the direct experience of pain (Ogino et al. 2007). Similar evidence is reported by Jabbi et al. (2008), whose analyses demonstrate that it is not only the activation of a brain region, but also its interaction with other brain areas which should be taken into account when one interprets functional neuroimaging data (see also Jabbi et al. 2008; Zaki et al. 2007). The existence of different pathways for top-down control is an important factor that should be taken into account when assessing individual differences in empathy as well as their link to prosocial behavior. For example, the responses of more “sensory-driven” persons might differ from those with better imaginative capacities.
These different processing styles might be assessed using emotional contagion questionnaires or by assessing personality traits such as extraversion/introversion.

**Future Directions**

The emergence of the field of social neuroscience has enabled substantial insight into the neural underpinnings of empathy in the past few years. Nevertheless, we are just starting to understand the neuronal and behavioral foundations of this complex psychological phenomenon. In this final section we will outline those domains and questions we believe should be the focus of future research efforts.

Even though empathy has been studied extensively by developmental psychologists in the past few decades (Eisenberg & Strayer 1987; Hoffman 2000; Zahn-Waxler et al. 1992), our knowledge about the role of the developing brain in the domains of emotions and empathy is very limited. Little information exists about how the above-mentioned bottom-up and top-down processes contribute to the development of empathy and related concepts such as emotional contagion, sympathy, and compassion, how these functions change across the life span, and to what extent these changes depend on the maturation and function of the prefrontal cortex in early childhood, adolescence, and later in life (but see, e.g., Blakemore 2008).

Social neuroscientists should also invest further effort in disentangling the mechanisms underlying individual differences in empathy. Are these differences attributable to variations in sensory-mediated emotion sharing and contagion, or do they result from differences in cognitive control and other top-down influences? How are they related to genetic variation, and what role do endocrine differences play in both cross-gender and within-gender comparisons? In a similar vein, it would be particularly interesting to learn how stable personality traits affect empathic responses. For example, despite obvious theoretical connections between these phenomena, we do not know how differences in attachment style or extraversion/introversion affect the neural correlates of empathy and prosocial behavior. Finally, the question of whether the phenomenon of empathy itself represents a trait rather than a situation-specific state variable is of great interest.

Further, empathy researchers need to develop new methods and approaches to assess the functional significance of shared activations. Are these representations valid indicators of simulation processes and what exactly is simulated by the observer? How domain-specific or domain-general are the observed empathic brain responses in anterior insula with respect to coding the qualities of pain, disgust, taste, or general aversion, respectively? Does the shared representations account hold for all emotions, including positive ones? Initial evidence suggests that it does (see the studies of pleasant tastes discussed above; see also Jabbi et al. 2007; van der Gaag et al. 2007, for examples in the realm of positive emotions), but future studies should investigate the sharing of positive emotions such as joy, pride, or elation—which are less directly related to primary sensory triggers and less well understood.

In addition, a great deal of the currently available findings could be largely explained by emotional contagion and self-centered aversive responses of the observer, such as personal distress or a generalized withdrawal response triggered by the sight of an aversive object. How do we separate these self-oriented from other-oriented responses? Are the observed activations in anterior insula the result of shared emotions, of personal distress, or rather of sympathy and compassion?

Another important question is how empathic brain responses and individual differences in empathy are linked to prosocial behavior. What is the role of personal distress versus sympathy in predicting helping and other forms of prosocial behavior? For example, is the absence of prosocial behavior due to deficits in affective sharing, insufficient regulation of high personal distress, or a combination of the two? An answer to this question requires a valid
distinction between these two processes and the development of ecologically valid measures of prosocial behavior.

Finally, and despite the long tradition of plasticity research in the field of neuroscience, very little is known about the malleability of the empathic and emotional brain. Can we train people to become more empathic, and which processing level (bottom-up or top-down) should be targeted in order for such a training to be most effective and persistent? For example, is it more effective to increase sensory awareness and low-level affective sharing, or should we develop emotion regulation strategies or positive compassionate feelings to reduce personal distress and withdrawal behavior? How does this interact with individual differences in trait empathy and related concepts, and at what age(s) should training take place? Apparently, the investigation of the latter questions requires large-scale longitudinal studies which could have enormous implications for education and society as a whole.

Philosophers and psychologists have long wondered about “this thing called empathy.” Only recently social neuroscience has begun to provide some support in this endeavor. Initial findings are encouraging that we will some day have a better understanding of why, when, and how we experience empathy and whether we can use that knowledge to increase prosocial behavior and an intersubjectivity that is grounded in a better understanding of ourselves and of others.

Acknowledgments

The research leading to these results has received funding from the University of Zurich (Research Priority Program on the Foundations of Human Social Behavior) and the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement no. 205557 [EMPATHICBRAIN]. Both authors contributed equally to this chapter.

Conflicts of Interest

The authors declare no conflicts of interest.

References


Singer, T., Seymour, B., O’Doherty, J., et al. (2004). Empathy for pain involves the affective but not the


