

c00016

Understanding Others: Brain Mechanisms of Theory of Mind and Empathy

Tania Singer

OUTLINE

s0120	Introduction	249			
s0020	<i>Social Neuroscience</i>	249		<i>The Role of Interoceptive Cortex in Feeling and Empathy</i>	257
s0030	<i>Social Neuroscience and its Relation to Neuroeconomics and Decision-making</i>	250		<i>Understanding Others in Psychopathology: Psychopathology, Autism, and Alexithymia</i>	258
s0040	Defining Concepts	252		<i>When do we Care About Others? Modulatory Factors of Empathy</i>	259
s0050	The Study of “Theory of Mind”	253			
s0060	<i>The History of Theory-of-mind Research</i>	253		Open Questions and Implications for Future Research in Neuroeconomics	260
s0070	<i>The Neural Foundation of Theory of Mind</i>	253		<i>The Automaticity Assumption and its Relationship to Social Preferences</i>	261
s0080	The Study of Empathy and Feelings	254		<i>The Link between Empathy, Fairness, “Theory of Mind,” and Prosocial Behavior</i>	261
	<i>The Mirror Neuron System and its Role for Action Understanding</i>	254		<i>References</i>	263
s0100	<i>Empathy: A Shared Network Hypothesis</i>	255			
s0110	<i>Individual Differences in Empathy</i>	256			

s0010

INTRODUCTION

s0020

Social Neuroscience

p0010

In the past few years, the field of neuroscience has shown increased interest in the study of the affective and social brain. A new interdisciplinary field called *social neuroscience* has emerged from a union between classical cognitive neuroscience and social psychology. In recent decades, classical cognitive neuroscience has considerably advanced our understanding of how the

brain processes shape, color, smells, and motion; recognizes objects; discriminates sounds; grasps actions; and enables us to perform higher-order cognitive operations like short- and long-term memory tasks, speech generation and recognition, and the executive functions involved in planning, multi-tasking, and self-monitoring. This approach, however, was governed by the implicit assumption that understanding a single brain is sufficient for understanding the behavior of all humans. Clearly, such an approach does not take into consideration the fact that humans

are inherently social. Brains do not exist in isolation, and their basic functioning reflects their participation in the social culture into which they were born. It is likely that the differences between our brains and those of apes and monkeys are associated with our outstanding skills in social cognition, such as our ability to represent and understand the abstract beliefs and desires of others.

p0020 In general, social neuroscience seeks to understand phenomena in terms of the complex interactions between social factors and their influence on behavior, the cognitive processes underlying behavior, and finally the neural and hormonal mechanisms subserving cognitive processes (see also Ochsner and Lieberman, 2001). A multi-level and multidisciplinary approach such as this also requires the use of a multi-method research strategy, including methods as varied as behavioral measures (e.g., questionnaires, reaction times), neuroscientific imaging techniques (e.g., functional magnetic resonance imaging, fMRI, electroencephalogram, EEG; transcranial magnetic stimulation, TMS), and autonomic measures (e.g., heart rate, galvanic skin conductance).

p0030 At the beginning, social neuroscience focused predominantly on the investigation of basic social abilities (for a review and overview papers, see Adolphs, 1999, 2003; Ochsner and Lieberman, 2001; Blakemore *et al.*, 2004). Several functional imaging studies, for example, have investigated the neural correlates of attending, recognizing, and remembering socially relevant stimuli, such as the facial expressions of fear, attractiveness, trustworthiness, and racial identity, and the faces of fair and unfair players (Morris *et al.*, 1996; Hart *et al.*, 2000; Winston *et al.*, 2002; O'Doherty *et al.*, 2003; Singer *et al.*, 2004).

p0040 Another important line of research focuses on our ability to understand other people's minds – that is, their beliefs, intentions, and feelings. This line of research will be discussed in detail in the following sections of this chapter.

p0050 Yet another stream of research in social neuroscience has started to investigate moral and social reasoning in various ways. Moral reasoning is studied using *moral dilemma tasks*, which involve situations in which all possible solutions to a given problem are associated with undesirable outcomes (Greene *et al.*, 2001, 2004; Moll *et al.*, 2002a, 2002b, 2006; Greene, 2007; Moll and de Oliveira-Souza, 2007). *Social dilemma tasks* are closely related to but still distinct from moral dilemma tasks. Typically, social dilemma tasks involve strategies that differ with respect to the social desirability of their outcomes (e.g., self-serving vs cooperative). Social neuroscientists have used social dilemma tasks such as the simultaneous and sequential prisoners' dilemma game and the ultimatum game, which

were developed in the framework of game theory, to investigate the neural underpinnings of social exchange and mutual cooperation. In studies employing these tasks, people play games for monetary payoffs using different playing strategies, some selfish and some cooperative, which allows for the investigation of social reasoning (figuring out what the other player will do; see, for example, Gallagher *et al.*, 2002; McCabe *et al.*, 2001; Rilling *et al.*, 2004), social emotions (emotional responses to fair and unfair play), and the interaction thereof (e.g., Montague *et al.*, 2002; Rilling *et al.*, 2002, 2007; Sanfey *et al.*, 2003; Singer *et al.*, 2004; King-Casas *et al.*, 2005). This is the stream of research that overlaps the most with or has even led to the emergence of the new field of neuroeconomics.

Social Neuroscience and its Relation to Neuroeconomics and Decision-making

s0030

Even though the fields of social neuroscience and neuroeconomics are still perceived as two distinct fields, the topics they are concerned with overlap substantially, both in content and in methodology. Thus, researchers in both fields are interested in understanding the nature of human social interaction and human decision-making and aim to determine the neural mechanisms underlying these complex social skills. Economic decision-making, for example, frequently takes place in the context of social interactions. Game theory, developed in economics, has come to provide a very effective quantitative framework for studying how different pieces of information, incentives, and social knowledge influence strategies optimal for social interaction. In game theoretical paradigms, people typically engage in economic exchange tasks in the laboratory. One prominent example of a game that has also frequently been used in neuroscientific investigations is the ultimatum game (Sanfey *et al.*, 2003; Knoch *et al.*, 2006). In this game, Mover One (M1) is given a certain amount of money and can then decide how much she wants to share with Mover Two (M2). M2 looks at the offer and can then decide whether she wants to accept or reject it. If the offer is rejected, no one receives any money. Such a move can be conceived of as a way to punish M1. However, if M2 is purely interested in money, she should accept any possible offer from M1, irrespective of whether this offer is deemed fair or unfair. Another such game is the dictator game, in which M2 is neither allowed to reject nor accept the offers made by M1, but just passively receives whatever is offered.

p0060

But why does understanding how we understand others' minds matter in economic exchange? Let's

p0070

go back to the above examples. When we compare the offers M1s typically make in ultimatum games to those made in dictator games, we find that M1s typically offer less in dictator games. The two games differ in that, in the ultimatum game, M2 is an active player who can influence the profits of both players. Thus, M1 has to construct a “theory of mind” of M2: What will M2 give me? Does M2 value fairness and, if so, which offer does M2 believe to be fair? How will M2 react if I give him X amount of money? More generally, the study of economic decision-making in the context of game theory is based on the assumption that people can predict other people’s actions when they understand their motivations, preferences, and beliefs (for a similar argument, see also Singer and Fehr, 2005; McCabe and Singer, 2008). However, economists still know little about the mechanisms that enable people to put themselves into other people’s shoes and how these mechanisms interact with decision-making in an economic context.

p0080

Social neuroscientists and neuroeconomists have recently started to clarify the neural mechanisms underlying our capacity to represent others’ intentions, beliefs, and desires (referred to as *cognitive perspective-taking*, *theory of mind*, *mind-reading*, or *mentalizing*) and to share others’ feelings (referred to as *empathy*). Even though our abilities to mentalize and to empathize are mostly used in concert when we try to understand other people’s intentions, beliefs, desires, and feelings, preliminary evidence from studies of populations of patients with marked social deficits, like those with autism or psychopathy, suggest that mentalizing and empathizing are actually two distinct abilities that rely on distinct neural circuitries (see also Figure 17.1; Blair, 2005; Singer, 2006). For example, patients with autistic spectrum disorders often have deficits in cognitive perspective-taking, while psychopaths are very good at understanding other people’s intentions and consequently at manipulating other people’s behavior. In contrast, psychopaths lack empathy, which may be the reason for their antisocial behavior.

p0090

Such a dissociation points to an important difference between our abilities to mentalize and to empathize. Whereas both abilities play an important role in drawing inferences about other people’s cognitive and emotional states, it has been suggested that empathy not only has an epistemological but also a motivational and social role (for a similar argument, see de Vignemont and Singer, 2006). Thus, empathy has very often been related to morality, altruism, justice, prosocial behavior, and cooperation (Batson and Shaw, 1991; Hoffman, 2000; Eisenberg and Morris, 2001). In the eighteenth century, the famous economist and

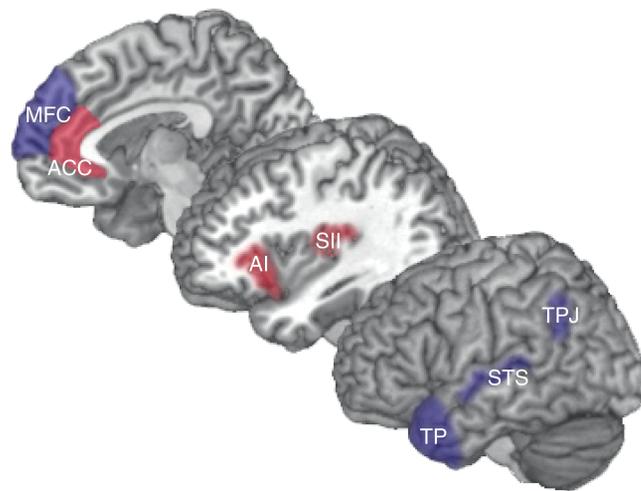


FIGURE 17.1 Brain networks involved in understanding others. Schematic representation of the brain areas typically involved in theory of mind (blue) and empathy (red) tasks. MPC, medial prefrontal cortex; ACC, anterior cingulate cortex; AI, anterior insula; SII, secondary somatosensory cortex; TP, temporal poles; STS, superior temporal sulcus; TPJ, temporo-parietal junction.

f0010

philosopher Adam Smith (1759/2004: 1) alluded to the altruistic potential of human nature when he wrote about compassion or what he termed “fellow feeling” in *The Theory of Moral Sentiments*:

How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it. Of this kind is pity or compassion, the emotion which we feel for the misery of others, when we either see it, or are made to conceive it in a very lively manner. That we often derive sorrow from the sorrow of others, is a matter of fact too obvious to require any instances to prove it; for this sentiment, like all the other original passions of human nature, is by no means confined to the virtuous and humane, though they perhaps may feel it with the most exquisite sensibility. The greatest ruffian, the most hardened violator of the laws of society, is not altogether without it.

Accordingly, empathy is also likely to render people less selfish because it enables them to share others’ emotions and feelings, which can help to motivate other-regarding behavior – in other words, behavior beneficial to another person and not only to oneself (e.g., helping someone).

p0100

Some behavioral evidence indeed suggests that people help others more when they report having empathized with them (Eisenberg and Morris, 2001). Empathy may also explain why people give charitable donations, an area of research which has also become the focus of neuroeconomic investigations (Chapter 20 of this volume; Moll *et al.*, 2006).

p0110

In the following, major findings concerning the neural mechanisms underlying our ability to

p0120

understand the beliefs, intentions, motives, and feelings of other people will be summarized and discussed in light of possible implications for social and economic decision-making.

s0040

DEFINING CONCEPTS

p0130

Clearly, the ability to understand other people's thinking and feeling is a fundamental component of our "social intelligence" and is needed for successful everyday social interaction. The literature very often refers to this as our capacity for *human empathy*. Even though, in lay terms, "empathy" usually refers to a unitary concept, a survey of the literature shows that empathy is a complex phenomenon composed of a variety of subskills and systems. It would go beyond the scope of this chapter to give a full account of existing definitions of empathy (for other relevant overviews from the fields of social neuroscience and psychology, see Wispe, 1986; Batson, 1987, 2008; Batson *et al.*, 1987; Eisenberg and Fabes, 1990; Hoffman, 2000; Preston and de Waal, 2002; Decety and Jackson, 2004; Keysers and Gazzola, 2006, 2007; Decety and Lamm, 2007). In this chapter, a neuroscientific perspective is taken, according to which three main systems rely on partially separable neural circuitries that all subserve our capacity to understand other people: (1) our ability to understand other people's motor intentions and action goals; (2) our ability to understand other people's beliefs and thoughts, which has been referred to as *theory of mind* (e.g., Premack and Woodruff, 1978), *mentalizing* (Frith and Frith, 2003), *mind-reading* (Baron-Cohen, 1995), or *cognitive perspective-taking*; and (3) our ability to understand other people's feelings, which is referred to as *empathy* or *emotional perspective-taking* (for a similar distinction, see Preston and de Waal, 2002; Gallese, 2003; Blair, 2005; de Vignemont and Singer, 2006; Singer, 2006; Decety and Lamm, 2007; Keysers and Gazzola, 2007).

p0140

De Vignemont and Singer (2006) define empathy as follows: we "empathize" with others when we have (a) an affective state, (b) which is isomorphic to another person's affective state, (c) which was elicited by observing or imagining another person's affective state, and (d) when we know that the other person's affective state is the source of our own affective state.

p0150

The first statement is important because it differentiates empathy from theory of mind, cognitive perspective taking, and mentalizing. The term "mentalizing" connotes a person's ability to cognitively represent the mental states of others, including their affective states, without becoming emotionally involved. The

term "empathizing" connotes the capacity to share other people's feelings. Accordingly, when one "empathizes" with another person who is in pain, one feels the other person's pain in one's own body. In contrast, when one understands someone else's thoughts, one does not feel the thought of the other in one's own body. There are no qualia attached to it. This difference may become clearer when we consider the characteristics of psychopaths: Whereas psychopaths do not have an impaired ability to understand other people's wishes, beliefs, intentions, and desires, it appears that they do lack the embodied feeling of empathy, a feeling which enables those without psychopathy to anticipate others' suffering, thereby often preventing them from harming others. Thus, although psychopaths do possess the ability to mentalize, they are not able to empathize (for a similar argument, see Blair, 2005).

The second statement is important for the differentiation between empathy and *sympathy* or *compassion*. In all three cases, we feel vicariously for another person. However, when we empathize, we share another person's feelings; when we sympathize or show compassion we do not necessarily share the same feeling with the other person. For example, when I empathize with a person who feels sad, I feel sad myself, whereas when I sympathize or feel compassion for a person who feels sad, I feel either pity or love for the person but not sad. Also, when I notice that someone is jealous of me, I can sympathize with or show compassion towards that person, but I do not feel jealous myself. Further, when one empathizes with another person, there does not have to be a prosocial motivation attached to it – that is, a wish to maximize the other person's happiness or alleviate the other person's distress; when we sympathize or show compassion for another person, there is. For example, a torturer may use empathy in order to sense how to increase his victim's suffering, thus showing no compassion for his victim. Experiencing too much empathy can also lead to selfish instead of other-regarding behavior. For example, sharing too much of the other person's distress can lead to a withdrawal from – instead of helping – the suffering person. In general, however, empathy is conceived to be a first necessary step in a chain that begins with affect sharing, a subsequent understanding of another person's feelings, which then motivates other-related concern and finally engagement in helping behavior. Empathy and prosocial decision-making are thus closely linked.

p0160

Finally, the last statement of the above definition of empathy is important for the differentiation between empathy and *emotional contagion*, whereby the latter connotes a reaction in which one shares an emotion with another person without realizing that the other

p0170

person's emotion was the trigger. For example, babies start crying when they hear other babies crying, long before they develop a sense of a self separate from others. Recently, Neil Harrison and colleagues (2006) found initial evidence for pupillary contagion. When subjects were presented with photos of sad faces with different pupil sizes, their own pupil size mirrored that shown in the photos. Here, emotional contagion engaged the Edinger-Westphal nucleus in the brainstem which controls pupil size. Phenomena such as pupillary contagion occur involuntarily, and may represent a precursor of empathy. However, they are not considered "empathic responses," because the subjects are not aware that they are vicariously feeling for another person.

task, which is also used frequently in the field of ToM research, is the *Sally-Anne task*, in which Sally puts a ball in a basket and then Anne takes the ball out of the basket while Sally is out of the room. A series of studies using either of these tasks showed that children age four and older start to correctly attribute false beliefs to others and give verbal explanations when asked. At age five, over 90% of children understand this task, and at age six all of them do so (Baron-Cohen *et al.*, 1985; Perner *et al.*, 1987; for a review, see Frith and Frith, 2003). When the task is simplified with a little game and does not use verbal report as a dependent measure, even children as young as three years of age seem to have an understanding of false beliefs (Clements and Perner, 1994). Research in the domain of autistic spectrum disorders suggests that the ability to mentalize is severely delayed in autism. The lack of a theory of mind in most autistic children could explain their observed failure in communication and social interaction (for a review, see Frith, 2001).

s0050

THE STUDY OF "THEORY OF MIND"

s0060

The History of Theory-of-mind Research

p0180

In 1978, Premack and Woodruff published a seminal paper in which they coined the term "theory of mind" (ToM) while discussing whether chimpanzees are capable of representing other primate's minds in terms of their desires, intentions, and beliefs. Despite extensive research conducted on this question after their paper was published, the debate about whether the capacity to have a theory of mind is uniquely human still has not been settled (Call, 2007). Overall, the literature appears to suggest that this ability is absent in monkeys, and only exists in a limited form in apes (Povinelli and Bering, 2002).

p0190

Around the same time, developmental psychologists also showed great interest in the study of the developmental time-course of our capacity to mentalize (for a review, see Frith and Frith, 2003). On the basis of a proposition by the philosopher Daniel Dennett (1978), who suggested that the most stringent test for the presence of ToM would be to see whether someone is able to predict someone else's actions on the basis of that person's false belief, Wimmer and Perner (1983) developed the false-belief paradigm to test children's mentalizing abilities. In the false-belief task, the following story is told: Maxi has some chocolate and puts it into a blue cupboard. Maxi leaves the room. Then his mother comes in and moves the chocolate to a green cupboard. Maxi comes back to get his chocolate. Where will Maxi look for the chocolate? A child who states that Maxi will look in the blue cupboard knows that he falsely believes the chocolate to be there. Control questions are posed to test whether the child understood the sequence of events: Where is the chocolate really? Do you remember where Maxi put the chocolate in the beginning? Another

The Neural Foundation of Theory of Mind

s0070

With the development of modern imaging techniques, the study of our capacity to reason about other people's minds has become the focus of cognitive neuroscience research. Imaging studies performed with healthy adults have used different paradigms to investigate which neural structures underlie our capacity to reason about other people's non-observable internal states. In these studies stories are typically told, on the basis of texts, abstract moving shapes, or cartoons, to subjects in the scanner, who are asked to understand the intentions, beliefs, and desires of the protagonist in the respective stories (for a review, see Gallagher and Frith, 2003). Theory-of-mind studies have consistently shown the involvement of a network comprising the posterior superior temporal sulcus (STS) extending into the temporoparietal junctions (TPJ), the medial prefrontal cortex (mPFC), and sometimes also the temporal poles (TP). A schematic representation of the mentalizing brain network is illustrated in Figure 17.1 in blue.

p0200

Frith and Frith (1999) suggested that the mPFC may represent mental states decoupled from reality, while the STS helps process a causal relationship between visual motion/action and another person's intended goals, and the temporal poles draw on encodings of past experience to "simulate" another person's experience. Recently, Rebecca Saxe suggested that different subcomponents of ToM have different developmental time-courses and rely on different brain regions. In line with earlier approaches in developmental

p0210

psychology and philosophy, she proposed that the ability to understand mental state concepts like desires, goals, and feelings develops earlier than the ability to represent the more abstract contents of mental states, such as beliefs, and that the former relies on functions of the mPFC, whereas the latter is specifically associated with TPJ functions (Saxe and Wexler, 2005; Saxe and Powell, 2006).

p0220 Game theoretical paradigms have also been used to investigate mentalizing (Gallagher *et al.*, 2002; McCabe *et al.*, 2001; Rilling *et al.*, 2004). Subjects are scanned while playing strategy games against someone sitting outside the scanning room. For example, Gallagher *et al.* (2002) and McCabe *et al.* (2001) compared the brain areas involved when subjects played against another person with those involved when subjects played against a computer. These studies have repeatedly demonstrated medial prefrontal lobe involvement.

p0230 The mPFC is not only involved when people mentalize about other people's thoughts, intentions, and beliefs, but also when people are reflecting on their own states (Mitchell *et al.*, 2005a). Jason Mitchell (Mitchell *et al.*, 2002, 2005b, 2006) recently conducted a series of interesting studies on mentalizing that suggest that there are functional differences between judging the mental states of similar and dissimilar others. A more ventral part of the mPFC was recruited when participants made self-judgments or judgments about people whom they perceived as being similar to themselves with respect to appearance or political attitudes. By contrast, a more dorsal part of the mPFC showed enhanced activation – close to the activation found in the mentalizing studies cited above – when subjects judged the mental states of people perceived as being dissimilar to themselves. This suggests that we may use two different strategies when inferring other people's mental states: With one strategy we simulate the other person on the basis of knowledge we have about ourselves; with the other strategy we infer the mental states of the other person on the basis of more abstract knowledge we have acquired about the world. The latter strategy may also involve knowledge about stereotypes, and raises the interesting question as to whether judging another person's mental state may be biased in different ways depending on whether we perceive them as similar or dissimilar to ourselves. *Egocentric bias*, the propensity to understand other people's states in terms of our own, may easily occur if we simulate others on the basis of ourselves while ignoring possible differences between ourselves and others. In addition, misattributions may occur when we judge other people's mental states on the basis of stereotyped or categorical knowledge that underestimates the similarity between the other person and ourselves. An

interesting twist for future research in this domain would be to explore whether brain data can predict what decisions are made during social exchanges based on whether ventral or dorsal mPFC is activated, which would reflect an unconscious perception of people being similar or dissimilar to ourselves.

THE STUDY OF EMPATHY AND FEELINGS

s0080

The Mirror Neuron System and its Role for Action Understanding

s0090

p0240 While theory-of-mind research focuses on complex inferences about abstract mental states such as another person's beliefs, another line of neuroscientific endeavor has focused on our ability to understand other people's goals and intentions by merely observing their actions. This line of research originated with the seminal discovery by a group in Parma, Italy, that neurons in the premotor cortex of macaque monkey brains were firing both when a monkey performed hand movements itself and when it merely observed another monkey or a human performing the same hand movements (Gallese *et al.*, 1996; Rizzolatti *et al.*, 1996; Ferrari *et al.*, 2003). These so-called *mirror neurons* were the first evidence for a brain mechanism which not only represents the subject's own world, but also another person's. It was suggested that mirror neurons may represent the basis for imitation. Thus, when imitating someone else's actions, we first have to transform what we see (action perception) into our own motor programs which allow us to generate a certain action sequence. The discovery of mirror neurons suggested the existence of such translation mechanisms in the primate brain.

Since the discovery of mirror neurons, several studies have demonstrated a similar common coding of the perception and generation of motor actions in the human brain using imaging techniques like PET and fMRI (for a review, see Grezes and Decety, 2001). In these studies, people were scanned while they watched movies depicting short motor actions. The observed activation was then compared to that observed when the scanned subjects performed the same motor action themselves. In line with the studies on monkeys, these studies revealed that the same circuitry was recruited when subjects merely observed another person performing an action and when they performed the same action themselves. These shared neuronal representations included brain regions in supplementary motor area (SMA), pre-SMA, premotor

p0250

cortex, the supramarginal gyrus, intraparietal sulcus, and superior parietal lobe.

p0260

Currently, researchers are debating about the exact function of the mirror neuron system and its role in social cognition. Some authors have suggested that the mirror neuron system might play a general role in understanding other people's intentions and goals by providing us with an automatic simulation of their actions (Gallese and Goldman, 1998; Fogassi *et al.*, 2005). This suggestion is in line with simulation theoretical accounts developed in philosophy in the context of theory-of-mind research (Gordon, 1986; Heal, 1986; Harris, 1989). Simulation theory holds that what lies at the root of our mature mind-reading abilities is the ability to project ourselves imaginarily into another person's perspective by *simulating* their mental activity using our own (for an overview, see Carruthers and Smith, 1996). Simulation theorists originally opposed the so-called "theory theorists" who suggest that we understand other people's minds because we have acquired abstract knowledge and a folk psychological theory about the world and about the structure and functions of our own and other people's minds (Wellman, 1990; Gopnik and Wellman, 1994). Simulation theoretical approaches have now been extended to the domain of actions and feelings. To understand what another person is doing, we simulate their movements using our own motor program; to understand what other people are feeling, we simulate their feelings using our own affective programs (see also de Vignemont and Singer, 2006; Keysers and Gazzola, 2006). Whereas these accounts see the specific role of mirror neurons in the understanding of others' motor actions and action-related intentions, but not their feelings, Iacoboni and colleagues have suggested a motor theory of empathy according to which mirror neurons have a much broader role in social cognition in that their role is not only to understand action but also to understand others' emotions and minds (Carr *et al.*, 2003; Dapretto *et al.*, 2006; Iacoboni and Dapretto, 2006). Others have criticized this broad interpretation of the role of mirror systems, believing that it overemphasizes the role of motor mirror neurons in social cognition. They maintain that mirror neurons may help us to understand other people's simple, observable action goals, but not their abstract beliefs as usually conceptualized in theory-of-mind tasks (Jacob and Jeannerod, 2005; Saxe, 2005).

expanded to include the ability to share not only motor actions but also feelings and sensations with others (Preston and de Waal, 2002; Gallese, 2003; Decety and Jackson, 2004; Decety and Lamm, 2006; de Vignemont and Singer, 2006). Thus, in addition to the ability to understand action intentions or more abstract mental states such as other people's beliefs or wishes, humans can also empathize with others – that is, share and understand feelings and emotions. Humans can feel empathy for other people in a wide variety of contexts – when others feel basic primary emotions and sensations such as anger, fear, sadness, joy, pain, and lust, as well as more culturally variable secondary emotions like embarrassment and jealousy. Inspired by earlier perception–action models (Prinz, 1990) in the domain of action understanding, Preston and de Waal (2002) proposed a neuroscientific model of empathy suggesting that observing or imagining another person in a particular *emotional* state automatically activates a representation of that state in the observer, with its associated autonomic and somatic responses. The term "automatic" in this case refers to a process that does not require conscious and effortful processing, but which can nevertheless be inhibited or controlled.

Indeed, fMRI studies in humans have provided evidence for a role of such shared neural networks that enable one to feel – by merely perceiving or imagining another person feeling pain, touch, or disgust in the absence of any stimulation to one's own body – what it feels like for the other person to be in pain, touched, or disgusted. For example, some studies have been able to demonstrate that similar neural responses in anterior insula cortex (see Figure 17.1) – a brain region involved in processing, among other sensations, disgust and taste – are elicited when subjects view pictures of disgusted faces and when they smell disgusting odors themselves (Wicker *et al.*, 2003), or when subjects view videos showing people sampling pleasant or unpleasant tastes and when they sample the different tastes themselves (Jabbi *et al.*, 2007). In contrast, another study found shared activation in secondary somatosensory cortices when subjects watched videos of people being touched and when they were being touched themselves (again, see Figure 17.1). These results are in line with the role of somatosensory cortices for the processing of touch (Keysers *et al.*, 2004).

p0280

The majority of studies on empathic brain responses have been conducted in the domain of pain (Morrison *et al.*, 2004, 2007; Singer *et al.*, 2004, 2006; Avenanti *et al.*, 2005, 2006; Botvinick *et al.*, 2005; Jackson *et al.*, 2005, 2006; Bufalari *et al.*, 2007; Cheng *et al.*, 2007; Gu and Han, 2007; Lamm *et al.*, 2007;

p0290

s0100

Empathy: A Shared Network Hypothesis

p0270

The idea that perception–action links in our brains enable us to understand others has recently been

Moriguchi *et al.*, 2007; Morrison and Downing, 2007; Saarela *et al.*, 2007). For example, in an early study, Singer and colleagues (2004) recruited couples and measured empathy *in vivo* by assessing brain activity in the female partner while painful stimulation was applied either to her own or to her partner's right hand via electrodes attached to the back of the hand. The male partner was seated next to the MRI scanner, and a mirror system allowed the female partner to see her own as well as her partner's hand lying on a tilted board in front of her. Before the experiment started, the couples were allowed to engage in social interaction to increase the feeling of being in a "real-life" situation. Differently colored flashes of light on a screen behind the board pointed to either the male or the female partner's hand, indicating which of them would receive painful and which would receive non-painful stimulation. This procedure enabled measurement of pain-related brain activation when pain was applied to the scanned subject (felt pain) or to her partner (empathy for pain). The results suggest that parts of the so-called *pain matrix* – the bilateral anterior insula (AI), rostral anterior cingulate cortex (ACC), brainstem, and cerebellum – were activated when subjects experienced pain themselves, as well as when they saw a signal indicating that a loved one had experienced pain. These areas are involved in the processing of the affective component of pain – that is, how unpleasant the subjectively felt pain is. Thus, both the experience of pain to oneself and the knowledge that a beloved partner is experiencing pain activates the same affective pain circuits, suggesting that if a beloved partner suffers pain, our brains also make us suffer from this pain. Activation in this network was also observed when subjects saw an unknown but likeable person suffering pain (Singer *et al.*, 2006), when subjects watched videos showing body parts in potentially painful situations (Jackson *et al.*, 2005, 2006), painful facial expressions (Lamm *et al.*, 2007), or hands being pricked by needles (Morrison *et al.*, 2004; for a review, see de Vignemont and Singer, 2006). The important role of AI in empathy was further corroborated by two studies focusing on the effects of expertise in compassion. Lutz *et al.* (2004) found greater AI activation in long-time meditators (Buddhist monks) as compared to untrained (control) meditators while they were engaging in meditation practices for the cultivation of compassion, which the monks had practiced for many years. Similarly, Lazar *et al.* (2005) found significant differences in cortical thickness in prefrontal cortex and right AI in experienced as compared to non-meditators. Figure 17.1 schematically summarizes the areas (marked in red) found to be typically activated in empathy tasks.

Individual Differences in Empathy

Evidence for individual differences in empathic skills has not only been found in studies of populations with a high degree of empathic skills stemming from long-time compassion meditation, but also in the abovementioned empathy studies of adults randomly selected from the normal population. As we all experience in our everyday lives, people are not equally empathic. Scientifically, individual differences in empathic capacity can be assessed using standard empathy questionnaires developed and validated by psychologists, such as the Empathic Concern Scale of the Interpersonal Reactivity Index (IRI; Davis, 1980) and the Balanced Emotional Empathy Scale (BEES; Mehrabian and Epstein, 1972). These scales measure psychological traits which are conceptualized as personality dispositions that are relatively stable over the lifespan of a given person. Analyses of empathic brain responses obtained while subjects were observing other people suffering – be it their loved ones or people the subjects liked (Singer *et al.*, 2004, 2006) – have revealed individual differences in activity in empathy-related pain-sensitive areas (ACC and AI), and that these differences co-vary with interindividual differences in IRI and BEES scores. The higher subjects scored on these questionnaires, the higher their activation in ACC and anterior insula. Interestingly, Jabbi *et al.* (2007) observed similar correlations between IRI subscales and empathic brain responses in the AI for subjects who had observed others tasting pleasant or unpleasant drinks associated with facial expressions of joy or disgust alternatively. Empathic brain responses are not only positively correlated with trait measures of empathy, but also with unpleasantness ratings which subjects give online after each trial of an empathy-inducing condition of a scanning session (Jackson *et al.*, 2005; Lamm *et al.*, 2007; Saarela *et al.*, 2007). Future research will have to clarify how these individual differences in empathic brain responses come about, and whether they are able to explain individual differences in prosocial behavior – two lines of research which have not yet been sufficiently addressed. In summary, findings of shared circuitries underlying one's own sensations and feelings and the observation of similar sensations and feelings in others suggest that we use neural representations reflecting our own emotional responses to understand how it feels for others to be in a similar state. They further suggest that our ability to empathize may have evolved from a system which represents our own internal feeling states and allows us to predict the affective outcomes of an event for ourselves and for other people (e.g., Singer *et al.*, 2004). Thus, deficits in

s0110

p0300

representing one's own feeling states should result in deficits in empathizing with others, a hypothesis that will be discussed in more detail below.

p0310 Furthermore, results from fMRI studies on empathy for taste, disgust, and pain suggest that AI cortices play a crucial role in empathy (see also Figure 17.1). Some authors also refer to insular cortex as the interoceptive cortex (e.g., Craig, 2002) because this brain region is involved in processing a variety of information about internal bodily states, including pain, taste, hunger, thirst, and arousal. In the following, a more detailed account of the possible functions of interoceptive cortex and its role for feelings in general and empathy in particular is provided.

s0120 The Role of Interoceptive Cortex in Feeling and Empathy

p0320 Whereas the beginning of affective and social neuroscience was characterized by a strong focus on exploration of the role of amygdala in emotional processing, that focus has now broadened to include another structure that plays a crucial role in processing feelings: the insular cortex and, in particular, the anterior insular (AI) cortex. It has been suggested that these regions represent a crucial part of the human interoceptive cortex (Craig, 2002) and subserve neural representations of internal bodily and feeling states (Damasio, 1994; Critchley, *et al.*, 2001, 2004).

p0330 Interoceptive models of emotions have a long tradition in psychology, and propose that cortical representations of internal bodily signals are at the origin of feeling states. In the late eighteenth century, William James and Carl Lange suggested with the now famous *James-Lange theory* that changes in bodily responses are a necessary condition for emotional experience to arise (James, 1894; Lange, 1885). Thus, we feel our hearts beating when we fall in love or experience fear; we feel our stomachs constricting when we are under stress and have to make a difficult decision; and we feel our face reddening with rage or blushing when we experience an embarrassing situation. They argued that emotions cannot be experienced in the absence of these bodily feelings.

p0340 Based on anatomical observations in non-human species, Bud Craig has elaborated on these notions and 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. In humans, this modality-specific sensory representation of the body's physiological condition in the posterior insula is initially re-represented in the anterior insula on the same side of the brain, and then, by

way of a callosal pathway, remapped to the other side of the brain in the right AI. Such a second-order re-representation in right AI is assumed to subserve subjective feelings, and was even proposed as the seat of our awareness of a physical self as a feeling entity (see also Damasio, 1994; Critchley *et al.*, 2001). At the same time, afferents also project by way of the medial dorsal thalamic nucleus to ACC to produce a motivation to engage in relevant behavior. Thus, direct activation of both the insula and the ACC may correspond to simultaneous generation of both a feeling and an affective motivation, with its attendant autonomic effects.

Indeed, imaging studies focusing on the relationship between peripheral measures of arousal and brain activity give robust evidence for the crucial role of rostral ACC and AI cortices in the representation of internal bodily states of arousal as well as the awareness of these states (Critchley *et al.*, 2001, 2003, 2004). The role of AI in interoceptive awareness was specifically highlighted by two studies conducted by Critchley's group. To study the effects of peripheral arousal feedback to the brain, they selected subjects with pure autonomic failure (PAF), which entails an inability to generate autonomic arousal due to specific peripheral denervation of the autonomic system. Using a fear-conditioning paradigm, they compared the brain responses of these subjects to those of normal controls when participants either consciously or unconsciously processed angry faces that had been paired with loud, aversive noise stimuli. The control subjects, in contrast to the PAF subjects, showed an autonomic response when exposed to the conditioned emotional stimuli; namely, enhanced activity in right AI. This suggests a sensitivity of right AI to autonomic feedback, which is absent in individuals with PAF. In addition, emotional awareness of the stimuli was manipulated using backward masking, a procedure in which two stimuli are presented in such rapid sequence that the second stimulus "masks" the first, thus disrupting the observer's opportunity to consciously evaluate the content of the first stimulus. In accordance with the theory suggesting a role for AI in the conscious experience of emotions, the researchers demonstrated, as in previous studies, an enhanced BOLD response in amygdala to unconsciously perceived threat stimuli, but only enhanced activation in AI when the conditioned face was consciously perceived (Critchley *et al.*, 2002). In a subsequent study, Critchley demonstrated that the activity and size of right AI were positively associated with the degree to which participants were aware of their own heartbeat (Critchley *et al.*, 2004). Overall, these and other findings suggest that interoceptive cortex plays an important role for the representation and awareness of feeling states arising from the body.

p0360 As the abovementioned results on empathic brain responses suggest, the very same structures (AI and ACC) which play a crucial role in representing our own feeling states also seem to be crucial in processing vicarious feelings. Based on this observation, Singer *et al.* (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 allow us to form subjective representations of feelings. These representations allow us not only to understand our 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 empathic 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, using fMRI, the anticipation of pain has been found to activate more anterior insular regions, whereas the actual experience of pain activates more posterior insular regions, which confirms 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). Similarly, in Singer *et al.*'s (2004) empathy study, activity in posterior insular cortices – contralateral to the stimulated hand – was only observed when participants were actually experiencing pain themselves, whereas activity in AI was observed when participants were experiencing pain themselves and when they were vicariously simulating it for others.

p0370 A model suggesting that the representation of one's own feeling states is necessary for empathy to arise would make two predictions. First, training the capacity to understand our own feelings would go hand in hand with training the capacity for empathy. Second, deficits in understanding one's own emotions should be associated with empathy deficits. Whereas evidence for the first hypothesis is still lacking, evidence for the second hypothesis is slowly accumulating. The next section of this chapter will focus on the role played by interoceptive cortex in pathologies associated with a lack of empathy and social cognition, such as autism and psychopathy.

s0130 **Understanding Others in Psychopathology: Psychopathy, Autism, and Alexithymia**

p0380 The study of patients with specific psychopathological deficits in social cognition is relevant not only

for the development of effective treatments for these conditions, but also for a better understanding of the mechanisms underlying social cognition in the healthy population. For example, at the beginning of this chapter, I stated that mentalizing and empathizing represent two different capacities that rely on different circuitries, and both result in a better understanding of other people's minds. Figure 17.1 illustrates the two different circuitries, marked in blue and red respectively. These two capacities usually work together, so normal individuals presumably activate both networks when confronted with tasks in which they have to draw inferences about the mental states of others. It would be difficult to prove a dissociation between these two systems. However, one could elegantly support the claim that there are two different pathways for mentalizing and empathizing by demonstrating a double dissociation in two different patient populations, such as autism and psychopathy.

Generally defined, autistic spectrum disorders (ASD) are pervasive developmental disorders characterized by abnormalities of social interaction, impairments in verbal and non-verbal communication, and a restricted repertoire of interests and activities. People with high functioning autism or Asperger syndrome differ from ASD patients mainly in that they have high intelligence and no impairments in verbal communication. People with ASD and Asperger syndrome have often been found to have difficulties in attributing mental states, such as beliefs, desires, or intentions, to others (reviewed by Frith, 2004). It is less clear, however, whether people with Asperger syndrome also lack an ability to empathize with others. Psychopathy, in contrast, is characterized as a personality disorder mainly marked by a lack of empathy, narcissism, impulsivity, selfishness, the instrumental use of others, and altered emotional sensitivity (for an overview, see Hare, 2006). However, it is not clear whether psychopaths also have deficits in their ability to mentalize. Thus, behavioral studies with psychopathic populations have found selective emotional dysfunction, such as impairment in aversive conditioning, autonomic response to threat, augmentation of the startle reflex to visual threat primes, impaired processing, and altered autonomic response to sad and fearful facial and vocal expressions (Hare, 1982; Levenston *et al.*, 2000; Blair, 2001; Flor *et al.*, 2002). Neurophysiological evidence has been reported of reduced amygdala response during emotional memory and aversive-conditioning tasks, as well as reduced gray matter in the amygdala and hippocampus (Kiehl *et al.*, 2001; Veit *et al.*, 2002; Blair, 2005). More importantly, Sterzer *et al.* (2007) investigated the structure of brains of adolescents with conduct disorder,

p0390

a psychiatric disorder which often develops into anti-social personality disorder after the age of 18 and is associated with over-aggressive behavior towards people, objects, and animals, and with frequent norm violations. Sterzer and colleagues found reduced activity in amygdala and AI in the adolescents with conduct disorder as compared to normal controls. Moreover, the empathy scores of adolescents with conduct disorder correlated negatively with gray-matter volume in bilateral AI, suggesting the important role of this structure for the correct representation of affective state and the subsequent appropriate empathic response. Birbaumer *et al.* (2005) has also reported hypoactivation in amygdala and AI in psychopathic individuals during a fear-conditioning task.

p0400

Findings such as these suggest a difficulty in emotional empathic responding without a clear indication of theory-of-mind impairment. As suggested above, these findings are indications of a possible double dissociation of empathizing and mentalizing skills in patients with autism and those with psychopathy. Patients with autism may lack mentalizing abilities, but may not be deficient in the ability to share others' feelings; in contrast, patients with psychopathy lack empathy, but are unimpaired in their understanding of other people's thoughts and beliefs. However, no fMRI study has ever demonstrated a clear dissociation between mentalizing and empathizing deficits in psychopathy and autism.

p0410

Our group has started to investigate a similar dissociation by studying interoceptive awareness and empathy in high-functioning patients with autism or Asperger syndrome (AS) and a condition named alexithymia. Whereas autistic spectrum disorders (ASD) are associated with severe deficits in social cognition and communication, alexithymia is a subclinical phenomenon involving a lack of emotional awareness or, more specifically, difficulty in identifying and describing feelings, and in distinguishing feelings from the bodily sensations of emotional arousal (Nemiah *et al.*, 1976). Alexithymia is thought to be present in 10% of the general population (Linden *et al.*, 1994; Salminen *et al.*, 1999), and was observed in 50% of high-functioning patients with autism or AS (Hill *et al.*, 2004). Even though we know that people with AS very often have difficulty attributing mental states to others, and that these deficits are associated with less activation in the mentalizing brain network described above (Frith and Frith, 2006), we know very little about their capacity to experience feelings from their own bodies or to empathize with the feelings of others. Therefore, in a recent fMRI study, we scanned subjects with AS and controls with varying degrees of alexithymia while they

performed a task that required them to experience their own feelings (Silani *et al.*, 2008). Specifically, subjects were asked to judge how they felt about emotionally arousing pictures. Results showed that the degree of severity in alexithymia, as measured by two different alexithymia scales, was correlated with less activation in anterior insula. However, a lack of activation in insular cortices during interoceptive awareness of emotions was not specific to the AS diagnosis, but was predicted entirely by the degree of alexithymia. Thus, controls with stronger alexithymic symptoms also showed less activation in interoceptive cortex. These data again indicate that AI plays a role in understanding one's own emotions. Interestingly, individual differences in the degree of alexithymia correlated negatively with individual differences in a questionnaire measuring people's empathic personality disposition, and levels of both alexithymia and empathy were predictive of brain activation in AI during interoception. These findings are perfectly in line with the prediction that deficits in understanding one's own emotions result in empathy deficits, and that both should be correlated with lesser activation in AI. Current investigations in our laboratory using a paradigm for the measurement of empathic brain responses to others' suffering and a classical mentalizing paradigm are extending these findings by scanning alexithymic and non-alexithymic subjects with and without a diagnosis of AS to show empathy deficits on the level of brain responses, as well as a double dissociation between empathy and theory of mind. Such a double dissociation would indicate that a lack of theory of mind, but not empathy, is characteristic of AS, but not alexithymia; and that a lack of interoceptive awareness and empathy is characteristic of alexithymia, but not AS without emotional deficits. This finding would be the first to disprove the implicit general notion of a global empathy deficit in autism, and would also provide clear evidence for the existence of two distinct roots to the understanding of other people's minds: a "cold" cognitive one and a "hot" empathic one.

When do we Care About Others? Modulatory Factors of Empathy s0140

In the previous section of this chapter, examples were given from patient populations with severe emotional and empathy deficits. However, as described above, there are also substantial individual differences with respect to empathy in the normal healthy population. In addition, the degree to which we have empathic feelings also varies as a function of situational factors. For example, it is

p0420

usually easier to empathize with someone who has treated one well than with someone who has treated one poorly. In the past 2 years, fMRI studies have embarked on an investigation of the modulatory factors of empathic brain responses. For example, with respect to empathy for pain, a subject's affective link to the other person (Singer *et al.*, 2004, 2006), the subject's appraisal of whether the reason the other person is suffering is justified (Lamm *et al.*, 2007), the frequency of a person's prior exposure to pain-inducing situations (Cheng *et al.*, 2007), and the intensity of the inflicted pain (seeing a needle pricking vs penetrating a muscle, Avenanti *et al.*, 2006) all seem to play a role in the modulation of the magnitude of empathic brain responses.

p0430 One example that is particularly relevant for the field of economics is a study conducted by Singer *et al.* (2006) demonstrating evidence for the modulation of empathic brain responses to another person's pain as a function of the perceived fairness of the other person. In this study, male and female volunteers first played a repeated trust game as Mover 1 with two confederates. In the game, Mover 1 can send her endowment of 10 points to Mover 2. Mover 2 can then respond to Mover 1's trust by sending 0 to 10 points back to Mover 1. Each point returned to Mover 1 is tripled. Mover 2 can thus opt for strategies with varying degrees of fairness. In Singer *et al.*'s (2006) experiment, one confederate was instructed to play fairly by reciprocating the subject's trust by returning fair amounts of points; the other was asked to play unfairly by selecting only self-interested choices and responding with no or only minimal returns. After that, an empathy-for-pain paradigm similar to the one reported by Singer *et al.* (2004) was used to measure the subject's empathic brain responses while either the subject or one of the confederates was receiving painful stimulation to her hand. To assess gender differences in empathy and its modulation, both men and women were scanned and paired with either two male or two female players. As in previous empathy studies, empathy-related activation in ACC and AI was observed for both genders when the fair, likeable player was in pain. However, men, but not women, showed an absence of such empathic activity when seeing an unfair player in pain. Instead, men showed increased activation in areas associated with reward (nucleus accumbens), which correlated positively with their desire for revenge as assessed by questionnaires after the scanning session. 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. This finding is in agreement

with results from a study conducted by de Quervain and colleagues (2004) showing similar reward-related activation when players were scanned while they were able to deliver punishment points to participants who had defected on them in previous games.

This pattern of results contributes to a microfoundation for theories of social preferences. These theories suggest that people's valuations of other players' payoffs depend on how fairly the other players have played in previous games (Fehr and Gächter, 2000): People tend to place a positive value on others' payoffs if the others have played fairly, but a negative value on others' payoffs if the others have played unfairly. This pattern of preferences implies that people prefer to cooperate with fair opponents and to punish unfair opponents. And indeed, a variety of experiments in behavioral economics were able to demonstrate that people are willing to punish players who had previously behaved unfairly in monetary exchange games, even though this behavior may seem irrational because they knew they would never see the other player again and had to pay to punish the other player. They were therefore working against optimizing their own income by spending fair amounts of money on punishing other people. Economists term this behavior *altruistic punishment*, due to its costly nature (for details, see Chapter 15 of this volume). Neuroscientific findings suggest that punishing people who have violated social norms activates reward circuitries usually engaged in processing primary rewards. This may explain why people are motivated to engage in altruistic punishment even if this behavior is costly: it is rewarding.

Further investigation of the factors that modulate empathic brain responses will be of great relevance for a better understanding of the conditions under which prosocial and other-regarding behavior, on the one hand, and revenge-driven or egoistic behavior, on the other, are more likely to occur.

OPEN QUESTIONS AND IMPLICATIONS FOR FUTURE RESEARCH IN NEUROECONOMICS

Recent efforts in social neuroscience and neuroeconomics have helped shed light on the mechanisms underlying our ability to mentalize and to empathize. Even though these two abilities have distinct features and seem to rely on different neuronal circuitries, both allow humans to represent other people's states – their intentions, beliefs, and thoughts, their emotions and sensations. In the following section, important open

questions will be raised and discussed in light of their implications for economic theory and future neuroeconomics studies.

s0160 **The Automaticity Assumption and its Relationship to Social Preferences**

p0470 One important feature of the mechanisms outlined is that representing other people's states and goals seems to happen mostly automatically and without awareness. For example, in the earlier empathy studies in the domains of pain, touch, and disgust, subjects were not even told that the goal of the study was to investigate empathic brain responses, but merely instructed passively to watch a scene or movie (Wicker *et al.*, 2003; Keysers *et al.*, 2004; Singer *et al.*, 2004, 2006). Nevertheless, even without explicit instruction, when we perceive others' feelings, the brain networks that represent our own feelings appear to be automatically activated. It seems that we automatically share other people's feelings. This observation is in line with earlier perception-action models of motor behavior and imitation, and with their extension to the domain of empathy (Preston and de Waal, 2002; Gallese, 2003). For example, Preston and de Waal (2002) proposed a neuroscientific model of empathy suggesting that observation or imagination of another person in a particular emotional state *automatically* activates a representation of that state in the observer with its associated autonomic and somatic responses. The term "automatic" in this case refers to a process that does not require conscious and effortful processing, but which can nevertheless be inhibited or controlled.

p0480 Such automatic perception-action mechanisms could have important implications for economic theory. Without thinking about or being consciously aware of doing so, we perceive others' feelings and motivational states; this activates the same brain networks that represent our own feelings and motivational states. Empathic resonance mechanisms such as these create a link between our own and other people's needs, desires, motivational states, and emotions. Since other people's emotions affect our own emotional state, and our own emotions are important determinants of our motives and actions, other people's feeling states will partly shape our own motives towards them. This implies that our motives are not only self-interested, but also automatically other-regarding due to resonance mechanisms such as these.

p0490 Figure 17.2 exemplifies how empathic resonance mechanisms may influence economic models for the prediction of actions in social exchange settings as conceived in game theoretical frameworks. Accordingly,

our actions in such games are determined by our preferences (e.g., desires, beliefs, needs) and our beliefs about the other player's preferences which, in turn, will determine the other player's actions. This chapter has reviewed evidence from social neuroscience about the mechanisms enabling us to draw inferences about other people's mental and motivational states, as well as about their actions and intentions. Empathic resonance mechanisms may establish a link between one's ability to predict others' motives and the nature of one's own motives. That is, other people's emotions may partly shape our own motives concerning them. This link is illustrated as a red arrow in Figure 17.2. To provide an example: when confronted with a person suffering from her parents' having passed away, many people will empathize with the person; this shared suffering may motivate them to try to alleviate the other person's suffering (and thereby feel better themselves), even though they will incur costs doing so and even though this other-regarding behavior may not have been produced were it not for empathic resonance mechanisms. Donations to charity may draw on such empathic mechanisms, when, for example, people are motivated to donate money to charities that operate in third-world countries after seeing a documentary about starving children.

s0170 **The Link between Empathy, Fairness, "Theory of Mind," and Prosocial Behavior**

p0500 Of course, the propensity to engage in other-regarding behavior is not determined exclusively by whether a person is equipped with such empathic resonance mechanisms, but by many other factors as well. One such factor is that emotions and motivational states are transient phenomena, so the chronological proximity of the empathy-inducing stimulus may matter.

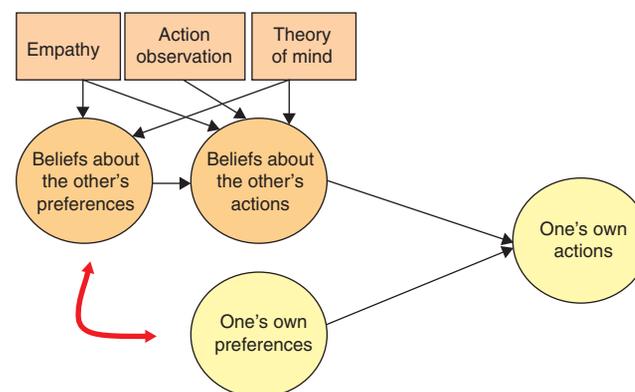


FIGURE 17.2 Economic model for the prediction of actions.

f0020

As we saw above, empathic brain responses have also been found to be modulated by factors such as the affective link or perceived fairness of another person. Thus, at least in men, empathic brain responses seem to be absent when people who were previously defecting in economic games are now in pain. These empathic brain responses were even replaced by their opposite: feelings of revenge and *Schadenfreude* – the joy of seeing another person suffering. Based on these and other observations, de Vignemont and Singer (2006) recently questioned the automaticity assumption of empathy, and proposed an alternative model in which empathy is not always merely the automatic consequence of the passive observation of emotional cues in others, but rather subject to an appraisal of the situation and modulation (de Vignemont and Singer, 2006). Accordingly, the magnitude of empathic brain responses is determined by a variety of modulating factors ranging from features associated with the emotional display in the suffering person (e.g., intensity of the emotions), the relationship between the empathizer and the other person (e.g., affective link and interdependence), to the features of the empathizer (e.g., gender, emotional experiences, and emotion regulation ability). A better understanding of the nature of these modulatory factors will also help us to understand the complex interplay of conditions determining prosocial behavior.

p0510 One prediction, however, that can easily be made is that people with a greater ability to empathize should display more other-regarding behavior. Even though the question concerning how empathy relates to prosocial behavior is crucial for our understanding of human nature and the development of methods to enhance prosocial behavior, surprisingly, this link has not yet been investigated in the context of neuroeconomics and social neuroscience research. This is partly due to the lack of a clear conceptualization of what we mean when we talk about prosocial behavior and the fact that we still lack adequate tasks for a well-controlled quantitative assessment of individual differences in prosocial behavior in a laboratory – tasks that can just as well be used in experiments performed in economics, neuroscience, and behavioral psychology. Thus, whereas economists have mainly used game theoretical paradigms such as the dictator game, the trust game, and the public good game to operationalize prosocial versus selfish behavior in economic exchange, social psychologists and developmental psychologists have mostly used ecologically valid tasks in which people are observed when confronted with situations in which they could help or refrain from helping. While the former measures allow for a more controlled quantitative assessment of

prosocial behavior (e.g., the number of monetary units sent to the other player), they suffer due to a lack of realism and demand effects in that either the task instructions or the task properties clearly indicate to the subjects how they are expected to behave. Thus, in most economic games, subjects are only given a choice between engaging in prosocial/fair or selfish/unfair behavior; this precludes the possibility of observing spontaneous, non-instructed prosocial behavior. More importantly, it is not clear whether such economic tasks assess the same prosocial behavior as the prosocial behavior measured in social psychology tasks that focus on the act of helping people in need. Prosocial behavior as assessed in game theoretical paradigms may well be based on fairness motives rather than empathic motivation. Thus, it may well be that individual differences in empathic brain responses may be a good predictor of helping behavior, but not of fairness-based cooperative behavior.

The distinction between fairness and empathic motivation and their relation to different forms of prosocial behavior has not been investigated to a sufficient extent – neither in the field of economics, nor in social sciences or neurosciences. It should be the focus of future neuroeconomics studies. Even though fairness and empathy may seem to be concepts closely related and important for the engagement in prosocial behavior, having one and not the other motivation may also lead to very different outcomes. Thus, studies in experimental economics have repeatedly provided evidence that fairness preferences lead to cooperative behavior, but also to a desire for revenge and retributivism when violated. People who have been treated unfairly in economic games are willing to punish unfair players (Fehr and Gächter, 2002; Fehr and Fischbacher, 2003), and recent neuroeconomics studies have provided evidence for reward-related activation in the brains of people who either actively (de Quervain *et al.*, 2004) or passively (Singer *et al.*, 2006) take revenge on these defectors. In contrast, purely empathic motivation may also result in cooperative behavior in similar games but to a lesser extent to revenge-related, that is, other-harming behavior, since this behavior is opposed to an empathic motivation.

Another interesting prediction that can be derived from neuroscientific investigations on empathy and cognitive perspective-taking or theory of mind is that empathic motivation is a better predictor of engagement in other-regarding behavior than cognitive perspective-taking is, because empathy engages motivational and emotional brain circuitries, while theory of mind relies on networks (mPFC, TPE, STS, temporal poles) typically believed to be less relevant for

motivation and emotions. As stated above, psychopaths may lack empathy, but not cognitive perspective-taking, explaining why they can engage in antisocial behavior while being very good at manipulating and fooling other people – an ability that requires an understanding of other people's beliefs and intentions.

p0540

Even though individual differences in empathy may be better predictors of the engagement in prosocial behavior, both capacities, emotional and cognitive perspective-taking, should help people to better predict other people's actions. An interesting question for future research is to determine the relative importance of our ability to empathize and to mentalize for the prediction of others' motives and actions in different situations, and to determine in which situations one is interfering with or even facilitating the other.

References

- Adolphs, R. (1999). Social cognition and the human brain. *Trends Cogn. Sci.* 3, 469–479.
- Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nat. Rev. Neurosci.* 4, 165–178.
- Avenanti, A., Buetti, D., Galati, G., and Aglioti, S. (2005). Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. *Nat. Neurosci.* 8, 955–960.
- Avenanti, A., Paluello, I.M., Bufalari, I., and Aglioti, S.M. (2006). Stimulus-driven modulation of motor-evoked potentials during observation of others' pain. *NeuroImage* 32, 316–324.
- Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.
- Baron-Cohen, S., Leslie, A.M., and Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition* 21, 37–46.
- Batson, C.D. (1987). Prosocial motivation: is it ever truly altruistic? In: L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, Vol. 20. New York, NY: Academic Press, pp. 65–122.
- Batson, C.D. (2008). These things called empathy. In: J. Decety and W. Ickes (eds), *The Social Neuroscience of Empathy*. Cambridge, MA: MIT Press, (in press).
- Batson, C.D., Fultz, J., and Schoenrade, P.A. (1987). Distress and empathy: two qualitatively distinct vicarious emotions with different motivational consequences. *J. Personality* 55, 19–40.
- Batson, C.D. and Shaw, L.L. (1991). Evidence for altruism: toward a pluralism of prosocial motives. *Psychological Inquiry* 2, 107–122.
- Birbaumer, N., Veit, R., Lotze, M. et al. (2005). Deficient fear conditioning in psychopathy: A functional magnetic resonance imaging study. *Arch. Gen. Psych.* 62, 799–805.
- Blair, R.J. (2001). Neurocognitive models of aggression, the antisocial personality disorders and psychopathy. *J. Neurol. Neurosurg. Psych.* 716, 727–731.
- Blair, R.J. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness Cogn.* 14, 698–718.
- Blakemore, S.J., Winston, J., and Frith, U. (2004). Social cognitive neuroscience: where are we heading? *Trends Cogn. Sci.* 8, 216–222.
- Botvinick, M., Jha, A.P., Bylsma, L.M. et al. (2005). Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. *NeuroImage* 25, 312–319.
- Bufalari, I., Aprile, T., Avenanti, A. et al. (2007). Empathy for pain and touch in the human somatosensory cortex. *Cerebral Cortex* 17, 2553–2561.
- Call, J. (2007). Past and present challenges in theory of mind research in nonhuman primates. *Prog. Brain Res.* 164, 341–353.
- Carr, L., Iacoboni, M., Dubeau, M.-C. et al. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc. Natl Acad. Sci. USA* 100, 5497–5502.
- Carruthers, P. and Smith, P. (eds) (1996). *Theories of Theories of Mind*. Cambridge: Cambridge University Press.
- Cheng, Y., Lin, C.P., Liu, H.L. et al. (2007). Expertise modulates the perception of pain in others. *Curr. Biology* 17, 1708–1713.
- Clements, W.A. and Perner, J. (1994). Implicit understanding of belief. *Cogn. Dev.* 9, 377–395.
- Craig, A.D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666.
- Critchley, H.D., Mathias, C.J., and Dolan, R.J. (2001). Neuroanatomical basis for first- and second-order representations of bodily states. *Nat. Neurosci.* 4, 207–212.
- Critchley, H.D., Mathias, C.J., and Dolan, R.J. (2002). Fear conditioning in humans: the influence of awareness and autonomic arousal on functional neuroanatomy. *Neuron* 33, 653–663.
- Critchley, H.D., Mathias, C.J., Josephs, O. et al. (2003). Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain J. Neurol.* 126, 2139–2152.
- Critchley, H.D., Wiens, S., Rotshtein, P. et al. (2004). Neural system supporting interoceptive awareness. *Nat. Neurosci.* 7, 189–195.
- Damasio, A.R. (1994). Descartes' error and the future of human life. *Scientific Am.* 271, 144.
- Dapretto, M., Davies, M.S., Pfeifer, J.H. et al. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nat. Neurosci.* 9, 28–30.
- Davis, M.H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Cat. Selected Doc. Psychol.* 10.
- Decety, J. and Jackson, P.L. (2004). The functional architecture of human empathy. *Behav. Cogn. Neurosci. Rev.* 3, 71–100.
- Decety, J. and Lamm, C. (2006). Human empathy through the lens of social neuroscience. *Scientific World J.* 6, 1146–1163.
- Decety, J. and Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist*, [doi:10.1177/1073858407304654].
- Dennett, D.C. (1978). Beliefs about beliefs. *Behav. Brain Sci.* 1, 568–570.
- de Quervain, D.J., Fischbacher, U., Treyer, V. et al. (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258.
- de Vignemont, F. and Singer, T. (2006). The empathic brain: how, when and why? *Trends Cogn. Sci.* 10, 435–441.
- Eisenberg, N. and Fabes, R.A. (1990). Empathy: conceptualization, measurement, and relation to prosocial behavior. *Motiv. Emotion* 14, 131–149.
- Eisenberg, N. and Morris, A.S. (2001). The origins and social significance of empathy-related responding. A review of empathy and moral development: Implications for caring and justice by M.L. Hoffman. *Social Justice Res.* 14, 95–120.
- Fehr, E. and Fischbacher, U. (2003). The nature of human altruism. *Nature* 425, 785–791.
- Fehr, E. and Gächter, S. (2000). Fairness and retaliation: the economics of reciprocity. *J. Econ. Persp.* 14, 159–181.
- Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature* 415, 137–140.
- Ferrari, P.F., Gallese, V., Rizzolatti, G., and Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and

- communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.* 17, 1703–1714.
- Flor, H., Birbaumer, N., Hermann, C. *et al.* (2002). Aversive Pavlovian conditioning in psychopaths: peripheral and central correlates. *Psychophysiology* 39, 505–518.
- Fogassi, L., Ferrari, P.F., Gesierich, B. *et al.* (2005). Parietal lobe: from action organization to intention understanding. *Science* 308, 662–667.
- Frith, C.D. (2004). Emanuel Miller Lecture: Confusions and controversies about Asperger syndrome. *J. Child Psychol. Psych.* 45, 672–686.
- Frith, C.D. and Frith, U. (1999). Interacting minds: a biological basis. *Science* 286, 1692–1695.
- Frith, C.D. and Frith, U. (2006). The neural basis of mentalizing. *Neuron* 50, 531–534.
- Frith, U. (2001). Mind blindness and the brain in autism. *Neuron* 32, 969–979.
- Frith, U. and Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Phil. Trans. R. Soc. Lond. B* 358, 459–473.
- Gallagher, H.L. and Frith, C.D. (2003). Functional imaging of “theory of mind”. *Trends Cogn. Sci.* 7, 77–83.
- Gallagher, H.L., Jack, A.I., Roepstorff, A., and Frith, C.D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage* 16, 814–821.
- Gallese, V. (2003). The manifold nature of interpersonal relations: the quest for a common mechanism. *Phil. Trans. R. Soc. Lond. B* 358, 517–528.
- Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609.
- Gopnik, A. and Wellman, H. (1994). The “theory theory”. In: L. Hirschfield and S. Gelman (eds), *Domain Specificity in Culture and Cognition*. New York, NY: Cambridge University Press, pp. 257–293.
- Gordon, R. (1986). Folk psychology as simulation. *Mind Language* 1, 158–171, Reprinted in 1995 in M. Davies and T. Stone (eds), *Folk Psychology: The Theory of Mind Debate*. Oxford: Blackwell Publishers.
- Greene, J.D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends Cogn. Sci.* 11, 322–323, author reply 323–324.
- Greene, J.D., Sommerville, R.B., Nystrom, L.E. *et al.* (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108.
- Greene, J.D., Nystrom, L.E., Engell, A.D. *et al.* (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44, 389–400.
- Grezes, J. and Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Hum. Brain Mapp.* 12, 1–19.
- Gu, X. and Han, S. (2007). Attention and reality constraints on the neural processes of empathy for pain. *NeuroImage* 36, 256–267.
- Hare, R.D. (1982). Psychopathy and physiological activity during anticipation of an aversive stimulus in a distraction paradigm. *Psychophysiology* 19, 266–271.
- Hare, R.D. (2006). Psychopathy: a clinical and forensic overview. *Psych. Clin. North Am.* 29, 709–724.
- Harris, P. (1989). *Children and Emotion*. Oxford: Blackwell Publishers.
- Harrison, N.A., Singer, T., Rotshtein, P. *et al.* (2006). Pupillary contagion: central mechanisms engaged in sadness processing. *Social Cogn. Affect. Neurosci.* 1, 5–17.
- Hart, A.J., Whalen, P.J., Shin, L.M. *et al.* (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *NeuroReport* 11, 2351–2355.
- Heal, J. (1986). Replication and functionalism. In: J. Butterfield (ed.), *Language, Mind, and Logic*. Cambridge: Cambridge University Press, Cambridge, Reprinted in 1995 in M. Davies and T. Stone (eds), *Folk Psychology: The Theory of Mind Debate*. Oxford: Blackwell Publishers.
- Hill, E., Berthoz, S., and Frith, U. (2004). Cognitive processing of own emotions in individuals with autistic spectrum disorders and in their relatives. *J. Autism Dev. Disorders* 34, 229–235.
- Hoffman, M.L. (2000). *Empathy and Moral Development: Implications for Caring and Justice*. Cambridge: Cambridge University Press.
- Iacoboni, M. and Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* 7, 942–951.
- Jabbi, M., Swart, M., and Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *NeuroImage* 34, 1744–1753.
- Jackson, P.L., Meltzoff, A.N., and Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage* 24, 771–779.
- Jackson, P.L., Brunet, E., Meltzoff, A.N., and Decety, J. (2006). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia* 44, 752–761.
- Jacob, P. and Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends Cogn. Sci.* 9, 21–25.
- James, W. (1894). Physical basis of emotion. *Psychological Rev.* 101, 2005–2010.
- Keysers, C. and Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Prog. Brain Res.* 156, 379–401.
- Keysers, C. and Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends Cogn. Sci.* 11, 194–196.
- Keysers, C., Wicker, B., Gazzola, V. *et al.* (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron* 42, 335–346.
- Kiehl, K.A., Smith, A.M., and Hare, R.D. *et al.* (2001). Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Biol. Psychiatry* 50, 677–684.
- King-Casas, B., Tomlin, D., Anen, C. *et al.* (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.
- Knoch, D., Pascual-Leone, A., Meyer, K. *et al.* (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829–832.
- Lamm, C., Batson, C.D., and Decety, J. (2007). The neural substrate of human empathy: effects of perspective taking and cognitive appraisal. *J. Cogn. Neurosci.* 19, 42–58.
- Lange, C. (1885). *The Emotions*. Reprinted in 1967 in C. Lange and W. James (Eds), *The Emotions*. New York, NY: Harner Publishing Co.
- Lazar, S.W., Kerr, C.E., Wasserman, R.H. *et al.* (2005). Meditation experience is associated with increased cortical thickness. *NeuroReport* 16, 1893–1897.
- Levenston, G.K., Patrick, C.J., Bradley, M.M., and Lang, P. (2000). The psychopath as observer: emotion and attention in picture processing. *J. Abnormal Psychol.* 109, 373–386.
- Linden, W., Wen, F., and Paulhaus, D.L. (1994). Measuring alexithymia: reliability, validity, and prevalence. In: J. Butcher, C. Spielberger, and N.J. Hillsdale (eds), *Advances in Personality Assessment*. Mahwah, NJ: Lawrence Erlbaum, pp. 125–143.
- Lutz, A., Brefczynski-Lewis, J.A., and Davidson, R.J. (2004). Loving-kindness and compassion meditation results in unique patterns

III. SOCIAL DECISION-MAKING, NEUROECONOMICS, AND EMOTION

- of fMRI activation and enhances the reactivity of the insula/cingulate neural circuitry to negative stimuli in meditators. Slide presentation presented at the 34th Annual Meeting of the Society for Neuroscience, San Diego, CA.
- McCabe, K. and Singer, T. (2008). Brain signatures of social decision making. In: C. Engel and W. Singer (eds), *Better Than Conscious: Exploiting the Capacity of Humans to Reach Decisions by Both Serial and Parallel Processing of Information*. Cambridge, MA: MIT Press, pp. 103–122.
- McCabe, K., Houser, D., Ryan, L. et al. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA* 98, 11832–11835.
- Mehrabian, A. and Epstein, N. (1972). A measure of emotional empathy. *J. Personality*, 40, 525–543.
- Mitchell, J.P., Heatherton, T.F., and Macrae, C.N. (2002). Distinct neural systems subserve person and object knowledge. *Proc. Natl Acad. Sci. USA* 99, 15238–15243.
- Mitchell, J.P., Banaji, M.R., and Macrae, C.N. (2005a). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *NeuroImage* 28, 757–762.
- Mitchell, J.P., Banaji, M.R., and Macrae, C.N. (2005b). The link between social cognition and self-referential thought in the medial prefrontal cortex. *J. Cogn. Neurosci.* 17, 1306–1315.
- Mitchell, J.P., Macrae, C.N., and Banaji, M.R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron* 50, 655–663, comment 531–534.
- Moll, J. and de Oliveira-Souza, R. (2007). Moral judgments, emotions and the utilitarian brain. *Trends Cogn. Sci.* 11, 319–321.
- Moll, J., de Oliveira-Souza, R., Bramati, I.E., and Grafman, J. (2002a). Functional networks in emotional moral and nonmoral social judgments. *NeuroImage* 16, 696–703.
- Moll, J., de Oliveira-Souza, R., Eslinger, P.J. et al. (2002b). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *J. Neurosci.* 22, 2730–2736.
- Moll, J., Krueger, F., Zahn, R. et al. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl Acad. Sci. USA* 103, 15623–15628.
- Montague, P.R., Berns, G.S., Cohen, J.D. et al. (2002). Hyperscanning: simultaneous fMRI during linked social interactions. *NeuroImage* 16, 1159–1164.
- Moriguchi, Y., Decety, J., Ohnishi, T. et al. (2007). Empathy and judging other's pain: an fMRI study of alexithymia. *Cerebral Cortex* 17, 2223–2234.
- Morris, J.S., Frith, C.D., Perrett, D.I. et al. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* 383, 812–815.
- Morrison, I. and Downing, P.E. (2007). Organization of felt and seen pain responses in anterior cingulate cortex. *NeuroImage* 37, 642–651.
- Morrison, I., Lloyd, D., di Pellegrino, G., and Roberts, N. (2004). Vicarious responses to pain in anterior cingulate cortex: is empathy a multisensory issue? *Cogn. Affect. Behav. Neurosci.* 4, 270–278.
- Morrison, I., Peelen, M.V., and Downing, P.E. (2007). The sight of others' pain modulates motor processing in human cingulate cortex. *Cerebral Cortex* 17, 2214–2222.
- Nemiah, J.C., Freyberg, H., and Sifneos, P.E. (1976). Alexithymia: a view of the psychosomatic process. In: O.W. Hill (ed.), *Modern Trends in Psychosomatic Medicine*, Vol. 3. London: Butterworths, pp. 430–439.
- Ochsner, K.N. and Lieberman, M.D. (2001). The emergence of social cognitive neuroscience. *Am. Psychologist* 56, 717–734.
- O'Doherty, J., Winston, J., Critchley, H. et al. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia* 41, 147–155.
- Perner, J.L., Leekam, S.R., and Wimmer, H. (1987). 2-year-olds' difficulty with false belief: the case for a conceptual deficit. *Br. J. Dev. Psychol.* 5, 125–137.
- Ploghaus, A., Tracey, I., Gati, J.S. et al. (1999). Dissociating pain from its anticipation in the human brain. *Science* 284, 1979–1981.
- Povinelli, D.J. and Bering, J.M. (2002). The mentality of apes revisited. *Curr. Dir. Psychol. Sci.* 11, 115–119.
- Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526.
- Preston, S.D. and de Waal, F.B. (2002). Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–20.
- Prinz, W. (1990). A common-coding approach to perception and action. In: O. Neumann and W. Prinz (eds), *Relationships Between Perception and Action: Current Approaches*. Berlin: Springer, pp. 167–201.
- Rilling, J.K., Gutman, D., Zeh, T. et al. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.
- Rilling, J.K., Sanfey, A.G., Aronson, J.A. et al. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage* 22, 1694–1703.
- Rilling, J.K., Glenn, A.L., Jairam, M.R. et al. (2007). Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biol. Psychiatry* 61, 1260–1271.
- Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* 3, 131–141.
- Saarela, M.V., Hlushchuk, Y., Williams, A.C. et al. (2007). The compassionate brain: humans detect intensity of pain from another's face. *Cerebral Cortex* 17, 230–237.
- Salminen, J.K., Saarijärvi, S., Aarela, E. et al. (1999). Prevalence of alexithymia and its association with sociodemographic variables in the general population of Finland. *J. Psychosomatic Res.* 46, 75–82.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A. et al. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.
- Saxe, R. (2005). Against simulation: the argument from error. *Trends Cogn. Sci.* 9, 174–179.
- Saxe, R. and Powell, L.J. (2006). It's the thought that counts: a specific brain regions for one component of theory of mind. *Psychological Sci.* 17, 692–699.
- Saxe, R. and Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia* 43, 1391–1399.
- Silani, G., Bird, G., Brindley, R. et al. (2008). Levels of emotional awareness and autism: an fMRI study. *Social Neurosci.*, DOI 10.1080/17470910701577020.
- Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neurosci. Biobehav. Rev.* 30, 855–863.
- Singer, T. and Fehr, E. (2005). The neuroeconomics of mind reading and empathy. *Am. Econ. Rev.* 95, 340–345.
- Singer, T., Seymour, B., O'Doherty, J. et al. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162.
- Singer, T., Seymour, B., O'Doherty, J.P. et al. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469.
- Smith, A. (2004). *The Theory of Moral Sentiments*. Kila, MT: Kessinger Publishing, (originally published in 1759).

III. SOCIAL DECISION-MAKING, NEUROECONOMICS, AND EMOTION

- Sterzer, P., Stadler, C., Poustka, F., and Kleinschmidt, A. (2007). A structural neural deficit in adolescents with conduct disorder and its association with lack of empathy. *NeuroImage* 37, 335–342.
- Veit, R., Flor, H., Erb, M. *et al.* (2002). Brain circuits involved in emotional learning in antisocial behavior and social phobia in humans. *Neurosci. Letts* 3283, 233–236.
- Wellman, H. (1990). *The Child's Theory of Mind*. Cambridge, MA: MIT Press.
- Wicker, B., Keysers, C., Plailly, J. *et al.* (2003). Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron* 40, 655–664.
- Wimmer, H. and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128.
- Winston, J.S., Strange, B.A., O'Doherty, J., and Dolan, R.J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat. Neurosci.* 5, 277–283.
- Wispe, L. (1986). The distinction between sympathy and empathy: to call forth a concept, a word is needed. *J. Pers. Social Psychol.* 50, 314–321.