Understanding Others: Brain Mechanisms of Theory of Mind and Empathy

Tania Singer and Anita Tusche

INTRODUCTION

Social Neuroscience

In the past few years, the field of neuroscience has shown increased interest in the study of the affective and social brain, taking into consideration the fact that humans are inherently social. A new interdisciplinary field called social neuroscience has emerged from a union between classical cognitive neuroscience and social psychology (for recent reviews, see Singer, 2012; Lieberman, 2012). 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 multilevel and multidisciplinary approach such as this 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., fMRI, EEG, or TMS) and autonomic measures (e.g., heart rate, galvanic skin conductance).

At the beginning, social neuroscience focused predominantly on the investigation of basic social abilities (for overview, see Adolphs, 1999, 2003; Blakemore et al., 2004; Ochsner and Lieberman, 2001). 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, racial identity, and faces of fair and unfair players (Hart et al., 2000; Morris et al., 1996; O’Doherty et al., 2003; Singer et al., 2004; Winston et al., 2002).

More recently, social neuroscience has addressed a broad range of questions concerning, for example, the processing of social rejection (Eisenberger, 2012; Eisenberger et al., 2003; Masten et al., 2011), the process of stereotyping (Nosek et al., 2009, 2011; Stanley et al., 2011) and the human ability to engage in emotion...
regulation (Ochsner and Gross, 2005; Wager et al., 2008; van’t Wout et al., 2010). Emotion regulation refers to processes by which people influence which emotions they have, when they have them, and how emotions are experienced and expressed (Gross, 2007). Recent advances have also been made concerning the effects of neuroendocrinology (e.g., steroid hormones or neuropeptides such as oxytocin and vasopressin) on social cognition and behavior (for reviews, see Bos et al., 2012; Insel, 2010; Meyer-Lindenberg et al., 2011).

Another important line of research has focused on our ability to understand other people’s minds, that is, their beliefs, intentions, and feelings. This line of research is the focus of the present chapter and will be elaborated on in the sections on Theory of Mind (ToM) (or mentalizing) and empathy (see Box 27.1).

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, 2007; Greene et al., 2001, 2004; Moll et al., 2002a,b; Shenhav and Greene, 2010; Sommer et al., 2010; for review, see Funk and Gazzaniga, 2009; Moll and de Oliveira-Souza, 2007; Moll et al., 2008). Social dilemma tasks are closely related to but still distinct from moral dilemma tasks. Social neuroscientists have used social dilemma tasks such as the simultaneous and sequential prisoner’s dilemma game and the ultimatum game (see Figure 27.1), which were developed within the framework of game theory, to investigate the neural underpinnings of social exchange and mutual cooperation. Studies employing these tasks involve people playing games for monetary payoffs and elicit the use of different playing strategies, some selfish and some cooperative, thereby allowing for the investigation of social reasoning (figuring out what the other player will do; e.g., Gallagher et al., 2002; McCabe et al., 2001; Rilling et al., 2004; Steinbeis et al., 2012), social emotions (emotional responses to fair and unfair play) and their interaction (Baumgartner et al., 2008, 2012; Fehr and

---

**BOX 27.1**

**GLOSSARY**

**Autistic disorder** (autism) refers to a pervasive neurodevelopmental condition associated with wide-ranging impairments in several areas of development, including social interaction and communication skills as well as markedly restricted repertoires of interests and activities.

**Asperger syndrome** (AS) refers to a pervasive developmental disorder that (like autistic disorder) is associated with marked impairments in social interaction and restricted behavior, while (unlike autistic disorder) language skills are not affected. It has been suggested that Asperger’s disorder is a milder form of autistic disorder.

**Compassion** refers to an emotional and motivation state that is associated with feeling concern for another’s suffering and desiring to enhance that individual’s welfare that can occur without the affective sharing by the observer. This affective state can be described as “feeling for” another person while empathy is characterized as “feeling with” someone.

**Empathy** refers to the ability to share the feelings of others. It can be defined as (i) an affective state which is isomorphic to another person’s affective state (ii) which was elicited by observing or imagining another person’s affective state (iii) when we know that the other person’s affective state is the source of our own affective state.

**Emotional contagion** refers to a phenomenon of an automatic adoption of an emotional state of another person. Compared to empathy, this state of affective sharing does not require knowledge about the origin of the affective experience (i.e., whether it is triggered by another person or lies within the observer).

**Psychopathy** is a disorder that is characterized by interpersonal behavior (e.g., pathological deception, manipulative/conning), affective responses (e.g., lack of remorse or guilt, lack of empathy), lifestyle (prone-ness to boredom/need for stimulation, parasitic lifestyle) and antisocial behavior (e.g., poor behavioral controls, juvenile delinquency).

**Schadenfreude** refers to a positive emotional state in the face of someone else’s misfortune (compared to envy that describes a negative emotional state in the face of another’s fortune).

**Theory of Mind** (ToM) or mentalizing describes the capacity to infer and to represent another person’s intentions, desires or beliefs. ToM differs from empathy in that the former does not denote a sharing of another person’s affective states, but rather a cognitive understanding of another person’s mental states.
Camerer, 2007; King-Casas et al., 2005; Montague et al., 2002; Rilling et al., 2002, 2007, 2008; Sanfey et al., 2003; Singer et al., 2004). This stream of research overlaps the most with, and has contributed to the emergence of, the new field of neuroeconomics. Note that, in addition to classic economic games, recent efforts have been made to develop novel paradigms that allow for the study of social encounters in a truly interactive manner by making use of virtual reality technologies (Schilbach et al., 2010a,b; Wilms et al., 2010). Moreover, the use of cross-correlational statistics to examine neural activation of several participants involved in a social interactive task has yielded initial promising results (Anders et al., 2011; Schippers et al., 2010; see also recent EEG studies on interbrain synchronization by Lindenberger et al., 2009; Kourtis et al., 2010).

Social Neuroscience and its Relation to Neuroeconomics and Decision Making

As introduced in the paragraph above, 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 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 like these (which are described in detail in Chapters 2, 11, and 25), people typically engage in economic exchange tasks in the laboratory. One example of a game that has also frequently been used in neuroscientific investigations is the ultimatum game (see Figure 27.1 and Chapters 2 and 11; Baumgartner et al., 2011; Knoch et al., 2006; Sanfey et al., 2003; Steinbeis et al., 2012). In this game, Mover One (M1) is given a certain amount of money and can then decide how much he/she wants to share with Mover Two (M2). M2 looks at the offer and can then decide whether he/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 were purely interested in money, he/she would accept any possible offer from M1, irrespective of whether this offer is deemed fair or unfair. Another such game that has been used in neuroscientific studies is the dictator game (see Figure 27.1 and Chapter 11; Moor et al., 2012; Steinbeis et al., 2012), in which M2 is neither allowed to reject nor accept the offers made by M1, but just passively receives whatever is offered.

But why does understanding how we understand others’ minds matter in economic exchange? To understand the answer to that question let us return to the examples discussed above: when we compare the offers M1s typically make in ultimatum games to those

FIGURE 27.1 Economic games. Subjects are assigned to the role of Mover One (M1) or Mover Two (M2). In the dictator game (DG), M1 (proposer) decides how to divide an initial endowment of x monetary units (MUs) between himself/herself and the other player. M2 (responder) is then passively presented with this offer and the endowment is allocated as proposed. The ultimatum game (UG) is similar to the DG except that M2 can accept or refuse the proposed offer. If the offer is accepted, the endowment is allocated as proposed by M1. If M2 refuses the offer, neither of the players gets any money. In the standard version of the prisoner’s dilemma game (PDG) both players simultaneously decide whether to trust the other player and to cooperate or to defect, without information on the other’s decision. Payoff matrices depend on both players, decisions, with the maximum individual payoff in the case of the other’s cooperation and one’s own defection.
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 for M2. Does M2 value fairness and, if so, which offer does M2 believe to be fair? How will M2 react if I give him/her x amount of money? This is the algorithmic process by which agents construct what are known in economic circles as beliefs. 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 McCabe and Singer, 2008; Singer and Fehr, 2005). However, economists still know little about (and have been classically uninterested in) 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.

Social neuroscientists and neuroeconomists have, thus, focused on clarifying the neural mechanisms underlying our capacity to represent others’ intentions, beliefs, and desires (referred to variously as Theory of Mind (ToM), cognitive perspective taking, mind reading, or mentalizing) and to share others’ feelings (referred to as empathy; see Box 27.1). 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 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; Eisenberg and Morris, 2001; Hoffman, 2000).

Accordingly, empathy is also likely to render people less selfish because it enables them to share others’ emotions and feelings, thereby motivating other-regarding behavior. Some behavioral and imaging evidence indeed suggests that people help others more when they report having empathized or show enhanced empathy-related brain activation with them (Eisenberg and Morris, 2001; Hein et al., 2010). Interestingly, despite the recent lack of interest in empathy by economists, the eighteenth century economist Adam Smith (2004: 1759) opened his second great volume, The Theory of Moral Sentiments by declaring:

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

The following section introduces central concepts in social neuroscience and summarizes major findings on the neural mechanisms underlying our ability to understand the beliefs, intentions, motives, and feelings of other people. Moreover, these findings will be discussed in light of possible implications for social and economic decision making.

**DEFINING CONCEPTS**

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 sub-skills 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 Batson, 1987, 2009; Batson et al., 1987; Decety and Jackson, 2004; Decety and Lamm, 2007; Eisenberg and Fabes, 1990; Hoffman, 2000; Keysers and Gazzola, 2006, 2007; Preston and de Waal, 2002; Wispe, 1986). In this article, 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: (a) our ability to understand other people’s motor intentions and action goals; (b) our ability to understand other people’s beliefs and thoughts, which is referred to as Theory of Mind (ToM), mentalizing, or cognitive perspective taking; and (c) our ability to understand other people’s feelings, which is referred to as empathy or emotional perspective taking (see also Blair, 2005; Decety and Lamm, 2007; de Vignemont and Singer, 2006; Keysers and Gazzola, 2007; Singer, 2006). Here we focus on outlining the latter two, cognitive perspective taking (or ToM) and empathy (see Box 27.1). A similar distinction has been proposed by James Blair who distinguishes between three main subsystems of empathy: cognitive, motor, and emotional empathy (Blair, 2005). Even though Blair’s conceptualization is very similar to the one proposed here, for purposes of clarity, empathy will be
used here to denote the capacity to understand other people’s feelings by sharing their affective states. Thus, empathizing with others does include an affective involvement. In contrast, ToM, cognitive perspective taking or mentalizing enables a person to represent the mental states of others, including their affective states, based on knowledge alone and without becoming emotionally involved.

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 (see Box 27.1), suggest that mentalizing and empathizing are actually two distinct abilities that rely on distinct neural circuitries (see Figure 27.2; 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, but not ToM, which may be the reason for their antisocial behavior (see also Blair, 2008). Thus, whereas psychopaths are apparently not impaired in their cognitive understanding of other people’s wishes, beliefs, intentions, and desires, it appears that they do not engage in empathizing with other and thus lack the feeling which could prevent them from harming other people as it would allow them to anticipate others’ suffering.

Empathy cannot be equated with affect sharing (experience of similar affective reactions) because affect sharing is also characteristic of emotional contagion, sympathy, personal distress, and compassion (see Box 27.1 and the Social Emotions section of this chapter). In line with other authors, de Vignemont and Singer (2006) have proposed a narrower definition of empathy that comprises three main components (also see Singer and Lamm, 2009). According to their definition, we empathize with others when we (i) have an affective state which is isomorphic to another person’s affective state and (ii) which was elicited by observing or imagining another person’s affective state (iii) when we know that the other person’s affective state is the source of our own affective state. The latter is important for differentiating empathy from emotional contagion in which affect sharing also takes place, but self-other distinction does not. Furthermore, empathy differs from sympathy and compassion in that, in the former, another person’s affect is shared, but no other-regarding concern or motivation takes place (for similar argument see Klimecki and Singer, 2012; Singer and Steinbeis, 2009). In other words, there is no motivation to maximize another person’s happiness or alleviate another person’s distress (see also the section Social Emotions for a more exhaustive coverage of emotional contagion and compassion). Note, however, that too much empathy can result in one’s own distress and consequently in a withdrawal from — instead of helping — the suffering person. Yet, in general, empathy is conceived to be a necessary first 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 prosocial behavior. Thus, empathy and prosocial decision making may be considered closely linked.

As stated above, ToM or mentalizing differs from empathy in that the former does not denote a sharing of another person’s affect but rather a cognitive understanding of another person’s intentions or beliefs. In the following, major findings and important streams of research on the study of theory of mind are presented.

### THE STUDY OF THEORY OF MIND

#### The History of Theory of Mind Research

In 1978, Premack and Woodruff published a seminal paper in which they coined the term Theory of Mind while discussing whether chimpanzees are capable of representing other primate’s minds in terms of their desires, intentions, and beliefs.
The Neural Foundation of Theory of Mind

With the development of modern imaging techniques, the study of our capacity to reason about other people’s minds has become a 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, subjects in a scanner are typically provided with stories based in text, abstract moving shapes, or cartoons and are asked to understand the intentions, beliefs, and desires of the protagonist in the respective stories (for a review, see Gallagher and Frith, 2003). Studies on ToM 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; for reviews and meta-analysis, see Amodio and Frith, 2006; Frith and Frith, 2010; Mitchell, 2009; Saxe et al., 2004; Van Overwalle, 2009). A schematic representation of the mentalizing brain network is illustrated in Figure 27.2 in blue. 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, while the temporal poles draw on encodings of past experience to “simulate” another person’s experience. Rebecca Saxe has suggested that different subcomponents of ToM have different developmental time courses and rely on different brain regions. In line with earlier approaches in developmental 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 Powell, 2006; Saxe and Wexler, 2005). Thereby, Saxe put forward the influential notion of a specific functional role of the (right) TPJ in inferring mental states and beliefs of others. However, this rather domain-specific view of the TPJ has been challenged by studies that linked TPJ activation to more domain-general, low-level computational processes underlying for example reorientation of attention or multi-sensory integration (Corbetta and Shulman, 2002; Decety and Lamm, 2007; Mitchell, 2008; however, for converse evidence, see Scholz et al., 2009; Young et al., 2010).

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 and colleagues (2002) and McCabe and colleagues (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.

Please note, however, that the mPFC is not only involved when people mentalize about other people’s thoughts, intentions, and beliefs, but also when people engage in self-referential processing such as introspection about one’s own mind, mental self-projection and mind-wandering (Buckner and Carroll, 2007; Christoff et al., 2009; Jenkins and Mitchell, 2011; Schooler et al., 2011; see also Chapters 8, 11, 13, 20 for evidence linking the mPFC to more general value-related processing). Interestingly, it has been suggested that people might use their own mental states as a starting point when inferring mental states of others, followed by an adjustment of these self-based inferences based on the perceived differences between the self and the other. Support for such an anchoring and adjustment view was recently provided by Tamir and Mitchell (2010). The authors found that activation in the mPFC related linearly to this self-other discrepancy when inferring the mental states of others. This finding builds on earlier work on mentalizing by Jason Mitchell (Mitchell et al., 2002, 2005, 2006) that suggested that there are functional differences between judging the mental states of similar and dissimilar others. One part of the mPFC was shown to be 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 them. 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 about whether judging another person’s mental state may be biased in different ways depending on whether we perceive them as similar or dissimilar. “Egocentric bias,” the propensity to understand other people’s states in terms of one’s 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 oneself.

THE STUDY OF EMPATHY AND FEELINGS

Empathy: A Shared Network Hypothesis

In addition to the ability to understand abstract mental states such as another person’s beliefs or desires, humans can also empathize with others, that is, share and understand another person’s feelings and emotions. Humans can feel empathy for other people in a wide variety of contexts: when others feel basic emotions and sensations such as anger, fear, sadness, joy, pain, and lust, as well as more complex emotions like embarrassment and social exclusion. 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 overviews, see de Vignemont and Singer, 2006; Keysers and Gazzola, 2006, 2009; Singer and Lamm, 2009). For example, some studies have been able to demonstrate that similar neural responses in the anterior insula (AI) cortex (see Figure 27.2) – 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 (see Figure 27.2) when subjects watched videos of people being touched and when they were being touched themselves (Keysers et al., 2004). These results are in line with the role of somatosensory cortices for the processing of touch (Ebisch et al., 2008; Keysers et al., 2004).
The majority of studies on empathic brain responses have, however, been conducted in the domain of pain (Avenanti et al., 2005, 2006; Botvinick et al., 2005; Bufalari et al., 2007; Cheng et al., 2007; Gu and Han, 2007; Jackson et al., 2005, 2006; Lamm et al., 2007a; Moriguchi et al., 2007; Morrison et al., 2004, 2007; Morrison and Downing, 2007; Singer et al., 2004, 2006; Saarela et al., 2007; for recent meta-analyses, see Lamm et al., 2011; Fan et al., 2011). 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 (see Figure 27.3). This imaging study was probably the first ever to involve two people in the usually nonsocial scanner environment. 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 would appear on a screen behind the board pointing to either the male or the female partner’s hand, indicating who would receive the painful and the non-painful stimulation. This procedure enabled the measurement of pain-related brain activation when pain was applied to the scanned subject (felt pain) or to her partner (empathy for pain).

FIGURE 27.3  Empathy for pain: experimental setup. Singer and colleagues (2004) recruited couples to measure brain responses in the female partners (placed within the scanner) while painful stimulation was applied via electrodes to either her own hand or to her partner’s hand which was lying on a tilted board in front of the female partner. Different colored arrows would appear on a screen behind the board pointing to either the male or the female partner’s hand, indicating who would receive the painful and the non-painful stimulation. This procedure enabled the measurement of pain-related brain activation when pain was applied to the scanned subject (felt pain) or to her partner (empathy for pain).

By using multi-voxel pattern techniques (Chapter 6), distributed activation patterns in the AI and the mACC have recently been shown to encode self-experienced pain as well as vicarious responses evoked when seeing another in pain as is consistent with shared networks accounts of empathy (Corradi-Dell’Acqua et al., 2011).

In summary, the consistency of findings of shared circuitries underlying one’s own sensations and feelings and the observation of similar sensations and feelings in others suggests that we use neural representations reflecting our own emotional responses to understand how it feels for others to be in a similar state. Furthermore, the evidence indicates that our ability to empathize may have evolved from a system that 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). In particular, brain regions such as the AI
cortices and the ACC (see Figure 27.2) have frequently been shown to play a central role in empathy for others. Given recent meta-analytic findings (Fan et al., 2011; Kurth et al., 2010; Lamm et al., 2011), this seems to hold true for empathic responses in various domains including emotional and physical pain, taste, and disgust and even for higher-order emotions such as embarrassment (Krach et al., 2011) and social exclusion (Masten et al., 2011; for recent review of shared networks for social and physical pain, see also Eisenberger, 2012).

The Role of Interoceptive Cortex in Feeling and Empathy

Whereas the beginning of affective and social neuroscience was characterized by a strong focus on the exploration of the role of amygdala in emotional processing, the focus has now broadened to include another structure that plays a crucial role in processing feelings: the insular cortex and, in particular, the AI. As introduced above, numerous findings from neuro-imaging studies on empathy for taste, disgust, and pain indicated that the AI cortices play a crucial role in empathy and feeling states. It has been suggested that these regions represent a crucial part of the human interoceptive cortex (with interoception referring to the sense of the physiological condition of the body; Craig, 2002) and subserve neural representations of internal bodily and feeling states, including pain, taste, hunger, thirst, and arousal (Critchley et al., 2001, 2004; Damasio, 1994). A special issue on the insula in the journal Brain Structure and Functions published in 2010 offers a commendable collection of current evidence on the variety of cognitive and affective functions of the insula cortex. Here, a more detailed account of the possible functions of the interoceptive cortex and its role in feelings in general and empathy in particular is provided.

Interoceptive models of emotions have had a long tradition in psychology and propose that cortical representations of internal bodily signals are at the origin of feeling states. In the late nineteenth 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 because we have to make a difficult decision; and we feel our face reddening with rage or blushing when we experience an embarrassing situation. Emotions cannot be experienced in the absence of these bodily feelings.

Based on anatomical observations in nonhuman species, Bud Craig (2002, 2009) 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 the 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 the right AI is assumed to subserve subjective feelings and was even proposed to be the seat of our awareness of a physical self as a feeling entity (see also Critchley et al., 2001; Damasio, 1994). At the same time, afferents also project by way of the medial dorsal thalamic nucleus to the ACC to produce behavioral drive. Thus, direct activation of both the insula (also referred to as “limbic sensory cortex”) and the ACC (referred to as “limbic motor cortex”) may correspond to a 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 the 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 (i.e., loss of nerve supply) 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 the right AI. This suggests a sensitivity of the right AI to autonomic feedback, which is absent in individuals with PAF. In addition, emotional awareness of the stimuli was manipulated using backward masking (refers to the phenomenon of decreased conscious processing of a usually briefly presented visual stimulus (“target”) when another visual stimulus (“mask”) is presented immediately afterwards). In accordance with the theory suggesting a role of the
AI in the conscious experience of emotions, the researchers demonstrated, as in previous studies, a sensitivity of amygdala to unconsciously perceived threat stimuli and a sensitivity of the AI to consciously perceived conditioned faces (Critchley et al., 2002). In a subsequent study, Critchley demonstrated that the activity and size of the 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 the interoceptive cortex plays an important role for the representation and awareness of feeling states arising from the body.

As the above-mentioned 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 and colleagues (2004) extended an interoceptive model of emotions to the domain of empathy and suggested that cortical re-representations in AI of bodily states may have a dual function (for similar argument, see also Singer et al., 2009; Lamm and Singer, 2010). First, they allow us to form subjective representations of feelings. These representations not only allow us 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 this context, it is noteworthy that the anticipation of pain (Ploghaus et al., 1999) or pleasant touch (Lovero et al., 2009) has been found to activate more anterior insular regions, whereas the actual experience of pain or pleasant touch also engages more posterior insular regions. This is in line with the above-mentioned 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. Similarly, in Singer and colleagues’ (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 both when participants were experiencing pain themselves and when they were vicariously experiencing it. 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 increasing the capacity for empathy. Second, deficits in understanding one’s own emotions would be associated with empathy deficits. Whereas evidence for the first hypothesis is still lacking, evidence for the second hypothesis is slowly accumulating (Bird et al., 2010; Silani et al., 2008).

Individual Differences in Empathy

So far, we have presented major findings on the neural substrate underlying the human ability to empathize with others and highlighted the relevant role of the interoceptive cortex in empathy-related processing. Yet as we all experience in our everyday lives, people are not equally empathic. Evidence for individual differences in empathic skills has been observed in the previously mentioned empathy studies of adults randomly selected from the normal population. 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). 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 covary with inter-individual differences in IRI and BEES scores. The higher subjects scored on these questionnaires, the greater the activation in the ACC and AI. Interestingly, Jabbi and colleagues (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 were not only positively correlated with trait measures of empathy, but also with unpleasantness ratings which subjects gave online after each trial of a scanning session (Jackson et al., 2005; Lamm et al., 2007a; Saarela et al., 2007; Singer et al., 2008; for an overview, see Lamm et al., 2011). Interestingly, recent meta-analytic evidence suggests that such online state measures of felt empathy or unpleasantness ratings might yield even more robust correlations with neural responses than trait measures of empathy (Lamm et al., 2011).

Empathic responses were also found to be modified by individual characteristics such as the degree of alexithymia. 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 Asperger syndrome (AS; see Box 27.1; Hill et al., 2004). Interestingly, individual differences in the degree of alexithymia have recently been shown to be negatively correlated with individual differences in trait empathy (Bird et al., 2010; Silani et al., 2008). Using fMRI, Silani and colleagues (2008) scanned subjects with AS and controls with varying degrees of alexithymia while they performed a task that required them to experience their own feelings. Specifically, subjects were to judge how they felt about emotionally loaded pictures. Results showed that the degree of severity in alexithymia, as measured by two different alexithymia scales, was associated with less activation in AI. 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 confirm again that the AI plays a role in understanding one’s own emotions. Interestingly, individual differences in the degree of alexithymia correlated highly negatively with individual differences in trait empathy, 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 the AI.

When do we Care About Others? Modulatory Factors in Empathy

As described in the last section, there are substantial individual differences with regard to empathy in the normal healthy population as well as in patient populations with severe emotional and social deficits. In addition to these person-specific differences in empathic responses, the degree to which we have empathic feelings also varies as a function of situational factors and our appraisal of the situation. For example, it is usually easier to empathize with someone who has treated one well than with someone who has treated one poorly. In recent 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., 2007a), the frequency of a person’s prior exposure to pain-inducing situations (Cheng et al., 2007), the intensity of the inflicted pain (Avenanti et al., 2006; Saarela et al., 2007) and group membership (Avenanti et al., 2010; Hein et al., 2010; Xu et al., 2009) all seem to play a role in the modulation of the magnitude of empathic brain responses.

For instance, in a recent fMRI study, Hein and colleagues (2010) investigated whether social group membership impacts the empathic responses to the suffering of another person and the willingness to engage in costly helping. Using an empathy-for-pain paradigm, soccer fans expressed increased empathic concern for the suffering of members of their favorite soccer team (ingroup) compared to members of the rivalry team (outgroup). In line with previous findings, the self-reported degree of empathic concern was reflected in neural responses in the AI. Participants’ empathy-related processing in the AI when exposed to another’s suffering predicted subsequent helping behavior: the decision to reduce the others’ pain by enduring half of that pain themselves. Importantly, this effect was much stronger for ingroup members than for outgroup members. The decision to refrain from helping an outgroup member, on the other hand, was related to activation in the ventral striatum (nucleus accumbens) when witnessing the other experience pain (see Figure 27.4). These findings indicate that contextual factors such as group membership can modulate empathic brain responses in the AI and can motivate prosocial decisions such as costly helping. Moreover, group membership was found to modulate reward-related brain responses when witnessing other’s suffering, which may be linked to a motivational system that opposes empathy-related motivation (i.e., reflects gloating) and decreases prosocial behavior.

Importantly, these findings are largely consistent with previous results by Singer and colleagues (2006) showing that empathic responses to another person’s pain are modulated by the perceived fairness of the other. In this study, male and female volunteers first played repeated sequential prisoner’s dilemma games (see Figure 27.1 and Chapter 2) as Mover 1 with two confederates. One confederate played fairly by reciprocating the subject’s trust by returning fair amounts of money; the other played unfairly by selecting only self-interested choices and responding with no or minimal returns. After that, an empathy-for-pain paradigm similar to the one reported by Singer and colleagues (2004; see Figure 27.3) was used to measure the subject’s empathic brain responses while either the subject or one of the confederates was receiving painful stimulation to his/her hand. To assess gender differences in empathy and its modulation, both men and women were scanned and paired with a pair of either female or male players. As in previous empathy studies,
empathy-related activation in the ACC and the 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, see Figure 27.4), 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. It suggests that punishing free riders activates reward circuitries usually engaged in processing primary rewards and may help to explain why people are motivated to engage in altruistic punishment even though this behavior may seem irrational and altruistic because it is costly.

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 hand, are more likely to occur.
SOCIAL EMOTIONS: EMOTIONAL CONTAGION, COMPASSION, ENVY, AND SCHADENFREUDE

So far, this chapter has mainly focused on empathy and the human ability to share the feeling of others. Yet emotional phenomena such as emotional contagion and compassion (see Box 27.1) that are closely related to empathy, play an important role in our social lives and crucially shape social interactions. In addition to outlining these social emotions, this section briefly addresses affective states such as schadenfreude and envy (see Box 27.1) that oppose empathic responses and prosocial motivation.

Emotional contagion describes the phenomenon of an automatic adoption of an emotional state of another person. Compared to the notion of empathy as introduced before (de Vignemont and Singer, 2006; Singer and Lamm, 2009), this state of affective sharing does not require knowledge about the origin of the affective experience (whether it is triggered by another person or lies within the observer). For example, long before babies develop a sense of a self, separate from others, they start crying when they hear other babies crying (Simner, 1971). Using fMRI and pupillometry, 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. Please note, however, that even though the concept is distinguished from empathy, emotional contagion is believed to frequently precede empathy (Singer and Lamm, 2009).

Compassion, on the other hand, refers to a state that is associated with “feeling concern for another’s suffering and desiring to enhance that individual’s welfare” (Keltner and Goetz, 2007), that can occur without the affective sharing by the observer (Klimecki and Singer, 2012; Singer and Steinbeis, 2009). As outlined in Klimecki and Singer (2012), this definition of compassion is closely related to the concepts of “empathic concern” as used by Batson and colleagues (1983) and “sympathy” as used by Eisenberg and Fabes (1990). In other words, empathic concern, but also compassion, can be described as “feeling for” another person while empathy is characterized as “feeling with” someone (Batson, 2009). In line with this notion of shared affective states in empathy, subjects have been shown to experience marked negative affect when witnessing the distress of another (Klimecki et al., 2012; Lamm et al., 2007b; Saarela et al., 2007). For instance, using a novel Socio-affective Video Task (SoVT), Klimecki and Singer (2012) presented subjects with film sequences of people in distressing situations. At pretest, subjects responded with strong negative affect to the task. Self-reported empathic responses in reaction to others’ suffering were accompanied by increased activation in brain regions that have frequently been linked to empathy for pain such as the ACC and AI (see Figure 27.2; for meta-analysis, see Lamm et al., 2011). Importantly, after subjects received brief compassion training and adopted a compassionate state during the task, positive affect increased when confronted with others’ suffering. Adopting a compassionate state when exposed to others’ distress was also associated with increased activation in the medial orbitofrontal cortex (mOFC), ventral tegmental area/substantia nigra (VTA/SN), pallidum and putamen (see Figure 27.5). Evidence from recent cross-sectional studies further supports the involvement of these structures in compassion: Using pictures of sad faces, Kim and colleagues (2009) found that taking a compassionate attitude towards another’s sad affect recruited the mOFC and VTA/SN. Likewise, adopting an attitude of love towards disabled people was found to involve the VTA, pallidum and the mid insula (Beauregard et al., 2009; see also Immordino-Yang et al., 2009; Lutz et al., 2008). Please note that the insula has also been involved in previous neuroimaging studies on the effects of expertise in compassion (Lazar et al., 2005; Lutz et al., 2004).

The finding that compassion training increased positive affect when witnessing others’ suffering (Klimecki et al., 2012) suggests that a cultivation of compassion might offer a new coping strategy which could help people to face distressing social situations with enhanced resilience. Given that both empathic responses and emotional contagion in such situations might lead to the experience of personal distress and, ultimately, result in withdrawal behavior to reduce one’s own negative emotions (Batson et al., 1983; Eisenberg and Fabes, 1990; Klimecki and Singer, 2012), this result of a compassion training is particularly striking. In addition to advantageous effects of compassion training on positive affect (Fredrickson et al., 2008), recent evidence also indicates that neuroendocrinological responses to stress might benefit from the adoption of this feeling state (Pace et al., 2009). Initial evidence even suggests that short-term training of compassion may increase prosocial behavior.
towards others in a task unrelated to the training (Leiberg et al., 2011).

Next, we turn to other social emotions such as envy or schadenfreude that seem to counteract empathic responses and prosocial motivation. Envy can be described as a negative emotional state in the face of another’s fortune, while schadenfreude refers to a positive emotional state in the face of someone else’s misfortune (Shamay-Tsoory et al., 2007). Participants who compared themselves with another person who was considered to possess desirable and self-relevant characteristics were found to respond with stronger envy and a corresponding increase of activation in the ACC (Takahashi et al., 2009; see also Mobbs, 2009). In contrast, when participants were presented with misfortune happening to the envied person, they reported stronger schadenfreude (Takahashi et al., 2009). The latter was reflected in brain responses in the ventral striatum (VS), supporting the notion of a rewarding nature of this affective state. In line with this finding, Dvash and colleagues (2010) found increased activation in the VS for participants that lost money when the other player in a game lost even more money, accompanied by stronger schadenfreude. Interestingly, reward-related activation in the ventral and dorsal striatum has also been linked to affective states such as a desire for revenge and altruistic punishment of defectors when fairness preferences have been violated (de Quervain et al., 2004; Singer et al., 2006). Thus, Singer and colleagues (2006) found participants’ activation in the VS to increase when seeing an unfair player in pain, which correlated positively with their desire for revenge as assessed by questionnaires after the scanning session. Empathic responses, on the other hand, were significantly reduced when they passively watched the unfair confederate receiving pain. These results suggest that motivational systems such as the desire for revenge can win over an empathic motivation when people are confronted with another’s suffering who they believe to deserve to be punished. Interestingly, Hein and colleagues (2010) showed that the decision to refrain from helping an outgroup member was related to activation in the VS when witnessing the other experiencing pain. Participants’ empathy-related processing in the AI for another’s suffering, on the other hand, predicted the decision to engage in costly helping. These findings indicate that the opposing motivational systems of empathy on one side and envy, schadenfreude or revenge on the other side can be predictive of engagement in prosocial or egoistic behavior. For a more exhaustive coverage of these latter studies and of fairness and revenge see also the section When do we Care About Others in this chapter, and Chapter 11 in this volume.

**FUTURE RESEARCH DIRECTIONS**

Recent efforts in social neuroscience and neuroeconomics have helped to shed light on the mechanisms underlying social cognition and social emotions such as empathy, compassion, our sense of fairness, revenge and gloating. Different neural networks have been identified that allow us to represent other people’s cognitive and emotional states and first steps have been made to specify the contributions of these social cognition networks in social decision making. Yet the field of social neuroscience is moving on, and so are the questions that will hopefully be addressed by future research. In the following section, important open questions and promising research directions will be addressed and discussed in light of their implications for the field of neuroeconomics.

Lately questions about the plasticity of socioaffective capacities have started to move into the focus
of social neuroscience. For instance, as reviewed above, studies have started to investigate the structural and functional neural effects of expertise in compassion and its impact on affect and neuroendocrinological responses (Fredrickson et al., 2008; Klimecki et al., 2012; Lazar et al., 2005; Lutz et al., 2004; Pace et al., 2009). Here, longitudinal training-designs using a multi-method approach promise to substantially broaden our knowledge of the neural basis of the malleability of socio-affective capacities and their relation to changes in health, subjective wellbeing and other-regarding behavior. Initial evidence has indicated that a short-term compassion training can augment prosocial behavior even in tasks unrelated to training (Leiberg et al., 2011). Considering this finding, the research direction of interventions to foster the acquisition and training of social skills might be potentially relevant for neuroeconomic studies on social decision making. Such studies may also speak to relevant questions in neuroeconomics regarding the malleability of social preferences which have typically been regarded as stable over time.

Another promising research line of social neuroscience focuses on developmental aspects of socio-affective capacities over the life course and its impact on decision making. This line of research follows a long tradition of behavioral research on the development of social cognition and emotions (Eisenberg, 2000; Eisenberg and Fabes, 1990; Leslie, 1987; Wellman et al., 2001; Wimmer and Perner, 1983; Zahn-Waxler et al., 1992). At present, more and more evidence is accumulating on the neural basis underlying ontogenetic changes of social cognitive abilities and prosocial behavior during childhood and adolescence (for reviews, see e.g. Frith and Frith, 2007; Blakemore, 2008, 2012). By examining the interaction of differential time courses of brain development and co-occurring, age-related changes in cognitive and affective psychological functions, this approach promises to add considerably to our current models of social cognition and emotions. Interestingly, it has been suggested that our ability to empathize and to infer others’ mental states undergoes differential ontogenetic changes due to the different developmental trajectories of the respective underlying neural structures (Singer, 2006). More precisely, it has been suggested that empathic capacities and associated limbic and para-limbic structures evolve earlier in life than ToM, which has been linked to later-maturing temporal and prefrontal structures in the brain. Longitudinal and cross-sectional neuroimaging studies that acquire structural and functional brain measures and explicitly assess empathic and mentalizing skills over a wide age range will be required to provide compelling evidence for this notion.

Initial evidence for the potential of a developmental approach in neuroeconomics is provided by a recent cross-sectional study that investigated developmental trajectories of fairness preferences and its impact on economic decision making. In this study, Steinbeis and colleagues (2012) showed that the late-maturing dorsolateral prefrontal cortex is tightly linked to age-related differences in strategic behavior in monetary social exchange games. In two behavioral experiments with 6 to 13 year-old children, prosocial choices were compared when children were playing an ultimatum game (where a rejection of the offer by the other player is possible; see Figure 27.1) to those in a dictator game (where no rejection by the other player is possible, see Figure 27.1). The difference in offers between these two economic games was used to assess strategic decision making. Overall, children of various ages were found to offer more when the other player was able to reject the offer (i.e., punish for the unfair offer) than when no such punishment option was available. More importantly, however, strategic behavior was found to increase with age and to be related to improved impulse control with age. Individual differences in impulse control and strategic decision making in turn were explained by individual differences in the structural and functional characteristics of the late-maturing dorsolateral prefrontal cortex. These findings highlight the research potential of age-dependent changes in social decision making that include developmental brain maturation and developmental changes in cognitive abilities (see also pioneering MRI studies on developmental brain maturation by Sowell et al., 1999; Giedd et al., 1999). Future studies will have to employ this approach to shed light on the relationship between observed changes in social cognitive abilities, such as mentalizing and empathizing, to age-dependent changes in the brain and their impact on social decision making. Making use of the proposed differential time courses of brain structures underlying mentalizing and empathizing, this approach might also help to clarify the differential contribution of both capacities to prosocial behavior.

Furthermore, advanced methods of fMRI data analysis may help to clarify the nature of neural computations in observed shared networks, beyond a mere analysis of overlap of neural activation across experimental conditions. In contrast to conventional univariate analysis of fMRI data that focuses on activation in single voxels, multivariate pattern analysis techniques routed in machine learning take advantage of information contained in multiple voxels distributed across space (see Chapter 6 for more on these methods or Haynes and Rees, 2006; Haxby et al., 2001; Kriegeskorte et al., 2006; Norman et al., 2006; Tusche et al., 2010). It has been proposed that this multivariate pattern-information reflects the developmental
content encoded in the brain and might even reflect the underlying neural population code (Kamitani and Tong, 2005). This approach might help to answer open questions about what exactly is encoded (and shared) in “shared networks,” such as the AI and the ACC, in first-hand and vicarious experience of pain (Corradi-Dell’Acqua et al., 2011). Moreover, multivariate pattern approaches might also examine the representational content encoded in areas such as the AI and ACC or TPJ across various social and non-social tasks that have implicated these regions in a multitude of cognitive and affective functions. Yet, to move interpretations and the whole field of social neuroscience successively towards causal models of social cognition and emotions, additional techniques such as transcranial magnetic stimulation and transcranial direct current stimulation, pharmacological interventions, cross-cultural designs and genetic-imaging approaches should be progressively employed and combined with each other. The soaring application of such a multi-method (and multi-level) approach will help to address open questions and challenges of social neuroscience and to foster the further development of this promising field.

References


REFERENCES


NEUROECONOMICS


Leiberg, S., Klimke, O., Singer, T., 2011. Short-term compassion training increases prosocial behavior in a newly developed prosocial game. PlosOne. 6, e7798.


are selective for theory of mind and exogenous attention. PLoS One. 4, e4869.


