Models, Mechanisms and Moderators Dissociating Empathy and Theory of Mind

Philipp Kanske, Anne Böckler and Tania Singer

Abstract Most instances of social interaction provide a wealth of information about the states of other people, be it sensations, feelings, thoughts, or convictions. How we represent these states has been a major question in social neuroscience, leading to the identification of two routes to understanding others: an affective route for the direct sharing of others' emotions (empathy) that involves, among others, anterior insula and middle anterior cingulate cortex and a cognitive route for representing and reasoning about others' states (Theory of Mind) that entails, among others, ventral temporoparietal junction and anterior and posterior midline regions. Additionally, research has revealed a number of situational and personal factors that shape the functioning of empathy and Theory of Mind. Concerning situational modulators, it has been shown, for instance, that ingroup membership enhances empathic responding and that Theory of Mind performance seems to be susceptible to stress. Personal modulators include psychopathological conditions, for which alterations in empathy and mentalizing have consistently been demonstrated; people on the autism spectrum, for instance, are impaired specifically in mentalizing, while spontaneous empathic responding seems selectively reduced in psychopathy. Given the multifaceted evidence for separability of the two routes, current research endeavors aiming at fostering interpersonal cooperation explore the differential malleability of affective and cognitive understanding of others.

Philipp Kanske and Anne Böckler are joint first authors/contributed equally.

P. Kanske (✉) · A. Böckler (✉) · T. Singer
Department of Social Neuroscience, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1a, 04103 Leipzig, Germany
e-mail: kanske@cbs.mpg.de

A. Böckler
Department of Psychology, Julius-Maximilians-Universität Würzburg, Röntgenring 11, 97070 Würzburg, Germany
e-mail: aboeckler@cbs.mpg.de; anne.boeckler@uni-wuerzburg.de

Curr Topics Behav Neurosci
DOI 10.1007/7854_2015_412

Keywords Empathy · Theory of Mind · Social cognition · Temporoparietal junction · Anterior insula

Contents

1 Introduction

Our everyday lives embed us in various social networks, from breakfast at the family table to formal meetings with colleagues, in which we constantly and flexibly need to adjust our actions to those of others. A crucial step in order to successfully master these different kinds of social encounters is the accurate understanding of others' emotional and mental states. This includes sensing another's mood, inferring how to interpret another's (ironic?) utterance, or identifying our interaction partner's expectations. How do we achieve such understanding of others and what are the underlying mechanisms? Social neuroscience has addressed this question during the last decades and identified two routes allowing for representing others' states (Frith and Frith 2005; Singer 2006): The affective route entails directly sharing an observed person's emotions (empathy) (de Vignemont and Singer 2006), while the cognitive route enables inferences of other's thoughts, goals, or intentions (Theory of Mind, mentalizing) (Frith and Frith 2005; Premack and Woodruff 1978).

This chapter will first introduce these routes and their neural underpinnings in more detail and then outline a recent approach to dissociate empathy and mentalizing on a behavioral and brain level. Subsequently, situational and personal modulators of empathic responding and Theory of Mind capacities will be presented and a final outlook sketches current investigations of the malleability of the two routes to understanding others.

2 Empathy

The affective route to understanding others has mainly been studied under the term empathy. The concept was introduced by Robert Vischer and Theodor Lipps, who emphasized the role of "inner imitation" what they called "Einfühlung" (see Carr et al. 2003). The current literature defines empathy as an affective state, that is, (1) isomorphic to another person's affective state, (2) elicited by observing or imagining another person's affective state, and (3) includes knowing that the observed person's state is the source of one's own affect (de Vignemont and Singer 2006). Isomorphism refers to the idea that empathizing with somebody means directly sharing, "feeling with," his or her emotions, such as pain, sadness, or joy. Complementarily, observing another in an emotional state may, of course, also elicit non-isomorphic emotions such as envy, schadenfreude, or compassion. Crucially, even though these are all vicarious affective states (Paulus et al. 2013), empathy is per definition shared rather than complementary. Another important distinction of empathy concerns emotional contagion, the reaction in which one shares an emotion with another person *without* realizing that the other's emotion was the trigger. An example is the observation that infants start crying when they hear other infants cry, even before they presumably develop a clear sense of being an independent agent (Singer and Lamm 2009).

In order to investigate the neural underpinnings of empathy, psychologists and neuroscientists have used paradigms in which both the participant and an observed other received painful stimulation. Critically, this allowed for a direct comparison of neural activity elicited by being the recipient and by being a mere observer of pain (e.g., Singer et al. 2004). Clear overlap between these two conditions was found in a network of areas including bilateral anterior insula and middle anterior cingulate cortex. Therefore, "shared" brain networks have been proposed as an underlying mechanism for our ability to empathize (Decety 2010). Such shared neuronal representations subserving both the first-hand as well as the vicarious experience of emotions have been reported in different domains ranging from disgust (e.g., Wicker et al. 2003), touch (e.g., Keysers et al. 2004 and facial emotional expressions (e.g., Carr et al. 2003) to higher-order emotions such as social exclusion (Masten et al. 2011) and embarrassment (Muller-Pinzler et al. 2015).

Recent meta-analyses corroborate the involvement of anterior insula and middle anterior cingulate cortex in empathy for pain (Lamm et al. 2011). When including different types of empathy paradigms and emotional stimuli, a more extensive neural network was identified, including anterior insula and anterior cingulate cortex, but also more dorsal regions of the medial frontal cortex, inferior frontal gyri and anterior/dorsal parts of the temporoparietal junction (Bzdok et al. 2012). Interestingly, this network seems independent of whether spontaneous or instructed empathizing is tested (Fan et al. 2011).

As pointed out earlier, a distinction between isomorphic and complementary emotions has been suggested in the conceptual specification of empathy (de Vignemont and Singer 2006). For instance, imagine running into an angry boss at the non-functioning copy machine. While your boss' anger may elicit anger in you too, you may also be intimidated and feel anxious. Empirical support for the distinction of isomorphic and complementary emotions in response to the same situation comes from recent neuroimaging studies on isomorphic empathic responding ("feeling with") and complementary compassion ("feeling for"), which can both be induced by the suffering of others. Compassion, the feeling of warmth, care, and affiliation, does not rely on "shared" neural networks but rather activates regions commonly associated with positive affect and reward such as ventral striatum and medial orbitofrontal cortex (Klimecki et al. 2013). It has been hypothesized that empathy and compassion also differ with regard to the roles they play in interpersonal behavior. While empathic responding to another's suffering does not per se entail a prosocial motivation (i.e., the wish to alleviate the other's suffering) and might even lead to withdrawal from the other due to empathic distress, compassion toward others includes the motivation to enhance their well-being. In line with this hypothesis, recent studies showed increased helping behavior related to compassion trainings (Leiberg et al. 2011).

3 Theory of Mind

The cognitive route to understanding others has been investigated under the terms perspective taking, mentalizing, or Theory of Mind. The latter term was originally coined by Premack and Woodruff (1978), discussing whether chimpanzees can represent others' mental states such as desires, intentions, and beliefs. Similarly, developmental research has been dedicated to the question at what age children "have" a Theory of Mind. This conceptual definition of the term has been closely linked to the usage of so-called false-belief tasks that measure whether participants can correctly predict the actions of somebody who holds a false belief that is different from the participants' true belief. More recent investigations approach the concept by focusing on the cognitive operations that are involved in reasoning about another's (different) mental states (see Apperly 2012) or perceptions (Bockler and Zwickel 2013). In conclusion, research defines Theory of Mind as the process of inferring and reasoning about the perceptions, beliefs, thoughts, or emotions of others (Frith and Frith 2005).

In order to investigate the neural underpinnings of this cognitive route to understanding others, paradigms require participants to judge (mis-)assumptions or (false) beliefs of others that differ from their own, requiring the inhibition of the own mental state (Wimmer and Perner 1983). For example, participants are presented with a short story about a boy who believes that his bike is in the garden, while in fact his mother had taken it into the garage. Contrasting such stories with true belief conditions or with stories about changes in the physical world (e.g., an outdated map) yields activation in temporoparietal junction, superior temporal sulcus, temporal poles, and anterior and posterior midline regions (Dodell-Feder et al. 2011).

Recent meta-analyses convincingly demonstrate the involvement of the described neural network when incorporating a wide variety of different Theory of Mind tasks (Bzdok et al. 2012). Testing more stringently for regions that are consistently activated in most of these different tasks reveals a core network comprising temporoparietal junction and medial prefrontal cortex (Schurz et al. 2014). Comparing the different types of Theory of Mind tasks shows a more differentiated pattern within the overall network, with specific activation clusters for (among others) false-belief tasks (e.g., Aichhorn et al. 2009), social animations (e.g., Blakemore et al. 2003), and rational action judgment tasks (e.g., Walter et al. 2004). While judging others' mental states in realistically complex social situations has been shown to lead to more widespread activation within the Theory of Mind network (Wolf et al. 2010), the above described separations may prove helpful for delineating the specific processes that are represented in different parts of the network and that contribute to full-blown Theory of Mind.

Regarding the role of Theory of Mind in interpersonal behavior, it can be speculated that the capacity to understand other's mental states allows for more prosocial decision making. Possible mechanisms could include a better understanding of others' needs and circumstances or, based on the capacity to flexibly inhibit one's own mental states, an enhanced ability to inhibit one's immediate selfish needs (Batson 2011). In fact, recent evidence in children shows that Theory of Mind performance predicts enhanced prosocial behavior and thereby leads to better peer group integration (Caputi et al. 2012).

4 Dissociating Empathy and Theory of Mind

As described in the previous sections, empathy and Theory of Mind have inspired different research traditions that have employed different types of experimental paradigms and identified largely distinct neural networks underlying the two functions. Spelling out the differentiation on a conceptual level, empathy denotes the embodied sharing of a sensory or affective state, while Theory of Mind refers to the propositional knowledge about the state of another (including others' affective states). Beyond the differences between the two routes to understanding others, however, both functions yield access to another person's inner state. Coming back to the initial examples of breakfast at the family table or meetings among colleagues, both empathy and Theory of Mind are necessary to adequately react to your stressed husband's ironic remark about the timeless beauty of your vintage morning gown or to your new colleague's mixed feelings about his first group presentation. It is therefore plausible to assume that both functions and the related neural networks are concurrently required and active in almost all everyday social interactions. Indeed, meta-analytical evidence suggests that the temporoparietal junction, which is a core region of the Theory of Mind network (Schurz et al. 2014), is also activated in empathy studies (Bzdok et al. 2012). In particular, a recent meta-analysis gives some indication that empathizing with another's pain activates not only the typical empathy network (anterior insula, middle anterior cingulate cortex), but also Theory of Mind-related areas (temporoparietal junction, temporal poles, superior temporal sulcus, anterior and posterior midline regions) in situations where the other's state is not immediately visible and needs to be inferred (Lamm et al. 2011). Research in social neuroscience, therefore, needs to investigate the specifics and the separability of the neural networks underlying empathy and Theory of Mind, especially in ecologically valid instances of realistically complex social understanding that require both routes.

While some studies have compared cognitive and affective aspects of Theory of Mind (i.e., mentalizing on others' cognitive or affective states; (e.g., Bruneau et al. 2012; Shamay-Tsoory and Aharon-Peretz 2007) or have studied empathy and Theory of Mind in separation (Dziobek et al. 2011), the simultaneous manipulation and assessment of empathy and Theory of Mind within individuals has only recently tested (Kanske et al. 2015). Of course, empathy and mentalizing processes may co-occur in many experimental depictions of social situations. For instance, when manipulating empathic requirements by means of observed social exclusion, participants likely also take the (cognitive) perspective of the displayed agents, which is reflected in concurrent activation of both empathy- and Theory of Mind-related networks (Masten et al. 2011). Critically, because the two processes are not manipulated independently, clear-cut attribution of empathy and mentalizing functions to the specific neural activations are impossible. In the following paragraphs, we will outline a paradigm that explicitly applied an orthogonal manipulation of empathy and Theory of Mind and delineate the results in greater detail.

The task, henceforth termed EmpaToM, confronts participants with brief video clips of a person reporting on an autobiographic episode (see Fig. 1a). These episodes vary in emotionality (entailing neutral events such as the preparation of a meal or negative events such as loss of a loved one) and in Theory of Mind demands (necessitating inferences about mental states such as a planned deception or about physical states such as weather conditions). As a behavioral measure of empathic responding, participants rate the valence of their own affective state after each video. A subsequent rating asks for participants' compassion for the person in the video. Theory of Mind performance is assessed with multiple choice questions that ask either about the mental states of the person in the video or about factual relations in the story (control condition). This design orthogonally manipulates empathy and Theory of Mind requirements, hence including a condition in which both processes are elicited.

Using the EmpaToM, the typical neural networks for empathy and Theory of Mind could be identified, including anterior insula, anterior cingulate/medial frontal cortex for empathy and temporoparietal junction, temporal poles, and anterior and posterior midline regions for Theory of Mind. The revealed activation overlapped with meta-analytical masks (Bzdok et al. 2012) and with established empathy and Theory of Mind paradigms assessed in the same individuals (Dodell-Feder et al. 2011; Klimecki et al. 2013); behavioral parameters of empathizing and mentalizing could also be validated with existing tasks. Hence, even in complex situations of

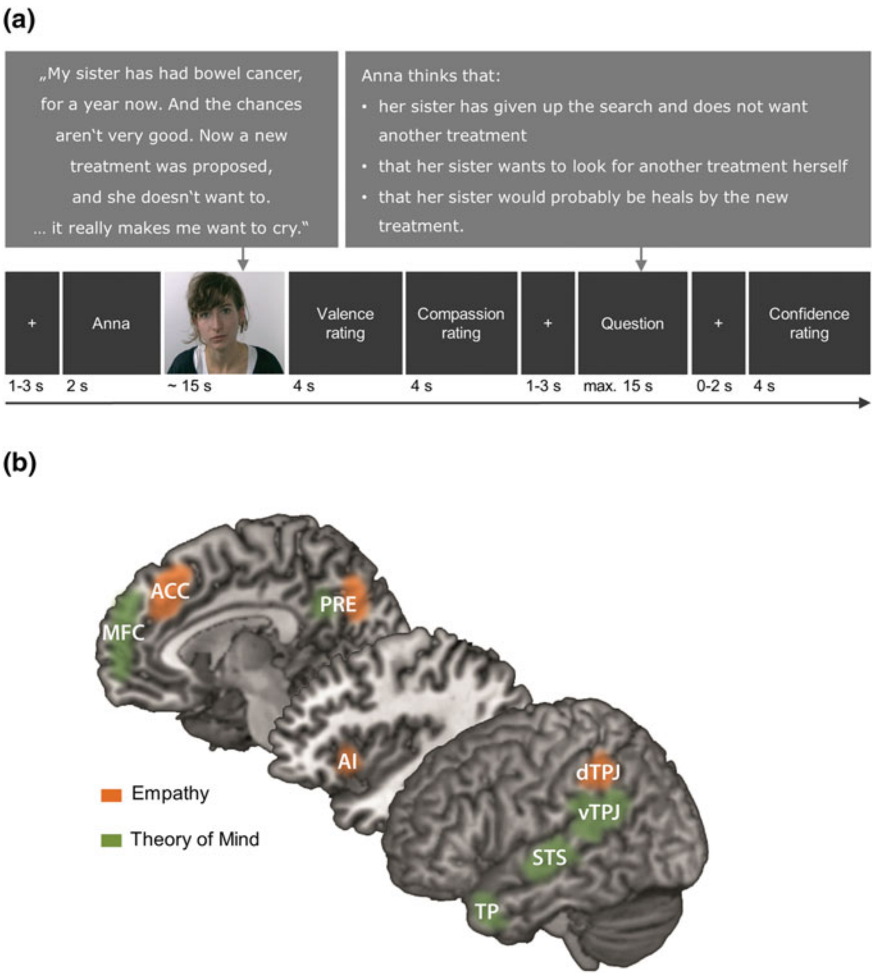

Fig. 1 **a** Overview of the EmpaToM task with an example video narration and corresponding question of the negative emotion/Theory of Mind condition. **b** Empathy and Theory of Mind networks (schematically depicted in the *left hemisphere*). Anterior cingulate cortex (ACC), anterior insula (AI), dorsal/ventral temporo-parietal junction (d/vTPJ), medial frontal cortex (MFC), precuneus/posterior cingulate cortex (PRE), superior temporal sulcus (STS), and temporal poles (TP). Adapted from Kanske et al. (2015)

social understanding that simultaneously require affect sharing and inferring others' mental states, the two networks can be clearly identified and mirror the results of studies investigating the functions separately (Lamm et al. 2011; Schurz et al. 2014). Assessing both functions simultaneously in the same individuals allows further investigation of the specifics by directly contrasting the activation related to empathy and Theory of Mind (see Fig. 1b for a schematic illustration). These specific contrasts replicated the typical networks, but also allowed differentiation of neighboring but distinct activations, for example, in temporoparietal junction. Specifically, a more dorsal anterior peak in the temporoparietal junction was associated with empathic responding, while mentalizing was related to more ventral parts of this region. Further supporting the distinction of empathy and Theory of Mind, the behavioral indices of empathy selectively correlated with activity in the empathy-related neural network, while Theory of Mind performance was only related to activity in the Theory of Mind network. Finally, the specific activation peaks for empathy and Theory of Mind were embedded in distinct task-free resting-state neural networks.

In sum, there is strong evidence for distinct neural routes underlying affective and cognitive understanding of others, even in situations that demand both functions concurrently. Open questions with regard to these two routes concern their interrelation, that is, whether and how activity in the distinct neural networks is orchestrated during online social understanding. For instance, some situations may call for a prioritization of one function over the other as has been demonstrated for cognitive control in emotional situations (Kanske et al. 2013), such as a physician who needs to infer whether a patient conceals a previous valium addiction, while treating the patient's painful injury. Another question on the interrelation of empathy and Theory of Mind is how the respective capacities are distributed inter-individually, that is, whether people with a particular proficiency in mentalizing are also strong empathizers.

5 Situational Influences on Empathy and Theory of Mind

In recent years, numerous situational modulators of empathic responding have been identified. It seems to be crucial for affect sharing, for instance, that attention is directed toward the painful aspects of a situation (Gu and Han 2007). When judging the painful consequences of situations, participants showed higher activation in empathy-related regions than when focusing attention on neutral aspects of the same stimulus (e.g., counting objects). Also, instructing participants to imagine being in the painful situation themselves clearly increased empathic responding on a neural level (Avenanti et al. 2005) and, accordingly, better perspective taking predicted enhanced early emotion detection (Kanske et al. 2013). A further top-down modulation concerns prior knowledge about the pain sensitivity of an observed person: Enhanced empathic responding was found when participants believed the other to actually feel a painful stimulation as compared to when they believed the other to be anesthetized (Lamm et al. 2007).

Interestingly, the social relation to the observed other also plays an important role for the strength of an empathic response. In particular, ingroup versus outgroup membership has been demonstrated to moderate empathic responding (Hein et al. 2010; Xu et al. 2009). Neural activation in anterior insula while observing others' pain has been found to be increased for racial ingroup as compared to outgroup members (Xu et al. 2009). Inviting fans of rivaling soccer teams, Hein et al. (2010) additionally showed that the anterior insula response predicted subsequent helping behavior for fans of the same soccer team (ingroup members). Besides group membership, it is also the sympathy toward another that shapes affect sharing. In a game show paradigm, participants showed activation of the ventral striatum in response to own gains, but also to gains of sympathetic versus unsympathetic others (representing an example of empathy for others' positive emotions) (Mobbs et al. 2009). An indication for the influence of interpersonal behavior on subsequent empathizing stems from evidence linking the fairness of an observed other in a previous interaction to the strength of empathic anterior insula activity (Singer et al. 2006).

A recent set of studies has probed how empathic responding is influenced by the observer's own emotional state (Silani et al. 2013). Incongruent affective stimulation of observer and observed (e.g., one receiving pleasant and the other receiving unpleasant touch) biases the empathic judgments of the observer toward the valence of his own stimulation (emotional egocentricity bias). Overcoming this bias is related to activity in an anterior/dorsal part of the temporoparietal junction that closely corresponds with the specific empathy peak that was identified in the previously described EmpaToM task (Kanske et al. 2015). Hence, the anterior/dorsal temporoparietal junction may play a role in separating one's own emotional state from the emotional state of the other, a function that seems crucial for many situations (and paradigms) involving empathic responding.

While the situational moderators of empathy have been studied in different domains, less is known about how the propensity and capacity to take another's perspective is shaped by contextual modulators. One critical factor for successful Theory of Mind performance seems to be the current psychosocial stress levels. Smeets et al. (2009) demonstrated that higher individual cortisol responses to the Trier Social Stress Test impair mentalizing in women, but enhance mentalizing in men. In developmental research, studies have isolated a number of sociodemographic and educational factors that predict Theory of Mind performance in children. In addition to low socioeconomic status, a parenting style relying strongly on power assertion (e.g., physical punishment) predicts worse performance in false-belief tasks, while communication focused parenting styles enhances childrens' capacity to understand others' perspectives (Cutting and Dunn 1999; Pears and Moses 2003).

6 Personal Influences on Empathy and Theory of Mind

In addition to situational influences on empathy and Theory of Mind, both functions are shaped by inter-individual differences in the personality and clinical domain. Gender differences are especially prominent in the literature, suggesting that women are more empathic and more prone to mentalize than men. The fact that gender effects have been mainly observed in self-reports (Rueckert and Naybar 2008) has been taken as an indication for demand characteristics in a sense that women feel they are expected to be more socially sensitive (Rueckert and Naybar 2008).

However, some studies have also reported behavior-based and neural activation differences, showing for instance that empathy in women is less susceptible to previous unfairness of the observed other (Singer et al. 2006) and that emotion recognition activates mirror regions to a stronger degree in women (Schulte-Ruther et al. 2008).

While deficits in both empathy and Theory of Mind have been described for a number of different mental disorders including depression (Wolkenstein et al. 2011) and schizophrenia (Mohnke et al. 2014), claims for selective deficits in one or the other function have also been made (Blair 2008). Theory of Mind deficits are a core component of autism spectrum disorder. Already during development, children on the autism spectrum pass classical false-belief tasks later than healthy children or than clinical control groups, even when controlling for language abilities (Baron-Cohen et al. 1985). Using ecologically valid and complex assessments of Theory of Mind, more recent studies show deficits also in high-functioning adult patients with autism spectrum disorder (Dziobek et al. 2006). Suggesting compensatory processes in high-functioning individuals, neuroimaging studies demonstrated hyperactivation during mentalizing in Theory of Mind-related regions (Mason et al. 2008). Interestingly, empathic responding seems preserved in autism when controlling for alexithymia (a trait characterized by difficulties in describing own emotional states (Bird et al. 2010). Further supporting this dissociation, the cortical structures subserving empathy do not differ between neurotypical and autistic individuals, whereas the structural network subserving Theory of Mind is hampered in autism (Bernhardt et al. 2014). The opposite pattern of empathy and Theory of Mind functioning has been described in psychopathy. While performance on cognitive and affective aspects of Theory of Mind seem to be intact in psychopaths (Blair et al. 1996), spontaneous empathic responding is largely reduced (Meffert et al. 2013).

7 What Now? Conclusions and Future Directions

The reviewed evidence strongly suggests the existence of two separable routes to understanding others. Based on the conceptualizations and underlying neural networks as well as on situational and personal moderators, empathy and Theory of Mind can be distinguished. In particular with regard to specific impairments of the functions in psychopathology, the question arises how (differentially) malleable the capacities to empathize and mentalize are. Therefore, one of the core tasks for future research will be to establish and evaluate targeted interventions aiming, for instance, at enhancing perspective taking skills in autism spectrum disorder and at improving the ability to share others' feelings in psychopathy.

Recent research has provided first indications for plasticity of empathy and compassion capabilities in healthy individuals (Klimecki et al. 2014). Participants underwent two one-week meditation-based trainings emphasizing either empathic or compassionate responding to other people's suffering. Neuroimaging results demonstrated a specific activation increase of anterior insula and middle anterior cingulate cortex after empathy training and of ventral striatum and medial orbito-frontal cortex after training compassion. This finding was mirrored by behavioral results pointing toward higher sharing of negative affect after empathy training and enhanced reports of positive affect after the compassion training. Plasticity of social understanding has also been revealed in the domain of Theory of Mind. For example, children with autism spectrum disorder showed specifically improved mentalizing performance after a 16-week Theory of Mind training (Begeer et al. 2011).

A hitherto unanswered question concerns the specificity of such interventions when applied in the same population. Hence, do trainings of the affective route to understanding others selectively enhance behavioral and neural markers of empathy while not influencing Theory of Mind performance, and vice versa? Or can shared mechanisms of social cognition be identified that enhance empathic responding even after specific trainings of mentalizing abilities? This and similar questions could be addressed with paradigms such as the EmpaToM task described earlier in this chapter since they manipulate and assess both functions of social cognition simultaneously.

We believe that the evidence outlined in this chapter demonstrates that research on the mechanisms of social understanding has moved beyond its infancy, revealing insight into the processes that are at work, for instance, during our family breakfasts and while we collaborate with our colleagues. The next steps must entail assessing the interplay of these affective and cognitive processes in controlled as well as in more complex interactive situations and probing their respective contributions to different types of social encounters, ranging from simple action coordination to large-scale cooperation.

References

Aichhorn M, Perner J, Weiss B, Kronbichler M, Staffen W, Ladurner G (2009) Temporo-parietal junction activity in theory-of-mind tasks: falseness, beliefs, or attention. J Cogn Neurosci 21 (6):1179–1192. doi:10.1162/jocn.2009.21082

Apperly IA (2012) What is “theory of mind”? Concepts, cognitive processes and individual differences. Q J Exp Psychol (Hove) 65(5):825–839. doi:10.1080/17470218.2012.676055

Avenanti A, Bueti D, Galati G, Aglioti SM (2005) Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. Nat Neurosci 8(7):955–960. doi:10.1038/nn1481

Baron-Cohen S, Leslie AM, Frith U (1985) Does the autistic-child have a theory of mind. Cognition 21(1):37–46. doi:10.1016/0010-0277(85)90022-8

Batson CD (2011) Altruism in humans. Oxford University Press, New York

Begeer S, Gevers C, Clifford P, Verhoeve M, Kat K, Hoddenbach E, Boer F (2011) Theory of mind training in children with autism: a randomized controlled trial. J Autism Dev Disord 41 (8):997–1006. doi:10.1007/s10803-010-1121-9

Bernhardt BC, Valk SL, Silani G, Bird G, Frith U, Singer T (2014) Selective disruption of sociocognitive structural brain networks in autism and alexithymia. Cereb Cortex 24(12):3258–3267. doi:10.1093/cercor/bht182

Bird G, Silani G, Brindley R, White S, Frith U, Singer T (2010) Empathic brain responses in insula are modulated by levels of alexithymia but not autism. Brain 133(Pt 5):1515–1525. doi:10.1093/brain/awq060

Blair RJ (2008) Fine cuts of empathy and the amygdala: dissociable deficits in psychopathy and autism. Q J Exp Psychol (Hove) 61(1):157–170. doi:10.1080/17470210701508855

Blair RJ, Sellars C, Strickland I, Clark F, Williams A, Smith M, Jones L (1996) Theory of mind in the psychopath. J Forensic Psychiatry 7(1):15–25. doi:10.1080/09585189608409914

Blakemore SJ, Sarfati Y, Bazin N, Decety J (2003) The detection of intentional contingencies in simple animations in patients with delusions of persecution. Psychol Med 33(8):1433–1441

Bockler A, Zwickel J (2013) Influences of spontaneous perspective taking on spatial and identity processing of faces. Soc Cogn Affect Neurosci 8(7):735–740. doi:10.1093/scan/nss061

Bruneau EG, Pluta A, Saxe R (2012) Distinct roles of the 'shared pain' and 'theory of mind' networks in processing others' emotional suffering. Neuropsychologia 50(2):219–231. doi:10.1016/j.neuropsychologia.2011.11.008

Bzdok D, Schilbach L, Vogeley K, Schneider K, Laird AR, Langner R, Eickhoff SB (2012) Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Struct Funct 217(4):783–796. doi:10.1007/s00429-012-0380-y

Caputi M, Lecce S, Pagnin A, Banerjee R (2012) Longitudinal effects of theory of mind on later peer relations: the role of prosocial behavior. Dev Psychol 48(1):257–270. doi:10.1037/a0025402

Carr L, Iacoboni M, Dubeau MC, Mazziotta JC, Lenzi GL (2003) Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. Proc Natl Acad Sci USA 100(9):5497–5502. doi:10.1073/pnas.0935845100

Cutting AL, Dunn J (1999) Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. Child Dev 70(4):853–865. doi:10.1111/1467-8624.00061

de Vignemont F, Singer T (2006) The empathic brain: how, when and why? Trends Cogn Sci 10 (10):435–441. doi:10.1016/j.tics.2006.08.008

Decety J (2010) To what extent is the experience of empathy mediated by shared neural circuits? Emot Rev 2(3):204–207. doi:10.1177/1754073910361981

Dodell-Feder D, Koster-Hale J, Bedny M, Saxe R (2011) fMRI item analysis in a theory of mind task. Neuroimage 55(2):705–712. doi:10.1016/j.neuroimage.2010.12.040

Dziobek I, Fleck S, Kalbe E, Rogers K, Hassenstab J, Brand M, Convit A (2006) Introducing MASC: a movie for the assessment of social cognition. J Autism Dev Disord 36(5):623–636. doi:10.1007/s10803-006-0107-0

Dziobek I, Preissler S, Grozdanovic Z, Heuser I, Heekeren HR, Roepke S (2011) Neuronal correlates of altered empathy and social cognition in borderline personality disorder. Neuroimage 57(2):539–548. doi:10.1016/j.neuroimage.2011.05.005

Fan Y, Duncan NW, de Greck M, Northoff G (2011) Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. Neurosci Biobehav Rev 35(3):903–911. doi:10.1016/j.neubiorev.2010.10.009

Frith C, Frith U (2005) Theory of mind. Curr Biol 15(17):R644–R646. doi:10.1016/j.cub.2005.08.041

Gu X, Han S (2007) Attention and reality constraints on the neural processes of empathy for pain. Neuroimage 36(1):256–267. doi:10.1016/j.neuroimage.2007.02.025

Hein G, Silani G, Preuschoff K, Batson CD, Singer T (2010) Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. Neuron 68 (1):149–160

Kanske P, Heissler J, Schonfelder S, Forneck J, Wessa M (2013a) Neural correlates of emotional distractibility in bipolar disorder patients, unaffected relatives, and individuals with hypomanic personality. Am J Psychiatry 170(12):1487–1496. doi:10.1176/appi.ajp.2013.12081044

Kanske P, Schonfelder S, Wessa M (2013b) Emotional modulation of the attentional blink and the relation to interpersonal reactivity. Front Hum Neurosci 7:641. doi:10.3389/fnhum.2013.00641

Kanske P, Böckler A, Trautwein FM, Singer T (2015) Dissecting the social brain: introducing the EmpaToM to separate empathy and theory of mind in brain and behavior. Neuroimage 122:6–19

Keysers C, Wicker B, Gazzola V, Anton JL, Fogassi L, Gallese V (2004) A touching sight: SII/PV activation during the observation and experience of touch. Neuron 42(2):335–346

Klimecki OM, Leiberg S, Lamm C, Singer T (2013) Functional neural plasticity and associated changes in positive affect after compassion training. Cereb Cortex 23(7):1552–1561. doi:10.1093/cercor/bhs142

Klimecki OM, Leiberg S, Ricard M, Singer T (2014) Differential pattern of functional brain plasticity after compassion and empathy training. Soc Cogn Affect Neurosci 9(6):873–879. doi:10.1093/scan/nst060

Lamm C, Nusbaum HC, Meltzoff AN, Decety J (2007) What are you feeling? Using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. PLoS ONE 2(12):e1292. doi:10.1371/journal.pone.0001292

Lamm C, Decety J, Singer T (2011) Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. Neuroimage 54 (3):2492–2502. doi:10.1016/j.neuroimage.2010.10.014

Leiberg S, Klimecki O, Singer T (2011) Short-term compassion training increases prosocial behavior in a newly developed prosocial game. PLoS ONE 6(3):e17798. doi:10.1371/journal.pone.0017798

Mason RA, Williams DL, Kana RK, Minshew N, Just MA (2008) Theory of mind disruption and recruitment of the right hemisphere during narrative comprehension in autism. Neuropsychologia 46(1):269–280. doi:10.1016/j.neuropsychologia.2007.07.018

Masten CL, Morelli SA, Eisenberger NI (2011) An fMRI investigation of empathy for 'social pain' and subsequent prosocial behavior. Neuroimage 55(1):381–388. doi:10.1016/j.neuroimage.2010.11.060

Meffert H, Gazzola V, den Boer JA, Bartels AAJ, Keysers C (2013) Reduced spontaneous but relatively normal deliberate vicarious representations in psychopathy, vol 136

Mobbs D, Yu R, Meyer M, Passamonti L, Seymour B, Calder AJ, Dalgleish T (2009) A key role for similarity in vicarious reward. Science 324(5929):900. doi:10.1126/science.1170539

Mohnke S, Erk S, Schnell K, Schutz C, Romanczuk-Seiferth N, Grimm O, Walter H (2014) Further evidence for the impact of a genome-wide-supported psychosis risk variant in ZNF804A on the theory of mind network. Neuropsychopharmacology 39(5):1196–1205. doi:10.1038/npp.2013.321

Muller-Pinzler L, Gazzola V, Keysers C, Sommer J, Jansen A, Frassle S, Krach S (2015) Neural pathways of embarrassment and their modulation by social anxiety. Neuroimage 119:252–261. doi:10.1016/j.neuroimage.2015.06.036

Paulus FM, Muller-Pinzler L, Westermann S, Krach S (2013) On the distinction of empathic and vicarious emotions. Front Hum Neurosci 7:196. doi:10.3389/fnhum.2013.00196

Pears KC, Moses LJ (2003) Demographics, parenting, and theory of mind in preschool children. Soc Dev 12(1):1–20. doi:10.1111/1467-9507.00219

Premack D, Woodruff G (1978) Does the chimpanzee have a theory of mind? Behav Brain Sci 515–526

Rueckert L, Naybar N (2008) Gender differences in empathy: the role of the right hemisphere. Brain Cogn 67(2):162–167. doi:10.1016/j.bandc.2008.01.002

Schulte-Ruther M, Markowitsch HJ, Shah NJ, Fink GR, Piefke M (2008) Gender differences in brain networks supporting empathy. Neuroimage 42(1):393–403. doi:10.1016/j.neuroimage.2008.04.180

Schurz M, Radua J, Aichhorn M, Richlan F, Perner J (2014) Fractionating theory of mind: a meta-analysis of functional brain imaging studies. Neurosci Biobehav Rev 42C:9–34. doi:10.1016/j.neubiorev.2014.01.009

Shamay-Tsoory SG, Aharon-Peretz J (2007) Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. Neuropsychologia 45(13):3054–3067. doi:10.1016/j.neuropsychologia.2007.05.021

Silani G, Lamm C, Ruff CC, Singer T (2013) Right supramarginal gyrus is crucial to overcome emotional egocentricity bias in social judgments. J Neurosci 33(39):15466–15476. doi:10.1523/JNEUROSCI.1488-13.2013

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(6):855–863. doi:10.1016/j.neubiorev.2006.06.011

Singer T, Lamm C (2009) The social neuroscience of empathy. Ann NY Acad Sci 1156:81–96. doi:10.1111/j.1749-6632.2009.04418.x

Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CD (2004) Empathy for pain involves the affective but not sensory components of pain. Science 303(5661):1157–1162. doi:10.1126/science.1093535

Singer T, Seymour B, O'Doherty JP, Stephan KE, Dolan RJ, Frith CD (2006) Empathic neural responses are modulated by the perceived fairness of others. Nature 439(7075):466–469. doi:10.1038/nature04271

Smeets T, Dziobek I, Wolf OT (2009) Social cognition under stress: differential effects of stress-induced cortisol elevations in healthy young men and women. Horm Behav 55(4):507–513. doi:10.1016/j.yhbeh.2009.01.011

Walter H, Adenzato M, Ciaramidaro A, Enrici I, Pia L, Bara BG (2004) Understanding intentions in social interaction: the role of the anterior paracingulate cortex. J Cogn Neurosci 16 (10):1854–1863. doi:10.1162/0898929042947838

Wicker B, Keysers C, Plailly J, Royet JP, Gallese V, Rizzolatti G (2003) Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. Neuron 40(3):655–664

Wimmer H, Perner J (1983) Beliefs about beliefs—representation and constraining function of wrong beliefs in young childrens understanding of deception. Cognition 13(1):103–128. doi:10.1016/0010-0277(83)90004-5

Wolf I, Dziobek I, Heekeren HR (2010) Neural correlates of social cognition in naturalistic settings: a model-free analysis approach. Neuroimage 49(1):894–904

Wolkenstein L, Schönenberg M, Schirm E, Hautzinger M (2011) I can see what you feel, but I can't deal with it: impaired theory of mind in depression. J Affect Disord 132(1–2):104–111. doi:10.1016/j.jad.2011.02.010

Xu X, Zuo X, Wang X, Han S (2009) Do you feel my pain? Racial group membership modulates empathic neural responses. J Neurosci 29(26):8525–8529. doi:10.1523/JNEUROSCI.2418-09.2009