Aus dem Zentralinstitut für Seelische Gesundheit
Abteilung Klinische Psychologie
Abteilungsleiter: Prof. Dr. phil. Peter Kirsch

Neural mechanisms of social cognition – the mirror neuron system and beyond

Inauguraldissertation
zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)
der
Medizinischen Fakultät Mannheim
der Ruprecht-Karls-Universität
zu
Heidelberg

vorgelegt von
Stephanie Nicole Lyn Schmidt

aus
Bad Soden am Taunus
2019

Dekan: Herr Prof. Dr. med. Sergij Goerdt

Referent: Herr Prof. Dr. phil. Peter Kirsch

TABLE OF CONTENTS

Page

LIST OF ABBREVIATIONS

ACC	Anterior Cingulate Cortex
AI	Anterior Insula
BA	Brodmann Area
c-system	Reflective System
DLPFC	Dorsolateral Prefrontal Cortex
FG	Fusiform Gyrus
fMRI	Functional Magnetic Resonance Imaging
IFG	Inferior Frontal Gyrus
IPL	Inferior Parietal Lobe
JTC	Jumping-To-Conclusions
MN	Mirror Neurons
MNS	Mirror Neuron System
mPFC	Medial Prefrontal Cortex
MVPA	Multivoxel Pattern Analysis
Nacc	Nucleus Accumbens
PFC	Prefrontal Cortex
PMC	Premotor Cortex
ROI	Region Of Interest
SMA	Supplementary Motor Area
STS	Superior Temporal Sulcus
sVx	Shared Voxels
TPJ	Temporoparietal Junction
TMS	Transcranial Magnetic Stimulation
ToM	Theory of Mind
vmPFC	Ventromedial Prefrontal Cortex
x-system	Reflexive System

Imagine your life and subtract every social interaction you ever had.
What is left?

1 INTRODUCTION

You are living proof of your own and your caregiver's social-cognitive skills. As babies cannot speak or feed themselves, non-verbal communication is vital. This is as true today as it has been in the past. Now that you are grown-up and, in contrast to our long-ago ancestors, living in a civilized world, you do not need other people to survive. And still, being rejected can be just as hurtful as physical pain (Eisenberger & Lieberman, 2004).

When thinking about our role in the world, we often think intelligence is the one outstanding factor differentiating us from other species. However, what might have given us the greatest evolutionary advantage and what might have fundamentally influenced who we are and how we live today is indeed our superior social cognition, which allowed us to find our ecological niche and grow our brains. Studies have shown strong correlations between brain size and social network size across and within species (Robin IM Dunbar, 1998; R. I. Dunbar, 2009), and nowadays this relationship even holds true for our social networks on the internet (Kanai, Bahrami, Roylance, & Rees, 2012).

Apparently, social interactions are critical for our lives, as is breathing and eating. Yet, while the exact mechanisms involved in breathing and eating are known, the seemingly simplest processes of social cognition are only poorly understood. Consequently, the main question for my doctoral thesis is one that also bothered other researchers for many decades: How exactly do our brains perform social cognition? My main focus will be on mirror neurons which cannot be directly measured in humans, but have to be assessed by indirect methods. Due to this challenge, the study of mirror neurons encouraged me, like other researchers, to also look into non-standard procedures, as will become evident throughout this thesis.

In study 1, I investigated by functional magnetic resonance imaging (fMRI) whether different social-cognitive processes share a common neural basis, and whether these are located in the putative mirror neuron system[1] (MNS). To find out whether the MNS

[1] I want to point out that *mirror neuron system* is not the most scientifically accurate term when referring to results obtained from neuroimaging studies, as we cannot make inferences about neurons. Some studies solve this by calling it the putative mirror neuron system. However, in this thesis, I choose to

is not only involved in different social-cognitive processes but also in the processing of different emotions, I conducted study 2 using an fMRI adaptation paradigm. Based on the assumption that the MNS alone is not sufficient in the case of ambiguous emotional faces, I conducted fMRI study 3 to investigate the involvement of the nucleus accumbens (Nacc) with its role in salience and reward for social decision-making.

In the introduction of my thesis, I will first explain the MNS, including the theory of embodied simulation which is particularly relevant to social cognition. After introducing different social-cognitive abilities on a conceptual level, I will present influential two-pathway models that attempt to explain social-cognitive functioning with respect to fast and automatic versus slow and deliberate processing. Using the context of the two-pathway models, I will explain neurobiological findings behind distinct processes central to social cognition, which are emotion perception, empathy and theory of mind. The last part of my introduction will be dedicated to the challenges researchers face when measuring the MNS. In the main part of this thesis, I will present studies 1 to 3. While studies 1 and 2 of my thesis are focused on the automatic processing associated with the MNS, study 3 additionally involves deliberate decision-making.

1.1 Mirror neurons

Mirror neurons "opened a window into the neural clockwork that allows us to understand other individuals".
(Keysers & Fadiga, 2008, p. 193)

More than 30 years ago, Italian researchers made an unexpected discovery that has substantially influenced our understanding of the mechanisms underlying social interactions. When studying the specialization of neurons to different types of grasping movements in macaques, di Pellegrino, Fadiga, Fogassi, Gallese and Rizzolatti (1992) discovered by chance that some neurons in the premotor cortex would fire not only when a macaque performed a specific movement, but also when the experimenter performed that movement and the macaque simply observed it. The existence of this phenomenon has been confirmed in many studies since (Caggiano et al., 2012; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Kraskov, Dancause, Quallo, Shepherd,

write about the "mirror system" when writing it out, and prefer to use the common abbreviation MNS, as I expect it allows for easier readability with MNS literature.

& Lemon, 2009), possibly even in birds (Prather, Peters, Nowicki, & Mooney, 2008), and these neurons are commonly known as mirror neurons (MN). MN in macaques have been shown to respond to action-specific, goal-oriented movements (Bonini et al., 2012; Bonini et al., 2011; Caggiano et al., 2012; Fogassi et al., 2005; for a review, see Ocampo & Kritikos, 2011; Rizzolatti & Craighero, 2004). They provide an explanation for motor learning by imitation (Rizzolatti & Craighero, 2004), which has been assumed to have contributed considerably to human evolutionary success (Dean, Kendal, Schapiro, Thierry, & Laland, 2012) as individuals can learn from others and build upon that knowledge.

In monkeys, MN are routinely measured using single-cell recordings. In humans however, this accurate but invasive technique can only be applied in patients undergoing brain surgery for medical reasons, so most MNS research was conducted using fMRI, which is non-invasive and has comparatively high spatial resolution. I will discuss the problems with assessing MN in more detail in section 1.7 of the introduction of this thesis and in my studies.

The first target area to identify mirror mechanisms in humans was the homologue structure to the macaque MN region, which is the pars opercularis of the inferior frontal gyrus (IFG). Studies in humans have confirmed that the IFG is activated both when observing someone else perform a specific action and when oneself performs the same action (Iacoboni et al., 1999; Kilner, Neal, Weiskopf, Friston, & Frith, 2009; Molnar-Szakacs, Iacoboni, Koski, & Mazziotta, 2005; Montgomery, Isenberg, & Haxby, 2007). Further studies indicated that the IFG, together with premotor cortex (PMC) and inferior parietal lobule (IPL), builds the core of the MNS (Cattaneo & Rizzolatti, 2009; Rizzolatti & Craighero, 2004). To date, only one publication reported the use of single-cell recordings in humans undergoing epilepsy surgery. It confirmed the existence of MN in various brain regions, including supplementary motor area (SMA), hippocampus, parahippocampal gyrus, and entorhinal cortex (Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010). Finally, a meta-analysis on the basis of 125 fMRI studies (Molenberghs, Cunnington, & Mattingley, 2012) revealed a large set of areas with reported mirror properties, including IFG and the adjacent ventral PMC, IPL, primary visual cortex, cerebellum, and parts of the limbic system (see Figure 1).

However, most intriguing about the MNS is that it might not only be relevant for the understanding of motor actions, but to be the key to social understanding in

everyday life (Gallese, 2007a), as I will explain in the next paragraph after introducing the umbrella term social cognition.

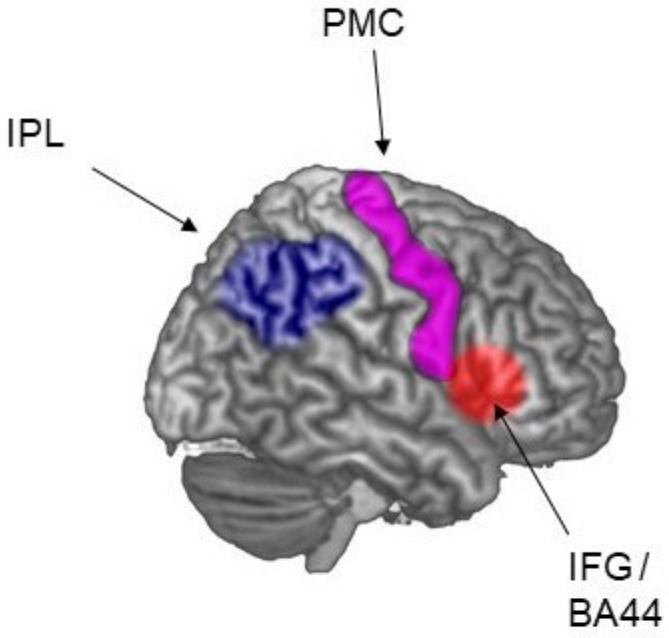


Figure 1: Brain regions associated with the MNS according to Molenberghs et al., 2012. IFG= Inferior Frontal Gyrus, IPL = Inferior Parietal Lobe, PMC = Premotor Cortex. Note: Only the regions identified in the meta-analysis that are thought to belong to the core MNS network (Cattaneo & Rizzolatti, 2009; Rizzolatti & Craighero, 2004) are depicted in this figure.

1.2 Social cognition and embodied simulation

Social cognition is an umbrella term, comprising many different psychological processes that allow us to understand our own as well as other individuals' emotional and mental states (Beer & Ochsner, 2006; Brothers, 2002). We can infer meaning from many different social signals, including gaze direction (Hamilton, 2016), facial configurations (Montagne, Kessels, Frigerio, de Haan, & Perrett, 2005), body language (Atkinson, Dittrich, Gemmell, & Young, 2004), and vocal prosody (Pereira, 2000). On top of that, in social interaction, as the word *interaction* indicates, one does not merely perceive these signals, but also react to the perceived ones and send out own signals.

Quickly, one gets an impression of the conversational partner, and while it seems to be an overwhelming amount of information when explicitly thinking about it, the information is processed automatically and with ease, most of the time without any awareness. To accomplish this, our brains rely on several basic mechanisms. First, attention needs to be directed to the faces and relevant facial features, which occurs automatically in healthy humans. In particular the eye area, but also the mouth, are majorly responsible for transmitting emotion information (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Calvo & Nummenmaa, 2008). This processing advantage for faces, and in particular for the eyes, has been shown to be present even in young children (Taylor, Edmonds, McCarthy, & Allison, 2001). Once the brain is attentive to the relevant signals, the actual social-cognitive processing takes place. However, there has been an ongoing debate on how this is accomplished. Many areas of visual perception are hardwired, such as color perception, for which we have different types of receptors, or simple shape orientations, which are represented by neurons in the visual cortex . Indeed, some researchers have suggested that perceiving an emotion, or more generally inferring a mental state, happens equally automatically, via MN (Gallese, 2003b, 2007a, 2007b; Gallese & Goldman, 1998; Keysers & Gazzola, 2009). This line of thought gave rise to the theory of embodied simulation, which states that we understand other persons' mental states because our brain activation is the same when we observe someone expressing an emotion, as when we experience that emotion ourselves. Obviously, this would solve the long existing mystery of how seemingly automatic social cognition actually works, and indeed, studies have since confirmed that processing or imitation of faces is also related to activation in the MNS (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003; Enticott, Johnston, Herring, Hoy, & Fitzgerald, 2008; Schulte-Rüther, Markowitsch, Fink, & Piefke, 2007). Shared brain activation for own and observed mental states has been reported for several emotions or actions. For example, own and observed positive facial affect have been associated with shared activation in insula, IFG and PMC (Hennenlotter et al., 2005).

Social cognition however does not end at a brain's automatic simulation of what others feel. In reaction to the perceived feelings of another person, additional emotions and thoughts arise in ourselves. We may react with empathy (see section 1.3.1), infer our conversational partners hidden intentions (see section 1.3.2), or even develop the urge to help. Whether different social-cognitive processes indeed rely on common activation in the MNS will be subject of study 1.

In the next subsections, I will first introduce the different subprocesses of social cognition on a conceptual level. After explaining the two-pathway models, I will present findings on the neural correlates associated with the distinct social-cognitive processes and connect them to the two-pathway models.

1.3 Introduction of social-cognitive subprocesses

Humans possess the ability to perceive emotional content via several modalities, including vocal intonation, body posture and facial configurations. My focus was on the latter, so for all my studies, I used facial configurations intended to express specific emotions. While one can imagine a possibly endless number of facial configurations and specific characteristics of emotional states, emotions are usually described either categorically or dimensionally. Categories commonly used include anger, fear, sadness and happiness (P. Ekman, 1992; Tracy & Randles, 2011), and the dimensional approach mainly builds on the central features valence, ranging from positive to negative, and arousal, ranging from low to high (Mehu & Scherer, 2015; Ortony & Turner, 1990).

For each emotional state we infer, there usually is someone who expresses and possibly feels an emotion. Of course, we are not merely inferring others' emotions, we are also reacting to them. On the one hand, we may feel with the other person, which is known as empathy. If the feeling of the other person's negative state becomes intensive and the focus is shifted from the other to the self, this would be called distress (for a detailed explanation of empathy, please refer to the next paragraph 1.2.2.1).

While many studies employ obvious stimuli, involving body parts in painful situations, to elicit empathy, similar processes are assumed to apply for subtler stimuli such as facial configurations that are perceived as being emotional. Besides empathy, also theory of mind (ToM) is considered a central social-cognitive ability. ToM requires the perceiver to take another individual's perspective and reason from that point of view, which may oppose own beliefs and perceptions (for a detailed explanation of ToM, please refer to 1.5.3).

1.3.1 Empathy

Empathy represents a central component of social interactions. Yet, even among researchers investigating empathy, there is no consensus on a definition, as exemplified in a literature review by Cuff and colleagues (2016) who identified 43 distinct definitions of empathy. These differences are also reflected in the wide range of tasks used to investigate empathy, impeding comparability between studies. As Preston and de Waal (2002) argue, this diversity arises from the problem that the mechanism of empathy is not truly understood. “Despite the various definitions of empathy, there is broad agreement on three primary components: (a) an affective response to another person, which often, but not always, entails sharing that person’s emotional state; (b) a cognitive capacity to take the perspective of the other person, and (c) emotion regulation” (Decety and Jackson, 2006, p. 54). Importantly, most researchers agree on the distinction of cognitive and affective empathy (Cuff, Brown, Taylor, & Howat, 2016), which is elaborately described in a publication by Walter (2012). He defines affective empathy as an affective state which is elicited by the assumed or inferred affective state of another person. This induced affective state is similar to that of the other individual and oriented towards them. Furthermore, the observer is aware of the causal relation of his or her own and the other’s affective state, including self-other distinction. Thus, affective empathy must be differentiated from emotional contagion, which is the adopting of another person’s emotions without clear self-other distinction. Cognitive empathy on the contrary does not necessitate an affective state in the observer, but only a cognitive understanding of a perceived affective state of another person. It is therefore more closely related to emotion perception or ToM (for details on ToM, please refer to section 1.5.3) which comprises the cognitive understanding not only of affective states but of mental states in general, extending to beliefs, intentions and desires (Walter, 2012). If the cognitive reasoning is directed towards another person’s affective state, this aspect of cognitive empathy can also be referred to as affective ToM. Cognitive ToM, in contrast, entails the cognitive reasoning process focusing on another person’s cognitive states (Walter, 2012).

1.3.2 Theory of Mind

ToM, often used synonymously with mentalizing, refers to our attribution of mental states, including intentions and desires, to other people (U. Frith & Frith, 2001). It has even been proposed to be the crucial mechanism setting us apart from other primates and central to advanced human abilities such as cooperation (Gallagher and Frith, 2003).

Traditionally, mentalizing ability or ToM is assessed using a false-belief paradigm. First developed by Wimmer and Perner (1983), the now most famous adaptation is known as Sally-Anne task (Baron-Cohen, Leslie, & Frith, 1985): Sally places a marble in a basket and leaves the room. Anne takes the marble and puts it in a box. The question is, where will Sally look for the marble when she comes back? Healthy adults know that she will look in the basket, where she had placed it, because she does not know that Anne put it somewhere else. So, even though we know the true location of the object, information that is salient in our thoughts, we can inhibit this knowledge and instead adopt Sally's perspective and reason about her wrong belief, whereas children younger than 3 to 4 years of age are not able to adapt Sally's perspective (A. M. Leslie, Friedman, & German, 2004).

Before presenting the neurobiological correlates of these social-cognitive processes, I want to introduce important conceptual models that go beyond individual brain regions or networks, but focus on the general processing of social information, thereby supporting the understanding of the mechanisms of social cognition.

1.4 Two-pathway models of social cognition

As is the case for many psychological processes, also for social cognition, two routes of information processing have been suggested. While the proposed names and details differ between scientists, the two pathways can be summarized as one being fast and automatic, and the other being slow and deliberate.

In the following sections, I will focus on three concepts: The first two, distinguishing implicit versus explicit processing mechanisms (e.g., C. D. Frith & Frith, 2008), or the so called c- and x-system (Satpute & Lieberman, 2006) are mainly adopted from cognitive psychology. The third has its foundation in (social-)cognitive

neuroscience and concerns the distinction of mirroring and mentalizing (Becchio et al., 2012; Van Overwalle & Baetens, 2009).

1.4.1 Implicit versus explicit processing

Cognitive processes can generally be explicit or implicit. Implicit processes are fast, inflexible, automatic and mostly unconscious, whereas explicit ones are slow, flexible and mentally effortful (C. D. Frith & Frith, 2008).

Interestingly, implicity versus explicity may not necessarily be intrinsic to a specific social-cognitive process, but both can occur in the same task, as will become evident in the following example for ToM. In the Sally-Anne task, the famous task introduced in section 1.3.2 that requires understanding of the false belief of a protagonist, children younger than about 4 years and individuals with autism will give the wrong answer (Baron-Cohen et al., 1985; Wimmer & Perner, 1983). Remarkably, non-verbal variations of the task indicate that also infants attribute the right beliefs to the protagonist (Buttelmann, Carpenter, & Tomasello, 2009; Onishi & Baillargeon, 2005). This contradictory finding has been explained by the difference of implicit and explicit ToM. For example, it has been found that very young children will (explicitly) tell the experimenter the wrong answer, but (implicitly) look to the correct location. However once they are older, children can hold explicit knowledge about the false belief and can articulate that accordingly (Clements & Perner, 1994). Interestingly, also in a large study on adults, explicit mentalizing during ToM questions was associated with activation in the same neural network that was implicitly activated when watching associated ToM videos (Kanske, Böckler, Trautwein, & Singer, 2015). The authors suggest that this finding supports the assumption that mentalizing may occur spontaneously and implicitly. Furthermore, the overlap of brain regions might reflect a close relationship of implicit and explicit mentalizing in healthy individuals. Individuals with autism, in contrast, seem to lack the implicit understanding, but can acquire the correct (explicit) reasoning for false believe situations (U. Frith, 2004). Generally, it is assumed that the ability to infer emotions, desires and intentions develops before the reasoning about beliefs (Saxe, Carey, & Kanwisher, 2004).

To summarize, in this section, using the example of ToM, it was shown that social-cognitive processes can occur implicitly and explicitly, a distinction that is not necessarily intrinsic to tasks, but may depend on other factors such as the

developmental stage of the individual. Another factor that is also central to the concept of implicit versus explicit processing, also plays a role in the next section: Speed. We know that explicit learning is fast, and implicit learning occurs slowly (McDougle, Bond, & Taylor, 2015). As we will see in the following, there are even certain sets of brain regions that are associated with either fast or slow processing, thereby building the basis of two systems proposed to underlie social cognition.

1.4.2 The x- and c-system by Satpute and Lieberman (2006)

In a seminal paper, Satpute and Lieberman (2006) propose a dual-process model of automaticity and control for social perception. The x-system (x for reflexive) is slow learning, fast operating, apt to parallel processing and does not need to be conscious. Due to its bidirectionality, it can process implicit semantic and evaluative symmetric relationships and it represents common cases. The phylogenetically likely younger c-system (c for reflective) is fast learning, slow operating and possesses symbolic computational ability, which helps to represent asymmetric relationships, exceptions and special cases. It is important for holding inferential goals in mind and integrating prior knowledge including situational constraint information; we often experience it as inner monologue and the feeling of agency (Satpute & Lieberman, 2006).

With regard to social cognition, Satpute and Lieberman (2006) propose as parts of the x-system, the amygdala, with its role in fear processing and its relationship to fight-and-flight responding, the basal ganglia, including Nacc involved in developing statistical models of the world and attaching emotional significance to them, ventromedial (vm) prefrontal cortex (PFC) which is strongly connected with basal ganglia, amygdala, and other limbic structures and included in the formation of intuition, and dorsal anterior cingulate cortex (ACC) which is involved in emotional distress. A further major region of the x-system, the lateral temporal cortex, comprises superior temporal sulcus (STS), temporal poles and lateral and inferior temporal lobes. The authors highlight the relevance of these regions for semantic and perceptual processes and their assumed involvement in constructing stereotypes, individual impressions and dispositional attributions. Specifically, they point out the key role of the STS in social cognition, its involvement in recognizing people, following eye-gaze, processing biological motion, understanding peoples' actions and goals, inferring intentions, ToM,

and perspective taking. As further support of the STS' placement in the x-system, the authors refer to studies showing that neuronal discharge in STS and behavior associated with STS function are very fast.

In contrast, the c-system (Satpute & Lieberman, 2006) is composed of lateral PFC, including dorsolateral PFC (DLPFC) which is involved in many demanding cognitive tasks, such as reasoning and logic, fluid intelligence, problem solving, emotion regulation and behavioral inhibition, the posterior parietal cortex, which is associated with working memory, reasoning, self-focused attention and perspective taking, ventral ACC, known for its role in the processing of emotional conflict and also the anticipation of pain, medial temporal lobe, which is involved in memory retrieval, and medial frontal cortex. Satpute and Lieberman (2006) suggest that processes like ToM could first rely on the c-system, but with more experience shift to the x-system.

To summarize, the x- and c-system are composed of defined brain regions and representative for different social-cognitive subprocesses. It is therefore a comprehensive concept, in which also the distinction of implicity and explicity could be integrated. Figure 2 illustrates areas of the x- and c-system that are relevant to this thesis. In addition, Table 1 (page 14) lists the brain regions and specifications of both systems. I will also refer to this distinction throughout the thesis, to integrate it with further research findings and also my own studies.

The theories on implicit versus explicit and x- versus c-system also fit the popular distinction of mirror versus mentalizing network, which I introduce as a third concept.

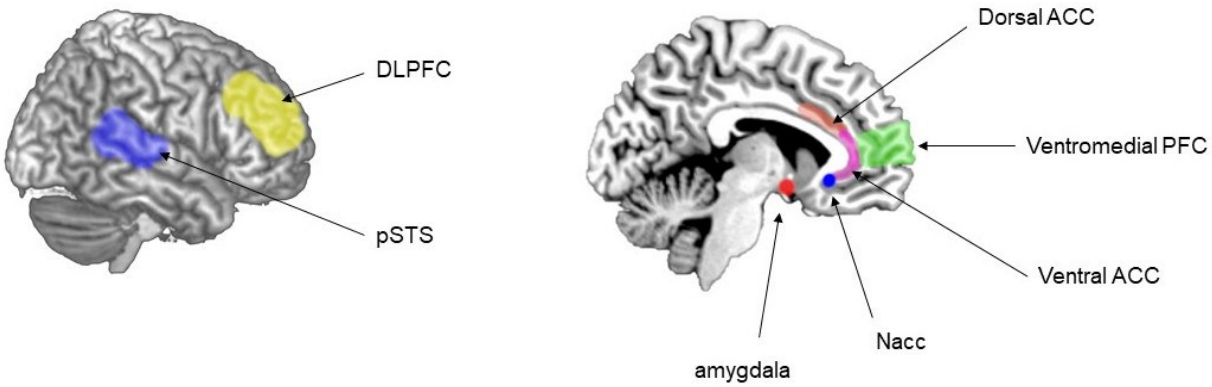


Figure 2: Illustration of brain regions that are involved in social cognition. These brain regions are mentioned in particular as part of the x- and c-system (1.4.2), and for empathy (1.5.2). ACC = Anterior Cingulate Cortex. DLPFC = Dorsolateral PFC. Nacc

= Nucleus accumbens which is part of the basal ganglia. PFC = Prefrontal Cortex. pSTS = posterior Superior Temporal Sulcus.

1.4.3 Mirroring versus mentalizing

While the MNS is a focus of this thesis and has been thoroughly introduced in 1.1, I will additionally summarize the most important aspects of another system important to social cognition, the mentalizing system. Mentalizing has been defined as "the capacity to understand ourselves and others in terms of intentional mental states, such as feelings, desires, wishes, attitudes and goals" (Luyten & Fonagy, 2015, p. 366). As explained in section 1.2, mirroring is assumed to occur automatically and intuitively and understanding of a mental state is based on a shared representation in our brain. Mentalizing, in contrast, is thought to require cognitive perspective taking and reasoning from this other point of view, which is likely based on previous experiences. Mentalizing is therefore a high-level social-cognitive skill, the development of which seems to be largely dependent on one's environment (for a literature overview, please refer to Luyten & Fonagy, 2015). In a meta-analysis, van Overwalle and Baetens (2009) summarize the regions and functions of the MNS and the mentalizing system. Specifically, the MNS is considered as centering around anterior intraparietal sulcus (sometimes referred to as rostral parietal lobule) and PMC (equivalent to BA44 and BA6). It responds rapidly, and mainly to observable motor actions that serve an obvious goal, matching motor representations in the observer. When the MNS fails because there is no appropriate representation of an action, for example when the action is abstract or unusual, such as opening the door using one's foot on the door handle, and also requires involvement of attentional focus, high-level processing is needed to make sense of the action goals. This is accomplished by the mentalizing system, consisting of temporoparietal junction (TPJ) and medial PFC (mPFC). While TPJ seems to be responsible for inferring temporary beliefs and intentions, mPFC is related to the deduction of stable trait characteristics (Van Overwalle, 2009). It is worth mentioning that this definition is not in complete agreement with the definition of Satpute and Lieberman (2006) who assumes that at least the ventral part of the mPFC contributes to the reflexive system. Critically, several studies associate the mentalizing system rather with posterior STS than the adjacent TPJ, and additionally include the temporal poles (U. Frith & Frith, 2003).

The meta-analysis by van Overwalle and Baetens (2009) also contributes to answering a question that has occupied researchers for a long time: What is the relationship between mirroring and mentalizing? Evidence suggests that the mirror system and the mentalizing system are never simultaneously active or dependent on one another, but rather complement each other (Van Overwalle & Baetens, 2009).

So, while the three two-pathway models are based on different ideas, their similarities are obvious, and in addition they complement each other thanks to their slightly different focus. For better understanding and as a reference, the brain regions associated with the MNS and the mentalizing system, as well as with the x- and the c-system are summarized in Table 1. Furthermore, the brain regions identified as regions of interest for studies 1, 2 and 3 are depicted in Figure 3 to allow the reader to keep the focus of my research in mind when reading the next part about the neurobiology of social cognition. These regions are IFG, IPL, amygdala, insula, STS and fusiform gyrus. The latter was not explicitly mentioned until now, but is part of the inferior occipital cortex and introduced in the following section.

Table 1: Overview over brain mechanisms according to the two pathway models with associated brain regions and literature references.

Mechanism	Brain Regions	Reference
x-system	ventromedial PFC, STS, temporal poles, lateral temporal lobes, inferior temporal lobes, dorsal ACC	Satpute & Lieberman (2006)
c-system	lateral PFC, medial frontal cortex, medial temporal lobe, posterior parietal cortex, ventral ACC, basal ganglia, amygdala	Satpute & Lieberman (2006)

Mirror neuron system	anterior intraparietal sulcus (or rostral parietal lobule), premotor cortex (composed of BA44 and BA6)	van Overwalle & Baetens (2009)
Mentalizing system	temporoparietal junction, medial PFC	van Overwalle (2009)
	posterior STS, temporal poles, medial PFC	Frith & Frith (2003)

Abbreviations: ACC = anterior cingulate cortex, PFC = prefrontal cortex, STS = superior temporal sulcus.

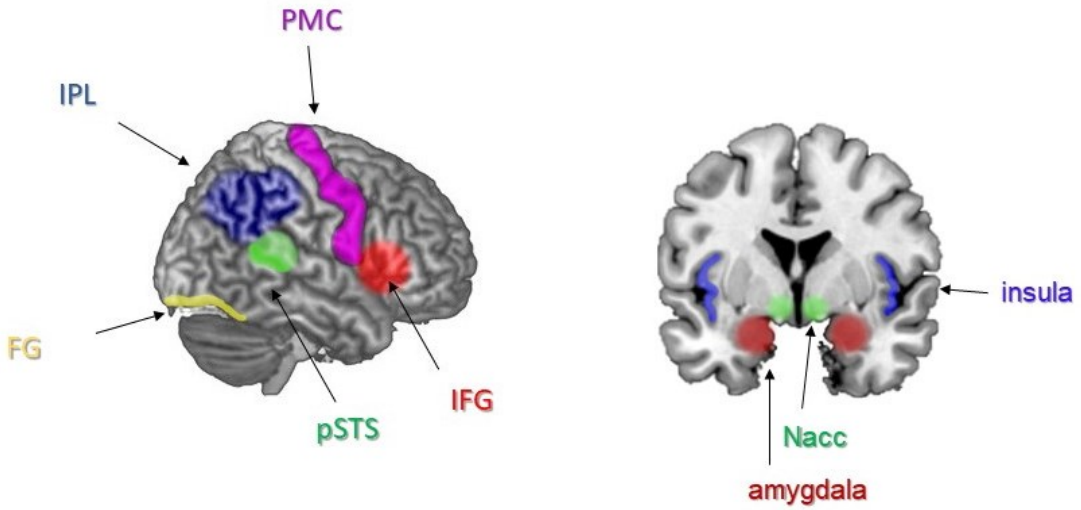


Figure 3: Exemplary illustration of the core regions of the MNS and the mentalizing system, providing a basis for the regions of interest in this PhD thesis. FG = Fusiform Gyrus. IFG = Inferior Frontal Gyrus. IPL = Inferior Parietal Lobule. Nacc = Nucleus Accumbens. PMC = Premotor Cortex. pSTS = posterior Superior Temporal Sulcus.

1.5 Neurobiology of social-cognitive processes

In the previous sections, I introduced the MNS with regard to social cognition in general and explained how processes of social cognition could be related to automatic versus deliberate processing. The next sections will provide more details on the neurobiology of the different subprocesses of social cognition. I will present some

important study results and discuss how they integrate with the two-pathway models, starting with findings on emotion perception, since this process will be central to all my studies.

1.5.1 Perception and expression of emotions

A vast amount of literature is based on a small set of emotional categories and the idea that each emotion is associated with a distinct facial configuration (e.g., Fusar-Poli et al., 2009). It is commonly assumed, that this association functions reliably and bidirectionally, meaning that each emotion has a specific facial configuration, and from each, we can infer a specific emotional state. Many studies support this assumption by showing that these categories exist in different countries and cultures throughout the world and even in infants (P. Ekman et al., 1987; Izard, 1994, but see also the review by Barrett et al., 2019), suggesting that they might be innate. The most influential categorization is the one by Ekman and colleagues (Paul Ekman, Friesen, & Ellsworth, 1972) who proposed the so called basic emotions: fear, anger, happiness, disgust, sadness and surprise. However, a recent publication (Barrett, Adolphs, Marsella, Martinez, & Pollak, 2019) critically points out a lack of evidence confirming that emotions are reliably and unambiguously expressed via facial configurations. They present studies showing that one emotion can be expressed with different facial configurations and that one facial configuration can be used to express different emotional states, dependent on the context. Furthermore, they argue, facial configurations can be interpreted differently, which is also majorly influenced by culture and therefore might be learned instead of innate. For the sake of scientific accuracy, I follow the authors' recommendation and use the term *facial configuration* instead of *emotional facial expression* throughout this thesis[2]. Importantly, and in agreement with Barrett and colleagues (2019), my studies' concept of emotions includes that emotions can be represented on a continuum of valence, going from negative to positive, and

[2] Barrett and colleagues (2019) also suggest using the term *emotion perception* or *emotion inference* instead of emotion recognition. While I agree with their reasoning and suggestion for many cases, in particular the natural inference of emotions in everyday situations, I think that *emotion recognition* can accurately be used to describe the act of determining an emotion that was expressed with the intention to represent one of the basic emotions. Most importantly, I will leave the disentangling of these terms to other studies, and use them interchangeably throughout this work, while being aware that what is *recognized* might not be the true emotion of the observed person, but only the usual categorization into an emotion category based on configurational features.

arousal, going from low to high intensity as proposed by the dimensional approach (Russell, 1980).

One brain region, that has been suggested central to the processing of emotions, including perception of others' and own emotions, is the amygdala (Costafreda, Brammer, David, & Fu, 2008; Critchley et al., 2000; Fitzgerald, Angstadt, Jelsone, Nathan, & Phan, 2006; Fusar-Poli et al., 2009; Gur et al., 2002; Habel et al., 2007; Sergerie, Chochol, & Armony, 2008). It might in fact be the amygdala's role in salience detection that makes it so central to emotion processing (Cunningham & Brosch, 2012; Liberzon, Phan, Decker, & Taylor, 2003; Santos, Mier, Kirsch, & Meyer-Lindenberg, 2011).

However, as emotion perception is a complex process, several additional regions are involved. Considering specifically the brain activation associated with facial stimuli, studies generally find increased activation in regions including the fusiform gyrus (FG), inferior occipital gyrus, amygdala, cingulate gyrus, medial frontal gyrus, middle frontal gyrus, precentral gyrus and insula for neutral faces (for a review and meta-analysis, see Fusar-Poli et al., 2009). In contrast to neutral faces, happy faces show even increased activation in amygdala, FG and ACC, sad faces in amygdala and lingual gyrus, fearful faces in amygdala, FG and medial frontal gyri, and angry faces in insula and inferior occipital gyrus (Fusar-Poli et al., 2009). In addition to the key role of the amygdala, these results demonstrate an association of different emotional categories with individual activation patterns. Importantly, the FG is associated with emotion perception and also essential for face processing (Geday, Gjedde, Boldsen, & Kupers, 2003; Haxby & Gobbini, 2011; Haxby, Hoffman, & Gobbini, 2000), which explains its presence in the majority of results from studies using facial stimuli.

To conclude, emotion perception seems to clearly involve regions of the fast x-system, as indicated by increased activation of amygdala, FG as part of the temporal lobe and medial frontal gyrus.

1.5.2 Empathy

As mentioned above, while empathy is investigated using a variety of definitions and tasks, one robust agreement among researchers is on the dissociation between cognitive and affective empathy, which is also supported by findings from personality and developmental disorders. For example, individuals with borderline personality

disorder give higher ratings for affective empathy but lower ratings for cognitive empathy than healthy individuals (Harari, Shamay-Tsoory, Ravid, & Levkovitz, 2010). In contrast, psychopathic traits are associated specifically with decreased affective empathy (Wai & Tiliopoulos, 2012). These findings, showing selective impairment in only one of the two empathic abilities, support the idea that cognitive and affective empathy are based on separate mechanisms and therefore likely involve different neural pathways.

To my knowledge, meta-analyses of fMRI activation during empathy exist only for empathy for pain studies, which represent a large subset of all empathy studies. I therefore report the results of an empathy for pain meta-analysis to provide an example of commonly found activation patterns during empathy. Lamm and colleagues (2011) suggested that while anterior insula (AI) and medial cingular cortex act as core regions of empathy, there is also paradigm-specific neural activation falling in two networks. One set of paradigms, referred to as *picture-based*, was based on pictures of body parts in painful situations. The corresponding network comprising supramarginal gyrus, inferior parietal cortex, and BA44, is overlapping with the regions commonly associated with the MNS. The other set of paradigms, described as *cue-based*, involved currently experienced pain by the participant or another person next to the participant. In contrast to the MNS activation in picture-based paradigms, cue-based paradigms activated regions associated with the mentalizing system, i.e. precuneus, medial PFC, posterior STS, TPJ, temporal poles. Functional connectivity studies support the existence of separate brain networks for cognitive versus affective empathy. While affective empathy was associated with stronger connectivity in social-emotional networks, centering around amygdala, orbitofrontal cortex, AI, ACC and temporal poles, cognitive empathy was related to stronger connectivity in social-cognitive and interoceptive networks, including STS, AI, brainstem and cerebellum (Cox et al., 2011).

Lesion studies reinforce the idea of a double dissociation, with lesions in the IFG being associated with deficits in affective empathy, and lesions in the vmPFC being related to impaired cognitive empathy (Shamay-Tsoory, Aharon-Peretz, & Perry, 2009). Another study also indicates that the IFG might be a key region in emotional empathy (Jabbi & Keysers, 2008). On top of that, the IFG is not only adjacent to but also functionally connected to the region of AI and frontal operculum, linking the IFG to the proposed core empathy network (Jabbi & Keysers, 2008).

While involvement of the IFG and SMA as part of PMC suggest a link to the MNS, involvement of dorsal ACC and vmPFC further suggests a dominant role of the x-system.

1.5.3 Theory of mind

ToM has been investigated using a wide variety of tasks (for a review, see Carrington & Bailey, 2009) which possibly involve different mechanisms, so the results of meta-analyses have to be interpreted with care. One meta-analysis on the most common ToM tasks, categorized into 6 groups, identified posterior TPJ and mPFC as core network over the different groups, in addition to group-specific activation (Schurz, Radua, Aichhorn, Richlan, & Perner, 2014). Schaafsma and colleagues (2015) emphasize that research should aim at gaining a true understanding of ToM and the distinct mental processes involved. They present an overview of tasks used, including on the one hand different stimulus types, such as verbal narratives or cartoons, and on the other hand different psychological assumptions, such as false belief attribution (see Sally-Anne task in section 1.3.2) or reading the mind in the eyes[3]. The authors point out the central question, whether ToM works by intuitively simulating the other person's mental state (embodied simulation) or is based on the construction of a theory about the other person's mind (theory-theory). While embodied simulation would be related to the MNS, theory-theory would be associated with the mentalizing system. Depending on the specific process at hand, additional cognitive processes needed for ToM would be activated, such as executive control processes. Schaafsma and colleagues (2015) suggest that we are likely inferring others' mental states using a mixture of these strategies. In particular, the authors highlight the need to define basic processes, such as decoupling, recursion, prediction, memory, motivation, and to investigate their involvement and associated brain activation in specific ToM tasks.

Accordingly, Schurz and Perner (2015) evaluated the results of their previous meta-analysis on ToM tasks (Schurz et al., 2014) with regard to 9 neurocognitive theories, thereby attempting to differentiate domain-general and domain-specific mechanisms subserving ToM. While most theories fail to predict results of the

[3] Pictures of the eye area of individuals are presented, and participants have to select out of four adjectives the one best describing the associated mental state, such as joking, insisting, amused and relaxed.

respective tasks, I will here summarize two of the theories which are largely in line with Schurz and Perner's (2015) predictions. This helps to better understand not only their work but also more generally some of the mechanisms underlying ToM. Please note that the suggested brain regions may deviate from the ones identified in the meta-analysis and also from those suggested by the x-/c-system, because on the one hand, the work of the authors of the following studies is older, on the other hand, their focus was not to find commonly activated regions, but instead differentiate individual mechanisms that are part of ToM. For example, based on the proposal by Perner and colleagues (J. Perner & Leekam, 2008; Josef Perner & Roessler, 2010), two proposed cognitive mechanisms of ToM, which are goal inference and perspective taking, should be associated with activation in posterior STS and the area comprising IPL and dorsal TPJ, respectively. Indeed, activation in IPL was increased only for tasks involving perspective taking, i.e. false-belief tasks and trait judgements. On the contrary, no perspective taking is needed and no IPL activation is found for social animations, strategic games, or reading the mind in the eyes task.

As another example, based on Gobbini and colleagues (2007), Schurz and Perner (2015) predict the activation patterns of overt versus covert mental states. Overt mental states are those that can be immediately observed, while covert mental states need to be inferred, such as the false belief of the true location of a marble. Based on the prediction, overt mental states, as present in the mind in the eyes task and social animations, should be related to increased activation in ventral posterior STS. In contrast, covert mental states, assessed by false belief tasks, strategic games and trait judgments, should not lead to increased activation in this area. While the prediction was true for most tasks, the rule regarding covert versus overt mental states apparently was not, as false belief tasks showed enhanced activation in posterior STS, but as a task involving a covert mental state should not have done so (for more details and a comprehensive analysis of all theory-task combinations, please refer to Schurz & Perner, 2015). The here presented results of the meta-analysis by Schurz and Perner (2015) illustrate the need for a better understanding of basic mechanisms, thereby supporting the argumentation by Schaafsma and colleagues (2015).

That different social-cognitive processes are related to each other, has been indicated by a study showing that ToM performance is closely related to that of emotion perception. Specifically, in a task using facial stimuli, both processes involve an overlapping set of brain regions, including STS, IFG reaching into the insula,

somatosensory cortex, amygdala and right middle frontal gyrus (Mier, Lis, et al., 2010). Indicating higher processing demands with increasing complexity, activation was reported to be stronger for ToM than for emotion perception.

To conclude, as would be expected, the mentalizing network plays an important role for ToM tasks, as indicated by the TPJ, STS and mPFC activation (please, see Figures 2 and 3). While STS activation is associated with fast, x-system processing, IPL as part of the posterior parietal cortex would rather be considered as belonging to the c-system. As we can see, even though mentalizing and ToM are associated with specific brain regions, they might require a mixture of fast automatic and slow deliberate processes. This idea is further supported by the study by Mier and colleagues (2010) who used facial stimuli, i.e. stimuli comparable to those of the studies in this thesis. In addition to other regions, they report activation in the fast responding amygdala, and the MNS key-region IFG. The task by Mier and colleagues (2010) however, can be defined as an affective ToM task. Differences in activation patterns between cognitive and affective ToM will be discussed in the following.

1.5.4 Cognitive and affective theory of mind

As also Schaafsma and colleagues (2015) pointed out, there is a strong need to understand the mechanisms underlying ToM, and accordingly several attempts have been made to integrate existing studies. In a review on the neuroanatomical bases of ToM, Abu-Akel and Shamay-Tsoory (2011) proposed that a “neurobiological model of ToM should minimally explain three basic mentalizing processes which include the ability to *represent* cognitive and affective mental states, *attribute* these mental states to self and other, and finally *apply* (or deploy) these mental states in a manner that allows one to correctly understand and predict behavior” (pp. 2971-2972). In their model, mental states are first represented in TPJ and then guided through STS or precuneus/posterior cingulate complex. For the subsequent processing, the authors distinguish the cognitive *cold* and the affective *hot* network. While cognitive ToM involves dorsal medial and lateral PFC, dorsal ACC, and dorsal striatum / dorsal temporal pole, affective ToM includes orbitofrontal cortex / vmPFC, inferolateral frontal cortex, ventral ACC and ventral striatum / amygdala / ventral temporal pole (Abu-Akel & Shamay-Tsoory, 2011). Their model, including the ACC’s role in directed attention and in the representation of self versus other mental states, and the connections with

the ventral and dorsal attentional systems, is in line with research finding performance in ToM-related tasks dependent on attentional inhibition (Bialystok & Senman, 2004).

The distinction between cognitive and affective ToM has also been suggested from lesion or psychopathology studies. For example, one interesting finding comes from Parkinson's disease, where early stages are associated with dopamine dysfunction in the dorsal striatum (Dauer & Przedborski, 2003; Kish, Shannak, & Hornykiewicz, 1988; Owen, 2004) and deficits in cognitive ToM (Roca et al., 2010), whereas later stages also impair ventral striatum and affective ToM (Bodden et al., 2010). In addition, inhibitory repetitive transcranial magnetic stimulation (TMS) over right DLPFC increased reaction times for cognitive, but not affective ToM, indicating selective specificity of the DLPFC for cognitive ToM (Kalbe et al., 2010).

Above cited results confirm the conclusion from the previous section that ToM includes brain regions from both x- and c-system, which might be particularly important when distinguishing affective versus cognitive ToM.

One recent fMRI study with a large sample size of 178 participants was designed to further investigate the different neural networks involved in affective versus cognitive understanding of others (Kanske et al., 2015). Participants watched videos of autobiographic narratives that varied in valence (negative or neutral) and regarding ToM related contents (ToM or no-ToM). Afterwards they rated a number of questions, including (1) how they felt (on a continuous scale from negative over neutral to positive) as a measure of empathy[4], (2) which thoughts the previously seen actor might have had (multiple choice) as a measure of ToM, or alternatively (3) as a control condition about factual knowledge that could be acquired from the video (multiple choice). This study's results indicate that subjective valence ratings were related to activation in empathy related brain regions, including dorsal AI, dorsal ACC/mPFC, IFG, supramarginal gyrus/dorsal TPJ. Cognitive ToM in contrast was related to stronger activation in mentalizing regions, including ventral TPJ, STS, temporal poles, precuneus and MPFC. Regarding the self-other distinction required for both tasks, Kanske and colleagues (2015) point out that enhanced TPJ activation was located more dorsally for affective empathy, but more ventrally for cognitive ToM. Using these regions as seeds in resting state functional connectivity analyses, they identified distinct networks that resembled the previously reported task-related networks. Taken

[4] It should be mentioned that according to the definition used for the empathy paradigm of study 1 this would not be considered affective empathy, but rather distress.

together, the results of that study support the assumption of two pathways for social understanding. The authors continue to discuss the resemblance of these two networks with other networks. Specifically, the empathy network seems to resemble the salience network (also called reactive/externally oriented network, task control network, or cingulo-opercular network), which is characterized by fast detection of salient stimuli and reactive orienting to them, subserving adequate reaction to the other's emotional state (Kanske et al., 2015). In contrast, the ToM network is comparable with the regions of the default mode network and is associated with distinguishing internal from external information as well as generating and contemplating on thoughts. The divergence of the findings by Kanske and colleagues' (2015) regarding ventral and dorsal networks from the model proposed by Abu-Akel and Shamay-Tsoory (2011) might be explained by differences in the definitions and tasks used, and consequently by the mechanisms that were targeted by the tasks.

To summarize, social-cognitive processes can be categorized into cognitive versus affective, and also into fast and automatic versus slow and deliberate. While the former distinction is man-made, the latter is biological. Some processes, such as ToM are complex and seem to be based on a mixture of different systems. Emotion perception in contrast, seems to rely on mainly one neural system, the fast and automatic one. Against this background, the question arises when the switch happens: What are the limits of automatic fast perception and when is more deliberate processing required? Study 3 represents one approach to finding the answer to this question: When ambiguous facial stimuli are used, automatic processing is not sufficient to resolve the conflict, so deliberation comes into play and decision making becomes necessary. In the next section (1.6), I will therefore explain some important basics of decision making and how this might be involved in resolving ambiguous facial configurations.

1.6 When automaticity and deliberation interact: ambiguous facial configurations

One task famously used to investigate probabilistic decision making is the so called beads task (Huq, Garety, & Hemsley, 1988). In this task, a participant is faced with the following problem: There are two jars, one filled with 80% red and 20% blue beads, the other one with 20% red and 80% blue beads. Given a sequence of presented beads, the participant has to determine out of which jar the beads were

taken. The process of gathering evidence and eventually making a decision, as exemplified in the beads task as well as its adaptation, the fish-in-the-lake task (Woodward, Munz, LeClerc, & Lecomte, 2009), can be divided into two stages: First, probabilistic reasoning is applied and the inner probabilistic model is consistently updated with the incoming information. At some point, this inner model reaches a point of certainty, leading to the final decision, i.e. which jar or lake best fits the evidence. While probabilistic reasoning during the task is associated with activation in regions of executive function, including DLPFC and parietal regions, the final decision was linked to activation in ventral striatum and ventral tegmental area (Esslinger et al., 2013). Most interestingly, individuals with delusions (McLean, Mattiske, & Balzan, 2017) seem to consider less evidence before coming to a conclusion, which is known as hasty decision making or jumping-to-conclusion bias. Regarding brain functioning during probabilistic decision making, individuals with schizophrenia showed reduced response in ventral striatum and ventral tegmental area (Rausch et al., 2014). The abnormal reasoning in patients can be explained by abnormal dopaminergic activity, which causes aberrant salience. In fact, aberrant salience might even be causal to the emergence of delusions, as otherwise unimportant stimuli come to seem relevant (Heinz & Schlagenhauf, 2010; Kapur, 2003).

The ventral striatum houses the Nacc, which is part of the dopaminergic mesolimbic pathway, well-known for its key role for motivation and reward (Kringelbach & Berridge, 2010). Social interactions, but also viewing attractive faces, or faces perceived as being positive are considered rewarding and have been shown to activate the Nacc (Aharon et al., 2001; Hahn & Perrett, 2014; Izuma, Saito, & Sadato, 2008; Krach, Paulus, Bodden, & Kircher, 2010; Spreckelmeyer et al., 2009). Importantly, the Nacc is also involved in salience attribution (Esslinger et al., 2013; Kapur, Mizrahi, & Li, 2005) and final decision making in probabilistic reasoning tasks (Rausch et al., 2014; Rausch et al., 2015). While the Nacc is part of the fast and automatic x-system, associated with salience and final decision making, DLPFC and parietal lobe are important regions for the slow and deliberate decision making, linked to the c-system

I therefore propose that the Nacc with its involvement in salience, reward and decision making is also central to social-cognitive decision making as occurring in emotion perception. Study 3 of my thesis is focused on the role of the Nacc for social cognition.

As briefly mentioned before, MN research in humans comes with its own challenges. Before presenting the objectives of my studies, I will therefore dedicate the next section to the challenges and possible solutions when investigating the MNS.

1.7 Challenges in the measurement of mirror neurons

While MN seem a promising substrate of social cognition (Gallese, 2007a; Gallese, Keysers, & Rizzolatti, 2004), we cannot measure them in humans due to ethical restrictions. Most studies therefore rely on fMRI, which has a good spatial resolution. Still, each voxel, the smallest unit of measurement, is commonly around 3 mm^3 in size, therefore containing approximately hundreds of thousands up to millions of neurons. Increased activation in a voxel could therefore stem from the same or from neighboring neuronal populations within that voxel, which makes a big difference regarding mirror neuron interpretations. Gazzola and Keysers (2009) propose the term *shared voxels* (sVx) to refer to voxels that show increased fMRI response during both observation and execution of an action. This term is chosen to more accurately describe what is measured, without implying mirror neuron activation where there might just be activation in neighboring neuronal populations. fMRI analysis usually follows standardized protocols which are optimized for the majority of data. However, in the case of the MNS, some of these otherwise good techniques might lead to false positives: Spatial smoothing might lead to the wrong impression of an overlap in activation between two conditions, when in fact the activation would fall in neighboring voxels without the smoothing. Likewise, group-level analysis might suggest that on average there was increased activation in one area for both conditions, when indeed, half of the subjects had increased activation in one condition, and the other half in the other condition.

Gazzola and Keysers (2009) overcome these challenges in their sVx analysis, which they base on a publication by Morrison and Downing (2007): (1) data are not smoothed (a method to enhance signal-to-noise ratio, but thereby reducing spatial specificity) during preprocessing and (2) only voxels, that show activation in all conditions *within* a participant are considered for the group level analysis. Using this method, Gazzola and Keysers reported that during the observation and execution of motor actions, there are more shared voxels than at chance-level in ventral and dorsal premotor, inferior parietal, supplementary motor, middle cingulate, somatosensory,

superior parietal and middle temporal cortex and cerebellum (Gazzola & Keysers, 2009). This method is applied in study 1 of my thesis using social-cognitive paradigms.

However, this approach allows no conclusions on whether the MNS also distinguishes between emotions or their valences, which would provide further information on the role of the MNS for social cognition. This question can be elegantly addressed using fMRI adaptation, another promising method, which is based on the simple biological fact that repeated stimulation of a neuron leads to a decreased response of that neuron. If the aspect to which the neuron is sensitive changes, this novelty leads to increased response of the specific neuron. Research confirms, that this adaptation effect can be seen in fMRI (Grill-Spector, Henson, & Martin, 2006).

Previous research using fMRI adaptation reported voxels in IFG (de la Rosa, Schillinger, Bulthoff, Schultz, & Uludag, 2016; Kilner et al., 2009) or IPL (Chong, Cunnington, Williams, Kanwisher, & Mattingley, 2008) that showed adaptation from observation to execution conditions and vice versa, which is considered an indicator of mirror function. First evidence also indicates that STS, amygdala and IFG show adaptation in response to facial configurations intended to express no emotion (neutral) or fear (Ishai, Pessoa, Bikle, & Ungerleider, 2004). In study 2 of my thesis, I applied fMRI adaptation to investigate whether the mirror system differentiates faces based on positive or negative valence.

1.8 Objective

The aim of this thesis is to further the understanding of neural mechanisms underlying social cognition. The underlying assumption is that social cognition relies on fast automatic and slow deliberate processes that are recruited depending on the complexity of the social-cognitive process. MN, assumed to build the neuronal basis of social cognition by automatic and fast responding to social stimuli however, cannot be measured directly in humans. When measuring the MNS indirectly, using fMRI, it is critical to adapt the processing routine to the specific requirements underlying the assumption of the MNS, including overlapping activation for different processes, not only across, but also within participants. In addition, previous studies have mainly focused on single social-cognitive processes. Evidence supporting the role of the MNS as a common neural basis of social cognition, however, would necessitate evidence for an involvement of the MNS in a range of social-cognitive processes.

Study 1 was therefore carried out to find a common neural basis of social cognition, and determine the role of the MNS, using different social-cognitive tasks and an analysis routine aimed at avoiding false positive results. While the results of study 1 allow conclusions on the similarities in brain activation between tasks, the question how exactly the MNS is involved in these processes is left unanswered. The theory of embodied simulation proposes that the MNS allows the understanding of another person's emotions (Gallese, 2007b). If this is true, the MNS would not only respond to facial dynamics perceived as expressing emotions in general, but also distinguish between emotions. As no study so far has investigated this question, study 2 tested the differential response of the MNS to facial configurations intended to express fear versus happiness. Importantly, study 2 was based on an fMRI adaptation paradigm, a method considered a gold standard for MNS research (Fuelscher et al., 2019).

In case of ambiguous facial configurations, the automatic processing associated with the MNS may be supported by additional brain regions contributing to deliberate processing. The aim of study 3 was therefore to investigate to role of decision making for social cognition. Specifically the question was, whether the findings from non-social probabilistic decision making tasks also apply to social cognition. In addition, the paradigm allowed to determine whether reward or salience drives decisions on facial configurations designed to express ambiguous emotions.

2 STUDY 1: IDENTIFICATION OF THE RELEVANCE OF MIRROR AREAS FOR DIFFERENT SOCIAL COGNITIVE PROCESSES

2.1 The human mirror neuron system – a common neural basis for social cognition?

2.1.1 Abstract

According to the theory of embodied simulation, mirror neurons (MN) in our brain's motor system are the neuronal basis of all social-cognitive processes. The assumption of such a mirroring process in humans can only be supported by results showing that within one person the same region is involved in different social cognition tasks.

We conducted an fMRI-study with 75 healthy participants who completed three tasks: imitation, empathy, and theory of mind. We analyzed the data using group conjunction analyses and individual shared voxel counts.

Across tasks, across and within participants, we find common activation in inferior frontal gyrus, inferior parietal cortex, fusiform gyrus, superior temporal sulcus, and amygdala.

Our results provide evidence for a shared neural basis for different social-cognitive processes, indicating that interpersonal understanding might occur by embodied simulation.

2.1.2 Introduction

Mirror neurons (MN) which might enable us to understand others people's emotions and even to infer their intentions (Gallese, 2007a) fascinate scientists and non-scientists alike. Could MN help us bond with other people, because we feel and know how they feel without them even saying a word? Could MN be a "hidden crystal ball" that allows us to see into the near future of social situations, and anticipate whether our interaction partner will be hitting or hugging us a few seconds later? Our study is the first to approach these questions by providing evidence for a shared neural basis for the three fundamental social-cognitive processes imitation, empathy, and theory of mind (ToM) both *within* and *across* the same participants.

A MN mechanism was first presented in 1992 (di Pellegrino et al., 1992) when the authors identified neurons in the monkey brain area F5 that fired not only when the monkey performed a hand movement but also when it observed the same movement performed by the human experimenter. Since then, many studies applying single cell recordings have revealed such neurons in the monkey brain (Gallese et al., 1996; Rizzolatti & Luppino, 2001). Importantly, it was shown that MN code not only for an action, but also for the goal of an action and thus allow prediction (Rizzolatti, Cattaneo, Fabbri-Destro, & Rozzi, 2014). The idea how this MN mechanism helps us to understand others is called embodied simulation (Gallese, 2007a); i.e. by motor resonance we feel how others feel and thus recognize their current state and even their intention. fMRI studies showed activity in inferior prefrontal cortex, premotor cortex, inferior parietal cortex and superior temporal sulcus (STS) in humans observing and imitating actions (Buccino, Binkofski, & Riggio, 2004; de la Rosa et al., 2016; Iacoboni, 2009). A meta-analysis of 125 fMRI studies identified several regions frequently associated with activity for execution and observation of actions (Molenberghs et al., 2012). Among the most frequent regions were inferior (Brodmann Area (BA) 44 and BA9) and middle frontal gyrus (BA6), inferior (BA40) and superior parietal lobe (BA7), as well as the insula (BA13). In particular the inferior prefrontal cortex with adjacent middle frontal gyrus, comprising BA44 and BA6, is considered a key region of the human MNS (e.g., Enticott et al., 2012; Rizzolatti & Craighero, 2004). Not only because it was found to be a structural homologue of the monkey mirror neuron area F5, but also because it is assumed to represent actions, and even the intentions and goals of actions (Iacoboni & Dapretto, 2006; Iacoboni et al., 2005). Interestingly, Buccino and colleagues showed that the human cortex topographically represents actions according to the body region with foot movements represented in BA6 and mouth movements rather in BA44 (Buccino et al., 2004), suggesting a homunculus of action mapping.

Beyond the mere representation of emotionally neutral actions, such as grasping, or finger tapping, the MNS seems to be involved in the recognition of emotions, in empathy and in ToM (Mier, Lis, et al., 2010; Molenberghs et al., 2012; Schulte-Rüther et al., 2007). Recognizing emotional facial expressions results in activation in the regions of the MNS (Mier, Lis, et al., 2010) as well as in the face processing network, including fusiform gyrus, superior temporal sulcus and amygdala (Haxby et al., 2000). A comparable pattern was found during imitation of emotional

facial expressions (K. R. Leslie, Johnson-Frey, & Grafton, 2004). Studies investigating additional social-cognitive processes, such as ToM and empathy, revealed a comparable activation pattern with activation in the face processing network and the MNS (Carr et al., 2003; Mier, Lis, et al., 2010; Schulte-Rüther et al., 2007). While Mier and colleagues (Mier, Sauer, et al., 2010) demonstrated that emotion recognition and ToM activate comparable brain regions, they also showed that activation in the MNS regions is stronger for ToM than for emotion recognition, suggesting enhanced activation in the MNS with increasing demands for social cognition. Thus, an interesting question is not only whether there is common activation for different social-cognitive processes, but also whether the brain differentiates between these different social-cognitive processes.

Despite the long line of research, studying the human MNS still suffers from a number of severe methodological problems. First, electrophysiological single-cell recordings, which are required for a clear-cut demonstration of MN properties, are not feasible in healthy humans (see (Mukamel et al., 2010) for a study in epileptic patients). Therefore, the majority of studies approaching MN in humans rely on methods with lower spatial or temporal resolution such as fMRI, which is one of the best choices when aiming for high spatial resolution, but still an indirect method relying on blood oxygenation (BOLD signal) and not directly neuronal activity.

Second, the validity of results of fMRI-based MN research may be compromised by two common steps of data processing: 1) Smoothing of brain activation smudges brain activity so that activation in neighboring voxels overlaps and thus becomes less discriminable. 2) Group analyses average over participants, so it is not even sure if activation of two tasks comes from one and the same person or is just reached by averaging over several people. To speak of mirror neuron activity however, it is essential for activation to take place in the same neurons *within* participants.

Finally, it is currently unclear whether different tasks of social cognition are being performed by a common network of brain regions, i.e. whether the MNS provides a unified substrate for all facets of social cognition. While fMRI studies showed activity in a number of recurring areas during several tasks involving observing and imitating actions (Buccino et al., 2004; de la Rosa et al., 2016; Iacoboni, 2009; Molenberghs et al., 2012), this provides only indirect evidence, as activity is not only compared between different participants, but even between different studies with different designs.

In this study, we attempt to identify a common network of social cognition in humans using fMRI recordings. To overcome the problems outlined above, we developed a new set of social-cognitive tasks that use the same stimulus materials (i.e. facial expressions) to test three fundamental processes of social cognition: imitation, empathy and ToM. Based on previous literature (Carr et al., 2003; Gallese, 2007a; Mier, Lis, et al., 2010; Schulte-Rüther et al., 2007), we hypothesize that all three tasks activate both the emotional face processing network (amygdala, fusiform gyrus and STS) and the MNS (BA44 and IPL). Furthermore, we expect to see differences in the amount of activation in these regions between the different tasks, e.g. increased activation in the MNS in ToM compared to emotion recognition (Mier, Sauer, et al., 2010) (Mier et al. 2010). We investigate these hypotheses by first analyzing activation within the individual tasks *across* participants with smoothed data. We then extend this approach by investigating shared activation across these tasks by group analyses on smoothed data *across* participants, as well as by analyzing the smoothed and unsmoothed data *within* participants. This last step, which is crucial to identify regions with MN properties in humans, is made possible by using the approach of shared voxels (sVx), published in a seminal paper by Gazzola and Keysers (Gazzola & Keysers, 2009), based on the work of Morrison and Downing (Morrison & Downing, 2007). The approach allows for comparison of activation from different tasks using the very same unaltered spatial position and puts the focus on shared activation *within* participants, thus overcoming the problems introduced by smoothing and group analyses. Since we assess different social-cognitive functions, we expect not only activation that is common to all of these functions, but also distinct activation patterns specific to the individual tasks and conditions within tasks. Thus, we additionally focus on differences between the different sub-processes of social cognition by comparing activation patterns within tasks.

2.1.3 Materials and Methods

2.1.3.1 Participants

We recruited 80 persons, 5 of which had to be excluded from the final analyses due to more than 3 mm translation or 3° rotation (N = 1), anatomical aberrations (N = 1) or technical issues (N = 3). Final sample for analyses consisted of 42 females and

33 males between age 18 and 36 years (mean: 23.45 years, ± 3.83) with higher education entrance certification. All participants reported no history of psychiatric or neurologic disease and fulfilled the inclusion criteria for MRI measurements.

2.1.3.2 Study Procedure

Participants were informed about study procedure and aims, signed written informed consent, and practiced all tasks on a laptop. The study was approved by the ethics committee of the Medical Faculty Mannheim, University of Heidelberg, and is part of a larger project on the human mirror neuron system. Participants joined two appointments, the first appointment with a simultaneous EEG-fMRI set-up and the second appointment with transcranial magnetic stimulation prior to fMRI scanning. Data reported in this manuscript refers to the fMRI-results of the first appointment.

2.1.3.3 Experimental Design

We used three experimental paradigms covering different processes of social cognition: An imitation task, an empathy task and a theory of mind (ToM) task. For all three tasks, we used pictures from the Karolinska Directed Emotional Faces stimulus set (Lundqvist, Flykt, & Öhman, 1998) of 5 females and 5 males, as well as control stimuli without social information. Task were implemented with Presentation Software (Version 18.1; www.neurobs.com) and presented via video goggles. Responses were given with a diamond shaped button device (Current Designs, Inc., Philadelphia, PA, USA). Task order 1. Imitation, 2. Empathy, 3. ToM was fixed for all participants.

2.1.3.3.1 Imitation

The imitation task (Figure 4-d1) had three experimental conditions: Imitation, Execution and Observation, as well as a control condition. At the beginning of each block the instruction cue ‘Observe’, ‘Imitate’, or ‘Execute’ was presented. The stimuli for the Observation and Imitation block were angry and fearful faces. In the Imitation block, participants had to imitate the facial expression as accurately as possible, in the Observation block to passively view the facial expression. In the Execution block,

participants read the word ‘anger’ or ‘fear’ and had to produce the corresponding facial expression. In the control condition participants had to pronounce the German letter ‘Ä’ or ‘A’ aloud. Ä and A were chosen to roughly resemble the facial expressions during anger and fear, respectively.

Experimental blocks contained 4 stimuli each and were alternated with blocks of 2 control stimuli. The instruction cues prior to each block were presented for 2 seconds, the face stimuli for 5 seconds, control stimuli for 3 seconds. Stimuli within the blocks were presented in pseudo-randomized order and were separated by an inter-stimulus-interval of 1-3 seconds. The instruction cues initiating a new block were preceded by an inter-block-interval of 4-6 seconds. Each experimental block was presented 5 times, resulting in 20 trials for each experimental condition and 30 trials for the control condition. Task duration was 13 minutes.

2.1.3.3.2 Empathy

The empathy task (Figure 4-d2) again consisted of three experimental conditions Affective Empathy, Cognitive Empathy and Distress, and one control condition. At the beginning of each block the instruction cue ‘How bad do I feel?’ (Distress), or ‘How bad does the presented person feel?’ (Cognitive Empathy), or ‘How much do I empathize with the presented person?’ (Affective Empathy), or ‘How big is the circle?’ (control condition) was shown. Participants were instructed to think about the cued question while watching fearful or angry faces. After each stimulus, the question was displayed again, together with a continuous visual analog scale from ‘not at all’ to ‘very much’ (control condition: ‘small’ to ‘large’) on which participants had to indicate their answer.

Analogous to the imitation task, we chose a design with experimental blocks of 4 stimuli alternating with a control block of 2 stimuli. The instruction cues prior to each block were presented for 2 seconds, the face and control stimuli for 3 seconds and the visual analogue scale for 4 seconds. Stimuli within the blocks were presented in pseudo-randomized order and were separated by a jittered inter-stimulus-interval of 1-3 seconds. The instruction cues initiating a new block were preceded by a jittered inter-block-interval of 4-6 seconds. Each experimental block was presented 5 times and each control block 15 times, making 20 trials for each experimental block and 30 total control trials. Total duration of the task was 17 minutes.

2.1.3.3.3 Theory of Mind (ToM)

The ToM task (Figure 4-d3) also had three experimental conditions Affective ToM, Emotion Recognition, Neutral Face Processing, and a control condition. The different conditions were implemented by different statements preceding the facial or control stimuli. The participants' task was to indicate by button press whether the picture matched the previous statement (yes, or no). Statements were the German versions of: "This person is about to bluster" and "This person is about to run away" for the Affective ToM condition, "This person is angry", and "This person is afraid" for the Emotion Recognition condition, "This person is female" and "This person is older than 29 years old" for the Neutral Face Processing condition, and "This is a circle" and "This is a triangle" for the control condition.

The ToM task was presented in an event-related design. Each statement was presented for 2 seconds and the subsequent stimulus for an additional 2 seconds. A jittered inter-trial interval of 1-3 seconds was applied. All trials were presented in pseudo-randomized order, with 20 trials per condition, making a total of 80 trials. All in all, this task took about 8 minutes.

2.1.3.4 fMRI Data Acquisition and Analysis

fMRI data was acquired using a 12 channel head coil in a 3T Siemens Magnetom Trio at the Central Institute of Mental Health in Mannheim, Germany. During the tasks, we used echo-planar imaging with 32 descending 3x3x3mm slices with 1mm gap, TR = 2000 ms, TE = 30 ms; flip angle = 80°, field of view = 192 mm; matrix = 64x64 . Prior to functional imaging, a MPRage was acquired of each participant (TR = 1570 ms, TE = 2.75 ms; flip angle = 15°, field of view = 256 mm; matrix = 256x256; voxel size 1x1x1 mm).

Data was analyzed with Statistical Parametric Mapping 8 (SPM8, http://www.fil.ion.ucl.ac.uk/spm/software/spm8/). Preprocessing included slice time correction, realignment to the mean image, normalization with coregistration to the MPRage and resampling with 3 x 3 x 3 mm voxel size. First-level analyses were run twice, once with unsmoothed data and once with smoothed data, using an 8 mm Gaussian kernel. For all first-level analyses, the face stimuli were modelled as events

(folding the HRF with a stick function) as regressors in the general linear model, and the according 6 movement parameters derived from realignment were used as regressors of no interest. For the Imitation task, we modelled the Imitation, the Observation, the Execution and the control condition, as well as the cues. For the Empathy task, Affective Empathy, Cognitive Empathy, Distress, as well as the control condition and the cues were used as regressors. For the ToM task, affective ToM, Emotion Recognition and neutral face processing, as well as the control condition were applied in separate regressors.

Contrast of interest for second level analyses for the Imitation task were: Imitation > control, Observation > control, Execution > control, Imitation > Observation, Imitation > Execution; for the Empathy task: Affective Empathy > control, Cognitive Empathy > control, Distress > control, Affective Empathy > Cognitive Empathy, Affective Empathy > Distress, Cognitive Empathy > Distress; and for the ToM task: ToM > control, Emotion Recognition > control, neutral face processing > control, ToM > neutral face processing, ToM > Emotion Recognition, and [ToM > control]> [Emotion Recognition > control] > [neutral face processing > control].

Significance threshold was set to $p < 0.05$ FWE corrected, $k = 10$ for the analyses within tasks. Significance threshold for the group conjunction analyses was set to $p < 0.001$ without a cluster size threshold. The threshold was chosen analogous to the threshold for the sVx analyses that is described below. Region of interest analyses were conducted for the IPL (left: 870 voxels, right: 868 voxels), BA44 (left: 252 voxels, right 255 voxels), STS (left 324 voxels, right 161 voxels), fusiform gyrus (left 617 voxels, right 627 voxels) and the amygdala (left 47 voxels, right 47 voxels). Masks for IPL, BA44 fusiform gyrus and amygdala were taken from WFU_pickatlas. Since no STS mask is available in the WFU_pickatlas, it was based on activity in a former study with the ToM task (Mier, Lis, et al., 2010) and has been successfully applied in further studies with the ToM task (Mier, Haddad, et al., 2014). Significance level for the ROIs was set to $p < 0.05$ small volume corrected (svc) with $k = 10$ for the single experiments, but without setting a cluster size threshold for the conjunction analysis.

Behavioral data was analyzed with IBM SPSS Statistics V20 (https://www.ibm.com/us-en/marketplace/spss-statistics), applying repeated measures ANOVAs, as well as post-hoc t-tests.

2.1.3.5 sVx analysis

We based the sVx analyses on the first level models described above, for 2 sets of contrasts: 1) (Imitation > control) & (Affective Empathy > control) & (ToM > control), and 2) (Imitation > neutral) & (ToM > neutral). The second set of contrasts was selected to account for face processing (the Empathy task is not included in the second set, because no social control condition is available). Based on Gazzolla and Keysers (2009), the significance threshold for the sVx analyses (with smoothed and with unsmoothed data) was set so that the probability to incorrectly define a voxel as sVx in a single participant was less than 0.001 (I.e. for each of the 3 contrasts in set 1, we set $p < 0.05$, corresponding to a total probability for the set of $0.05^3 = 0.000125$, and for each of the 2 contrasts in set 2 we set $p < 0.01$, corresponding to a total set probability of $0.01^2 = 0.001$ to incorrectly define a voxel as sVx). Based on Boolean maps of the single contrasts, we calculated the logical '&' to obtain the sVx maps containing the intersections of voxels over the contrasts. These sVx maps served to count the number of sVx in the ROIs as well as the whole brain. For each participant, the individual brain/ROI volume was taken as a reference for the required number of sVx to surpass chance level. The required number of sVx was based on a cumulative binomial distribution function with a voxelwise level of 0.001 and a threshold of $p < 0.05$ for finding the returned number of voxels by chance. The number of participants with sVx and the number of participants with sVx above chance level is reported.

2.1.4 Results

We used three experimental paradigms covering different processes of social cognition: An Imitation task, an Empathy task and a Theory of Mind (ToM) task. For all three tasks, we used pictures from the Karolinska Directed Emotional Faces stimulus set (Lundqvist et al., 1998) of 5 females and 5 males, as well as control stimuli without social information. Since by nature, fMRI data is high-dimensional and requires correction for multiple testing, in addition to whole brain analyses, we conducted analyses limited to predefined ROIs to account for possible type-I-errors. Our regions of interest (ROIs) for all three tasks were BA44, IPL, STS, fusiform gyrus and amygdala.

2.1.4.1 Imitation is linked to activation in the mirror neuron system

The Imitation task (Figure 4-d1) had three experimental conditions: Observation, Imitation and Execution, as well as a control condition. In the Observation and Imitation condition, participants were shown angry and fearful faces, which they should observe or imitate, respectively. In the Execution condition, participants read the word 'anger' or 'fear' and had to produce the corresponding facial expression. In the control condition participants had to pronounce the German letter 'Ä' or 'A' aloud. Ä and A were chosen to roughly resemble the facial expressions during anger and fear, respectively.

Whole brain analyses for the comparison of Imitation with the control condition mainly revealed activity in inferior parietal, frontal and temporal regions, in premotor cortex and visual cortex, as well as in the basal ganglia. The comparison of Imitation with Observation and with Execution revealed similar patterns in both cases. Small-volume correction for our ROIs confirmed significantly higher activation in all ROIs for Imitation than for the other conditions. Comparison of Execution with control mostly resulted in activation in cerebellum and inferior temporal lobe. ROI analyses showed higher activation for Execution than control in fusiform gyrus and STS bilaterally, as well as in left IPL and left BA44. Observation > control revealed mainly activation in visual cortex, in orbitofrontal cortex and in parahippocampal gyrus, reaching into the amygdala. ROI analyses showed significant activation in bilateral fusiform gyrus, in bilateral amygdala and in right STS. The comparison of the experimental conditions with the control conditions, including the overlap between conditions, is displayed in Figure 4-r1. Detailed results of the Imitation task can be found in supplementary tables 4 and 5, for whole brain and ROI analyses, respectively.

2.1.4.2 Empathy is linked to activation in the mirror neuron system

The Empathy task (Figure 4-d2) consisted of three experimental conditions: Affective Empathy, Cognitive Empathy and Distress, and a control condition. After the presentation of either an emotional (fearful or angry face, experimental conditions) or neutral stimulus (circle, control condition), participants were instructed to answer the questions 'How bad do I feel?' (Distress), 'How bad does the presented person feel?'

(Cognitive Empathy), ‘How much do I empathize with the presented person?’ (Affective Empathy), or ‘How big is the circle?’ (ontrol condition).

Whole brain analyses revealed increased activation in several cortical regions for Affective Empathy compared to control, including superior temporal sulcus, inferior parietal cortex, inferior frontal gyrus and visual cortices, in medial frontal gyrus and precuneus, as well as in the amygdala. A comparable picture occurred when comparing Cognitive Empathy with control and Distress with control. All empathy conditions in comparison to the control condition resulted in enhanced activation in all ROIs. Figure 4-r2 displays activity in the experimental conditions in comparison to the control condition, including the overlap between conditions.

Comparison of the empathy conditions showed significantly higher activity for Distress than for Affective and Cognitive Empathy in the temporoparietal junction (TPJ) bilaterally, as well as in the precuneus. ROI-analyses additionally showed that Distress resulted in enhanced activation in BA44 right, IPL left and right, STS left and right compared to Cognitive Empathy. ROI-analyses comparing Distress and Affective Empathy revealed significantly higher activation during Distress in bilateral fusiform gyrus, as well as in bilateral IPL and STS. Affective Empathy was linked to higher activation in left TPJ than Cognitive Empathy. Accordingly, ROI-analyses showed significantly higher activation in left IPL for Affective Empathy compared to Cognitive Empathy. Cognitive Empathy led to stronger activation than Affective Empathy in the executive control network, including regions of parietal and frontal cortex. ROI-analyses for Cognitive compared to Affective Empathy showed significant activity in right BA44, left STS and bilateral IPL, but in a more dorsal part of the IPL than for Affective compared to Cognitive Empathy. Detailed results of the Empathy task can be found in supplementary tables 6 and 7, for whole brain and ROI analyses, respectively.

2.1.4.3 ToM is linked to activation in the mirror neuron system

The ToM task (Figure 4-d3) also had three experimental conditions: Affective ToM, Emotion Recognition, Neutral Face Processing, and a control condition. The different conditions were implemented by different statements preceding the stimuli. The participants’ task was to indicate by button press whether the picture matched the previous statement (yes, or no). Statements were the German versions of: “This person is about to bluster” and “This person is about to run away” for the Affective ToM

condition, “This person is angry”, and “This person is afraid” for the Emotion Recognition condition, “This person is female” and “This person is older than 29 years old” for the Neutral Face Processing condition, and “This is a circle” and “This is a triangle” for the control condition.

Whole brain regression analysis showed that activation in STS and inferior frontal gyrus was highest for ToM, medium for Emotion Recognition, and lowest for neutral faces. Whole brain analyses of ToM in comparison to control, as well as in comparison to neutral faces mainly showed activation in STS, inferior frontal gyrus reaching into the insula, visual regions, and premotor cortex. ROI analyses for both ToM compared to control and ToM compared to neutral revealed significantly increased activation for ToM in all ROIs. Whole brain comparison of ToM to Emotion Recognition mainly revealed enhanced activation in the TPJ region. The corresponding ROI-analyses comparing ToM with Emotion Recognition showed higher activation for ToM in all ROIs, except for BA44. Whole brain analyses for Emotion Recognition compared to control and to neutral both revealed a similar pattern as the comparison of ToM with these conditions (i.e. inferior frontal gyrus, STS, premotor cortex). The corresponding ROI-analyses for Emotion Recognition compared to control and to neutral showed activation in all ROIs, except for the IPL in the comparison with control, and the amygdala in the comparison with neutral. The comparisons of the experimental with the control condition, including overlaps between conditions, are depicted in Figure 4-r3. Detailed results of the ToM task can be found in supplementary tables 8 and 9, for whole brain and ROI analyses, respectively.

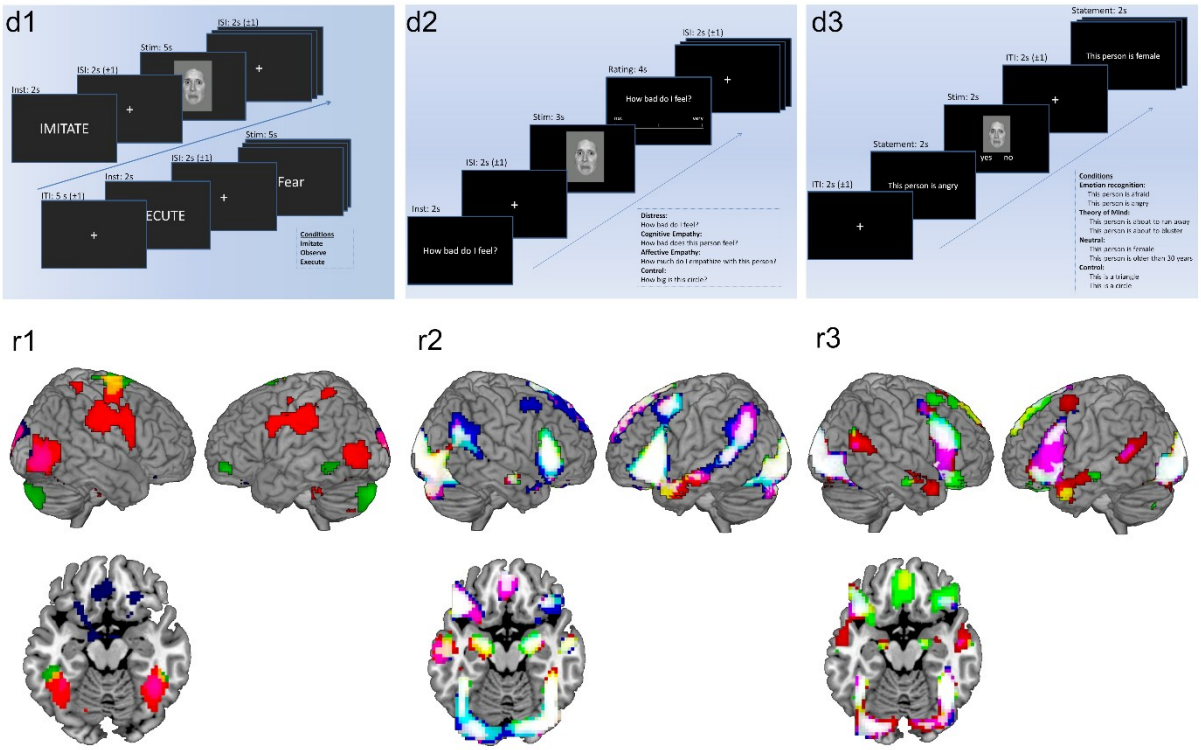


Figure 4. Task designs and fMRI activation for each of the three paradigms:

d1: Imitation task: Imitation task flow showing order and lengths of all events, exemplarily for two conditions.

r1: fMRI activation in the Imitation task: fMRI activation of all conditions of the Imitation task compared to the non-social control condition, slice coordinate Z=56. Red: Imitation > Control. Blue: Action > Control. Green: Observation > Control. Pink: overlap of red and blue. Yellow: overlap of red and green. Significance threshold p < 0.05, FWE-corrected, minimal cluster size k = 10.

d2: Empathy task: Empathy task flow showing order and lengths of all events, exemplarily for one condition. All questions of the instruction and the rating of the four conditions are shown in the lower right corner.

r2: fMRI activation in the Empathy task: fMRI activation of all conditions of the Empathy task compared to the non-social control condition, slice coordinate Z=56. Red: Affective Empathy > Control. Blue: Distress > Control. Green: Cognitive Empathy > Control. Pink: overlap of red and blue. Yellow: overlap of red and green. Cyan: overlap of blue and green. White: overlap of all three contrasts. Significance threshold p < 0.05, FWE-corrected, minimal cluster size k = 10.

d3: ToM task: ToM task flow showing order and lengths of all events. All statements of the four conditions are shown in the lower right corner.

r3: fMRI activation in the ToM task: fMRI activation of all conditions of the ToM task compared to the non-social control condition, slice coordinate Z=56. Red: ToM > Control. Blue: Emotion Recognition > Control. Green: Neutral > Control. Pink: overlap of red and blue. Yellow: overlap of red and green. Cyan: overlap of blue and green. White: overlap of all three contrasts. Significance threshold $p < 0.05$, FWE-corrected, minimal cluster size $k = 10$.

2.1.4.4 Common activation in the mirror neuron system exists across tasks

A whole brain conjunction analysis to investigate activation *across* participants and tasks including ToM, Affective Empathy and Imitation, each compared to control mainly revealed bilateral activation in amygdala, fusiform gyrus and STS, as well as activation in inferior frontal gyrus and premotor cortex (Figure 5 a). ROI-analyses confirmed significant activation in all of these regions, except for right BA44. To assure that these effects are not merely representing the processing of the faces that were used in all conditions, we also conducted a conjunction analysis on ToM compared to Neutral and Imitation compared to Observation. This conjunction analysis (comparing both task conditions with their social control condition) revealed common activation in STS and IPL bilaterally, as well as in bilateral inferior frontal gyrus, and in premotor cortex (Figure 5 b). The corresponding ROI-analyses revealed activation in bilateral STS and IPL, as well as in left BA44. Detailed results of the conjunction analyses can be found in supplementary tables 10 and 11, for whole brain and ROI analyses, respectively.

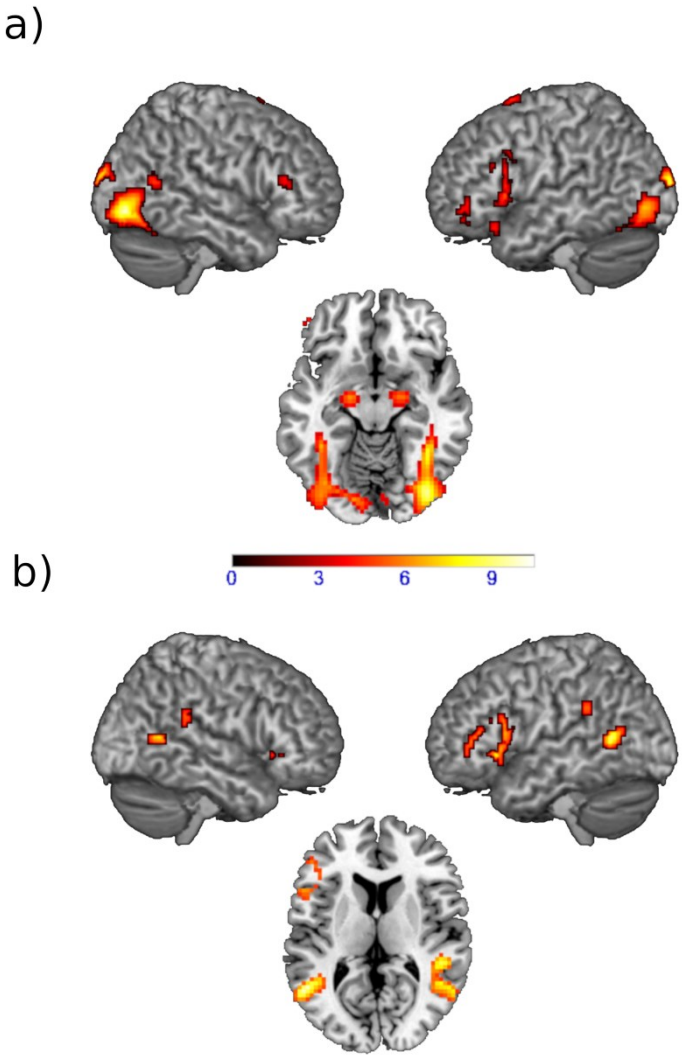


Figure 5. fMRI activation of the conjunction analyses: fMRI activation of the conjunction analyses. a) of all three tasks ((Imitation > Control) & (Affective Empathy > Control) & (ToM > Control), smoothed data), slice coordinate Z=60, b) of the Imitation and the ToM task ((Imitation > Observation) & (ToM > Neutral), smoothed data), slice coordinate Z=82. Color bar indicating t-values. Significance threshold p < 0.001, uncorrected, minimal cluster size k = 10.

2.1.4.5 Shared voxel counts show activation in the mirror neuron system within participants

To investigate activation *within* participants across tasks sVx were counted. As shown in Table 2, almost all participants had sVx across tasks. For the smoothed data, 82.7% of participants, and for the unsmoothed data, 92% of participants, had more sVx than predicted at chance level (i.e. more than 80-85 voxels, depending on individual brain size). We find a significant number of participants with sVx across all three tasks in our ROIs, with the highest number of sVx in fusiform gyrus and STS for the sVx analysis across all three tasks. For the sVx count of Imitation compared to observation and ToM compared to neutral, again sVx were revealed in all of our ROIs (Table 3). However, in this case number of sVx was reduced for amygdala and fusiform gyrus. For both analyses, more participants with sVx were revealed for the unsmoothed data in contrast to the smoothed data. In Figure 6, brain renders and slices overlayed with the sVx counts from unsmoothed data over all three tasks compared to control are shown.

Table 2. Number of participants with shared voxels (sVx) in the regions of interest. Numbers in brackets indicate number of participants with number of sVx greater than chance level. Contrasts: (Imitation > Control) & (Affective Empathy > Control) & (ToM > Control). Number of sVx at chance level: [1]: 80-85 depending on brain size. [2]: 0. [3]: 1. [4]: 2, [5]: 3. Note: BA44: Brodmann Area 44, IPL: inferior parietal lobe, STS: superior temporal sulcus, FG: fusiform gyrus.

		amygdala[2]	**BA44[3]**	**IPL[5]**	**STS[3]**	**FG[4]**	**brain[1]**
smoothed	left	12 (12)	29 (24)	15 (8)	34 (31)	62 (52)	75 (62)
	right	12 (12)	19 (14)	16 (8)	43 (40)	66 (61)	
unsmoothed	left	6 (6)	36 (20)	53 (20)	57 (45)	66 (58)	75 (69)
	right	9 (9)	22 (16)	44 (19)	58 (49)	73 (55)	

Table 3. Number of participants with shared voxels (sVx) in the regions of interest. Numbers in brackets indicate number of participants with number of sVx greater than chance level. Contrasts: (Imitation > Observation) & (ToM > Neutral). Number of sVx at chance level: [1]: 80-85 depending on brain size. [2]: 0. [3]: 1. [4]: 2, [5]: 3. Note: BA44: Brodmann Area 44, IPL: inferior parietal lobe, STS: superior temporal sulcus, FG: fusiform gyrus.

		amygdala [2]	**BA44[3]**	**IPL[5]**	**STS[3]**	**FG[4]**	**brain[1]**
smoothed	left	2 (2)	25 (24)	22 (17)	33 (29)	17 (9)	71 (58)
	right	2 (2)	19 (13)	30 (16)	18 (15)	16 (6)	
unsmoothed	left	3 (3)	28 (17)	44 (11)	40 (27)	31 (6)	75 (58)
	right	1 (1)	33 (17)	43 (19)	27 (8)	31 (5)	

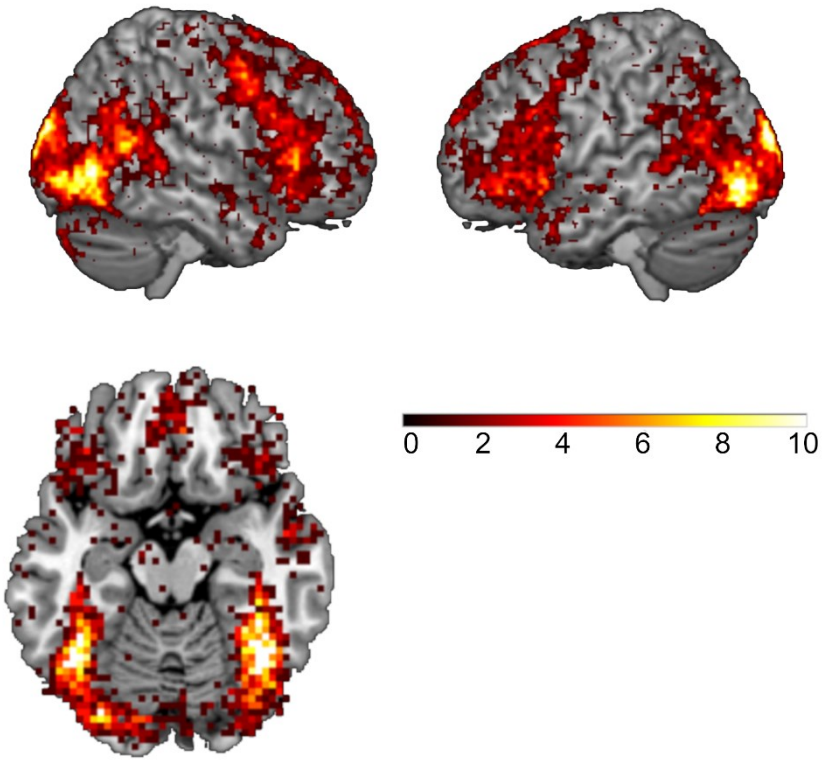


Figure 6 sVx counts for all three tasks (Imitation > Control) & (Affective Empathy > Control) & (ToM > Control): Number of participants with sVx for each voxel, unsmoothed data, slice coordinate Z=56. Color bar indicating the number of participants.

2.1.5 Supplementary Material

Table 4. Functional brain imaging results for the imitation task (p < 0.05 FWE-corrected, k = 10). Note: Subcluster peaks are inserted.

Imitation > Observation			MNI			
Area	BA	Cluster	x	y	z	t-value
Precentral Gyrus	6	18.871	-45	-13	37	22.76
Precentral Gyrus	6		54	-7	37	21.20
Precentral Gyrus	6		45	-13	37	20.81

Imitation > Control			MNI			
Area	BA	Cluster	x	y	z	t-value
Middle Occipital Gyrus	18	659	-9	-97	13	15.55
Lingual Gyrus	18		0	-85	1	14.65
Cuneus	18		9	-94	16	13.14
Precentral Gyrus	4	1.769	60	-13	37	14.67
Inferior Frontal Gyrus	9		54	5	31	10.36
Middle Frontal Gyrus	6		33	-4	70	9.82
Fusiform Gyrus	37	884	42	-46	-14	13.83
Inferior Temporal Gyrus			48	-73	-2	10.51
Postcentral Gyrus	3	932	-54	-22	40	13.11
Inferior Frontal Gyrus	9		-51	5	34	9.37
Middle Frontal Gyrus	6		-27	-7	52	8.43
Fusiform Gyrus	37	204	-39	-55	-14	8.75
Inferior Temporal Gyrus	37		-42	-46	-17	8.49
Fusiform Gyrus	19		-36	-70	-14	5.71
Amygdala		248	18	-10	-11	8.44
Amygdala			-18	-10	-11	7.13
Putamen			24	5	7	6.37
Middle Temporal Gyrus	37	196	-54	-67	7	8.15
Middle Occipital Gyrus	19		-45	-82	4	5.96
Cerebellum		42	-24	-34	-35	6.14
Cerebellum			-33	-43	-41	5.88
Cerebellum			-42	-37	-29	5.00
Caudate		19	15	-28	28	6.02

Cerebellum		45	-21	-70	-41	5.96
Cerebellum			-30	-64	-50	5.85
Thalamus		14	0	-4	22	5.64
Parahippocampal Gyrus	28	16	12	-19	-32	5.46
Cerebellum			3	-22	-38	5.32
Putamen		11	-24	2	10	5.41

Imitation > Action			MNI			
Area	BA	Cluster	x	y	z	t-value
Lingual Gyrus	18	2.763	0	-85	1	19.18
Cuneus	18		6	-94	13	15.36
Fusiform Gyrus	37		39	-52	-14	14.66
Inferior Frontal Gyrus	9	4.334	45	5	28	11.70
Postcentral Gyrus	4		60	-16	34	10.79
Postcentral Gyrus	3		-57	-22	40	10.41
Insula	13	57	-36	-4	16	9.38
Thalamus		236	9	-16	7	8.95
Thalamus			-6	-19	7	7.45
Thalamus			9	-28	-5	5.27
Amygdala		66	-21	-7	-11	6.97

Action > Control			MNI			
Area	BA	Cluster	x	y	z	t-value
Superior Frontal Gyrus	6	113	27	-10	76	8.41
Middle Frontal Gyrus	6		36	-4	70	7.70
Precentral Gyrus	6		45	-7	64	6.32
Cerebellum		646	-30	-88	-32	7.27
Cerebellum			30	-79	-38	6.64
Cerebellum			24	-91	-38	5.96
Fusiform Gyrus	37	74	-48	-37	-14	7.05
Middle Temporal Gyrus	21		-66	-52	-8	5.48
Middle Temporal Gyrus	37		-60	-46	-11	5.32
Middle Frontal Gyrus	10	16	-45	56	-11	6.92
Caudate		97	18	-37	25	6.49
Insula	13		27	-46	19	6.22

Cingulate Gyrus	31	59	-18	-40	25	6.49
Middle Occipital Gyrus	18	13	-9	-100	10	5.77
Fusiform Gyrus	37	18	48	-40	-17	5.68
Superior Frontal Gyrus	6	10	-12	11	73	5.51
Thalamus		11	0	-4	22	5.43

Observation > Control			MNI			
Area	BA	Cluster	x	y	z	t-value
Middle Occipital Gyrus	18	658	-6	-97	10	15.13
Cuneus	19		9	-94	19	14.03
Lingual Gyrus	18		0	-85	1	13.73
Fusiform Gyrus	37	405	42	-49	-14	11.10
Inferior Temporal Gyrus			48	-73	-2	9.37
Inferior Occipital Gyrus	19		42	-79	-8	9.26
Anterior Cingulate	25	200	0	-1	-8	7.50
Middle Frontal Gyrus	11		24	32	-17	7.26
Parahippocampal Gyrus	34		-12	-1	-20	6.58
Fusiform Gyrus	37	29	-36	-46	-14	7.05
Rectal Gyrus	11	100	-9	35	-20	6.73
Orbital Gyrus	11		3	41	-20	6.66

Table 5. Functional brain imaging results for the Imitation task, small volume corrected for the regions of interest (p < 0.05 small volume corrected, k = 0). Note: Subcluster peaks are inserted.

Imitation > Control			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	259	-51	8	25	6.73
			-54	11	10	5.19
			-42	-1	13	3.63
BA44	right	325	54	8	25	9.46
Amygdala	left	22	-21	-10	-11	5.42
			-18	-7	-14	5.37
Amygdala	right	30	21	-10	-11	7.19

			18	-7	-14	7.11
			30	-1	-20	4.14
Fusiform Gyrus	left	244	-42	-46	-17	8.49
			-39	-55	-17	8.41
			-33	-67	-14	5.43
Fusiform Gyrus	right	275	42	-49	-17	11.90
			21	-70	-17	4.30
Superior Temporal Sulcus	left	125	-54	-64	7	7.84
			-45	-73	16	4.46
Superior Temporal Sulcus	right	72	54	-61	10	6.70
			45	-58	16	5.07

Observation > Control			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
Amygdala	left	14	-18	-4	-20	4.36
Amygdala	right	6	18	-7	-14	3.14
Fusiform Gyrus	left	76	-36	-46	-14	7.05
Fusiform Gyrus	right	137	42	-49	-17	9.79
			36	-52	-14	8.89
Superior Temporal Sulcus	right	112	54	-67	16	4.77

Imitation > Observation			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	488	-57	-1	25	16.74
BA44	right	578	60	2	22	17.68
			39	8	4	10.81
Amygdala	left	24	-27	-4	-14	6.92
			-21	-10	-11	5.95
Amygdala	right	34	27	-4	-14	8.73
			24	-10	-11	7.27

			30	-1	-20	7.17
Fusiform Gyrus	left	256	-21	-64	-17	14.89
			-36	-58	-20	9.82
			-42	-37	-29	6.01
Fusiform Gyrus	right	254	21	-64	-17	15.93
			36	-49	-23	6.80
			39	-55	-23	6.74
Superior Temporal Sulcus	left	128	-54	-64	7	7.81
			-51	-61	10	7.74
			-45	-73	16	4.44
Superior Temporal Sulcus	right	30	57	-58	7	5.35
			51	-58	10	5.30

Action > Control			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	342	-48	5	25	5.83
			-39	2	13	5.49
			-42	-4	13	5.12
BA44	right	558	45	2	25	10.09
			51	8	25	9.64
			39	2	13	8.64
Amygdala	left	23	-21	-10	-11	6.32
			-24	-4	-14	6.29
Amygdala	right	34	24	-4	-14	8.51
			30	-1	-20	4.65
Fusiform Gyrus	left	198	-39	-58	-17	8.77
			-36	-55	-14	8.46
			-33	-61	-14	7.81
Fusiform Gyrus	right	244	39	-49	-17	13.78
Superior Temporal Sulcus	left	47	-45	-76	19	4.11

			-54	-61	1	3.99
			-54	-64	7	3.95
Superior Temporal Sulcus	right	49	51	-55	10	3.52
			45	-58	16	3.21

Action > Control			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	85	-57	17	4	4.01
Fusiform Gyrus	left	29	-30	-91	-26	5.58
Fusiform Gyrus	right	59	48	-40	-17	5.68
Fusiform Gyrus	right	32	27	-91	-26	4.11
			21	-94	-26	4.03
Superior Temporal Sulcus	left	107	-63	-46	-8	5.22
			-63	-61	-2	3.84
Superior Temporal Sulcus	right	42	51	-61	10	3.60
			45	-64	16	2.97

Table 6. Functional brain imaging results for the empathy task (p < 0.05 FWE-corrected, k = 10). Abbreviations: Cognitive = cognitive empathy, affective = affective empathy. Note: Subcluster peaks are inserted.

Distress > Control			MNI			
Area	BA	Cluster	x	y	z	t-value
Lingual Gyrus	17	3.088	-15	-94	-5	17.21
Middle Occipital Gyrus	18		-24	-91	-5	16.54
Cuneus	17		21	-91	1	16.18
Inferior Frontal Gyrus	47	1.318	-45	32	-8	14.03
Inferior Frontal Gyrus	47		-42	23	-14	12.46
Middle Frontal Gyrus	8		-42	14	46	11.44
Inferior Frontal Gyrus	47	686	51	29	-5	13.18
Amygdala			21	-10	-14	9.16

Inferior Frontal Gyrus	47		36	26	-17	8.95
Precuneus	7	414	-3	-58	37	12.73
Precuneus	31		0	-49	31	12.57
Supramarginal Gyrus	40	1.175	-45	-58	31	11.74
Supramarginal Gyrus	40		-60	-49	28	11.61
Middle Temporal Gyrus	22		-60	-43	4	11.05
Superior Frontal Gyrus	6	1.561	-3	17	64	10.54
Superior Frontal Gyrus	10		-6	59	31	10.42
Superior Frontal Gyrus	8		-6	26	58	10.11
Orbital Gyrus	11	186	0	41	-20	10.50
Superior Temporal Gyrus	39	746	57	-58	25	9.68
Superior Temporal Gyrus	22		51	-37	1	9.22
Superior Temporal Gyrus	13		60	-49	22	9.12
Amygdala		52	-18	-10	-14	7.97
Middle Frontal Gyrus	6	50	48	8	55	6.51

Distress > Cognitive			MNI			
Area	BA	Cluster	x	y	z	t-value
Supramarginal Gyrus	40	123	63	-43	31	7.21
Supramarginal Gyrus	40		60	-55	37	6.36
Supramarginal Gyrus	40	208	-57	-49	31	6.61
Inferior Parietal Lobule	40		-51	-55	43	5.84
Inferior Parietal Lobule	40		-57	-46	40	5.79
Precuneus	7	21	-3	-61	43	5.79

Cognitive > Control			MNI			
Area	BA	Cluster	x	y	z	t-value
Cuneus	18	2.875	18	-97	13	18.08
Middle Occipital Gyrus	18		24	-94	4	16.98
Cuneus	18		-27	-94	-5	15.77
Inferior Frontal Gyrus	47	1.576	-45	29	-5	13.10
Inferior Frontal Gyrus	45		-51	23	13	12.07
Superior Temporal Gyrus	22		-54	-43	4	9.07
Amygdala		192	21	-10	-14	12.19
Thalamus			21	-28	-2	6.86

Inferior Frontal Gyrus	45	411	57	29	1	11.57
Inferior Frontal Gyrus	47		36	29	-17	6.72
Precuneus	31	260	0	-49	31	11.22
Parahippocampal Gyrus	28	166	-18	-13	-14	9.78
Amygdala			-27	-4	-20	6.69
Thalamus			-21	-28	-2	6.42
Superior Frontal Gyrus	6	692	0	17	70	8.81
Superior Frontal Gyrus	9		-9	56	34	8.22
Superior Frontal Gyrus	8		-3	17	58	7.55
Middle Frontal Gyrus	6	160	-42	8	52	8.35
Superior Temporal Gyrus	22	330	48	-37	4	8.24
Superior Temporal Gyrus	39		54	-58	22	8.17
Superior Temporal Gyrus	22		42	-52	19	6.24
Middle Temporal Gyrus	21	76	54	-10	-14	7.88
Rectal Gyrus	11	87	0	38	-23	6.74
Medial Frontal Gyrus	11		0	53	-14	6.74

Distress > Affective			MNI			
Area	BA	Cluster	x	y	z	t-value
Superior Parietal Lobule	7	90	36	-70	49	6.87
Supramarginal Gyrus	40	119	60	-46	34	6.75
Precuneus	7	109	-3	-61	49	6.57
Superior Parietal Lobule	7		-15	-67	55	5.35
Angular Gyrus	39	34	-39	-73	25	6.46

Affective > Control			MNI			
Area	BA	Cluster	x	y	z	t-value
Inferior Occipital Gyrus	18	4.844	36	-85	-8	16.33
Cuneus	17		21	-94	1	14.90
Cuneus	18		15	-100	13	14.36
Precuneus	31	305	0	-49	31	14.87
Superior Frontal Gyrus	9	1.049	-9	56	31	10.97
Superior Frontal Gyrus	9		-9	50	40	10.26
Superior Frontal Gyrus	6		-6	20	70	8.48
Inferior Frontal Gyrus	47	258	54	29	1	10.41

Inferior Frontal Gyrus	47		36	29	-17	6.75
Amygdala		124	-21	-13	-14	9.84
Amygdala			-30	-7	-17	7.70
Orbital Gyrus	11	187	0	41	-20	9.70
Amygdala		111	21	-10	-11	9.42
Amygdala			30	-4	-20	7.80
Superior Frontal Gyrus	8	110	-39	17	55	7.87
Nodule		15	3	-55	-38	6.63
Superior Temporal Gyrus	38	19	42	17	-32	5.94

Affective > Cognitive			MNI			
Area	BA	Cluster	x	y	z	t-value
Angular Gyrus	39	32	-54	-64	34	5.65

Cognitive > Affective			MNI			
Area	BA	Cluster	x	y	z	t-value
Superior Parietal Lobule	7	84	33	-58	52	5.71
Superior Parietal Lobule	7		36	-67	49	5.57
Inferior Parietal Lobule	40		39	-52	58	5.49
Middle Frontal Gyrus	6	36	27	11	52	5.56
Superior Parietal Lobule	7	15	-15	-67	55	5.53
Precentral Gyrus	9	22	39	8	31	5.25
Middle Frontal Gyrus	9		51	14	31	5.24

Table 7. Functional brain imaging results for the empathy task, small volume corrected for the regions of interest (p < 0.05 small volume corrected, k = 0). Abbreviations: affective = affective empathy, cognitive = cognitive empathy, Neutral = Neutral face processing. Note: Subcluster peaks are inserted.

Affective > Control			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	136	-48	20	19	10.50
BA44	right	56	57	20	7	5.49
			48	20	19	4.16

Amygdala	left	47	-21	-10	-11	8.11
			-27	-7	-17	7.03
			-30	-4	-20	6.52
Amygdala	right	46	21	-10	-11	9.42
			30	-4	-20	7.80
Fusiform Gyrus	left	191	-36	-46	-17	8.73
			-39	-61	-17	6.44
			-27	-88	-26	5.99
Fusiform Gyrus	left	32	-42	-4	-29	6.59
Fusiform Gyrus	left	1	-42	-28	-20	4.15
Fusiform Gyrus	right	224	42	-49	-17	10.69
			24	-82	-23	7.57
			27	-88	-26	7.49
Fusiform Gyrus	right	35	60	-4	-29	3.74
Superior Temporal Sulcus	left	200	-45	-55	22	8.12
			-54	-61	22	7.82
			-63	-49	7	6.78
Superior Temporal Sulcus	right	139	60	-58	22	5.97
Superior Temporal Sulcus	right		57	-64	13	5.30

Cognitive > Control			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	145	-48	20	16	11.88
BA44	right	74	48	20	22	6.92
			57	20	7	5.86
Amygdala	left	46	-21	-10	-11	8.09
			-27	-4	-20	6.69
Amygdala	right	47	21	-7	-14	10.74

Fusiform Gyrus	left	226	-39	-46	-17	9.71
			-42	-67	-20	7.01
			-24	-85	-20	6.69
Fusiform Gyrus	right	236	42	-49	-17	10.55
			42	-67	-20	8.14
			24	-79	-20	7.00
Superior Temporal Sulcus	left	212	-45	-58	22	7.50
			-60	-49	10	7.19
			-63	-46	-2	4.31
Superior Temporal Sulcus	right	145	54	-58	22	8.17

Distress > Control			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	143	-48	20	19	9.83
			-54	20	10	9.74
BA44	right	87	57	20	4	8.81
			48	20	19	6.80
Amygdala	left	47	-18	-7	-14	7.16
			-21	-10	-11	7.11
Amygdala	right	47	21	-10	-11	8.86
Fusiform Gyrus	left	220	-39	-46	-17	9.52
			-42	-61	-17	8.71
			-30	-82	-20	8.27
Fusiform Gyrus	left	33	-45	-4	-29	4.50
Fusiform Gyrus	right	230	39	-52	-17	11.62
			27	-82	-23	11.32
			42	-70	-20	9.67
Superior Temporal Sulcus	left	233	-54	-61	22	10.26
			-63	-49	10	9.38
			-60	-49	22	8.97

Superior Temporal Sulcus	right	148	57	-52	22	8.79
			54	-58	22	8.77

(Cognitive & Affective) > Distress			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
Fusiform Gyrus	right	41	33	-43	-17	4.03

Table 8. Functional brain imaging results for the ToM-task (p < 0.05 FWE-corrected, k = 10). Abbreviations: ToM = affective Theory of Mind, Emo = Emotion recognition, Neutral = Neutral face processing. Note: Subcluster peaks are inserted.

(ToM > Control) > (Emo > Control) > (Neutral > Control)			MNI			
Area	BA	Cluster	x	y	z	t-value
Inferior Frontal Gyrus	45	249	-51	32	7	8.02
Inferior Frontal Gyrus	45		-48	17	16	5.68
Superior Temporal Gyrus	22	242	-54	-52	13	7.65
Superior Temporal Gyrus	22	338	48	-37	4	7.00
Superior Temporal Gyrus	22		60	-49	16	6.52
Superior Temporal Gyrus	22		54	-10	-11	5.59
Inferior Frontal Gyrus	45	55	54	29	4	6.40
Superior Temporal Gyrus	38	18	48	14	-26	5.39

ToM > Control			MNI			
Area	BA	Cluster	x	y	z	t-value
Middle Occipital Gyrus	18	3.222	12	-97	16	18.08
Lingual Gyrus	17		12	-91	1	16.96
Cuneus	18		-15	-97	7	16.67
Inferior Frontal Gyrus	45	1.554	-54	26	10	13.64
Inferior Frontal Gyrus	47		-42	23	-2	13.34
Middle Frontal Gyrus	46		-45	20	25	12.42
Inferior Frontal Gyrus	45	923	48	20	25	12.87
Inferior Frontal Gyrus	46		51	29	16	12.11

Inferior Frontal Gyrus	47		33	23	-2	11.88
Medial Frontal Gyrus	8	1.030	-3	17	52	11.56
Superior Frontal Gyrus	9		-9	56	34	11.54
Superior Frontal Gyrus	9		9	56	40	7.34
Superior Temporal Gyrus	22	400	48	-37	4	10.15
Superior Temporal Gyrus	22		51	-58	19	8.65
Superior Temporal Gyrus	22		57	-49	13	7.73
Superior Temporal Gyrus	38	247	45	14	-29	9.88
Superior Temporal Gyrus	22		51	-13	-11	9.86
Nodule		61	0	-55	-32	8.31
Amygdala		45	18	-10	-14	8.25
Amygdala			30	-7	-17	5.45
Superior Temporal Gyrus	22	319	-60	-49	10	8.16
Middle Temporal Gyrus	39		-42	-58	22	5.60
Orbital Gyrus	11	127	3	41	-20	7.92
Rectal Gyrus	11		3	29	-26	7.70
Amygdala		46	-18	-10	-11	7.51
Precuneus	7	44	0	-61	37	6.57
Thalamus		30	9	-31	1	6.08
Thalamus			21	-31	1	5.73
Thalamus		11	-6	-16	7	6.01
Middle Temporal Gyrus	21	20	36	-1	-35	5.87
Middle Frontal Gyrus	6	25	51	8	49	5.58
Caudate		13	12	8	4	5.56

ToM > Neutral			MNI			
Area	BA	Cluster	x	y	z	t-value
Inferior Frontal Gyrus	45	615	-48	32	7	12.96
Inferior Frontal Gyrus	45		-48	20	16	8.63
Superior Temporal Gyrus	22	531	-60	-52	10	10.39
Supramarginal Gyrus	40		-54	-49	19	8.70
Middle Temporal Gyrus	22		-51	-37	1	7.16
Superior Temporal Gyrus	22	88	48	-34	4	10.14
Superior Temporal Gyrus	22		57	-52	16	9.37
Superior Temporal Gyrus	22		57	-43	10	8.91
Inferior Frontal Gyrus	45	162	57	32	4	9.57

Cerebellum		148	24	-76	-32	7.33
Cerebellum			9	-82	-26	6.42
Cerebellum			15	-79	-32	6.29
Cerebellum		55	-21	-79	-32	7.24
Middle Temporal Gyrus	21	31	-54	-4	-11	6.46
Superior Frontal Gyrus	6	42	0	11	61	6.23
Cingulate Gyrus	32		-6	17	46	5.45
Precuneus	7	14	-6	-67	40	5.97
Caudate		10	21	-7	31	5.68

ToM > Emo			MNI			
Area	BA	Cluster	x	y	z	t-value
Superior Temporal Gyrus	22	339	63	-52	19	7.72
Superior Temporal Gyrus	22		60	-61	25	7.51
Angular Gyrus	39		51	-64	28	7.46
Middle Temporal Gyrus	39	363	-45	-61	28	7.29
Supramarginal Gyrus	40		-54	-49	31	6.35
Superior Temporal Gyrus	38	64	-45	17	-29	6.85
Superior Temporal Gyrus	38		-33	20	-29	6.36
Middle Temporal Gyrus	21	18	57	-7	-14	5.91
Superior Temporal Gyrus	38	13	45	14	-32	5.79
Middle Temporal Gyrus	21	17	-60	-7	-11	5.57
Superior Frontal Gyrus	9	12	-6	56	31	5.34

Neutral > Control			MNI			
Area	BA	Cluster	x	y	z	t-value
Middle Occipital Gyrus	18	2.757	12	-97	13	18.50
Cuneus	17		9	-94	4	16.55
Lingual Gyrus	17		-9	-91	-2	15.66
Inferior Frontal Gyrus	47	330	36	32	-11	12.40
Inferior Frontal Gyrus	47		33	23	-2	10.99
Middle Frontal Gyrus	11		30	38	-23	7.20
Inferior Frontal Gyrus	46	542	45	32	16	12.02
Middle Frontal Gyrus	9		45	14	28	10.90
Middle Frontal Gyrus	9		57	26	34	8.46

Rectal Gyrus	11	394	3	38	-20	11.69
Medial Frontal Gyrus	11		0	50	-17	9.78
Rectal Gyrus	11		3	14	-26	7.52
Superior Frontal Gyrus	9	1.027	-6	59	34	10.49
Medial Frontal Gyrus	8		0	29	46	10.25
Medial Frontal Gyrus	8		3	20	49	10.25
Middle Frontal Gyrus	9	225	-42	11	31	9.39
Middle Frontal Gyrus	46		-45	20	25	9.04
Inferior Frontal Gyrus	47	472	-39	23	-17	9.22
Inferior Frontal Gyrus	47		-30	20	-5	9.21
Inferior Frontal Gyrus	47		-39	23	-5	9.08
Amygdala		37	18	-10	-14	8.64
Uncus	28	38	30	-7	-32	8.06
Nodule		21	0	-55	-32	7.31
Precuneus	31	44	0	-52	28	6.80
Inferior Temporal Gyrus	21	21	63	-10	-20	6.74
Cerebellum		10	-33	-70	-44	6.58
Angular Gyrus	39	12	48	-61	28	6.48
Cerebellum		11	-9	-79	-32	6.44
Amygdala		15	-18	-10	-14	6.31
Middle Frontal Gyrus	6	33	36	14	61	6.24
Mammillary Body		10	0	-13	-8	6.21
Middle Temporal Gyrus	21	15	-63	-16	-14	5.86

Emo > Neutral			MNI			
Area	BA	Cluster	x	y	z	t-value
Inferior Frontal Gyrus	45	403	-51	32	4	10.39
Inferior Frontal Gyrus	45		-48	20	16	6.70
Superior Temporal Gyrus	22	218	-51	-52	10	9.60
Superior Temporal Gyrus	41	163	45	-40	10	6.79
Superior Temporal Gyrus	39		48	-52	10	6.62
Superior Temporal Gyrus	22		54	-40	10	6.41
Inferior Frontal Gyrus	45	80	54	29	4	6.67
Inferior Frontal Gyrus	47		48	26	-2	6.59
Cerebellum		15	27	-76	-38	5.81
Cingulate Gyrus	32	20	-6	17	46	5.76

Superior Frontal Gyrus	6		-3	11	58	5.69

Emo > Control			MNI			
Area	BA	Cluster	x	y	z	t-value
Middle Occipital Gyrus	18	2.831	12	-97	16	18.10
Lingual Gyrus	17		-21	-91	-8	15.93
Lingual Gyrus	17		12	-91	1	15.52
Inferior Frontal Gyrus	47	918	33	23	-2	13.25
Middle Frontal Gyrus	46		54	32	16	11.37
Inferior Frontal Gyrus	45		48	17	25	11.34
Inferior Frontal Gyrus	47	1.121	-45	23	-2	12.37
Inferior Frontal Gyrus	47		-54	23	4	11.14
Inferior Frontal Gyrus	45		-45	23	19	11.11
Cingulate Gyrus	32	538	-3	20	46	11.50
Superior Frontal Gyrus	6		-6	20	70	6.63
Superior Frontal Gyrus	8		-9	41	58	6.06
Nodule		28	0	-55	-32	8.04
Middle Temporal Gyrus	22	88	-48	-40	4	7.65
Superior Temporal Gyrus	22		-60	-49	13	5.95
Superior Frontal Gyrus	9	35	-9	59	34	6.69
Amygdala		14	18	-10	-14	6.38
Thalamus			9	-13	7	6.33
Superior Temporal Gyrus	22	25	48	-58	16	6.31
Superior Temporal Gyrus	22	19	48	-16	-11	6.23
Cerebellum			-15	-79	-35	6.12
Superior Temporal Gyrus	22	28	51	-40	10	6.05
Rectal Gyrus	11	28	-3	38	-23	6.00
Rectal Gyrus	11		3	29	-29	5.55

Table 9. Functional brain imaging results for the ToM task, small volume corrected for the regions of interest (p < 0.05 small volume corrected, k = 0). Abbreviations: ToM = affective Theory of Mind, Emo = Emotion Recognition, Neutral = Neutral face processing. Note: Subcluster peaks are inserted.

Neutral > Control	MNI

Area	Hemisphere	Cluster	MNI x	y	z	t-value
BA44	left	94	-48	20	22	7.54
			-45	20	4	5.66
BA44	right	110	48	20	22	8.80
			51	14	25	6.95
			39	20	4	4.63
Amygdala	left	37	-21	-10	-11	5.56
			-18	-7	-14	5.21
			-27	2	-20	3.47
Amygdala	right	38	18	-7	-14	8.25
			30	-1	-20	3.41
Fusiform Gyrus	left	167	-36	-46	-17	9.26
			-36	-55	-14	8.11
			-27	-82	-20	5.59
Fusiform Gyrus	right	212	39	-49	-17	12.90
			33	-70	-14	6.98
			42	-28	-20	4.50
Superior Temporal Sulcus	right	52	51	-61	22	4.91

ToM > Emo			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
Amygdala	left	35	-18	-7	-14	3.68
			-21	-10	-11	3.45
Amygdala	right	25	18	-7	-14	2.81
			27	-7	-14	2.75
			24	-10	-11	2.73
Fusiform Gyrus	right	25	54	-4	-29	3.70
Superior Temporal Sulcus	left	290	-45	-58	22	6.09
			-54	-58	22	6.06
			-57	-49	1	5.14

Superior Temporal Sulcus	right	158	63	-52	19	7.72
			57	-61	22	6.86
			45	-52	22	6.08

Emo > Neutral			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	261	-48	20	16	6.70
			-54	17	4	6.36
BA44	right	130	57	20	4	4.63
			51	20	19	4.08
Fusiform Gyrus	left	22	-42	-40	-17	4.14
Fusiform Gyrus	right	17	42	-46	-17	3.69
Superior Temporal Sulcus	left	203	-51	-55	10	9.34
Superior Temporal Sulcus	right	98	51	-52	10	6.01

ToM > Neutral			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	271	-48	20	16	8.63
			-51	20	4	7.68
BA44	right	149	57	20	4	5.21
			51	20	19	5.02
Amygdala	left	21	-30	-4	-20	3.02
			-27	-7	-14	2.98
Amygdala	right	16	27	-7	-14	3.29
			30	-4	-20	3.19
Fusiform Gyrus	left	44	-42	-43	-17	4.56
Fusiform Gyrus	left	34	-27	-82	-23	3.88
Fusiform Gyrus	right	51	21	-82	-23	4.71
			27	-82	-23	4.69

Superior Temporal Sulcus	left	267	-60	-52	10	10.39
			-54	-49	19	8.70
Superior Temporal Sulcus	right	154	57	-52	16	9.38

Emo > Control			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	184	-48	20	19	10.38
			-48	20	4	10.06
BA44	right	145	48	20	22	10.62
			39	20	4	6.57
			57	20	4	6.14
Amygdala	left	30	-27	-1	-23	4.13
			-21	-10	-11	3.85
			-18	-7	-14	3.11
Amygdala	right	32	21	-10	-11	5.63
			18	-7	-14	5.34
			30	-1	-23	3.83
Fusiform Gyrus	left	163	-39	-46	-17	10.93
			-27	-82	-20	5.90
			-33	-70	-14	4.27
Fusiform Gyrus	right	212	39	-49	-17	14.90
			39	-58	-17	14.11
			33	-64	-14	6.46
			24	-82	-20	4.39
Superior Temporal Sulcus	left	121	-60	-49	13	5.95
Superior Temporal Sulcus	right	128	48	-58	16	6.31

ToM > Control			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	179	-48	20	22	11.91
			-51	20	10	11.33
BA44	right	141	48	20	22	12.42
			54	20	4	6.09
			39	20	4	5.21
Amygdala	left	46	-21	-10	-11	6.72
			-18	-7	-14	6.65
			-27	-4	-23	4.65
Amygdala	right	41	21	-10	-11	7.95
			18	-7	-14	7.59
			30	-4	-20	4.67
Fusiform Gyrus	left	188	-39	-46	-17	11.25
			-24	-85	-20	7.65
			-33	-70	-14	5.89
Fusiform Gyrus	left	22	-39	-10	-29	4.23
Fusiform Gyrus	right	234	39	-49	-17	15.67
			33	-70	-14	6.78
			24	-82	-20	6.32
Fusiform Gyrus	right	28	39	-10	-32	4.37
Superior Temporal Sulcus	left	216	-60	-49	10	8.16
			-42	-58	22	5.60
Superior Temporal Sulcus	right	151	51	-58	19	8.65
			57	-58	19	8.57

Table 10. Functional brain imaging results for the conjunction analyses (p < 0.05 FWE-corrected, k = 0). Abbreviations: ToM = affective Theory of Mind, affective = affective empathy, Neutral = Neutral face processing. Note: Subcluster peaks are inserted.

Imitation (>Observation) & ToM (>Neutral) [smoothed data]			MNI			
Area	BA	Cluster	x	y	z	t-value
Superior Temporal Gyrus	22	305	48	-34	4	7.03
Superior Temporal Gyrus	39		48	-52	10	5.28
Superior Temporal Gyrus	22		48	-22	-5	4.78
Superior Temporal Gyrus	39	121	-51	-55	10	6.58
Inferior Frontal Gyrus	47	226	-42	26	-2	6.04

Imitation(>Control) & Affective (> Control) & ToM (> Control) [smoothed data]			MNI			
Area	BA	Cluster	x	y	z	t-value
Inferior Occipital Gyrus	19	571	39	-79	-8	10.36
Fusiform Gyrus	37		39	-49	-14	10.30
Cuneus	18	858	-15	-100	13	8.55
Fusiform Gyrus	37		-39	-46	-14	8.01
Middle Occipital Gyrus	18		12	-97	16	7.91
Amygdala		46	-18	-10	-11	6.00
Amygdala		61	18	-10	-14	5.64
Superior Frontal Gyrus	8	91	-6	17	55	5.43
Superior Frontal Gyrus	6		-6	14	73	4.87
Inferior Frontal Gyrus	45	148	-54	17	4	5.16
Middle Frontal Gyrus	46		-42	23	22	5.10
Superior Temporal Gyrus	22	30	57	-61	13	4.72

Table 11. Functional brain imaging results for the conjunction analyses, small volume corrected for the regions of interest (p < 0.05 small volume corrected, k = 0).

Abbreviations: ToM = affective Theory of Mind, affective = affective empathy, Neutral = Neutral face processing. Note: Subcluster peaks are inserted.

Imitation (>Observation) & ToM (>Neutral) [unsmoothed data]			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value
BA44	left	384	-36	5	4	65.99
			-42	8	4	60.72
			-57	-4	25	55.39
BA44	right	388	63	2	22	59.72
			57	-4	25	57.68
			63	8	22	52.34
Amygdala	left	13	-27	-4	-17	21.29
Amygdala	right	21	27	-7	-20	23.26
			27	-1	-23	17.67
Fusiform Gyrus	left	19	-21	-67	-17	70.50
			-30	-70	-14	12.46
		2	-39	-55	-23	20.72
Fusiform Gyrus	right	7	21	-64	-17	92.12
Superior Temporal Sulcus	left	118	-57	-61	13	30.20
			-60	-58	10	28.35
			-48	-64	16	27.83
Superior Temporal Sulcus	right	30	54	-55	10	32.08
			60	-55	10	24.42
			57	-61	10	20.42

Imitation(>Control) & Affective (> Control) & ToM (> Control) [smoothed data]			MNI			
Area	Hemisphere	Cluster	x	y	z	t-value

BA44	left	138	-54	20	22	46.05
			-57	17	16	37.80
			-48	20	22	29.87
BA44	right	13	51	-7	4	36.75
		50	48	20	22	27.96
			60	17	10	18.87
Amygdala	left	21	-18	-7	-14	31.84
			-27	-4	-17	23.94
Amygdala	right	26	18	-7	-14	59.33
			30	-4	-20	27.72
Fusiform Gyrus	left	18	-39	-49	-17	65.33
Fusiform Gyrus	right	9	36	-52	-14	58.37
		1	36	-49	-23	16.27
		3	27	-79	-20	12.32
Superior Temporal Sulcus	left	132	-54	-61	19	27.89
			-45	-61	19	27.63
			-42	-55	16	27.31
Superior Temporal Sulcus	right	93	57	-64	13	44.74
			57	-58	10	38.47
			45	-58	16	30.91

2.1.6 Discussion

Using a combination of three social-cognitive tasks (Imitation, Empathy, and ToM) based on emotional facial expressions and three different methods of analysis (analysis of individual tasks, conjunction analysis and sVx counts), we found converging evidence for a common network of brain regions underlying the different social-cognitive processes. Importantly, the combination of different analyses allowed us to confirm that common activation was present both *across* and *within* participants (Gazzola & Keysers, 2009), a result which has not been presented so far for processes

of social cognition. While the present study does not proof the existence of mirror neurons in the human brain, it strongly emphasizes both the existence of MN and their central involvement in social cognition (Molenberghs et al., 2012; Mukamel et al., 2010). Furthermore, we were able to show that activation in the network differentiates between different social-cognitive processes, in particular in the TPJ region.

2.1.6.1 Common activation across tasks

Analyses of the single tasks revealed activation in BA44 and IPL, as well as in fusiform gyrus, STS and amygdala. These results are in agreement with earlier studies showing an involvement of the emotional face processing network, as well as the MNS in different aspects of social cognition (Carr et al., 2003; Mier, Lis, et al., 2010).

This pattern of common activation from the single task analyses was statistically confirmed by the conjunction analyses that revealed shared activation across tasks in all regions of interest, except right BA44, strongly suggesting a common neural basis for different social cognitive-processes. Importantly, when restricting the analyses to ToM and Imitation (that each allowed contrasting with a social control condition), only MNS regions were revealed as regions with common activation, further emphasizing the importance of the MNS for social cognition. This analysis also suggests that in contrast to the common activation in regions for emotional face processing, namely fusiform gyrus and amygdala, the MNS effects were not merely driven by the processing of the facial stimuli, but are specific for social-cognitive processing. Both BA44 and IPL are homologues to monkey brain areas containing motor neurons (Rizzolatti, Luppino, & Matelli, 1998) and have been reported to show a pattern of activation in humans (Kilner et al., 2009; Molenberghs et al., 2012; Rizzolatti & Craighero, 2004; Rizzolatti et al., 1998; Thomas, De Sanctis, Gazzola, & Keysers, 2018) that has been described from monkey research on MN. In contrast, the STS is not thought to have motor neurons (Rizzolatti & Craighero, 2004). It is however a key region for processing variable aspects of the face, to process biological motion and has been shown to be substantial for intention recognition (Carr et al., 2003; Liu, Harris, & Kanwisher, 2010). A meta-analysis by van Overwalle (Van Overwalle, 2009) pointed out involvement of the STS in forwarding visual input to the MNS. Thus the STS can not only be attributed to the extended face processing system (Haxby et al., 2000), but

also provides visual input to the MNS, and seems to have an important role in social cognition that exceeds mere face processing.

In line with the study of Keysers and Gazzola (2009), the conjunction analyses were supported by the sVx analyses that revealed sVx *within* participants in each of our regions of interest. This was true when analyzing sVx across all three tasks with non-social stimuli as control condition. When analyzing activation across those two tasks with a facial expression as control condition (Imitation compared to observation and ToM compared to neutral face processing) sVx also occurred in all ROIs, but the number of sVx in amygdala and fusiform gyrus was reduced, again supporting a special role of the MNS for social cognition. Those analyses with unsmoothed data resulted in a higher number of sVx, reflecting the obvious fact that a better spatial resolution of the signal is beneficial for detecting sVx. Interestingly, in the conjunction analysis with smoothed data, no common activation was found in right BA44 - one of the core regions of the MNS. Thus, the sVx analysis seems to be a valuable and maybe even necessary additional step to account for interindividual variance in the MNS.

2.1.6.2 Differences between the task conditions

While all tasks lead to activation in the face processing network and in the MNS, there were also differences between different conditions of the tasks. For the Imitation task, the Execution condition led to activation in regions with motor neurons, while the observation condition resulted in activation in amygdala and fusiform gyrus. The activation in the Execution task can be explained by the motor performance of the participants. The lack of activation in regions of the MNS for the observation condition, however, is in contrast to previous studies on the observation and imitation of emotional facial expressions (van der Gaag, Minderaa, & Keysers, 2007) and contradicts the major assumption of the simulation theory: the notion that motor simulation occurs automatically without cognitive influence (Gallese, 2003b; Gallese & Goldman, 1998). This lack of significant activation in regions of the MNS during observation is discussed in the *limitations and perspectives* section.

In the Empathy task, differences between conditions were found in several ROIs, suggesting a differential involvement of the regions of the emotional face processing network and the MNS in these aspects of empathy. Possibly most importantly, Distress resulted in higher activity in comparison to Cognitive and Affective

Empathy in the precuneus and also the TPJ bilaterally which links posterior superior temporal with inferior parietal regions. Affective Empathy was also linked to higher left TPJ activation than Cognitive Empathy. TPJ and precuneus are key regions of the default mode network (Greicius, Krasnow, Reiss, & Menon, 2003) that has been linked to self-referential processing. While distress ("how bad do I feel") is focused on the self, Affective Empathy ("how much do I feel for the person") demands focus on both the self and the other person, and cognitive empathy ("how bad does the other person feel") puts the focus completely on the other person. Since the default mode network highly overlaps with our regions of interest that are involved in social cognition (Mars et al., 2012) it is plausible that the Affective Empathy and the Distress conditions elicit strong TPJ activation. In particular, the left TPJ seems to play an important role for self-reference during social cognition.

In agreement with earlier studies using the same ToM task, activation in the MNS regions is highest during ToM, followed by Emotion Recognition (Mier et al., 2016; Mier, Haddad, et al., 2014; Mier, Lis, et al., 2010; Mier, Sauer, et al., 2010), suggesting that social-cognitive demands are reflected in the height of activation. The result of higher activation for distress than Cognitive Empathy in the Empathy task in the TPJ region suggests that further regions can be integrated to the social-cognitive process, depending on the social-cognitive demands. Thus, the "social brain" seems to differentiate between different social-cognitive processes by strength of activation within certain regions, as well as by integrating further regions.

To summarize, whereas all tasks have common activation in MNS and face processing networks, single task conditions can be characterized by distinct features. While the Distress condition of the Empathy task involves DMN activation, reflective of self-referential processing, the Emotion Recognition and ToM conditions of the ToM task mainly differ in their strength of MNS activation.

2.1.6.3 Limitations and perspectives

By combining different tasks and different methods of analysis within the same study, we were able to overcome some of the methodological problems that are inherent in fMRI research. However, it has to be clearly acknowledged that we still rely on a measure that assesses mass signals of neurons converted to changes in blood flow and oxygen saturation. Thus, all our results are still indirect and rather coarse-

grained. Facing these limitations, it becomes an urgent matter to enhance the theoretical understanding of the neuronal response underlying the BOLD signal. In particular, local, spiking network models of individual brain regions (Hass, Hertag, & Durstewitz, 2016) could allow for such a deepened understanding, when constrained by data from human fMRI. The development of such models is an integral aspect of our ongoing project to understand the human MNS.

The methodological limitations outlined above may also provide explanations for some of the unexpected results. In particular, we did not find significantly increased MNS activity during the Observation of emotional faces, which challenges the idea of embodied simulation. Such a null result in fMRI does not necessarily imply the lack of such activation – it could also mean that activity of MN was too short or too weak to elicit the compensatory blood flow that generates a detectable BOLD signal. Another interesting interpretation is opened by a study by the group around Keysers and Gazolla testing the influence of the participants' levels of responsibility on their empathy for pain (Cui, Abdelgabar, Keysers, & Gazzola, 2015). Interestingly, the empathic brain response was reduced when participants were not responsible for the observed pain. These results suggest that MNS activation can be reduced depending on the context. In our tasks, the Observation condition was the only condition in which the participant had no further task than to observe a facial expression. Similar to the pain study summarized above, neural activation might be reduced if no active involvement is required. Or framed differently, it has been shown that activation in the MNS is stronger when intentionality comes into play. This is especially evident in studies showing that MNS activation is higher for meaningful actions than for meaningless actions (Iacoboni et al., 2005; Koski et al., 2002; Rizzolatti et al., 2014) and in studies showing that MNS activation can be modified by motivation (Cheng, Meltzoff, & Decety, 2006).

The sVx analysis, although in best support of a common ground of social cognition, also suffers from methodological problems that need discussion. In particular, it needs to be emphasized that a voxel of $3mm^3$ contains, depending on the specific region and calculation, several hundreds of thousands up to more than a million neurons. Consequently, to be counted as a sVx of the MNS, a voxel needs to contain many thousands MN. The method is thus too coarse to prove the influence of a subset of neurons within a region. This limitation may be the reason why, albeit significant, only part of the participants and only few voxels within of the MNS regions

had sVx properties. Furthermore, the tasks are not the strongest to elicit activation in motor neurons, because movement is neither observed, nor expressed (except in the imitation task). However, our aim was to probe social-cognitive skills, and commonly facial expressions are very subtle in comparison to large body movements. Thus, tasks that rely on rather large-scale motions (e.g. finger tapping, or hand movements) might more easily find more sVx. In addition, there is evidence that the MNS differentiates between different actions, meaning that different actions are represented in different parts of the MNS (Buccino et al., 2004). Thus, another option is that we found fewer shared voxels due to a more fine-grained response in the MNS that also differs between different emotions and social-cognitive processes. Future studies might use experimental designs that target MNS processing of different emotions, or that use repetition suppression designs to investigate the human MNS (Fuelscher et al., 2019).

2.1.7 Conclusions

This is the first study investigating a variety of social-cognitive processes within the same participants, allowing the assessment of a shared neural response to social stimuli. Conjunction, as well as sVx analyses revealed common neural activation in amygdala, superior temporal sulcus, fusiform gyrus, IPL and BA44 across tasks, suggesting an involvement of the emotional face processing network and the MNS for social cognition. Our findings support the assumption that the MNS is at the heart of our interpersonal understanding. To conclude, we propose that the answer to the question we raised in the introduction whether MN could be our “hidden crystal ball” that allows us to see into the near future of social situations, and to anticipate whether our interaction partner intends on hitting or hugging us, is a tentative yes.

2.2 Summary

According to the theory of embodied simulation, we understand other persons’ mental states, because they are represented in our brain in the same way as when we ourselves experience the mental state. Using pictures of facial configurations intended to express emotions and including a large sample (n=75), the results of study 1 showed that the different social-cognitive processes imitation, affective empathy, and affective

ToM all show individual patterns of activation, and also shared neural correlates, both within and across participants.

The data allowed two types of comparison: (1) Imitation, affective empathy and ToM could be compared to a non-social control condition, (2) for imitation and ToM, there was additionally a facial control condition. The main result is that a significant number of participants presented a significant number of sVx in the ROIs, indicating that the sVx analysis might be a good approach for future MNS studies, considering its increased spatial accuracy.

Interestingly, for the two-task analysis that used a face as a control condition, the number of participants with a significant number of sVx was considerably reduced for regions associated with emotional face processing, but not in the core MNS regions, as I present in Table 12 for easier comparability.

Table 12. Percentage of participants with a significant number of shared voxels for the two-task analysis relative to the three-task analysis. Tables with exact results for each analysis are presented in study 1.

		amygdala	**BA44**	**IPL**	**STS**	**FG**
smoothed	**left**	17%	100%	213%	94%	17%
	right	17%	93%	200%	38%	10%
unsmoothed	**left**	50%	85%	55%	60%	10%
	right	11%	106%	100%	16%	9%

Even though the numbers are significant, they might not be in full support of the assumption of MN as a basic mechanism involved in understanding others, simply because a majority of participants present no or only a limited number of sVx. While there are many physiological explanations possible, such as voxels being composed of different types of neurons and MN not reaching the required threshold or canceling each other out, the challenge of accurately measuring MN in humans remains.

Overall however, these results suggest that processing of facial configurations intended to express fear or anger is related to activation in a shared set of regions that are commonly activated during the processing of facial configurations, mirroring and mentalizing tasks, including STS, IFG and IPL. Considering that participants were not simply looking at the faces, but explicitly instructed to perform specific social-cognitive

tasks while looking at the faces, it seems reasonable that the results show activation in regions associated with the fast mirroring processes and also the slower mentalizing network.

The findings from study 1 mainly point to common regions for social cognition. However, study 1 did not address whether the MNS is also involved in differentiating between social information.

In study 2, I implemented an fMRI adaptation design, again using facial stimuli to investigate whether regions of the MNS are sensitive to emotional valence which would suggest a role of the MNS in the differentiation between emotions.

The task that was applied in study 2 should mainly evoke automatic processing, and therefore activate the MNS and regions of facial-affective processing. To examine whether the MNS distinguishes the valence of facial affect, an adaptation design was applied: When two stimuli are presented sequentially, neurons will fire similarly strong for both stimuli only if the feature to which they are specific changes. For example, when a red apple follows a green apple, neurons recognizing it as *apple* will not be excited upon the repetition, so their signal will be weaker for the second compared to the first apple. In contrast, neurons responsible to detect the color, will respond equally strong to both stimuli, since they perceive the novelty of the feature to which they are specific. The same effect could be expected for facial configurations intended to express emotions and will be subject of study 2.

3 STUDY 2: INVESTIGATING WHETHER MIRROR AREAS DISTINGUISH EMOTIONAL VALENCE

3.1 fMRI adaptation reveals: The human mirror neuron system discriminates emotional valence

3.1.1 Abstract

Our ability to infer other individuals' emotions is central for successful social interactions. Based on the theory of embodied simulation, our mirror neuron system (MNS) provides the essential link between the observed facial configuration of another individual and our inference of that emotion by means of common neuronal activation. However, so far it is unknown, whether the MNS differentiates the valence of facial configurations.

To increase the precision of our fMRI measurement, we used an adaptation design, which allows insights into whether the same neuronal population is active for subsequent stimuli of facial configurations. 76 participants were shown congruent, or incongruent consecutive pairs of facial configurations expressing fear or happiness.

Significant activation for changes in emotional valence from adaptor to target was revealed in fusiform gyrus, superior temporal sulcus, amygdala, insula, inferior parietal lobe and Brodmann area 44. In addition, activation change was higher in superior temporal sulcus, insula and inferior frontal gyrus for a switch from happiness to fear than for fear to happiness.

Our results suggest an involvement of the MNS in valence discrimination, and a higher sensitivity of the MNS to negative than positive valence. These findings point to a role of the MNS that goes beyond the mere coding of a motor state.

3.1.2 Introduction

Three girls standing on a stage, anxious expressions on their faces, their eyes opened wide with fear. You turned on your TV just to accidentally witness Heidi Klum's decision on which of the girls is going to make it into the next round of Germany's next top model. Suddenly, one of the girl's fearful expressions turns into a wide smile. You can see, even feel, her relief, her joy. After a second-long fight to hide her sadness, another girl breaks into tears. You do not need to hear Heidi's words to know what she just said.

Facial emotional expressions / facial configurations[5] are an important means of communication, carrying information about the individual's current state, intentions, evaluation of the situation, and relationship to other individuals On a neural level, the amygdala plays a central role, as it is involved in the processing and perception of different emotions (Costafreda et al., 2008; Fitzgerald et al., 2006; Sergerie et al., 2008) and also in the discrimination of emotions (Critchley et al., 2000; Gur et al., 2002; Habel et al., 2007). Emotional stimuli, in particular when they have negative valence, are considered salient for us humans as social beings (Kret, Sinke, & de Gelder, 2011; Santos et al., 2011), suggesting the amygdala's strong involvement in emotion processing might be explained as indicative of salience (Liberzon et al., 2003; Santos et al., 2011). A meta-analysis of 105 fMRI studies confirmed the role of the amygdala for emotion perception, and also highlights the importance of further areas: the insula for disgust and anger, and the fusiform gyrus (FFG) for joy and fear (Fusar-Poli et al., 2009). Most interestingly, the neural processing of observed facial configurations is not only restricted to brain areas associated with emotional processing and face processing, but it additionally seems to include a network known as the MNS (Keysers & Gazzola, 2009). Whereas a few studies demonstrated the involvement of the MNS in emotion perception (Enticott et al., 2008; Mier, Lis, et al., 2010), so far, it is unknown, whether the MNS also differentiates between emotions.

Mirror neurons (MN) were first described in macaque monkeys, in which some neurons responded both when they performed a specific movement and when they

[5] For a very recent publication on the scientific validity of studies investigating facial emotion expression and perception, please refer to (Barrett et al., 2019). Following their suggestions, we use the expression "facial configurations" when referring to what are commonly called "facial emotional expressions", and substitute "recognition" by "perception". Furthermore, we will discuss our findings with regard to their objections and suggestions.

observed the same movement being performed by the experimenter (di Pellegrino et al., 1992). Due to the limited possibility to perform single-cell recordings in humans, so far only one study confirmed the existence of mirror neurons in humans (Mukamel et al., 2010). However, several studies have investigated mirror properties in humans using fMRI. Based on these studies, the human MNS has been proposed to be composed of three central areas, (1) the superior temporal sulcus (STS) which processes incoming visual information, (2) the inferior parietal lobe (IPL) which holds a motor representation of the specific movement, and (3) the inferior frontal gyrus (IFG) reaching into the premotor cortex, which is associated with the goal-directedness of a movement (Iacoboni & Dapretto, 2006). Based on these findings, Gallese proposed the theory of embodied simulation (Gallese, 2003a, 2007a): During social interaction, an individual mirrors the emotional state of their counterpart, so that on a neural level, the same brain regions become activated upon observation of a specific expression as to one's own experience of the same emotion. This process occurs automatically and allows one to understand the other person's actions, expressions and intentions (Gallese, 2003b). Indeed, fMRI studies provide evidence for the theory of embodied simulation, confirming that the MNS is involved in the processing of own, as well as observed emotions (Bastiaansen, Thioux, & Keysers, 2009). However, whether the MNS differentiates between different emotions is still unclear.

In a previous study using facial configurations commonly associated with happiness and fear, we showed that regions of the MNS are involved in social-cognitive processes of different complexity, and that across and within participants, there are shared voxels for different social-cognitive processes (Schmidt et al., unpublished results). Now, we aim to extend these findings by investigating whether the MNS is not only involved in social cognition in general, but also sensitive to the valence of facial configurations. Actual measurement of MN would necessitate single-cell recordings and firing of the same neurons under different conditions. As this remains impossible using fMRI, we aim at achieving high accuracy of our results by applying a neuronal adaptation design. Usual fMRI analysis has coarse spatial resolution, so the activity of one voxel could be based on firing of the same or of neighboring neuronal populations. fMRI adaptation is considered one means to allow greater accuracy with regard to spatial conclusions: Neurons fire upon detection of a novel stimulus to which they are specific. Repeated presentation of the same stimulus leads to adaptation, i.e. reduced firing of that neuron. If several neurons respond with adaptation, this is reflected by

decline of the BOLD response (Grill-Spector et al., 2006). In other words, the neuron will only show a consistently increased response if the aspect to which it is specific, changes and therefore remains novel. Otherwise other neuronal populations that are responsive to the changed aspect will show increased firing. As only the global firing rate modulates the fMRI-BOLD response, changing of the stimulus leads to an increased activation if there are neuronal populations sensitive for different stimulus aspects within that region. Thus, if regions of the MNS show reduced adaptation to changes in facial configurations chosen to reflect positive or negative valence in comparison to the same valence presented after each other, it can be concluded that different neuronal populations within the MNS respond to different valences.

So far, few studies successfully applied fMRI adaptation designs to MNS investigations. Clear support of MN would be cross-modal adaptation, i.e. adaptation of MN regions from execution to observation and vice versa. Using such observation / execution tasks, de la Rosa and colleagues (de la Rosa et al., 2016) and Kilner and colleagues (Kilner et al., 2009) identified voxels in IFG that showed cross-modal adaptation, and Chong and colleagues (Chong et al., 2008) identified cross-modal adaptation in the right IPL. Whereas these three studies support MN mechanisms in humans, a study by Lingnau and colleagues (Lingnau, Gesierich, & Caramazza, 2009) did not find cross-modal adaptation. Finally, an adaptation study on facial configurations by Ishai and colleagues showed neural adaptation in inferior occipital gyrus, lateral FFG, STS, amygdala, IFG and insula in response to facial configurations expressing neutral and fearful emotional states (Ishai et al., 2004). The authors conclude that emotional valence and relevance of the task are crucial for the adaptation response. However, whether regions of the MNS show adaptation in response to different emotional valences is currently an open question.

We hypothesize that amygdala and STS are sensitive to the emotional valence inferred from a facial configuration and thus show greater activation when perceiving different valences in consecutive facial configurations compared to the repetition of the same facial configuration. An open question is whether IFG and IPL also respond to specific valences or show adaptation. An additional exploratory question is whether the MNS reacts differently to negative versus positive valence; i.e. whether regions in the MNS respond more strongly to a switch from a facial configuration expressing happiness to one expressing fear rather than vice versa.

3.1.3 Methods

3.1.3.1 Procedure

Data were collected as part of a larger project on the mirror neuron system. Participants were invited to two appointments: During the first appointment, they were extensively informed about goals and procedure of the study in oral and written form and gave written informed consent. They also filled in a set of questionnaires and gave a saliva sample. At the second appointment, we explained the tasks to the participants in detail and let them practice the tasks. Afterwards, participants spent about one hour in the MRI. At first, we performed a T1-weighted anatomical scan, lasting about 5 minutes, followed by the functional measurements. The here presented emotion adaptation paradigm came second after an about 8-minute long imitation task, and preceded two more tasks. For their participation, participants received 15€ for each of the two appointments, or student credits. In a reward paradigm at the end of the second appointment, all participants could additionally win up to 20€.

The study was conducted in agreement with the declaration of Helsinki and approved by the ethics committee of the medical faculty of the University of Heidelberg (Germany).

3.1.3.2 Sample

We invited 81 participants between age 18 and 35 years to our study. 2 had to be excluded because of brain anatomical aberrations, 3 because of high scores in the depression questionnaire (BDI) – despite a prior telephone screening to exclude mental disorders, 1 because of major genetic anomalies, and 1 because s/he did not complete the study. Our final sample includes 74 right-handed (45 women, mean age 22.4 years; 30 men, mean age 23.2 years), German participants with higher education entrance qualification and no self-reported history of neurological or mental disorder participated in the study. All participants fulfilled the general requirements for MR measurements.

3.1.3.3 Stimulus Set

The pictures of 3 women and 3 men, showing facial configurations intended to express fear or happiness, were taken from the Karolinska Directed Emotional Faces (KDEF) Stimulus Set (Lundqvist et al., 1998), and are validated for valence and arousal (Adolph & Alpers, 2010). Using GNU Image Manipulation Program (GIMP, Version 2.8.22, 2017), we removed the hair and neck from the pictures, converted the color space to grey-scale, resized the faces to a uniform size within the stimulus set, cut the total picture to uniform size, and corrected luminance levels, with separation of background and foreground pixels, using the MATLAB toolbox SHINE (Willenbockel et al., 2010). These steps were taken to avoid possible adaptation effects to physical features irrelevant to our task.

3.1.3.4 Emotion Adaptation Paradigm

In the paradigm, each trial consisted of two consecutive pictures of faces of the same identity. There were six possible combinations of first and second stimulus (please also see Table 13). Each of the six trial types was presented 12 times, leading to 72 trials in total, presented in pseudo-randomized order. For the first stimulus, also known as adaptor, a facial configuration associated with fear or happiness was presented for 3 seconds. After an inter-stimulus-interval (ISI) of 500-1000 milliseconds, the second stimulus, known as target, was presented for 1 second and 25% larger than the adaptor to control for low-level visual adaptation effects. The target was again a facial configuration associated with fear or a happiness, resulting in congruent (adaptor and target show same emotional expression) and incongruent (adaptor and target show different emotional expressions) conditions. In one third of cases, the target face was inverted, i.e. upside-down, and participants had to respond as quickly as possible with a button press. This control target therefore served to maintain participants' attention (Mohamed, Neumann, & Schweinberger, 2011), and is not considered as a target of interest. Between all trials, there was an inter-trial-interval (ITI) of 8 to 12 seconds. For the duration of the ISI and ITI, a white fixation cross was presented on a black background. Our choice of stimulus presentation time, ISI, ITI and different size of the stimuli was based on different publications, especially those by Schweinsteiger and colleagues (e.g., (Herzmann, Schweinberger, Sommer, & Jentzsch, 2004; Kaiser,

Walther, Schweinberger, & Kovacs, 2013)). We programmed and presented the paradigm using the software Presentation (Version 18.1, Neurobehavioral Systems Inc.).

Table 13 shows the six trial types and durations of the individual trial segments. Each trial type was presented 12 times, making 72 trials in total, presented in pseudo-randomized order. Between each adaptor and target stimulus, there was a fixation cross presented for 500-1000 ms as inter-stimulus interval, and before a new trial, there was a fixation cross presented for 8-12 seconds as inter-trial interval. For better readability, we here shorten the expression "facial configuration inteded to express a [specific emotion] state" to "[emotion] face". Note HH: happy adaptor face, happy target face; HF: happy adaptor face, fearful target face; HI: happy adaptor face, inverted happy target face; FF: fearful adaptor face, fearful target face; FH: fearful adaptor face, happy target face; FI: fearful adaptor face, inverted fearful target face.

Condition	*Adaptor (3s)*	*Target (1s)*
Congruent_HH	Happy face	Happy face
Incongruent_HF	Happy face	Fearful face
Control_HI	Happy face	Inverted happy face
Congruent_FF	Fearful face	Happy face
Incongruent_FH	Fearful face	Fearful face
Control_FI	Fearful face	Inverted fearful face

3.1.3.5 fMRI data acquisition

We recorded the fMRI data using a 3T Siemens Magnetom Trio (Siemens Inc., Erlangen, Germany). For the anatomical measurement, we first recorded a T1-weighted MPRage with TR = 2300 ms, TE = 3.03 ms, flip angle α = 9°, field of view = 192x192 mm, 192 layers and 1x1x1 mm voxel size. For the functional data, we used a T2*-weighted echo-planar-imaging (EPI) sequence with TR = 2000, TE = 28 ms, flip angle α = 80°, field of view = 192x192 mm, 33 layers and a voxel size of 3x3x3 mm with 1 mm gap. For the emotion adaptation paradigm, 277 scans were recorded for each participant.

Button presses were registered using a 4-button diamond Current Design response pad (Current Design Inc., Philadelphia, USA). A MR-compatible face cam (MRC Systems GmbH, Heidelberg, Germany) together with the software IC-Capture (Version 2.3.394.1917, The Imaging Source LLC), was used to ensure that participants do not fall asleep during the tasks and to verify that they followed the task instructions in the other paradigms that required facial imitation.

3.1.3.6 fMRI data processing

Data were processed using the MATLAB toolbox Statistical Parametric Programming (SPM12, The FIL Methods Group). For the preprocessing, we performed slice time correction, realignment to the mean image, coregistration to the MPRage, normalization to a standard brain with resampling to a voxel size of 3x3x3 mm and smoothing with an isotropic Gaussian kernel of 8x8x8mm using a full-width at half-maximum (FWHM) filter. For the general linear model of the first-level analyses, using an event-related design (folding the HRF with a stick function), we set the 6 movement parameters from the realignment as regressors of no interest, and as regressors of interest we set one regressor to each stimulus type (stimulus types and abbreviations are described in Table 13).

To test our hypotheses, we created three contrasts for second-level analyses, always using the target face event:

1. Incongruent > Congruent (Incongruent_HF and Incongruent_FH trials minus Congruent_FF and Congruent_HH).
2. (Incongruent_FH > Congruent_HH) > (Incongruent_HF > Congruent_FF)
3. (Incongruent_HF > Congruent_FF) > (Incongruent_FH > Congruent_HH)

Significance threshold at the whole-brain level was set to $p<0.05$, FWE-corrected, minimal cluster size $k>=10$. For small volume correction, the whole-brain threshold was set to $p<0.05$ uncorrected, $k>=10$ and a peak-level threshold of $p<0.05$, FWE-corrected.

Regions of interest were BA44, STS, IPL, amygdala and insula. The masks for all regions except the STS were taken from the WFU PickAtlas tool (Maldjian, Laurienti, & Burdette, 2004; Maldjian, Laurienti, Kraft, & Burdette, 2003). As WFU PickAtlas does

not provide an STS-mask, this mask was based on activity in a former study using a theory of mind task (Mier, Lis, et al., 2010).

3.1.4 Results

3.1.4.1 Adaptation effects in the MNS

Contrasting all incongruent target faces with all congruent target faces, whole-brain analyses showed significantly higher activation in fusiform gyrus (please see *Figure 7* and Table 14). Small-volume correction for the regions of interest confirmed increased activation in bilateral fusiform gyrus, and additionally in left IPL, bilateral BA44, right amygdala and bilateral insula (please see Table 15a).

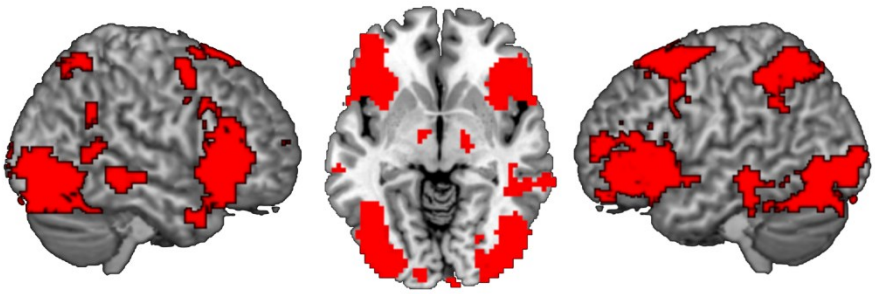


Figure 7. fMRI activation of the adaptation effects. Incongruent > Congruent, slice coordinate Z=67. Significance threshold for illustration purposes p<0.005, uncorrected, minimal cluster size k=10.

3.1.4.2 Adaptation effects depending on emotional valence

Whole brain corrected analyses revealed no significant effects. Small-volume correction however showed that activation was stronger for a switch from positive to negative valence than from negative to positive valence in right BA44, STS and insula (please see Table 15b).

Table 14. Functional brain imaging results for incongruent>congruent, FWE-corrected at whole-brain level, with p < 0.05 FWE-corrected, k = 10.

Incongruent > congruent			MNI			
Area		Cluster size	x	y	z	t-value
Fusiform gyrus	(right)	122	42	-49	-19	6.49
			45	-70	-10	6.31
			42	-82	-10	6.17
Fusiform gyrus	(left)	13	-39	-49	-22	5.78

Table 15. Functional brain imaging results for the different contrasts, small-volume corrected for the regions of interest with p < 0.05 small-volume corrected, k = 10. Note: BA44 = Brodmann Area 44, STS = Superior temporal sulcus

a)

(incongruent_HF>congruent_FF) - (incongruent_FH>congruent_HH)			MNI			
Area		Cluster size	x	y	z	t-value
BA44	(right)	123	39	11	8	3.858
Insula	(right)	97	39	11	8	3.858
STS	(right)	36	60	-64	20	3.108

b)

incongruent>congruent			MNI			
Area		Cluster size	x	y	z	t-value
BA44	(left)	139	-45	11	11	3.830
			-42	20	5	3.796
			-48	20	5	3.753

BA44	(right)		163	48	20	20	4.614
				51	20	5	4.124
Amygdala	(right)		32	21	-10	-13	3.930
				30	-1	-19	3.517
				24	-4	-16	3.456
Inferior parietal lobe		(left)	403	-45	-37	44	4.177
				-36	-55	56	4.088
				-33	-61	47	4.058
Fusiform gyrus		(left)	343	-39	-49	-22	5.777
				-39	-58	-16	5.280
				-45	-61	-22	5.187
Fusiform gyrus		(right)	341	42	-49	-19	6.491
				30	-85	-22	4.925
				36	-82	-22	4.837
Insula	(left)		120	-33	14	-4	4.792
				-45	11	11	3.830
Insula	(right)		112	36	14	-4	3.768

3.1.5 Discussion

Mirror neurons provide an intriguing explanation how we understand others. Using an adaptation design, our study provides evidence that the MNS is sensitive to the valence of emotions, and that MNS regions respond more strongly to a change from positive to negative than from negative to positive valence of facial configurations.

Almost 30 years of MN research brought us a deeper understanding of MN. It was conclusively shown that monkeys have neurons with mirror properties that are involved in action understanding and even allow predictions. While ethical restrictions prevent MN research in humans, the regions that are supposed to contain MN can be investigated non-invasively, showing that activation in the MNS goes beyond mere action understanding, but is also involved in social cognition, such as emotion recognition and theory of mind (Mier, Lis, et al., 2010). MNS activation during emotion perception, or during imitation of facial configurations goes in concert with activation in amygdala and insula. It is assumed that the MNS conveys information on the motor state of another person via insula to the amygdala which adds the emotional meaning (Carr et al., 2003). In agreement with our hypotheses, amygdala and insula responses were enhanced when adaptor and target were of different valences than when of the same valence / emotion, indicating adaptation effects. In addition, our results point to an influence of emotional valence on MNS activation, suggesting that the MNS is not merely representing a motor state, but might also be involved in coding the valence thereof. Adding to this assumption, MNS regions responded more strongly to a change from positive to negative than from negative to positive, which further suggests a higher sensitivity of MNS regions to negative emotions. Furthermore, a study by Cheng and colleagues (Cheng et al., 2006) reported a modulation of MNS response by motivation, and Enticott and colleagues (Enticott, Kennedy, Bradshaw, Rinehart, & Fitzgerald, 2010) found activation in the MNS only for meaningful, but not meaningless gestures. Considering these findings together, it can be assumed that a common feature influencing the MNS response is the salience of a stimulus.

We found higher activation for facial configurations expressing fear following ones intended to express happiness than vice versa, however not in all of our ROIs, but restricted to BA44, insula and STS. With regard to these effects, or partial lack thereof, there are several aspects to consider: While the IFG, including BA44, is thought to be involved in the prediction of an action, and mainly coding the goal of an action, rather than the explicit motor state (El-Sourani, Wurm, Trempler, Fink, & Schubotz, 2018; Iacoboni et al., 2005; Johnson-Frey et al., 2003; Newman-Norlund, van Schie, van Hoek, Cuijpers, & Bekkering, 2010; Nicholson, Roser, & Bach, 2017; Wurm, Hrkac, Morikawa, & Schubotz, 2014), the IPL seems to represent the explicit motor state (Fogassi et al., 2005; Fogassi & Luppino, 2005). It might be possible that this pure motor representation is not sensitive to valences. While the amygdala is

mainly linked to negative emotions, several studies showed amygdala activation for positive emotions, too, suggesting a broader role in emotion processing (Adolphs, 2010; Habel et al., 2007; Sergerie et al., 2008). Taken together, our results indicate that regions of the MNS, as well as amygdala and insula are sensitive to emotional valence. Additionally, BA44, STS and insula seem to react preferentially to a change from positive to negative valence. Another interesting result is the sensitivity of the fusiform gyrus to valence. It is well-known and also confirmed by adaptation designs that the fusiform gyrus responds to the identity of a person (Axelrod & Yovel, 2015; Grill-Spector, Knouf, & Kanwisher, 2004; Haxby et al., 2000; Winston, Henson, Fine-Goulden, & Dolan, 2004). Our results are in line with research reporting participation of the fusiform gyrus in emotion perception (Kawasaki et al., 2012; Monroe et al., 2013), suggesting that it is also involved in coding emotional valence. In the meta-analysis by Fusar-Poli and colleagues (Fusar-Poli et al., 2009), the fusiform gyrus was linked to both, happiness and fear.

A limitation is that results based on the fMRI BOLD technique cannot distinguish the responses of closely neighboring neuronal populations as these populations are smaller than voxel size, as each voxel contains the signal of several hundred thousand to millions of neurons. This might not only explain the absence of a differential response in fusiform gyrus, but might have confounded our results in general. Even though the adaptation design allows targeting different neuronal populations within a brain region, resulting in higher spatial precision, it is possible that adjacent neuronal populations responded to the different valences, one population only to fear, the other only to happiness, resulting in activation within the same voxels. Consequently, higher activation in MNS regions to a switch from positive to negative might be caused by more neuronal populations responding to fear (or negative emotions in general) than to happiness. Thus, albeit the adaptation approach is highly promising, it does not solve the spatial resolution deficits of fMRI. Related to this aspect, we want to mention that we only investigated regions assumed to contain MN – current fMRI techniques, no matter how advanced experimental designs or analysis methods are, allow no direct conclusions about neurons, but only about oxygen consumption in brain regions. Nevertheless, future studies might use facial configurations intended to express different negative emotions (e.g. anger and fear) to investigate the sensitivity of MNS regions not only to valence, but to differences in emotions. While valence is considered a central aspect differentiating emotions, our findings can only support a sensitivity to

valence, not to a specific emotion. Furthermore, Barrett and colleagues (Barrett et al., 2019) point out that specific emotional states can be associated with diverse facial expressions, and that perception of facial configurations varies widely between individuals, usually going beyond the set of basic emotions that underlie most scientific studies. In addition, the authors argue that while people learn to associate stereotypical facial configurations with certain emotions, previous studies could more convincingly provide evidence for our reliable expertise at discriminating positive from negative valence in facial configurations. Therefore, our focus on valence might have been more robust and in better agreement with current literature than when focusing on different types of emotions. When referring to the recent review by Barrett and colleagues (2019) another objection could be that we cannot be sure that our participants perceived smiling faces as expressing happiness, or wide-eyed faces as expressing fear. However, a possible categorization different from the one intended by us, would have led to smaller neuronal adaptation effects.

3.1.6 Conclusions

Our fMRI adaptation study shows that regions of the MNS are sensitive to the valence of facial configurations, in particular to the switch from positive to negative emotions, rather than vice versa. These results further support the role of the MNS in concert with amygdala and insula in social cognition and encourage the use of fMRI adaptation designs. Together, our findings suggest that regions of the MNS not only represent the observed motor state, but might process the affective meaning of the state, helping us to differentiate between emotions. The philosopher Gallagher suggests that we might have a direct understanding of others without any deeper, cognitive processing (Gallagher, 2008). The MNS might be essentially involved in such a direct understanding.

3.1.7 Acknowledgements

Our heartfelt thanks go to Julian Schlierkamp and Zhimin Yan for their assistance in data acquisition and in preparation of the results as well as to the participants, without them this project would not have been possible. We are grateful for funding of this project by the Heidelberger Akademie der Wissenschaften.

3.2 Summary

In study 2, an fMRI adaptation design was used to investigate whether the MNS is sensitive to the emotional valence of facial configurations. In an independent sample of 76 participants, changes in emotional valence were reflected in several regions, including amygdala, fusiform gyrus, insula, IFG, STS, and IPL. Interestingly, when a happy face preceded a fearful face, activation was even stronger in IFG, insula and STS than for the reverse contrast. Thus, all areas of the MNS seem to be sensitive to the emotional valence of faces. In addition, the other regions that were involved in the task support the assumption of the task requiring automatic, fast or implicit processing, by being directly part of the MNS or associated with the x-system. If the MNS is indeed sensitive to emotional valence, how will the decision on an emotion be made in case of ambiguous stimuli, presenting more than one emotion?

While studies 1 and 2 focused on the MNS, the focus is broadened for study 3, because in addition to the automatic processing of the affective facial information, probably slower and more deliberate cognitive processes come into play when facial configurations are ambiguous. In study 3, I investigated the role of probabilistic reasoning regions, such as DLPFC and parietal cortex, as well as of amygdala and Nacc for decision making regarding ambiguous facial stimuli. By again using facial configuration intended to express fear or happiness, additional insight can be gained into whether salience or reward drives the final decision.

4 STUDY 3: DEFINING THE ROLE OF DECISION MAKING FOR SOCIAL-COGNITIVE PROCESSES

4.1 Nucleus accumbens activation is linked to salience in social decision making

4.1.1 Abstract

Objective: Aberrant salience may explain hasty decision making and psychotic symptoms in schizophrenia. In healthy individuals, final decisions in probabilistic reasoning tasks are related to Nucleus accumbens (Nacc) activation. However, research investigating the Nacc in social decision making is missing. Our study aimed at investigating the role of the Nacc for social decision making and its link to (aberrant) salience attribution.

Methods: 47 healthy individuals completed a novel social jumping-to-conclusion (JTC) fMRI-paradigm, showing morphed faces simultaneously expressing fear and happiness. Participants decided on the 'current' emotion after each picture, and on the 'general' emotion of series of faces.

Results: Nacc activation was stronger during final decisions than in previous trials without a decision, particularly in fear rather than happiness series. A JTC bias was associated with higher Nacc activation for last fearful, but not last happy faces.

Conclusions: Apparently, mechanisms underlying probabilistic reasoning are also relevant for social decision making. The pattern of Nacc activation suggests salience, not reward, drives the final decision. Based on these findings, we hypothesize that aberrant salience might also explain social-cognitive deficits in schizophrenia.

4.1.2 Introduction

Daily, we are faced with the task to recognize other people's emotions. Whereas sometimes, the emotion is very clear and easy to recognize, at other times the facial expression is more subtle or ambiguous, requiring an active decision about the perceived emotion. While this can present a challenge even for people without mental disorders, it is especially difficult for patients with schizophrenia who have impairments in emotion recognition (Kohler et al., 2003; Kohler, Walker, Martin, Healey, & Moberg,

2010) and decision making (Heerey, Bell-Warren, & Gold, 2008; Moritz & Woodward, 2005). Based on the dopamine hypothesis of schizophrenia (Howes & Kapur, 2009), we hypothesize that decision making for emotions can be disturbed by aberrant Nucleus accumbens (Nacc) activity. To test this assumption we developed a new experimental paradigm that combines emotion recognition with decision making and applied it to a group of healthy participants.

Nacc and the fronto-parietal network, including parietal cortex and dorsolateral prefrontal cortex (DLPFC) are key regions for decision making (Matthews, Simmons, Lane, & Paulus, 2004; Philiastides, Auksztulewicz, Heekeren, & Blankenburg, 2011; St Onge, Ahn, Phillips, & Floresco, 2012; Zalocusky et al., 2016). Final decisions during probabilistic reasoning tasks are related to increased activation in ventral tegmental area (VTA) and Nacc in healthy participants, whereas schizophrenia patients (SZ) have reduced activation in these areas (Rausch et al., 2015). The Nacc, which is a part of the ventral striatum, has a high density of dopamine receptors and is a central region for motivation, reward and pleasure (Kringelbach & Berridge, 2010) with a major role for reward anticipation (Kringelbach & Berridge, 2010; Sabatinelli, Bradley, Lang, Costa, & Versace, 2007) and salience attribution (Berridge, 2006; Esslinger et al., 2013; Kapur et al., 2005). Dysfunction of the dopaminergic system appears to build the foundation of deficits characteristic for SZ which led to the “dopamine hypothesis of schizophrenia” (Howes & Kapur, 2009). In particular chaotic dopaminergic signaling in the Nacc has been proposed to be causal to the aberrant salience attribution in SZ, and hypersalience, i.e. enhanced salience attribution to seemingly neutral objects, has been assumed to cause delusions (Kapur et al., 2005).

Hasty decision making is known to occur in schizophrenia (Moritz & Woodward, 2005). A recent meta-analysis confirmed that people with psychosis decide based on significantly less evidence than healthy, as well as clinical populations without psychosis. Importantly, it was shown that a JTC bias is specifically linked to delusions (Dudley, Taylor, Wickham, & Hutton, 2016). In decision making, hypersalience may put too much weight on current information, leading to insufficient data gathering and thus hasty decisions, also called jumping-to-conclusion (JTC) bias (Speechley, Whitman, & Woodward, 2010). While hypersalience may cause the JTC bias in (non-social) decision making tasks, it may lead to wrong attributions of emotions and mental states to others in emotion recognition (Blackwood, Howard, Bentall, & Murray, 2001), which again may support the emergence of delusions (Mier & Kirsch, 2017).

Interestingly, there is evidence that the deficit in emotion recognition in SZ is most pronounced for ambiguous or neutral facial expressions (Kohler et al., 2010; Mier, Lis, et al., 2014). Ambiguous facial expressions are defined by the existence of more than one emotion, making a decision process for emotion recognition necessary. Since neutral facial expressions are defined by the absence of any emotion, false emotion recognition always implies the false perception of an emotion; i.e. a false positive decision for the existence of an emotion. Thus, hasty decision making and hypersalience might have a special role for biases in the recognition of ambiguous and neutral facial expressions, possibly suggesting an interaction of disturbed decision making and emotion recognition.

Aberrations in Nacc activity in SZ have not only been shown for decision making (Rausch et al., 2014), but present a stable finding for reward anticipation in SZ (Juckel et al., 2006; Morris et al., 2012). Thus, it is an interesting question whether reward or salience is the mechanism causing enhanced Nacc activity during final decision making, and consequently aberrant Nacc activity in SZ. Esslinger and colleagues found no differences in Nacc activation between rewarded and unrewarded final decisions, and concluded that Nacc activity reflects salience rather than the rewarding impact of the last stimulus (Esslinger et al., 2013). However, more studies directly investigating factors influencing Nacc activity during the final decision are necessary. Additionally, our knowledge on Nacc activation in decision making is based on “non-social” decision making tasks, leaving the question of the role of the Nacc for decision making during emotion recognition, i.e. in social decision making. Usually, emotion recognition has been associated with activation in the amygdala (Sergerie et al., 2008), best known for its role in fear processing, including recognition of fearful faces (Gläscher, Tüscher, Weiller, & Büchel, 2004; Hariri, Tessitore, Mattay, Fera, & Weinberger, 2002; Öhman, 2005). Moreover, the amygdala can reflect the salience of facial expressions (Santos et al., 2011). To a lesser extent than in fear, the amygdala is also activated in other negative and even positive emotions, including happiness (Phan, Wager, Taylor, & Liberzon, 2002). The anticipation of reward is linked to Nacc activation, for both monetary and social reinforcers (Izuma et al., 2008; Spreckelmeyer et al., 2009). Watching happy and attractive faces is considered rewarding, and activates the Nacc (Aharon et al., 2001; Hahn & Perrett, 2014; Phan et al., 2002). Facial expressions with a negative valence (e.g. fear, anger) however are not considered rewarding, but instead indicating threat and thus aversive conditions

(Anderson et al., 2007; Bishop, Duncan, & Lawrence, 2004). They are detected more quickly and accurately than happy faces in face-in-the-crowd tasks, suggesting higher salience for negative than positive facial expressions (LoBue, 2009; Pinkham, Griffin, Baron, Sasson, & Gur, 2010). Importantly, the Nacc maintains connections to the amygdala (Jackson & Moghaddam, 2001), suggesting interactions of amygdala and Nacc that might also be relevant for emotion recognition. Thus, both fear and happiness can be salient, but evidence suggests that fear is more salient than happiness and usually fear is not rewarding. We apply the knowledge of tasks using emotional stimuli to our task with the following logic: If increased Nacc activity during final decision making is associated with reward rather than salience, we expect it to be more prominent in the case of happy than fearful final stimuli.

Taken together, findings on final decision making in schizophrenia suggest reduced activation in the Nacc (Rausch et al., 2014), while based on the dopamine hypothesis (Kapur, 2003) enhanced Nacc activation would be predicted. Further, until now it is not clear whether this reduced Nacc activation during final decision making in SZ is based on aberrant salience, or aberrant reward anticipation. Since patients with SZ show impaired decision making (Dudley et al., 2016), as well as deficits in social cognition (Savla, Vella, Armstrong, Penn, & Twamley, 2012), investigating the interaction of these processes seems highly warranted. In our fMRI study, we include healthy participants to investigate whether findings from non-social JTC tasks can be replicated in a novel social JTC paradigm, which requires emotion recognition in mixed (morphed) facial expressions. We hypothesize that probabilistic decision making for emotion recognition leads to activation in the fronto-parietal network and that the final decision of a probabilistic reasoning process is linked to Nacc activation. By using faces showing fear and happiness in varying degrees, we want to explore whether salience or reward is linked to Nacc-activity during final decision making. If Nacc activation is related to reward rather than salience, Nacc activation should be stronger for happiness than fear. To get first evidence of the link between schizophrenia pathology and activation in the Nacc during final decision making in this social decision making task, we a) assess personality traits (schizotypy), and measures of social functioning (social network size and diversity), and b) compare participants according to their decision behavior.

4.1.3 Materials and Methods

4.1.3.1 Participants

47 healthy, right-handed Caucasian individuals with a general qualification for university entrance (29 women, 18 men; mean age 23.4 years (± 3.6), range 18-33 years) underwent functional magnetic resonance imaging in a Siemens Magnetom Trio 3T (Central Institute of Mental Health in Mannheim, Germany). Exclusion criteria were all assessed based on self-report and comprise a history of neurologic or psychiatric disease and presence of other diseases which require constant medication, as well as the general exclusion criteria for fMRI.

4.1.3.2 Study Procedure

The experiment was conducted as part of a study that was approved by the ethics committee of the University of Heidelberg and in agreement with the declaration of Helsinki. Participants were informed about study aims and procedures, signed written informed consent, received oral and written instruction on the paradigm, and completed a battery of questionnaires. Before the MR session, each participant practiced the paradigm until it was familiar and clear. Practice runs entailed the same identities as those used in the experiment. In contrast to experimental stimuli, which were based on fearful and happy facial expressions, practice stimuli were morphs between angry and happy, or between disgusted and happy faces. In the MR scanner, participants held a Current Designs 4-button diamond device in their right hand and watched the paradigm via video goggles. Prior to the experimental task and measurement, an MPRage anatomical measurement was performed, during which a nature movie was shown, so participants could get acquainted to the MR environment. Participants were reimbursed with 15€.

4.1.3.3 Experimental Design

In the style of the classical beads task (Huq et al., 1988) and the modified JTC task (Esslinger et al., 2013; Woodward et al., 2009), we developed a social JTC

paradigm (“Jemo”), which combines recognizing emotions in emotionally ambiguous faces with decision making. The happy and fearful facial expressions of 6 Caucasians (3 women, 3 men) of the NimStim Face Stimulus Set (http://www.macbrain.org/resources.htm, Tottenham et al., 2009) were selected, and for each individual, the happy and fearful face were morphed in 5 % steps, ranging from 0% (0% fearful, 100% happy) to 100% (100% fearful, 0% happy). The morphed pictures were taken from Matzke and colleagues (2014). In a pilot study, 25 healthy students judged which of the two emotions was predominant in each picture. Based on these ratings, we determined 7 morphs per stimulus person with the 4th morph being close to a 50/50 rating across participants, and the other 6 morphs having an increasing percentage of fear (3) or happiness (3) (see Figure 8 for examples).

In the Jemo paradigm, the most ambiguous 4th morph is presented as the first stimulus in a series of maximum 5 pictures. Each of the following morphs is less ambiguous, either more happy or more fearful. Every series of 5 stimuli has one incongruent stimulus, in which the recessive emotion prevails. On average, the incongruent morph consists to 77% (range: 61-92%) of the recessive emotion. The incongruent trial appears in 2nd, 3rd, or 4th position. The task of the participants is to identify a) the emotion of each stimulus (referred to as current emotion), and b) the predominant emotion in a series (referred to as general emotion) as soon as possible. If the participant correctly identifies the incongruent stimulus, the current emotion is correct. However, if the participant wrongly determines the incongruent stimulus to reflect the prevailing emotion within the series, this is considered an incorrect decision on the general emotion. Each picture is presented for 2 seconds, after which participants have 2 seconds to decide on the emotion displayed in this picture (current emotion), indicate their certainty about the decision within 4 seconds, and decide within 2.5 seconds whether they want to see another picture or already know the general emotion; in the latter case, they subsequently have 2 seconds to decide on the general emotion. Stimuli within a trial are presented with a jittered inter-stimulus-interval of 1s±0.5s, distinct series are separated by a jittered inter-trial-interval of 2s±1s. A fixation cross is presented during the inter-stimulus/trial-intervals. There are 24 trials in total (6 identities with 2 emotional directions, all presented twice). Duration of the experiment was dependent on the number of stimuli participants needed for a decision. They were told however, that the experiment takes around 15 minutes and were not aware that taking fewer stimuli to decide would reduce experimental time.

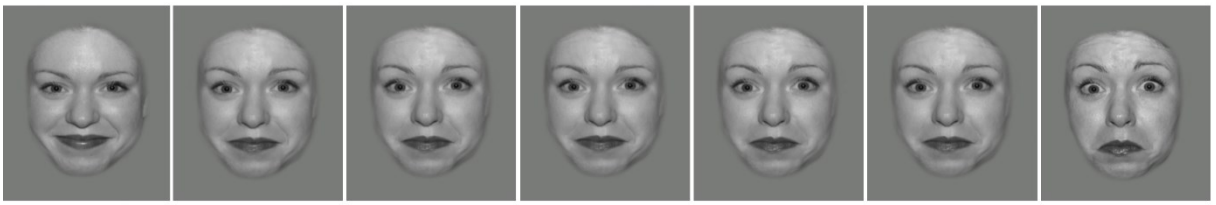


Figure 8. Example of stimuli: left: most happy picture, middle: 50/50 morph fearful-happy, right: most fearful picture, increments in between.

4.1.3.4 Questionnaires

Participants completed questionnaires assessing schizotypy (Schizotypal Personality Questionnaire, SPQ (Raine, 1991)) and social network behavior (Social Network Index, SNI (Cohen, Doyle, Skoner, Rabin, & Gwaltney, 1997)). Schizotypy refers to a combination of personality traits that largely overlap with symptoms of schizophrenia, both behaviorally and neurobiologically (Ettinger et al., 2015). The SPQ consists of 9 subscales, and includes the central aspects of schizoptypy such as constricted affect, unusual perceptual experiences and suspiciousness. The SNI assesses social ties in the private and professional environment, and is evaluated regarding three subscales: 1. Network Diversity, which reflects the number of social roles, in which the individual has regular contact, e.g. parent, child, spouse, employee, neighbor. 2. People, which counts the total number of people an individual is in regular contact with. 3. Roles, which reflects the number of different network domains in which an individual is active, which is based on the number of high-contact people in each network, e.g. family, work, neighbors. Previous studies suggest reduced social networks already in people with subclinical psychotic experiences (Gayer-Anderson & Morgan, 2013).

4.1.3.5 fMRI Data Acquisition and Analysis

fMRI data was acquired using a 12 channel head coil in a 3 T Siemens Magnetom Trio at the Central Institute of Mental Health in Mannheim, Germany. During the tasks, we used echo-planar imaging with 32 descending 3x3x4 mm slices including

1 mm gap, TR=2000 ms, TE=30ms, flip angle=80°, field of view=192mm, matrix=64x64.

Data were analyzed using SPM8 (www.fil.ion.ucl.ac.uk/spm/software/spm8/), with preprocessing including slice time correction, realignment, normalization to MNI space with resampling to 3x3x3mm voxels, and spatial smoothing with a 8mm full-width half-maximum Gaussian filter. We used a high pass filter of 512 seconds. First-level analysis included 7 regressors (last-fearful-face, last-happy-face, previous-fearful-faces, previous-happy-faces, happy-block, fearful-block, key presses) in a hybrid design modelling tonic (i.e. blocks of probabilistic reasoning) and phasic activity (i.e. events of final decision making and events without a final decision), according to our earlier publications with a non-social decision making paradigm (Esslinger et al., 2013; Rausch et al., 2014). The purpose of this hybrid design was two-fold: a) it allows analyzing phasic as well as tonic responses occurring in the experiment, and b) activation revealed with event-modulation is attributable to phasic effects under control of tonic effects, while the opposite is true for block-modulation. The contrasts for probabilistic reasoning were blocks with faces increasing in happiness and blocks with faces increasing in fear (> baseline fixation cross; block modulation). The contrast for final decision making was the difference in activation between the last face and all previous faces (event modulation). We also analyzed the interaction of brain activation during fearful last versus previous stimuli in comparison to happy last versus previous stimuli (event modulation).

In second-level random-effects group analyses, we applied t-tests to the contrasts of interest. Our regions of interest (ROI) included BA40 and BA7 (parietal cortex), BA46 and BA9 (DLPFC) for probabilistic reasoning, and Nacc and amygdala for final decision making. The masks were taken from the wakeforest university pickatlas (WFU Pickatlas, http://fmri.wfubmc.edu/software/PickAtlas). The Nacc mask was drawn according to an anatomic atlas and has already been successfully applied in our earlier studies with a JTC design (Esslinger et al., 2012). The significance threshold for whole-brain analyses was set to $p<0.05$, corrected for multiple testing using family-wise error (FWE), and a minimal cluster size of k=5 voxels. ROI significance was set to $p<0.001$, uncorrected, k=5 with $p<0.05$ small volume correction (svc) of the peak voxel. The Nacc mask had a size of 128 voxels on the left, and 93 voxels on the right side.

Questionnaires were analyzed with IBM SPSS Statistics 22 (Chicago, IL, US). Correlations of brain activation with questionnaires and behavior were calculated based on contrast estimates extracted from the Nacc ROI for the contrast all last faces > all previous faces (to assure the same number of voxels for eigenvariate extraction across participants, no significance threshold was set; i.e. p=1). In addition, behavioral subgroups of participants were compared with regard to their Nacc contrast estimates. Behavioral data were analyzed by repeated measures ANOVAs and t-tests.

4.1.4 Results

4.1.4.1 Behavior

Both, in blocks with increasingly happy and in blocks with increasingly fearful faces, subjects watched on average 3 pictures (happy: mean=3.02, SD=0.94; fearful: mean=2.97, SD=1.01; t(46)=1.04, p=0.306, d=0.05). There was no significant difference in performance between recognizing the current emotion in happy or fearful faces (happy: mean=62.74%, SD=22.92; fearful: mean=64.81%, SD=21.31; t(46)=0.37, p=0.71, d=0.09). Also, correctness of the decision on the general emotion within a block was not significantly different between the two emotion conditions (happy: mean=54.34%, SD=30.80; fearful: mean=60.72%, SD=30.91; t(46)=0.96, p=0.34, d=0.21). As illustrated in Figure 9, both accuracy and certainty of decisions increased with the number of stimuli considered. We performed a repeated measures general linear model to test the main effect number of stimuli on accuracy and certainty in fear and happiness blocks. There was a significant effect of stimulus number within a series on correct decisions in happiness blocks: $F(4, 116)=32.27$, $p<0.001$, $\eta_p^2=0.53$; on correct decisions in fear blocks: $F(4, 124)=58.48$, $p<0.001$, $\eta_p^2=0.65$; on certainty in happiness blocks: $F(2.77, 80.39)=55.08$, $p<0.001$, $\eta_p^2=0.66$; on certainty in fear blocks: $F(3.25, 100.69)=21.40$, $p<0.001$, $\eta_p^2=0.41$. Reaction times were not significantly different in the happy and fearful series (happy: mean=733ms, SD=142. fearful: mean=724ms, SD=127. t(46)=0.68, p=0.500, d=0.07). Experimental duration varied between participants depending on the number of stimuli they considered, and was on average 12.76 minutes (SD=3.84). The number of draws-to-decision (DTD) correlated significantly with the accuracy of decisions for the current (happiness block: r=0.512,

p<0.001; fear block: r=0.312; p=0.033) and for the general emotion (happiness block: r=0.481, p<0.001; fear block: r=0.453; p=0.0014).

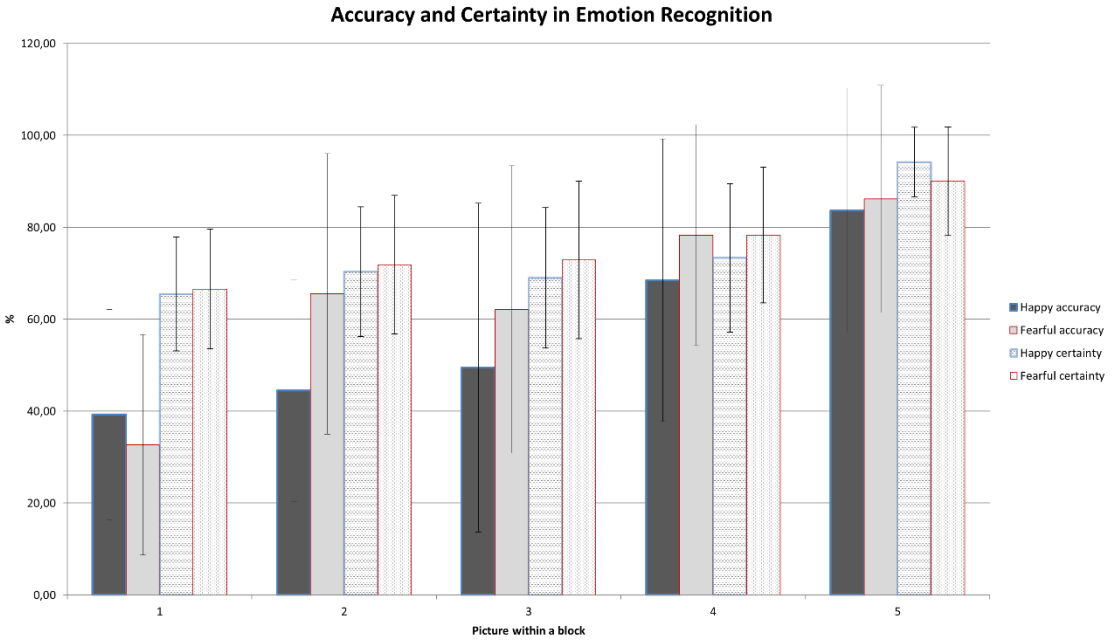


Figure 9. Mean percentages of correctly recognized emotions as bars with standard deviation; mean certainty as lines.

Exploratorily, we additionally analyzed whether the incongruent faces within each trial affected the decision process. On average, participants saw incongruent faces in 73.3% of trials (SD=23.4s). When participants decided on the general emotion immediately after the presentation of an incongruent face, performance was below chance level (25.4% correct responses, SD=32.2) with happiness as the dominant emotion within a series, and 34.8% correct responses (SD=35.9) with fear as the dominant emotion within a series. However, a decision on the general emotion of a series was only made after presentation of an incongruent face in 25.7% of series (SD=14.4).

4.1.4.2 Brain activation

Whole brain analyses revealed activation in the visual association cortex, and in parietal (BA7) and frontal (BA6 and BA44) lobe for fear blocks. During the happiness block, there was also enhanced activation in visual association cortex, parietal (BA7) and frontal (BA6) lobe. ROI analyses confirmed these results from the whole-brain

analyses . In both the fear block and the happiness block, there was activity in the DLPFC, and parietal cortex ROIs.

During all last faces compared to all previous faces, activation was increased in the bilateral putamen, with the cluster reaching into Nacc, and the anterior cingulate cortex (ACC; Figure 10). ROI-analyses confirmed enhanced Nacc activity for the last faces in comparison to all previous faces. This activation pattern was mainly driven by the fearful series comparing last to previous faces, but not by the contrast happy last versus happy previous. The interaction contrast (Figure 11) comparing the last fear face to all previous fear faces in comparison to the last happy face to all previous happy faces, revealed no significance at the whole brain corrected threshold, but ROI analyses showed stronger Nacc activation for fearful rather than happy last stimuli. None of the contrasts showed significant activation differences in the amygdala. Results of the whole brain analyses are presented in Table 16, ROI-analyses in Table 17.

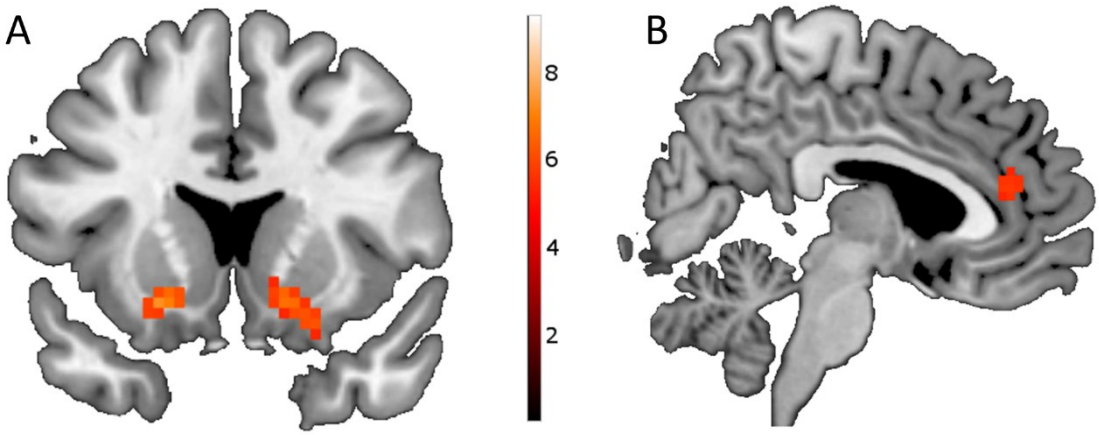


Figure 10. Whole brain activation for all last > all previous faces, p < 0.05, FWE-corrected, k = 5, at coordinates: x, y, z = 2, 14, -11. A) Nucleus accumbens, B) Anterior Cingulate Cortex.

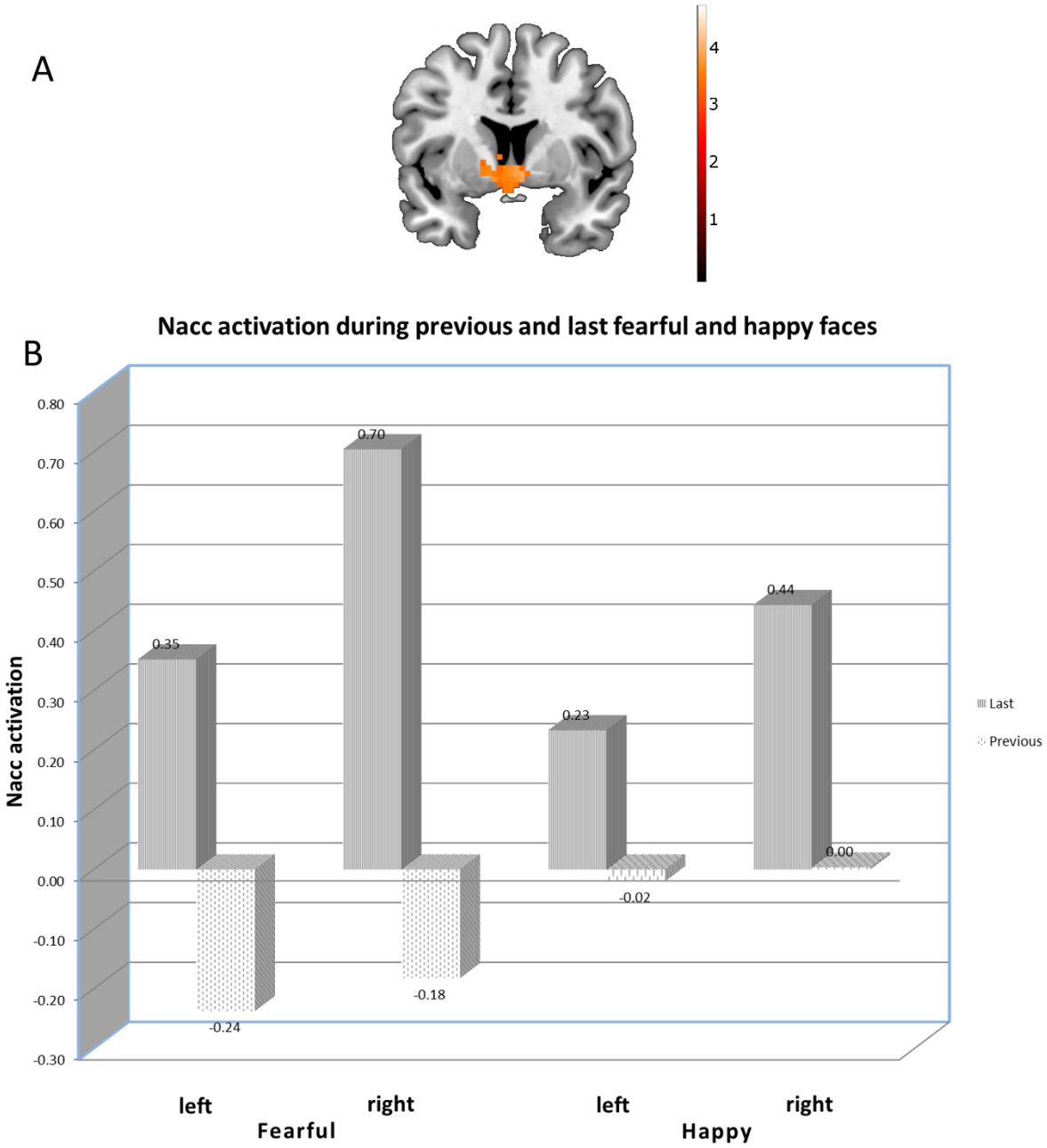


Figure 11. Interaction between last fearful and last happy face. A) Interaction (last fearful face greater than previous fearful faces compared to last happy face greater than previous happy faces), p < 0.001 uncorrected for display purposes., k = 5, (x, y, z = -9, 8, -8). B) Bars showing mean left and right Nacc activation for each of the conditions.

4.1.4.3 Brain-Behavior Associations

Our sample included an extreme group of 5 participants, who on average looked at less than 2 faces before deciding on the general emotion (1.45±0.37) which is

considered as JTC bias (Dudley et al., 2016; Rausch et al., 2015). We compared them to the other 42 participants (3.18±0.83; difference between groups: t(10.1)=8.30, p<0.001, d=2.70). We refer to the group with an average number of less than 2 faces as “L2” and the group with more than 2 faces as “M2”.

Even though L2 looked at fewer faces than M2 (happy: L2: 1.51 (±0.35), M2: 3.20 (±0.82), t(10.2)=8.33, p<0.001, d=2.68. fear: L2: 1.38 (±0.39), M2: 3.16 (±0.89), t(9.9)=7.90, p<0.001, d=2.59), they did not perform significantly worse in the decision on the general emotion within a series (happy: L2: 46.20% (±42.49), M2: 55.31% (±29.65); t(45)=0.62, p=0.538, d=0.25. fear: L2: 51.80% (±43.12), M2: 61.79% (±29.66), t(45)=0.68, p=0.501, d=0.27). However, comparing contrast estimates of L2 to M2 revealed higher activity during the last fearful face compared to the previous fearful faces in left Nacc (L2: 1.19 ± 0.55, M2: 0.24 ± 0.36, t(45)=5.22, p<0.001, d=2.04, see Figure 12), but not for the last happy face in comparison to the previous happy faces (L2: 0.10 ± 0.55, M2: 0.03 ± 0.37, t(45)=0.38, p=0.71, d=0.15).

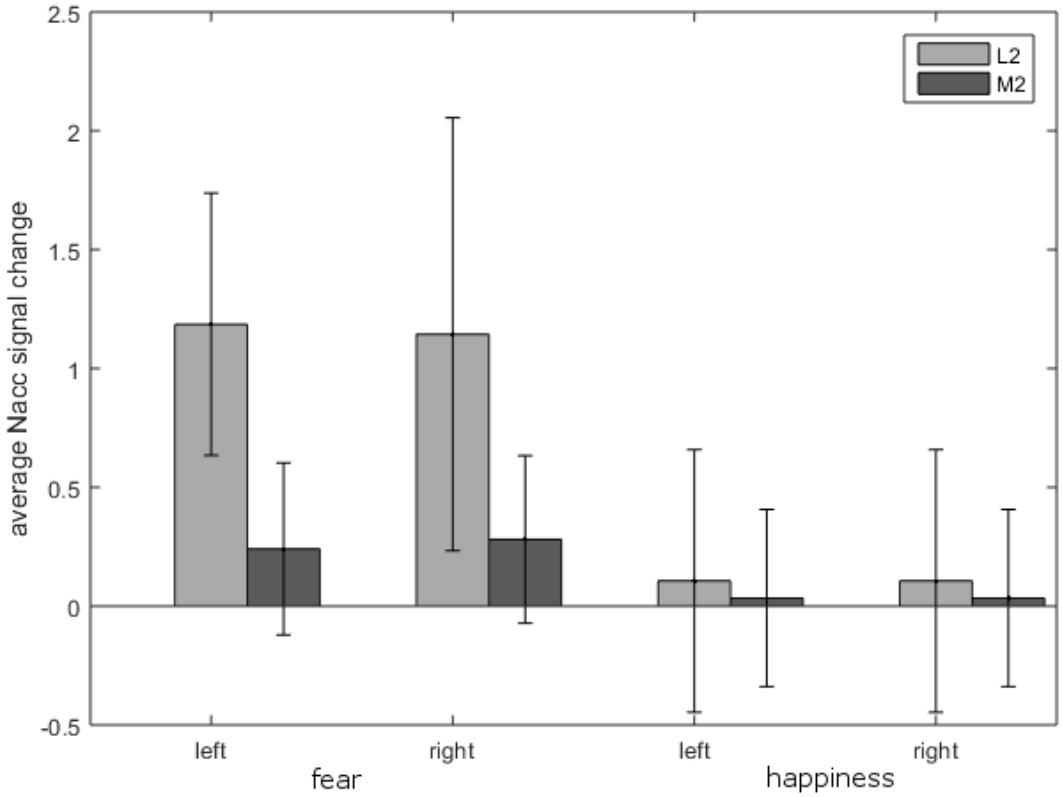


Figure 12. Bars showing mean left Nacc activation during the last fearful compared to all previous fearful faces and last happy compared to all previous happy faces for the extreme groups L2 (n = 5) and M2 (n = 42). Lines indicate standard deviation.

Exploratorily, to allow a correlation approach, instead of using the extreme group, we performed a median split of the whole group. Here, we excluded one outlier, who did not affect the fMRI results across the whole sample, but who drove many of the brain-behavior correlations that were no longer significant after excluding the person. With the median split, 23 persons had looked at less than 2.825 faces per block (L3), and 23 persons had considered more (M3). Analogous to L2 and M2, we refer to the groups as “L3” and “M3”.

As illustrated in Figure 13, the groups showed opposing correlations between Nacc activity for the last compared to all previous faces and the number of faces to reach a decision with a negative correlation in L3 (left: r=-0.48, p=0.02, right: r=-0.40, p=0.06) and a tendency for a positive correlation in M3 (left: r=0.35, p=0.10; difference of correlation strength to L3: z=-2.81, p=0.005, right: r=0.41, p=0.05, difference of correlation strength to L3: z=-2.7, p=0.003).

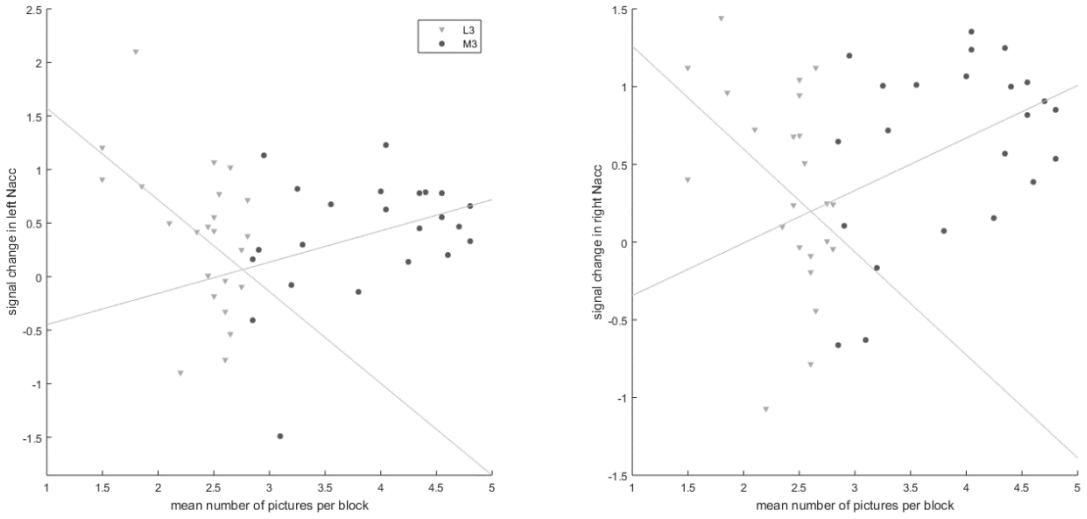


Figure 13. Correlation of Nacc activation during the last compared to the previous faces with the mean number of faces considered per block. Left: left Nacc. Right: right Nacc. Blue: subjects looking at less than 2.825 pictures per average block (L3), red: subjects looking at more than 2.825 faces per average block (M3).

4.1.4.4 Correlations of brain activity with questionnaires

Nacc activation during all last compared to all previous faces was negatively correlated with the number of social roles, assessed with the SNI (left: r=-0.42, p=0.004. right: r=-0.37, p=0.012), with the number of people one is in contact with (left: r=-0.27,

p=0.066, right: r=-0.31, p=0.034), and with the SPQ constricted affect scale (r=-0.32; p=0.029). No other correlations between Nacc activation for all last faces > all previous faces were significant. For the sake of comprehensiveness, it should be mentioned that no significant correlation between Nacc activation and schizotypy occurred when a four factor solution (Stefanis et al., 2004), instead of the nine factor solution of the SPQ was applied.

4.1.5 Supplementary Material

Table 16. Areas with significant activation during our contrasts of interest at whole-brain level, p < 0.05 FWE-corrected, k = 5. Note: The comparison of the last happy faces with all previous happy faces revealed no differences in brain activation at the given significance threshold and is thus not listed.

Contrast	**Area**	**Brodmann area**	**Cluster size k**	**MNI-coordinates**			**t-value**
				x	**y**	**z**	
All last > previous	Putamen left		41	-21	14	-11	7.23
	Putamen right		45	15	11	-11	6.74
	Anterior cingulate cortex	BA32	28	0	41	16	5.21
Fear last > previous	Anterior cingulate cortex	BA32	45	6	41	16	6.79
	Putamen left		27	-15	8	-11	6.58
	Putamen right		21	12	11	-8	6.09
Fear block	Visual Association Cortex, Occipital Lobe	BA18	176	12	-73	-5	10.29
	Pre-Motor and Supplementary Motor Cortex, Frontal Lobe	BA6	427	-21	-1	61	8.79
			125	27	-1	58	8.37
	Somatosensory Association Cortex, Parietal Lobe	BA7	21	-9	-61	58	6.21

	Inferior Frontal Gyrus, Frontal Lobe	BA44	5	-54	5	19	5.89
	Somatosensory Association Cortex, Parietal Lobe	BA7	8	15	-64	58	5.75
Happiness block	Visual Association Cortex, Occipital Lobe	BA18	157	12	-73	-5	10.24
	Pre-Motor and Supplementary Motor Cortex, Frontal Lobe	BA6	109	27	-1	58	8.29
			304	-21	-4	58	8.22
	Somatosensory Association Cortex, Parietal Lobe	BA7	10	-12	-64	55	5.76
			8	12	-61	58	5.58

Table 17. All significant ROI results for the examined contrasts, p < 0.001, k = 5, peak-level p<0.05 svc.

[1] (Last fearful > previous fearful) > (Last happy > previous happy).

Contrast	**Area**	**Hemisphere**	**Cluster size k**	**MNI-coordinates x**	**y**	**z**	**t-value**
All last>previous	Nacc	left	75	-21	14	-11	7.23
		right	76	15	11	-11	6.74
Fear last>previous	Nacc	left	96	-15	8	-11	6.58
		right	80	12	11	-8	6.09
Happiness last>previous	Nacc	left	10	-21	17	-8	3.96
		right	12	18	14	-11	4.23
Interaction pos[1]	Nacc	left	14	-9	8	-8	3.75
		right	9	6	8	-5	3.94
Fear block	BA7 and BA40	left	147	-9	-61	58	6.21
			62	-42	-34	43	5.70
		right	81	15	-64	58	5.75

			39	48	-31	46	4.34
	DLPFC	left	21	-57	5	31	4.91
			4	-39	32	31	4.16
		right	14	36	35	34	4.37
Happiness block	BA7 and BA40	left	123	-12	-64	55	5.76
			41	-42	-34	46	5.76
		right	68	12	-61	58	5.58
			36	48	-31	46	4.12
	DLPFC	left	18	-57	8	28	4.73
			2	-39	32	31	4.13
		right	9	36	35	34	4.14

4.1.6 Discussion

Our study aimed to investigate the neural correlates of social decision making and its link to (aberrant) salience attribution. Further, we planned on gaining evidence whether salience or reward is the driving factor for Nacc activation during final decision making. To this end, we used an emotion recognition task with morphed pictures, simultaneously expressing fear and happiness to varying degrees.

Our results successfully replicate the findings from previous studies using non-social stimuli. In line with Esslinger et al. (Esslinger et al., 2013), we found activity in the fronto-parietal network during probabilistic reasoning, as well as activity in the Nacc during final decision making. Interestingly, the enhanced activation of the Nacc was accompanied by activation in the anterior cingulate cortex (ACC). The ACC was shown to be involved in reward processing (P. Kirsch et al., 2003), emotional conflict resolution (Amit Etkin, Tobias Egner, Daniel M. Peraza, Eric R. Kandel, & Joy Hirsch, 2006), guiding voluntary choices (Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006), decision making (Rogers et al., 2004), and has a general role in regulating emotional and cognitive processing (Bush, Luu, & Posner, 2000). Thus in our emotion recognition task, the ACC response might reflect the conflict resolution and according decision for one of the emotions displayed in the morphed facial expressions.

Regarding Nacc activity, we were not only interested in the question whether we can replicate previous findings from JTC tasks without social stimuli, but also whether reward or salience is the driving factor during final decision making. Esslinger

and colleagues (Esslinger et al., 2013) concluded from the comparison of rewarded and unrewarded final decisions that salience, but not reward, results in Nacc activity during final decision making, whereas Sabatinelli et al. (Sabatinelli et al., 2007) found activation in Nacc and medial prefrontal cortex to be positively related to pleasantness and reward-value of pictures, but not to unpleasant pictures, salience or arousal. The interaction contrast, i.e. the activation during the last fearful face compared to the previous fearful faces in contrast to the last happy versus previous happy faces, revealed bilaterally enhanced Nacc activation. Hence, in support with the conclusion of Esslinger and colleagues (Esslinger et al., 2013), we assume salience, but not reward, to be the driving factor for final decision making. Interestingly, the ACC has also been shown to be involved in salience detection (Davis et al., 2005), strengthening this interpretation.

As schizophrenia (SZ) is associated with aberrant salience (Grace, 1991; Kapur, 2003), persons with SZ would be expected to have increased Nacc activity during final decision making. Accordingly, healthy participants who showed a JTC bias in our study indeed had enhanced Nacc activity during final decisions in fearful series. Additionally, analyses of the median-split groups hint toward a possible opposite pattern of Nacc activation between those with and without a JTC bias tendency, suggesting that fear is more salient to individuals looking at fewer pictures. Further, we revealed a positive association between DTD and performance across participants, linking impaired emotion recognition with hasty decision making. Thus, our results from healthy participants give first evidence that the aberrant salience hypothesis might be extended to explain biased emotion recognition.

Referring to the model of persecutory delusions by Freeman and colleagues (Freeman, Garety, Kuipers, Fowler, & Bebbington, 2002), it can be assumed that anomalous experiences and arousal are at the heart of the emergence of delusions. The authors define persecutory delusion as a threat belief that results from the interplay of various factors, such as emotions and beliefs, and general cognitive biases with these anomalous experiences / arousal. In our present study we demonstrate a link between aberrant salience (as reflected by enhanced Nacc activation) and a JTC-bias in emotion perception in healthy participants. We assume that a hasty decision about another person’s emotion can lead to wrong attributions during emotion recognition. The correlation between DTDs and accuracy supports this assumption, albeit not causally, but only on an associative level. While we are aware of evidence for a

negative bias during emotion recognition in SZ (Mier, Lis, et al., 2014), but not of evidence for a positive bias, we propose that aberrant salience does not have to lead to biased negative perceptions, but could also result in biased positive perceptions of emotions. Therefore, aberrant salience could result in anomalous experiences and arousal which would influence the interpretation of an emotion as positive or negative, dependent on the own current emotion and general biases. Since negative emotions, and in particular social and general anxiety, play a huge role in SZ (Achim et al., 2009), the probability for a false negative perception (i.e. a negative bias) is high, and with this the vulnerability for delusions with negative content, such as persecutory delusions, is increased. Thus, further studies are needed to show whether aberrant salience underlies the specific form of biased emotion recognition that occurs as negative bias for neutral facial expressions in schizophrenia (Mier, Lis, et al., 2014), or causes a general emotion bias.

Persons with SZ are known to have reduced social networks (Gayer-Anderson & Morgan, 2013). In agreement, correlational analyses with the questionnaires revealed a negative association between the diversity of social roles (spouse, neighbor, close friend, child, coworker, etc.), as well as the network size of individuals and their Nacc activation during the final decision. The additional negative association with constricted affect, as assessed with the SPQ, however, seems at first glance contradictory. Constricted affect belongs to the negative syndrome of schizotypy, and researchers assume opposing effects of negative and positive pathology on social cognition (C. D. Frith & Corcoran, 1996). Importantly, aberrant salience attribution and enhanced Nacc activity is linked to positive pathology, and in particular delusions (Kapur, 2003), while negative pathology has been found to go along with reduced Nacc activity (Wacker, Dillon, & Pizzagalli, 2009). However, since a reduced network size might also reflect social withdrawal and negative pathology, the correlational results should be interpreted carefully and warrant replication, especially, because the reported correlations are not corrected for multiple testing.

Several further questions should be addressed in future studies. We found differential activation of the amygdala neither for all last in comparison to all previous faces, nor for the last fearful faces in comparison to the last happy faces, suggesting a reduced role of the amygdala for final decision making. However, since the amygdala is an important brain region for emotion recognition (Mier, Lis, et al., 2010; Sergerie et al., 2008) and has been shown to be involved in salience processing (Santos et al.,

2011), further studies are needed to investigate the specific role of the amygdala in social decision making. In addition, to investigate the importance of different brain regions in social decision making more comprehensively, the presented social JTC-task might be analyzed with regard to the right temporoparietal junction and particularly its functional connectivity with the left hippocampus, which recently has been shown to be important for social decision making and social learning in the context of an iterated prisoner's dilemma game (Bitsch, Berger, Nagels, Falkenberg, & Straube, 2018).

To learn more about brain activation and networks involved in and relevant to the task, it is a necessary next step to invite persons with diagnosed schizotypy or SZ to complete the task; also, inviting healthy participants depending on their self-reported positive schizotypy symptoms, as well as comparing SZ patients with and without delusions would be of high interest. It is noteworthy that patients with SZ did not show Nacc hyperactivation in earlier studies with a non-social JTC task, but Nacc hypoactivation (Rausch et al., 2014). If patients also respond with reduced Nacc activation in the social JTC-task, it would be intriguing to find the tipping point in the course of the disease, or within the SZ spectrum, which separates increased from reduced Nacc activation and associated behavioral measures. However, since in our earlier studies we found hypoactivation in our non-social JTC task, not only for patients with SZ (Rausch et al., 2014), but also for individuals in an at-risk-mental state (Rausch et al., 2015), it can be rather assumed that the pattern of hypo- versus hyperactivity is stable across the course of the disease. This is an especially interesting and indeed controversial finding, because in the dopamine hypothesis of SZ, aberrant salience is clearly linked to enhanced subcortical dopamine responding and hyperfunctioning of the Nacc (Kapur et al., 2005; Maia & Frank, 2017). In agreement with an integrative framework of dopamine functioning for SZ (Maia & Frank, 2017), one explanation could be that positive pathology and in particular delusions are characterized by aberrant salience in the form of hypersalience and enhanced Nacc activation which would be linked to hasty decision making. On the contrary, aberrant salience in the form of hyposalience and diminished Nacc responding could be linked to slow decision making and negative pathology, such as apathy. Further, it should be mentioned that Nacc hypoactivity has not only been found during final decision making in SZ (Rausch et al., 2014), but is also a highly stable finding for reward anticipation in SZ that has been linked to deficient salience processing (Esslinger et al., 2012; Nielsen et al., 2012), reduced prediction error (Juckel et al., 2006), and the intake of typical antipsychotics

(P Kirsch, Ronshausen, Mier, & Gallhofer, 2007). Thus, our findings from the healthy sample are in agreement with predictions of the dopamine hypothesis and the theoretical framework of hasty decision making and hypersalience in SZ (Dudley et al., 2016; Kapur et al., 2005; Speechley et al., 2010), while the findings from SZ patients are not. Future studies should examine whether the proposed association between hasty decision making and delusions that has been confirmed on the behavioral level (Dudley et al., 2016) is also evident in studies investigating Nacc activation during final decision making for social stimuli in SZ.

A limitation is that we did not include a non-social control task, so we cannot directly compare social and non-social probabilistic decision making. Future studies including patients with SZ should test both, social and non-social decision making to examine the possibility of divergent activation patterns. Further, based on the observation that antipsychotic medication fails to normalize social cognition and emotion recognition abilities in patients (Kucharska-Pietura & Mortimer, 2013), it would be interesting to compare brain activity during the task in medicated vs. non-medicated patients. This might provide new insights into the specific effects of the medication with respect to social cognition, and hint towards requirements for drug improvement. In addition, future studies with a focus on the association between delusions and hasty social decision making, might use the emotions happiness and anger, instead of happiness and fear, because anger might be more suitable to cover the perceived threat in paranoid psychosis than fear. A further possible drawback is the usage of stimuli displaying disgust or anger and happiness for the practice trials. We aimed to avoid presenting stimuli that are used in the experiment. This however, might have led to higher salience for fear than happiness. Still, this neither explains the enhanced activation for the last versus previous fear faces, nor the interaction effect with the emotion (since lower salience should not only occur for the last happy face, but also for all previous happy faces for which we controlled when comparing the last fear with the last happy face). Finally, there was a large variability in block lengths within and between subjects, lasting from almost 20s to over 70s depending on the number of stimuli considered before deciding on the general emotion. Also, as the block number was fixed, the duration of the experiment depended on the number of stimuli considered. However, participants did not know they could influence the duration of the experiment with their choices. Thus, the measured Nacc-signal in the group with a JTC-bias might be more noisy (due to less trials for averaging the response to the

previous faces, or due to inferior model fit), but should not reflect aberrations in task motivation. In addition, there is evidence suggesting that activation in the Nacc is positively linked with the willingness for task effort (Green, Horan, Barch, & Gold, 2015; Schmidt, Lebreton, Cléry-Melin, Daunizeau, & Pessiglione, 2012) what is in disagreement with the assumption of reduced motivation causing the higher activation in Nacc and ACC in response to the last face in comparison to all previous faces. Still, we cannot rule out the possibility of reduced motivation influencing the perseverance during each block, and therefore block- and task-length, as well as brain activation.

4.1.7 Conclusions

We presented results from a social JTC paradigm that allows investigating the neural correlates of social decision making. We show for final decisions during emotion recognition that the Nacc a) together with the ACC shows strong differential activation, b) has higher activity in fear than in happiness series and c) has higher activity in fear series in participants with a JTC bias. Based on this first evidence from healthy participants, we suggest that the aberrant salience hypothesis of schizophrenia may be extended to explain biased social cognition. Future studies focusing on the impact of dopamine and salience attribution on social cognition in schizophrenia are highly warranted.

4.2 Summary

Study 3 was conducted to investigate the neural mechanisms involved in deciding on emotions in ambiguous facial configurations. Automatic processing associated with embodied simulation might not be sufficient to resolve the conflict of facial configurations with inconsistent emotions, so I expected that also brain regions associated with deliberate reasoning would be involved. 47 participants completed the social-cognitive JTC paradigm, which was an adaptation of the fish-in-the-lake task that had previously shown an involvement of fronto-parietal regions during probabilistic decision making and an importance of the ventral striatum for the final decision (Esslinger et al., 2012). Remarkably, patients with schizophrenia had reduced activation in the ventral striatum during reward anticipation which might be due to aberrant salience (Rausch et al., 2014).

Study 3 with ambiguous facial configurations, replicated the relevance of fronto-parietal regions during probabilistic decision making. Also in line with previous studies (Esslinger et al., 2013; Rausch et al., 2014), the final decision was associated with striatal activation in putamen reaching into Nacc. Importantly, activation in Nacc was stronger for fearful than for happy final faces, supporting the role of salience for Nacc activation in decision making.

These results, as well as the results from the other two studies, will be discussed in detail in the general discussion.

5 GENERAL DISCUSSION

In the present thesis, I aimed at deepening the understanding of neurobiological mechanisms underlying social cognition. Social cognition research gets its attractiveness not only from the fact that the topic is so central to our everyday lives and our evolution as a species, but also with regard to mental disorders that are often associated with complications in social interactions. Understanding the mechanisms therefore helps to better understand central functions of the brain, and might also build the foundation for therapy-oriented research.

In the three presented fMRI studies using pictures of facial configurations intended to express emotions, the focus was on the fast automatic processing which seems to be at least partially represented in the MNS. The greatest challenge is that we cannot measure MN in humans, but have to rely on indirect non-invasive methods instead. In study 1, I implemented a sVx analysis, which is considered more accurate than standard fMRI processing routines for application on the MNS (Gazzola & Keysers, 2009). The goal of the study was to determine whether imitation, affective empathy and theory of mind share a common neural basis and whether this basis is indeed located in the regions of the MNS. The results show activation in sVx in regions of the MNS over the three processes, suggesting a common neural basis for social cognition. To determine whether the regions of the MNS are also sensitive to the different valences in facial configurations, I implemented an fMRI adaptation design (de la Rosa et al., 2016; Winston et al., 2004) for study 2. The results indicate that regions associated with the MNS indeed differentiate between emotional valences. In real life, facial configurations are often ambiguous. The purpose of study 3 was to investigate which additional neurobiological mechanisms are involved in the processing of ambiguous facial configurations. Here, the focus was on the role of the Nacc, which is associated with directing salience (Berridge, 2006; Esslinger et al., 2013; Kapur et al., 2005), and the experience of reward (Knutson, Adams, Fong, & Hommer, 2001; Kringelbach & Berridge, 2010). Based on these results, we conclude that salience is a determining factor when deciding on the emotional content of a facial configuration.

The discussion section in publications is often limited due to restrictions on word count, so I will summarize and discuss the results of studies 1 – 3 adding further implications and conclusions in the following sections (5.1 to 5.3). In addition, I will integrate the results of my studies in two models, depicting the relationships between

social-cognitive processes and associated brain regions (5.4), discuss the implications for future research (5.5) and present final conclusions (5.6).

5.1 Shared voxels in MNS regions are involved in different social-cognitive processes

Extending on the results from previous studies, Gazzola and Keysers (2009) showed that the observation and execution of hand movements was related not only to activation in the same brain regions over all participants, but even within voxels and within participants. These sVx for hand actions were located in diverse regions, including ventral premotor (BA6/BA44), inferior parietal, middle temporal and somatosensory cortex.

I performed such a sVx analysis using facial stimuli intended to express anger or fear, which were presented in tasks requiring participants to perform imitation, affective empathy or theory of mind. So while the stimuli were the same for all conditions, the instructions and therefore the mental processes of the participants differed and led to differences in brain activation. While there are distinct activation patterns for each condition, there are also overlaps with sVx in several regions. In the next subsection, I will focus on the main findings of study 1, which concern the common neural basis of social cognition. Afterwards, I will discuss the additional findings from these tasks, which provide interesting insights, and build a foundation for future studies on the role of the MNS for social cognition.

5.1.1 A common neural basis of imitation, affective empathy and theory of mind

As my goal was to identify a common neural basis to social cognition, I considered it important to cover a wide spectrum of social-cognitive processes. While the tasks included even more processes, for the sVx and conjunction analyses I decided to select a mixture of processes of different levels of complexity. Specifically, these were imitation, affective empathy and ToM. Imitation, which can be considered a signature condition of MNS research (Molenberghs, Cunnington, & Mattingley, 2009), is thought to build the basis for other social-cognitive processes (Andrew N Meltzoff, 2002; Santiesteban, White, et al., 2012), and I expected MNS activation, because participants were simultaneously observing an emotional face and bringing

their own face into the same configuration. Affective empathy requires a shared affective state of observer and observed person (Decety & Jackson, 2006), and is therefore a hot emotional process (Abu-Akel & Shamay-Tsoory, 2011). Evidence indicates that affective empathy leads to even stronger MNS activation than cognitive empathy (Nummenmaa, Hirvonen, Parkkola, & Hietanen, 2008). ToM is considered a high-level social-cognitive skill (Santiesteban, White, et al., 2012) and can also be considered an affective process in our paradigm (Mier, Lis, et al., 2010). One important distinction between the empathy and ToM paradigms of study 1 is that to empathy, felt emotions are central, while to ToM, perspective taking is more crucial.

In line with Gazzola and Keysers (2009) for the execution and observation of actions, all three social-cognitive processes compared to a non-social control have a significant number of sVx in BA44 and IPL, which are commonly considered key regions of the MNS (Cattaneo & Rizzolatti, 2009; Rizzolatti & Craighero, 2004; Van Overwalle & Baetens, 2009). This supports the assumption that the MNS builds a basis for social cognition. Furthermore, the data revealed increased bilateral activation also in all other ROIs, namely amygdala, FG and STS. Amygdala and FG are thought to belong to the emotional face processing network (Haxby & Gobbini, 2011; Haxby et al., 2000), and the STS is assumed to be a crucial region of the mentalizing network (Carrington & Bailey, 2009; U. Frith & Frith, 2003), and also of the emotional face processing network (Haxby & Gobbini, 2011). In addition, amygdala, fusiform gyrus and STS are part of the fast and automatic x-system, proposed by Satpute and Lieberman (2006).

To confirm the findings of a shared neural basis using standard fMRI analysis methods, I performed a conjunction analysis, which is in agreement with the findings from the sVx analyses and further shows activation in regions adjacent to the ROIs, such as inferior occipital gyrus, which is also involved in face processing (Haxby & Gobbini, 2011; Pitcher, Walsh, & Duchaine, 2011). Interestingly, the occipital lobe is not considered in the model of the x- and c-system. Due to its role in face processing, I assume, that the inferior occipital gyrus would be part of the automatic x-system, as is also the fusiform gyrus.

The analysis against non-social control stimuli allows valuable insights. Still, it is crucial to analyze which of the shared activation goes beyond the processing of facial configurations perceived as emotional. Since two of the tasks in study 1 additionally contained a facial control, I also performed sVx and conjunction analyses for imitation

and ToM in contrast to the respective facial control condition. Again, for all ROIs, a significant number of participants shows a significant number of sVx. Remarkably, while the number of participants with sVx in bilateral BA44 and right IPL (in the smoothed analysis also left STS) is almost the same as for the three-task analysis, the numbers in amygdala and FG (for unsmoothed data also in STS) are considerably reduced (see p.73, Table 12). Conjunction analyses are in line with the sVx results and confirm increased activation in the area of BA44 and STS.

These results allow several conclusions regarding both the methodological and the neurobiological level. With regard to the method, the results from the sVx analysis demonstrate the impact of smoothing of fMRI data. Since not merely the same region, but the same voxel needs to be involved in all tasks within participants, smoothing can have a strong effect on voxel counts. This becomes obvious when looking at the results for IPL and STS, which yield almost twice the number of participants with sVx in the smoothed analysis compared to the unsmoothed data. Therefore, I would rather rely on the more conservative results of the unsmoothed sVx data as a basis for interpretation and would recommend researchers to base sVx analyses on unsmoothed data, too. With regard to the neurobiological implications, the reduced number of participants with sVx in FG in the comparison of the three-task with the two-task analysis can be explained by the fact that we controlled for face processing in the two-task analysis, and therefore both conditions contained face stimuli and differed only in the social-cognitive task applied to the face. The remaining percentage of participants with sVx might be explained by the finding that FG is also involved in the processing of emotions (Fusar-Poli et al., 2009; Geday et al., 2003), which likely also explains the remaining percentage in amygdala (Costafreda et al., 2008; Fitzgerald et al., 2006; Fusar-Poli et al., 2009; Habel et al., 2007; Sergerie et al., 2008). Interestingly, right IPL seems to be equally important for the two-task analysis as for the three task analysis, and left STS seems to contain more sVx than right STS. While I should refrain from overinterpreting these results, right IPL has been suggested to be particularly important for the MNS (Chong et al., 2008). Further research indicates a role of the right, but not the left IPL for self-other discrimination (Uddin, Molnar-Szakacs, Zaidel, & Iacoboni, 2006), an important element of social cognition. Regarding STS, I need to mention that the masks were based on the activations of a previous study. In these masks, the left STS contained almost twice the number of voxels than the right STS,

therefore offering greater opportunity for sVx, so I will not interpret the hemispheric differences for this ROI.

Overall, a larger number of participants having sVx in IPL and BA44 than in other ROIs, also when controlling for face processing, indicates that the intention with which individuals observe a face, i.e. the intention to imitate or to infer the displayed person's intentions, might be the decisive factor underlying involvement of these two regions.

While several regions involved in social-cognitive processing can be considered part of the automatic x-system, there is a need for further studies to disentangle the automaticity of single processes, and consequently the role of central regions, such as IFG.

To sum up, in line with previous studies (Carr et al., 2003; Mier, Lis, et al., 2010), the different social-cognitive tasks showed common activation in all ROIs which are associated with the processing of facial configurations perceived as expressing emotions (amygdala, FG), the MNS (BA44, IPL) and mentalizing system (here mostly represented by STS). These results were substantiated also with the sVx analysis, emphasizing that common activation patterns can not only be observed across but also within participants. The results therefore confirm those by Gazzola and Keysers (2009) and add to the applicability of the theory of embodied simulation to the social information perceived from faces.

5.1.2 Findings from individual task conditions

During imitation compared to social and non-social control, whole-brain analyses revealed that activation was also increased in BA 6, which includes PMC and supplementary motor cortex and lies adjacent to BA44. The activation pattern is line with a previous study on facial imitation (K. R. Leslie et al., 2004). Furthermore, a meta-analysis suggested, that the MNS is not restricted to BA44, but also includes BA6, combining them to the PMC region of the MNS (Van Overwalle & Baetens, 2009).

Challenging the idea of embodied simulation is the lack of activation in the MNS during the observation of facial configurations intended to express emotions compared to a non-social control task. In theory, the MNS would automatically activate upon detection of an emotional facial configuration in order to allow embodied simulation and therefore understanding of the other's mental state (Gallese et al., 2004).

However, the MNS key region BA44 shows no increased activation for observation compared to control. One possible explanation could be that the observation condition might have felt like “doing nothing” compared to the other conditions in the imitation task, which all involved moving one’s own face. Therefore, in comparison, the MNS might have been less activated. Especially, in the control condition, participants had to move their face (saying aloud German letters “A” or “Ä”), thereby naturally involving activation in the motor areas of the brain, which are overlapping with the MNS. So even if the observation of a face had actually relied upon the MNS, it might just not be visible in the applied contrasts due to lack of a suitable control condition for this question. Another possible interpretation for this lack of BA44 activation could be that the MNS is modulated by motivation and social relevance. As participants simply had to watch these pictures of faces, they might not have felt this situation socially relevant and were not motivated to infer mental states. This question was addressed in a follow-up study, designed to investigate the impact of motivation and intention on activation in the MNS. Analyses of this study are ongoing.

Another interesting finding was that of increased activation in the area of supramarginal gyrus, TPJ, and precuneus in distress compared to cognitive and affective empathy. Both, TPJ and precuneus are associated with self-referential processing and are part of the default-mode-network (Greicius et al., 2003). Indeed, the distress condition is the only empathy condition requiring subjects to explicitly shift their focus only to themselves, so it is expected that self-referential processing and associated brain activation is highest. The TPJ is also known for its role in self-other distinction (e.g., Kanske et al., 2015), and is involved in perspective taking (Costa, Torriero, Oliveri, & Caltagirone, 2008). Again, this might serve as an explanation for its involvement in the distress condition, as participants had to look at the facial configuration intended to express negative emotions of another individual while judging the strength of their own negative feelings. Dissolving this discrepancy can be considered highly demanding and might have strongly relied on the self-other distinguishing role of the TPJ. Interestingly, our findings regarding distress are in line with the proposal by Satpute and Lieberman (2006), who reviewed that self-focused processing was associated with activation in posterior parietal regions, possibly relying on symbolic representation to allow self-other distinction, and consequently part of the deliberative c-system.

Remarkably, and confirming previous results by Mier and colleagues (2010) who first developed the ToM paradigm, the results of study 1 revealed a stepwise increase of brain activation in STS and IFG from non-social control, over social control, emotion perception, to ToM. The fact that both regions differentiate between the experimental conditions and also in comparison to the control condition, indicate that solving the task might have relied on a combined process of mirroring and mentalizing. Possibly, IFG was spontaneously activated by the social-cognitive demands of the conditions. As participants had to explicitly decide whether or not the picture matched a previously shown sentence, there might have additionally been increased activation in STS. STS activation might have been strongest in the ToM condition, because participants should indicate the intention of the presented individual, which is likely to involve perspective taking. While the higher-order processes associated with ToM and STS would suggest a categorization of the STS in the slow c-system, it has previously been shown to respond very rapidly within 200ms, and therefore suggested to be part of the reflexive x-system (Satpute & Lieberman, 2006). One possible solution to this seeming contradiction would be that the STS is involved in both early and late processing, with different stages being linked to different social-cognitive mechanisms. So, for example, while gaze may be quickly processed in the STS, higher-order intention inference might occur at a later stage.

5.1.3 Implications from study 1 for future research

Overall, the sVx analysis presents a valuable approach in the context of MNS studies and I would recommend future studies with a special focus on common activations in the MNS basing their conclusions on comparably accurate methods. The tasks that were applied are thought to require different levels of social cognition, going from low-level imitation, thought to rely mainly on the MNS (Molenberghs et al., 2009), to high-level ToM, which is associated with the mentalizing system (U. Frith & Frith, 2003; Santiesteban, White, et al., 2012). Importantly, the study results indicate that different social-cognitive processes going beyond the mere processing of faces indeed rely on the same or at least closely neighboring neuronal populations. In addition, there are condition-specific differentiations between the activation of diverse empathic processes, observation of neutral faces, emotion perception and ToM. While most or all of the ROIs show increased activation in all conditions, the exact extent of the

activated area differs between conditions, and these differences in neighboring areas seem to represent important differentiations between the specific processes.

In a follow-up study, it might be interesting to be able to compare all social-cognitive conditions with each other to gain even more insights with regard to activation differences between processes. One study that comes close to this idea included emotion perception and ToM (Mier, Lis, et al., 2010), and already provided valuable insights into the relationship between these two processes. For example, and in line with the findings for a modification of the task that was used in study 1, reaction times were longer, and activation in regions including STS, IFG reaching into the insula, somatosensory cortex, amygdala and right middle frontal gyrus was stronger in ToM compared to emotion perception. Another study investigated empathy and ToM using a promising novel task design, called EmpaTom (Kanske et al., 2015) which uses video sequences of persons narrating emotional autobiographic events, the contents of which can be evaluated with regard to empathy or ToM elements. However, as mentioned in the introduction of this thesis, there exists a wide variety of definitions, tasks, and stimuli used for social-cognitive tasks, in particular empathy and ToM. So, future studies could for example aim to investigate an adaptation of the EmpaTom task, including even more social-cognitive concepts, such as the differentiation of affective and cognitive empathy, in addition to ToM. While study 1 included all these social cognitive processes, the design did not allow direct comparisons between some processes within one task, for example between empathy and ToM. In a future adaptation of the study, one might consider adopting a design which would additionally allow, for example, contrasting imitation with ToM, and thereby substantiating the assumed differences with regard to their social-cognitive complexity, or allowing the comparison of the empathy conditions with neutral faces or with ToM. Such as design could use sessions for the different tasks, including null events to statistically control for differences in session means. Obviously, each paradigm provides advantages, as does each type of stimulus. So possibly, to achieve a true understanding of social cognition, indeed a multitude of high quality studies, as the ones by Mier and colleagues (2010), Kanske and colleagues (2015) and study 1 of my PhD thesis (Schmidt et al, submitted) may be required, with the results complementing each other.

In order to better investigate the differences between the tasks and conditions, MVPA would also be a promising approach that I therefore plan to implement on the current data, as well as in future studies. MVPA has been famously applied in memory

research, where neural network classifiers were able to determine whether participants looked at pictures of faces, objects or locations, based on their fMRI activation (Norman, Polyn, Detre, & Haxby, 2006). With regard to social cognition, one study using MVPA showed that different emotion categories show individual multivoxel activation patterns in medial PFC and STS, regardless of expressing modality (face, body, voice) (Peelen, Atkinson, & Vuilleumier, 2010). MVPA could help specify distinctive activation patterns between social-cognitive processes, thereby furthering the understanding of mechanisms of social cognition. In addition, a major field of application would be in research on patient groups with social-cognitive deficits. It might be imaginable, that MVPA not only allows predicting whether an activation was recorded during cognitive or affective empathy, but also whether the individual had schizophrenia or another mental disorder.

One limitation that is common to imaging studies is that they allow no conclusions on whether activated regions are causal to the functioning of processes. In the introduction, I included results from lesion studies, which allow better understanding of the importance of lesioned regions. For example, while lesions in the IFG were associated with deficits in affective empathy, lesions in vmPFC were observed with impaired cognitive empathy (Shamay-Tsoory et al., 2009). Since patients with lesions are rare and researchers have no influence on the lesion, simulated lesions using transcranial magnetic stimulation (TMS) or transcranial direct-current stimulation (tDCS), which can transiently enhance or inhibit specific cortical regions, provide an elegant solution. For example, tDCS was used to confirm a role of the TPJ in self-other processing (Santiesteban, Banissy, Catmur, & Bird, 2012). And indeed, in the project around study 1, I also used inhibitory TMS over right IFG to determine its suggested role as a key region of the MNS for social cognition (Schmidt*, Popova* et al., in preparation).

To add to the existing literature, future studies could investigate the modulating effects of several individual factors. For example, in a sample with a greater age range, or even in a longitudinal study, researchers could investigate whether the common neural basis is stable over life time, or how activation patterns of individual tasks change. It would be particularly interesting to assess children and determine the age at which this common neural basis of social cognition is developed. While imitation can be observed in infants already a few days after their birth (Field, Woodson, Greenberg, & Cohen, 1982; A. N. Meltzoff & Moore, 1977), ToM ability is known to develop during

childhood (U. Frith & Frith, 2003). In general, one could assume that our social-cognitive skills are enhanced throughout our whole lives or dependent on the frequency and intensity of our social interactions. Indeed, it has been suggested that one reason for poor social functioning in individuals with autism might also be social isolation and therefore less exposure to learning opportunities regarding social cognition (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012). To account for these individual factors in my study, participants completed a comprehensive set of questionnaires, including social network index, autism quotient, and schizotypy personality questionnaire. For example, analyses of this data revealed that activation of posterior STS to neutral faces in our healthy participant sample was related to aspects of schizotypy and also to a genetic variant associated with schizophrenia, indicating an important role of posterior STS functioning for clinical considerations (Yan & Schmidt et al., submitted). Future analyses of this dataset will also include hypotheses regarding these connections. However, they are limited to a cross-sectional approach. As another example, gender and sexual orientation have been proposed to influence empathy ratings and brain activation (Perry, Walder, Hendler, & Shamay-Tsoory, 2013), and might therefore influence the results. It might therefore be valuable to control for these variables. Several other factors have also been proposed to influence empathy. Genetic variations explain over one third of variance, and are further modulated by environmental factors (Knafo et al., 2009; Knafo, Zahn-Waxler, Van Hulle, Robinson, & Rhee, 2008). In addition, in humans as well as in other animals, empathy is enhanced by familiarity, similarity, past experience, learning and salience (Preston & de Waal, 2002). The role of familiarity becomes especially obvious when considering different cultures, which was the subject of another study, I was involved in that revealed differences in social-cognitive processing between Chinese and German participants (Yan, Schmidt, et al., submitted).

Furthermore, all tasks in study 1 used anger and fear as emotions. On the one hand, it would be interesting to further differentiate the activation patterns for the emotions separately, on the other hand, it would be interesting to include different valences, and also investigate whether our findings could be replicated using positive instead of negative emotions.

As a first step to answer the question whether activation in the MNS distinguishes emotions, I conducted study 2, which I will summarize and discuss in the next section.

5.2 The MNS distinguishes emotional valence

In study 2, sequential pairs of facial configurations were presented in an fMRI adaptation paradigm to investigate whether the MNS can distinguish between emotional valences of facial configurations intended to express happiness and fear. Some scientists even consider fMRI adaptation or repetition suppression as the gold standard of MNS investigations (Fuelscher et al., 2019).

Before discussing the results of the MNS regions, it is important to mention that whole-brain analyses of the data show increased activation in FG when the valence within a facial pair changed, in comparison to when the same emotion was repeated. Also previous publications have suggested that emotional valence may modulate activation in the fusiform area (Geday et al., 2003). While in study 2, FG always showed higher activation when the emotion of the stimulus changed, it did not seem to be sensitive to the emotional direction. Possibly, this general response to a changing and therefore novel stimulus might be explained by attention; this explanation is supported by the results from a previous study, suggesting attention as a modulator of FG functioning (Vuilleumier, Armony, Driver, & Dolan, 2001).

ROI-analyses revealed increased activation for incongruent versus congruent valences in bilateral FG, in regions of the MNS; i.e. bilateral BA44, left IPL, and additional areas previously identified for the processing of facial configurations perceived as expressing emotions; i.e. bilateral insula, and right amygdala. Both, amygdala and insula are well known for their roles in emotion processing (Heinzel et al., 2005; Sergerie et al., 2008; Viinikainen et al., 2010). Most importantly, the results of study 2 provide first evidence that the MNS, represented by BA44 and IPL, is also involved in the discrimination of emotional valences. This indicates, that the MNS might not only represent the basis for a shared representation, but might indeed help humans understand this representation and therefore the emotion of the observed person.

Additionally, a negative valence following a positive one was related to stronger activation in right BA44, STS and insula than a switch from negative to positive valence. At first glance it seems unexpected that neither amygdala nor FG are part of the brain regions that show a stronger response for this contrast. First, the amygdala is particularly known for its involvement in the processing of fearful stimuli (Adolphs, 2008; Costafreda et al., 2008; Öhman, 2005). Second, FG maintains connections to amygdala (Frank, Costa, Averbeck, & Sabatinelli, 2019; Herrington, Taylor, Grupe,

Curby, & Schultz, 2011), and has been shown to have greater activation for fearful stimuli than for neutral ones (Vuilleumier et al., 2001). The discrepancy between these and my results however could be explained by the fact that study 2 did not contrast fearful and neutral but fearful and smiling facial configurations, and these regions respond more strongly to emotional stimuli than neutral ones.

The stronger effect for fear than for happiness in BA44 reaching into insula could be explained by greater salience. So possibly, salience is the driving factor not only for the Nacc, which also showed an increased response for fear than happiness in study 3, but also for the MNS. A possible connection between both phenomena could be the neurotransmitter dopamine. Nacc is a key region of the dopaminergic system (Salgado & Kaplitt, 2015), and dopamine plays an important role for motor functions (Ayano, 2016). Motor areas of the brain, in turn, build the core of the MNS (Cattaneo & Rizzolatti, 2009; Rizzolatti & Craighero, 2004). In addition, studies confirm that striatal dopamine is involved in the processing of emotions (Badgaiyan, 2010). As studies 1 and 2 were part of projects which entailed genetic analyses of all participants, a future publication will be devoted to the effects of genetic variations regarding the dopaminergic system and their influence on the MNS.

The possible importance of dopamine for the MNS also becomes obvious when considering its role for motivation and reward-seeking behavior (Ayano, 2016; Bromberg-Martin, Matsumoto, & Hikosaka, 2010; Hamid et al., 2016), together with the modulating effect of motivation on the MNS. For example, one study found that the motivation to eat was associated with increased activation to eating related stimuli in regions of the MNS and mentalizing system, namely IFG, STS and superior parietal cortex (Cheng et al., 2006). Lacking motivation in participants might also explain why the results of study 1 show no increased MNS activation for the observation of faces, but only for more engaging experimental conditions. To determine the role of motivation for MNS activation during the processing of emotional faces, I implemented a novel reward paradigm, which allows identification of the role of monetary motivation and social intention. This paradigm was part of study 2 using fMRI, and additionally presented to an independent sample of 80 participants using EEG. The analyses are still ongoing and will be part of a future publication.

To summarize, the results of this fMRI adaptation paradigm on emotional valences confirm that the MNS distinguishes positive and negative valence, thereby providing further evidence for the assumption of embodied simulation as a mechanism

to understand the emotions of our interaction partner. In addition, and in agreement with the proposed role of dopamine for MNS functioning, as well as with studies showing an influence of motivation on MNS activation, the results of study 2 suggest salience as modulating factor of MNS activation.

5.2.1 Implications from study 2 for future research

Study 2 was restricted to the observation of only two emotions, because even with this simplicity, the task took about 10 minutes to complete and was tiresome for participants due to the cognitively unchallenging nature of the task. Adding only one more emotion to the task and adding stimulus pairs for this emotion with fear and happiness, would have increased the number of conditions and consequently experiment time to at least 20 minutes. Future studies should investigate whether MNS regions are sensitive not only to the valence of a facial stimulus, but also to the specific emotion. For example, stimuli could include facial configurations intended to express anger, fear or sadness. Based on the theory of embodied simulation and the results of study 2, I would expect that the MNS can indeed distinguish all so called basic emotions, and possibly even many more common facial configurations, such as those associated with contempt, guilt or embarrassment, that have been proposed in more recent accounts of emotion classification, including a publication by Ekman (P. Ekman & Cordaro, 2011) who had previously put forward the concept of the six basic emotions.

In addition, the task in study 2 was not designed to identify MNS regions involved in perceiving and expressing emotion. Instead, my analysis was based on known MNS regions and required participants to simply observe the pictures, so emotion perception occurred implicitly, if at all; since participants neither explicitly performed facial movements themselves nor were asked to focus on their emotions, I can draw no conclusions on this side of the mirroring process. However, several studies (Enticott et al., 2008; Mier, Lis, et al., 2010; Schulte-Rüther et al., 2007; van der Gaag et al., 2007), including study 1 have established activation in the core MNS regions, also for facial stimuli. In addition, the project around study 2 contained another paradigm which included both observation and imitation in an fMRI adaptation design. If MNS response is suppressed even though the modality changes, i.e. from observation to imitation or vice versa (crossmodal), this supports the assumption of the mirror mechanism. The

results of that study will be subject of a future publication, because analyses are still ongoing.

Importantly, and in line with the recommendations of a very recent publication (Fuelscher et al., 2019), I am planning on implementing MVPA also for our crossmodal fMRI adaptation paradigm in addition to our regular analysis, as it has been suggested to be more sensitive to MNS activity. Applying MVPA to a motor-execution/-observation fMRI adaptation paradigm, Fuelscher and colleagues (2019) identified voxels in the area of anterior intraparietal sulcus which satisfied all criteria for mirror neuron involvement. Specifically, these criteria included increased activation during both observation (O) and execution (E), and both within and across modalities (O-O, E-E, O-E, E-O) characteristic patterns are shared for the repetition of the same stimulus (adaptation condition), but not when different stimuli were used (control condition). However, their study, as many studies in the field, had small sample size of only 12 participants, and MVPA results were significant, but only moderately above chance level, and not surviving correction for multiple testing, so their conclusions need to be confirmed in larger studies, as could be accomplished by the data set which includes the same participants as study 2.

Finally, while single-cell recordings are undoubtedly the most accurate method and the only way to measure MN, research can also benefit from exploiting existing methods. Other than applying specific analysis methods, as I did in studies 1 and 2, one can also combine different measurement techniques. For example, simultaneous EEG-fMRI combines good temporal with good spatial resolution and has provided interesting insights in other domains. For example, low-frequency EEG oscillations in the theta and alpha band, associated with recollection, are correlated with the connectivity of hippocampus with PFC and striatum, which are related to retrieval success, so the authors concluded that the EEG-recorded oscillations may represent a binding mechanism for these brain regions (Herweg et al., 2016). Remarkably, to date I know of no studies that used simultaneous EEG-fMRI to investigate the mirror neuron system or social cognition, besides our own (Schmidt et al., in preparation). However, during the preparation of the manuscript, challenges with data preprocessing occurred which I am currently aiming to solve, including replicating the results with an independent sample of 30 participants that had an EEG-measurement without simultaneous fMRI.

For some questions, more advanced methods help find the answers. For other questions, the paradigm alone seems sufficient to provide new insights. Study 3, discussed in the following section sheds light on the brain processes involved in perceiving an emotion based on ambiguous facial configurations.

5.3 Nucleus accumbens helps resolve ambiguous facial configurations

The assumption that emotion perception is a social-cognitive skill which is accomplished via embodied simulation with the automatic response of the MNS (Gallese, 2007a, 2007b; Gallese & Goldman, 1998; Gallese et al., 2004) is central to my PhD thesis. However, as the MNS preferentially processes familiar movements (Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005; Calvo-Merino, Grèzes, Glaser, Passingham, & Haggard, 2006), one might expect that the same is true for facial configurations. Consequently, ambiguous facial configurations that involve more than one emotion and thus might be less familiar could rely on more than automatic MNS processing. In real life situations, one might take the context into account, which has been shown to strongly influence the evaluation of facial configurations (Carroll & Russell, 1996). However, also without additional information, one should be able to resolve the conflict of ambiguous facial features.

One interesting approach to investigate how ambiguous facial features are processed and how a decision on one emotion is achieved, is adopted from research on decision making in schizophrenia, a mental disorder associated with abnormal dopaminergic signaling (Howes & Kapur, 2009) and difficulties in social interaction (Kohler et al., 2010; Mier & Kirsch, 2016). Previous non-social tasks had shown increased activation in DLPFC and parietal regions during probabilistic decision making, as required when deciding to which of two lakes with fixed color ratios a sequence of colored fish belongs (Rausch et al., 2014). While in healthy individuals, the final decision was associated with increased activation in VTA and Nacc, activation was comparably lower in patients with schizophrenia (Rausch et al., 2014). The Nacc is a key region of the dopaminergic system, involved in salience, reward and motivation (Kringelbach & Berridge, 2010; Sabatinelli et al., 2007; Salgado & Kaplitt, 2015). Interestingly, salience, not reward, seemed to be responsible for increased Nacc activation during decision making (Esslinger et al., 2012).

Study 3 replicated these findings using ambiguous emotional faces. Series of faces were presented, in which the first face was perceived as showing equal percentages of fear and happiness. The subsequent faces were less ambiguous. Participants had to decide for each stimulus, whether the currently dominant emotion was fear or happiness; as soon as they were certain, they should make a decision on the emotion dominant in the current series of pictures. In line with the fish-in-the-lake studies (Rausch et al., 2014), study 3 revealed increased activation in parietal and frontal lobe during the probabilistic reasoning process. During the final decision, activation was increased in putamen reaching into Nacc, as well as in ACC. In the following, I will first discuss the impact of the different frontal and parietal regions, afterwards the results regarding the Nacc.

Interestingly, in addition to DLPFC, which is central to probabilistic reasoning (Esslinger et al., 2012), another frontal region that showed increased activation during the probabilistic processing of facial configurations intended to express fear, was the MNS key region BA44 and its neighbor BA6 which has also been proposed to be an MNS region (Van Overwalle & Baetens, 2009). Increased BA44 activation for fearful but not for happy blocks is in line with the results from study 2, in which BA44 also responded with increased activity to facial configurations intended to express fear compared rather than happiness. Maybe for both studies this result is due to increased salience for fear than happiness. As expected, the MNS is supported by parietal/occipital areas during the processing of ambiguous faces. One of these areas was the somatosensory association cortex (BA7), which also showed increased activation in a previous study in which participants had to decide which of two schematic faces expressed greater sadness (Viviani, Dommes, Bosch, Stingl, & Beschoner, 2018). In the light of studies on pain, that also find an involvement of the somatosensory association cortex, Viviani and colleagues (2018) suggest that this area was active, because participants did not actually share an inner experience with the individuals they were looking at, but basing their decisions mainly on the visual features of the presented faces. Likewise, in study 3, participants may not have been sufficiently able to share the emotional experience of the depicted face, because it was a morphed face with emotional ambiguity. Consequently, participants might have also tried to identify the dominant emotion by analyzing the visual features.

The ACC, showing increased activation for final decisions, has previously been shown to be involved in emotional conflict resolution (A. Etkin, T. Egner, D. M. Peraza,

E. R. Kandel, & J. Hirsch, 2006), decision making (Kennerley et al., 2006) and reward processing (P. Kirsch et al., 2003), which might well explain its involvement in the final decisions of our emotional JTC task.

While the fish-in-the-lake studies had shown that the final decision was associated with activation in the ventral striatum, of which Nacc is a major part, the whole-brain results of study 3 additionally suggest an involvement of the putamen, which lies next to the Nacc in the dorsal striatum, and is also involved in dopaminergic signaling, and in reward and decision making (Balleine, Delgado, & Hikosaka, 2007).

Importantly, Nacc activation during final decisions was even stronger for fear than for happiness. As facial configurations intended to express fear can be considered salient, but not as rewarding (Elsherif, Sahan, & Rotshtein, 2017; Zheng et al., 2017) as happy faces (Spreckelmeyer et al., 2009), increased Nacc activation during fearful faces could indicate that salience was the decisive factor. Interestingly also, this effect of increased Nacc activation was enhanced in participants who had a greater tendency to a JTC bias, i.e. considering less evidence before making the final decision. Altered activation of the ventral striatum had also been reported in patients with schizophrenia (Rausch et al., 2014), who are known to exhibit a JTC bias, especially when they have delusions (Moritz & Woodward, 2005). Possibly, also our healthy participants who showed a JTC bias, perceived increased salience of these stimuli due to altered dopaminergic signaling, as would be expected in patients with schizophrenia or psychosis (Esslinger et al., 2012; Kapur et al., 2005).

With regard to the two-pathway models, the results indicate an involvement of the MNS during the processing of the ambiguous faces during probabilistic reasoning. Supporting the assumption that the MNS alone cannot solve the ambiguous face task, DLPFC, as well as parietal areas, which are considered part of the slow c-system (Satpute & Lieberman, 2006), also showed increased activation. During the final decision, Nacc reaching into putamen and dorsal ACC showed enhanced activation, all of which are part of the x-system (Satpute & Lieberman, 2006). So, while the probabilistic reasoning is supported by brain regions generally associated with slow and deliberate processing, the final decision relies on regions that support fast and automatic processes. This involvement of fast and automatic processes during the final decision might also explain how aberrant salience can reduce the amount of evidence considered and therefore fosters a JTC bias.

5.3.1 Limitations and future implications of study 3

Study 3 provides a basis for the research on decision making on ambiguous emotions. However, the study could be adapted in several ways to further deepen the understanding of decision making and emotion processing. On the one hand, one could adapt the paradigm, on the other hand, one could include a different participant sample. Study 3 only included healthy participants and looked at specific schizophrenia traits. The logical next step regarding the sample is to apply the paradigm to patients with schizophrenia in comparison to a healthy control group, which we are currently doing. With regard to the paradigm, it would be interesting to confirm the findings with other emotions. For example, one could use happy-angry morphs, to see whether the salience theory also holds true for angry faces, which would be expected since facial configurations intended to express anger are an indicator of threat. Further, one could investigate more deeply the role of flexibility and stability of reasoning with respect to the decision on the emotions. In particular, it would be interesting to determine the factors associated with sticking to one's initial assumption despite evidence supporting the contrary, versus quickly changing one's rule, based on comparatively little evidence. While the latter seems closely associated with the JTC bias linked to dopaminergic signaling, the former might either just lie on the opposite end of the spectrum, and therefore also be explained by dopamine, but it could also involve a completely different factor, such as education (C. E. Evans, Kemish, & Turnbull, 2004), or testosterone level (K. L. Evans & Hampson, 2014). In addition, in samples clearly associated with dopamine dysfunctioning, which besides schizophrenia might include Parkinson's disease, one could also investigate the effect of medication status. In this regard, it is important to mention that schizophrenia medication does not seem to alleviate the problems in social cognition (Kucharska-Pietura & Mortimer, 2013), which indicates that it remains essential to investigate the core mechanisms of social cognition especially with regard to neurochemistry.

Another clear next step to advance the understanding of the neural mechanisms underlying (disturbed) social decision-making is to complement the fMRI activation analyses with functional connectivity analyses, which I am implementing for our current study with the schizophrenia sample. Altered functional connectivity has already been linked to the social-cognitive characteristics of individuals with autism (Supekar et al., 2013) and also schizophrenia (Mukherjee et al., 2014) which is a disorder closely linked

to disconnection (Friston & Frith, 1995). Based on the findings from this study, we expect altered connectivity of the Nacc in the schizophrenia group.

In sum, all three studies provide further insights into the neural correlates of social-cognitive processes. While different neural processes share a common neural basis, there is, as one could expect, also individual task activation. Importantly, while regions of the MNS seem to be central to social-cognitive processing, they are not the only regions showing increased activation and might not be sufficient for all tasks. In the next sections, I will integrate the different findings of my PhD work into models of social cognition.

5.4 Models of social cognition

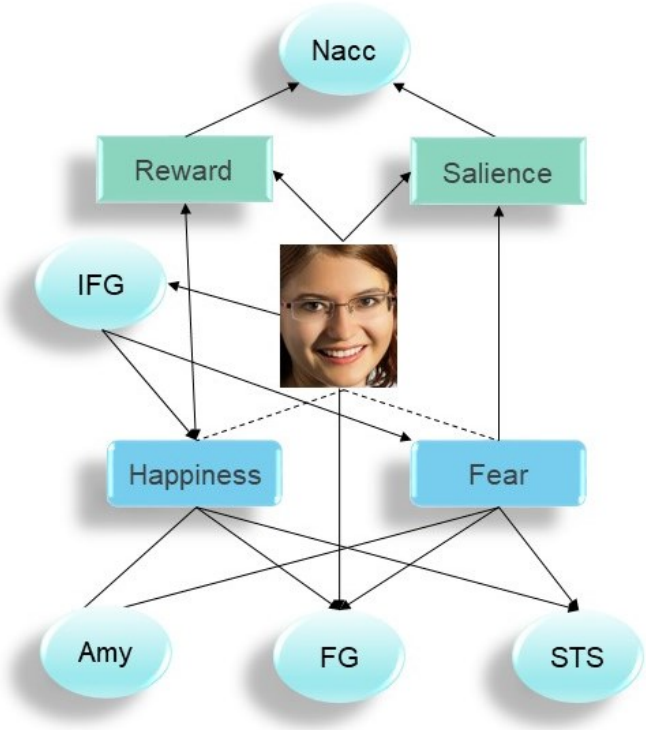
To illustrate the findings of my PhD work, I created model 1, combining the findings of studies 2 and 3 with regard to the processing of fearful and smiling faces. Model 2, being a generalization of model 1 integrates different social-cognitive processes that can be active when interacting with others, including affective processes such as empathy and ToM.

5.4.1 Model 1: Neural correlates of the perception of fearful and smiling faces

One core assumption of this thesis is that the MNS provides a fundamental mechanism underlying social cognition. Therefore, it is essential to understand how exactly the MNS is involved in social-cognitive processes. While study 2 was concerned with the MNS' role in valence discrimination, study 3 focused on ambiguous faces, which were expected to comprise additional regions besides the MNS. In model 1, I aimed at combining these findings, taking the theory of embodied simulation as a theoretical foundation.

When looking at facial configurations intended to express emotions, fusiform gyrus, amygdala and MNS regions show increased activation and allow the understanding of the basic facial features and the underlying emotion. Thanks to the shared representation in the MNS, one knows whether the other feels happiness or fear. In case of ambiguous emotions, Nacc is involved in the final decision, as fearful faces evoke stronger activation possibly due to their increased salience. Facial configurations perceived as expressing happiness are experienced as rewarding and

also activate amygdala and Nacc, but possibly to a lesser degree than fear. The fusiform gyrus is not only involved in the basic perception of faces, but also quickly identifies changes in emotional valence, i.e. when someone looks at a smiling face following a fearful one, and vice versa. In this latter case, when a facial configuration perceived as being fearful follows a smiling one, activation is increased in STS and IFG reaching into insula, i.e. MNS and mentalizing system. Again, this could be explained by the increased salience of the fearful stimuli. As discussed in section 5.2., one could assume that the dopamine associated with Nacc function also modulates the MNS. The connection of dopamine and the MNS is evident, because the MNS is typically spatially overlapping with motor areas, involved in movements, and dopamine is a key neurotransmitter for motor control (for a more detailed explanation, please refer to section 5.2).



Model 1. Brain regions involved in the perception of fear and happiness in facial configurations. Amy = Amygdala, FG = Fusiform Gyrus, IFG = Inferior Frontal Gyrus, Nacc = Nucleus accumbens, STS = Superior Temporal Sulcus.

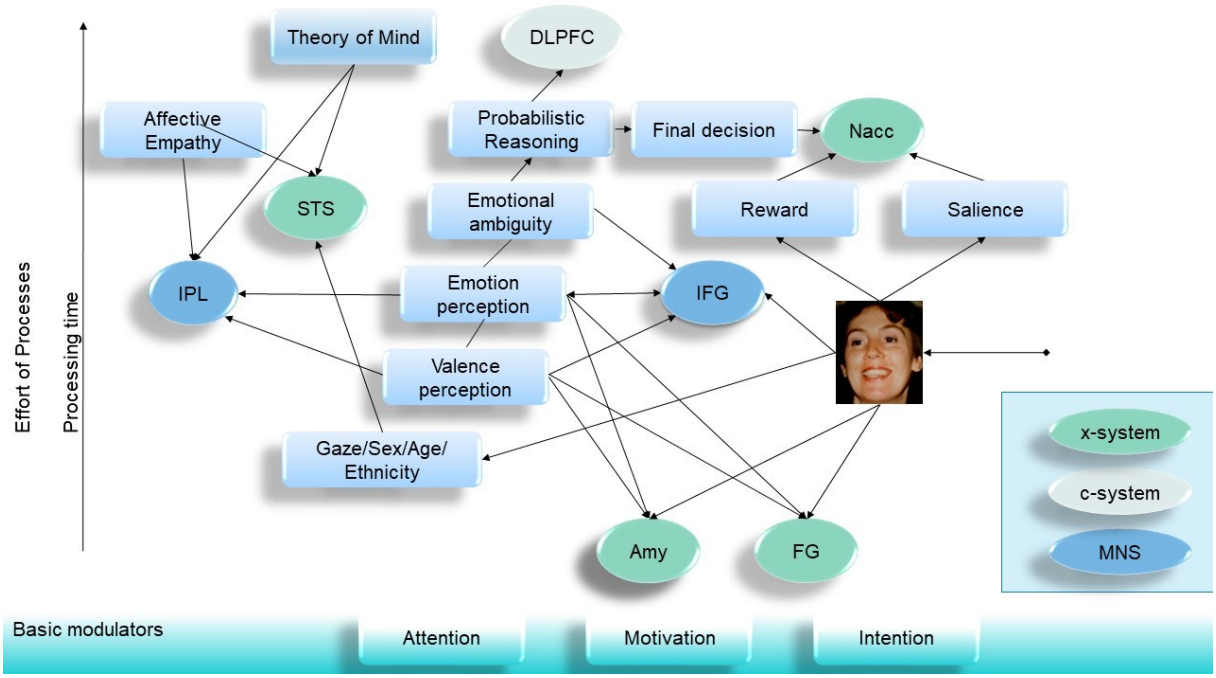
5.4.2 Model 2: Processing and reacting to an emotional face

While model 1 focusses on specific emotions, model 2 depicts a more general level, integrating different social-cognitive skills with different processing demands and associated brain regions. Since this model is mainly based on the paradigms of studies 1 to 3, all being centered around facial stimuli, face perception is at the core of the model and builds the basis for all further processes. While automatic processes are prevalent, there are also cognitive ones which may involve deliberate reasoning. In addition, with increasing complexity of the processes, processing time also increases.

It is important to discuss the relevance of x- and c-system (Satpute & Lieberman, 2006) to my thesis. The ROIs of studies 1 to 3 were chosen based on previous research on social cognition and included regions of the MNS, mentalizing system and processing of facial configurations perceived as being emotional. Many of these ROIs are also associated with the x-system, explaining the prominence of the x-system in my discussion. However, one should not conclude that these social-cognitive skills do not include regions of the c-system, because we did not explicitly test for it. With regard to the ROIs, the results of my PhD thesis indicate that the fast x-system comprises many regions identified as a common neural basis of social cognition, including amygdala and STS. Additionally, fusiform gyrus is part of the temporal lobe, which is also associated with the fast x-system. Also the ambiguous faces in study 3 activated regions associated with the x-system, such as Nacc and putamen during decisions on the general emotion. As expected, DLPFC and parietal regions, considered part of the slow and deliberate c-system, are involved in probabilistic reasoning during the processing ambiguous facial configurations. The whole brain analyses additionally reveal increased activation in regions that are not included in the ROIs, which can inform future studies, including possible studies aiming to distinguish automatic versus controlled processes in social cognition. For example, also medial temporal gyrus and medial frontal gyrus, both considered part of the c-system, show increased activation during neutral face processing, emotion perception and cognitive empathy.

To truly confirm their categorization into c- and x-system, Satpute and Lieberman (2006) recommend researchers to design dual tasks that require parallel processing of social-cognitive skills for which automaticity is to be investigated and specific cognitive tasks for which the associated brain functioning is well-established. Consequently, future studies should follow their suggestions to confirm the

automaticity of the MNS and associated regions, so they could be included in an updated model of the x- and c-system, deepening our understanding of the neural mechanism of social cognition.



Model 2. Social cognitive processes and associated brain regions. Brain regions are colored with respect to their categorization into x-system, c-system or MNS. Note: The STS is part of both x-system and mentalizing system. Amy = Amygdala, DLPFC = Dorsolateral Prefrontal Cortex, FG = Fusiform Gyrus, IFG = Inferior Frontal Gyrus, IPL = Inferior Parietal Lobe, Nacc = Nucleus accumbens.

5.5 Limitations and implications for future research

In my studies, brain activation was measured while participants performed the tasks, most of which also required button presses. While one can infer from button presses and brain activation that participants followed the instructions, additionally applying eye-tracking or recording facial movements would be of advantage. I let the participants practice the tasks until they felt familiar with them, and I attended this practice to make sure they solved the tasks correctly. Still, it would be useful to confirm that they focused their attention and behaved as intended by the study design. Especially in study 1, when participants were told to imitate faces, it would be good to have the recorded proof that they really did. When asked, participants reported to have

followed the instruction and imitated the presented faces. In addition, the movement artefacts support this at least for part of the subjects. Also with regard to methodological aspects, study 1 provided these valuable insights, so in study 2 I additionally used a face cam to monitor and record participants' faces during the tasks. Besides ensuring their compliance with tasks such as the imitation task, it also helps ensuring their wakefulness in paradigms that do not require button presses.

As briefly mentioned above, studies 1 and 2 included genetic analyses, which will help us gain more insights on the influence of genetic predispositions, e.g. variations in genes regulating the dopaminergic pathway, on social cognition. Dopamine has received comparably little attention in the field of social cognition research, but as mentioned above, especially with regard to the assumed role of the MNS for social cognition, dopamine appears as a fruitful candidate neurotransmitter. Dopamine is also related to mental disorders, which are associated with abnormal social functioning. Finally, dopamine is related to mechanisms which might modulate social-cognitive processes, such as direction of attention, salience or motivation, as also suggested by the results of study 2 and 3. One example of a well-studied common genetic variant with respect to the dopamine system is the single-nucleotide polymorphism (SNP) rs4680 (Meyer-Lindenberg et al., 2006). It affects the COMT gene which codes for the enzyme that breaks down prefrontal dopamine. Studies point to differences in prefrontal fMRI-BOLD related to this SNP (Mier, Kirsch, & Meyer-Lindenberg, 2010). A multitude of further findings exists regarding rs4680, including increased risk for schizophrenia (Egan et al., 2001; Shifman et al., 2002). Interestingly, COMT genotype in schizophrenia seems also related to differences in distress and empathy (Poletti et al., 2013). The imaging genetics results of my studies, including an elaborate discussion of the value and explanatory power of these analyses will be subject of a future publication (Schmidt et al., in preparation).

As stressed as a challenge in this thesis, there are limits to the investigation of MN using indirect methods, and the conclusions we can draw from them. Having adopted several approaches to nevertheless measure mirror neuron populations as accurately as currently possible using non-invasive techniques, I want to mention one very important and rising research area which can add substantially to the understanding of neural mechanisms, including MN. This field of research is computational neuroscience, which, roughly speaking, aims to computationally model the brain or selected networks and mechanisms. For example, Hass and colleagues

(2016), successfully established a model of the PFC, which replicated in vivo electrophysiological behavior. In collaboration with Sadeghi and Hass, the data from studies 1 and 2 are also used to inform computational models of the MNS.

While the computational aspects of data analyses are becoming stronger, the computational settings, in which social-cognitive paradigms are commonly placed, might soon be augmented by more realistic settings. This development, known as second person neuroscience, might provide more reliable insights in social cognition, as social processes would not be a mere response to the picture or video of a face, but instead including a social interaction (Redcay & Schilbach, 2019). This could be accomplished by measuring the brain activation of two persons interacting with each other. However, even if one of these persons were the experimenter, the behavior could not be replicated identically between subjects, simply because the experimenter is human, and might be in a different mood, or show slight deviations in facial configuration or tone of voice. Another option would be to use avatars in a virtual reality, which could be programmed to exhibit the same behavior to each participant. Obviously, research on social cognition can greatly benefit from the advancements in these new technologies. Ultimately, these approaches might also help to identify the basic mechanisms of social cognition, as demanded by Schaafsma and colleagues (2015).

5.6 Conclusions

The results of the presented studies point to a shared neural basis of different social-cognitive skills, mainly in regions associated with the MNS. Studies 1 to 3 add to a set of few studies using only pictures of faces as stimuli, and modulating the social-cognitive processes by task instruction. The distinct activation patterns, in addition to the common regions, indicate a successful implementation of this strategy. Importantly, sVx analyses reveal common activation within and across participants across tasks, and also fMRI adaptation is a valuable method for the investigation of the MNS using fMRI. Tasks too complex to rely solely on automatic MNS processing, additionally activate structures previously identified in probabilistic decision-making tasks that used non-social stimuli. In particular, salience of facial configurations intended to express fear drives Nacc activation when deciding on the dominant emotion in ambiguous facial stimuli.

To come back to the consideration from the introduction, the MNS might be the underlying network that allowed the successful non-verbal communication between you and your caregivers when you were a baby, and research keeps advancing the understanding of the neural functioning.

6 SUMMARY

In my PhD thesis, I present three functional magnetic resonance imaging studies aimed at investigating neurobiological mechanisms underlying social cognition. My thesis focuses on fast and automatic processes that are proposed to build the basis of social understanding, and might be activated in parallel to more effortful deliberate mechanisms. The proposed neural substrate of fast and automatic processes are mirror neurons, which according to the theory of embodied simulation allow humans to understand other individuals' actions, and even emotions and intentions. Since non-invasive techniques cannot be applied to measure mirror neurons, but only neural populations assumed to constitute the mirror neuron system, experimental paradigms and analysis routines that allow approximation of mirror neuron functions need to be developed.

In study 1, I demonstrated that different social cognitive skills, including imitation, affective empathy and theory of mind share a common neural basis, located in regions associated with the mirror neuron system. In addition to standard analyses, a shared voxel analysis was applied that revealed common activation for social-cognitive processes not only across, but also within participants.

Study 2 was set up to investigate whether the mirror neuron system can distinguish the valence of facial configurations. The use of a functional magnetic resonance imaging adaptation paradigm allowed to determine neural populations sensitive to emotional valence. While the fusiform gyrus was sensitive to changes from fearful to smiling faces and also from smiling to fearful faces, Brodmann area 44 reaching into insula, and superior temporal sulcus, i.e. regions more commonly associated with the mirror neuron system and with the so called mentalizing network, showed particularly increased activation for switches from smiling to fearful faces.

Study 3 was dedicated to the investigation of decision making in the context of ambiguous facial configurations. While probabilistic decision making on these facial configurations lead to activation in the executive control network, final decisions for an emotion resulted in nucleus accumbens activation. In addition, perceiving fear in a face lead to higher nucleus accumbens activation during final decisions than perceiving happiness. This finding can be linked to salience processing in the nucleus accumbens.

In conclusion, all three studies show an involvement of fast and automatic processing regions for different social-cognitive processes. Study 3 additionally examined the interaction with slower and more deliberate processes, as involved in probabilistic decision making on ambiguous faces. The mirror neuron system seems to be critically involved in different social-cognitive tasks and also sensitive to emotional valence. In cases when automatic processing is not possible, as when presented with ambiguous facial configurations, brain regions commonly associated with probabilistic decision making assist, and the nucleus accumbens, possibly by directing salience, is involved in the final decision. These results deepen the understanding of the mechanisms of social cognition and encourage the use of sophisticated methods in experimental paradigms and analysis.

7 REFERENCES

Abu-Akel, A., & Shamay-Tsoory, S. (2011). Neuroanatomical and neurochemical bases of theory of mind. *Neuropsychologia, 49*(11), 2971-2984. doi:10.1016/j.neuropsychologia.2011.07.012

Achim, A. M., Maziade, M., Raymond, É., Olivier, D., Mérette, C., & Roy, M.-A. (2009). How prevalent are anxiety disorders in schizophrenia? A meta-analysis and critical review on a significant association. *Schizophrenia bulletin, 37*(4), 811-821.

Adolph, D., & Alpers, G. W. (2010). Valence and arousal: a comparison of two sets of emotional facial expressions. *Am J Psychol, 123*(2), 209-219. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20518437

Adolphs, R. (2008). Fear, faces, and the human amygdala. *Curr Opin Neurobiol, 18*(2), 166-172. doi:10.1016/j.conb.2008.06.006

Adolphs, R. (2010). What does the amygdala contribute to social cognition? *Ann N Y Acad Sci, 1191*, 42-61. doi:10.1111/j.1749-6632.2010.05445.x

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron, 32*(3), 537-551. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11709163

Anderson, I. M., Del-Ben, C. M., McKie, S., Richardson, P., Williams, S. R., Elliott, R., & Deakin, J. F. (2007). Citalopram modulation of neuronal responses to aversive face emotions: a functional MRI study. *Neuroreport, 18*(13), 1351-1355. doi:10.1097/WNR.0b013e3282742115

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception, 33*(6), 717-746. doi:10.1068/p5096

Axelrod, V., & Yovel, G. (2015). Successful decoding of famous faces in the fusiform face area. *PLoS One, 10*(2), e0117126. doi:10.1371/journal.pone.0117126

Ayano, G. (2016). Dopamine: receptors, functions, synthesis, pathways, locations and mental disorders: review of literatures. *J Ment Disord Treat, 2*(120), 2.

Badgaiyan, R. D. (2010). Dopamine is released in the striatum during human emotional processing. *Neuroreport, 21*(18), 1172-1176. doi:10.1097/WNR.0b013e3283410955

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *J Neurosci, 27*(31), 8161-8165. doi:10.1523/JNEUROSCI.1554-07.2007

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*(1), 37-46. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/2934210

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J Child Psychol Psychiatry, 42*(2), 241-251. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11280420

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychol Sci Public Interest, 20*(1), 1-68. doi:10.1177/1529100619832930

Bastiaansen, J. A., Thioux, M., & Keysers, C. (2009). Evidence for mirror systems in emotions. *Philos Trans R Soc Lond B Biol Sci, 364*(1528), 2391-2404. doi:10.1098/rstb.2009.0058

Becchio, C., Cavallo, A., Begliomini, C., Sartori, L., Feltrin, G., & Castiello, U. (2012). Social grasping: from mirroring to mentalizing. *Neuroimage, 61*(1), 240-248. doi:10.1016/j.neuroimage.2012.03.013

Beer, J. S., & Ochsner, K. N. (2006). Social cognition: a multi level analysis. *Brain Res, 1079*(1), 98-105. doi:10.1016/j.brainres.2006.01.002

Berridge, K. C. (2006). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl), 191*(3), 391-431. doi:10.1007/s00213-006-0578-x

Bialystok, E., & Senman, L. (2004). Executive processes in appearance-reality tasks: the role of inhibition of attention and symbolic representation. *Child Dev, 75*(2), 562-579. doi:10.1111/j.1467-8624.2004.00693.x

Bishop, S. J., Duncan, J., & Lawrence, A. D. (2004). State anxiety modulation of the amygdala response to unattended threat-related stimuli. *J Neurosci, 24*(46), 10364-10368. doi:10.1523/JNEUROSCI.2550-04.2004

Bitsch, F., Berger, P., Nagels, A., Falkenberg, I., & Straube, B. (2018). The role of the right temporo-parietal junction in social decision-making. *Hum Brain Mapp, 39*(7), 3072-3085. doi:10.1002/hbm.24061

Blackwood, N. J., Howard, R. J., Bentall, R. P., & Murray, R. M. (2001). Cognitive neuropsychiatric models of persecutory delusions. *Am J Psychiatry, 158*(4), 527-539. doi:10.1176/appi.ajp.158.4.527

Bodden, M. E., Mollenhauer, B., Trenkwalder, C., Cabanel, N., Eggert, K. M., Unger, M. M., . . . Kalbe, E. (2010). Affective and cognitive Theory of Mind in patients with parkinson's disease. *Parkinsonism Relat Disord, 16*(7), 466-470. doi:10.1016/j.parkreldis.2010.04.014

Bonini, L., Ugolotti Serventi, F., Bruni, S., Maranesi, M., Bimbi, M., Simone, L., . . . Fogassi, L. (2012). Selectivity for grip type and action goal in macaque inferior parietal and ventral premotor grasping neurons. *J Neurophysiol, 108*(6), 1607-1619. doi:10.1152/jn.01158.2011

Bonini, L., Ugolotti Serventi, F., Simone, L., Rozzi, S., Ferrari, P. F., & Fogassi, L. (2011). Grasping neurons of monkey parietal and premotor cortices encode action goals at distinct levels of abstraction during complex action sequences. *J Neurosci, 31*(15), 5876-5886. doi:10.1523/JNEUROSCI.5186-10.2011

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron, 68*(5), 815-834. doi:10.1016/j.neuron.2010.11.022

Brothers, L. (2002). The social brain: a project for integrating primate behavior and neurophysiology in a new domain. *Foundations in social neuroscience, 367*, 385.

Buccino, G., Binkofski, F., & Riggio, L. (2004). The mirror neuron system and action recognition. *Brain Lang, 89*(2), 370-376. doi:10.1016/S0093-934X(03)00356-0

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci, 4*(6), 215-222. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10827444

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition, 112*(2), 337-342. doi:10.1016/j.cognition.2009.05.006

Caggiano, V., Fogassi, L., Rizzolatti, G., Casile, A., Giese, M. A., & Thier, P. (2012). Mirror neurons encode the subjective value of an observed action. *Proc Natl Acad Sci U S A, 109*(29), 11848-11853. doi:10.1073/pnas.1205553109

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an FMRI study with expert dancers. *Cereb Cortex, 15*(8), 1243-1249. doi:10.1093/cercor/bhi007

Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Curr Biol, 16*(19), 1905-1910. doi:10.1016/j.cub.2006.07.065

Calvo, M. G., & Nummenmaa, L. (2008). Detection of emotional faces: salient physical features guide effective visual search. *J Exp Psychol Gen, 137*(3), 471-494. doi:10.1037/a0012771

Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc Natl Acad Sci U S A, 100*(9), 5497-5502. doi:10.1073/pnas.0935845100

Carrington, S. J., & Bailey, A. J. (2009). Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Hum Brain Mapp, 30*(8), 2313-2335. doi:10.1002/hbm.20671

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *J Pers Soc Psychol, 70*(2), 205-218. doi:10.1037//0022-3514.70.2.205

Cattaneo, L., & Rizzolatti, G. (2009). The mirror neuron system. *Arch Neurol, 66*(5), 557-560. doi:10.1001/archneurol.2009.41

Cheng, Y., Meltzoff, A. N., & Decety, J. (2006). Motivation modulates the activity of the human mirror-neuron system. *Cereb Cortex, 17*(8), 1979-1986. doi:10.1093/cercor/bhl107

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends Cogn Sci, 16*(4), 231-239. doi:10.1016/j.tics.2012.02.007

Chong, T. T., Cunnington, R., Williams, M. A., Kanwisher, N., & Mattingley, J. B. (2008). fMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Curr Biol, 18*(20), 1576-1580. doi:10.1016/j.cub.2008.08.068

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive development, 9*(4), 377-395.

Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M., Jr. (1997). Social ties and susceptibility to the common cold. *JAMA, 277*(24), 1940-1944. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9200634

Costa, A., Torriero, S., Oliveri, M., & Caltagirone, C. (2008). Prefrontal and temporo-parietal involvement in taking others' perspective: TMS evidence. *Behav Neurol, 19*(1-2), 71-74. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18413921

Costafreda, S. G., Brammer, M. J., David, A. S., & Fu, C. H. (2008). Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies. *Brain Res Rev, 58*(1), 57-70. doi:10.1016/j.brainresrev.2007.10.012

Cox, C. L., Uddin, L. Q., Di Martino, A., Castellanos, F. X., Milham, M. P., & Kelly, C. (2011). The balance between feeling and knowing: affective and cognitive empathy are reflected in the brain's intrinsic functional dynamics. *Soc Cogn Affect Neurosci, 7*(6), 727-737. doi:10.1093/scan/nsr051

Critchley, H., Daly, E., Phillips, M., Brammer, M., Bullmore, E., Williams, S., . . . Murphy, D. (2000). Explicit and implicit neural mechanisms for processing of social information from facial expressions: a functional magnetic resonance imaging study. *Hum Brain Mapp, 9*(2), 93-105. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10680766

Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A Review of the Concept. *Emotion Review, 8*(2), 144-153. doi:10.1177/1754073914558466

Cui, F., Abdelgabar, A. R., Keysers, C., & Gazzola, V. (2015). Responsibility modulates pain-matrix activation elicited by the expressions of others in pain. *Neuroimage, 114*, 371-378. doi:10.1016/j.neuroimage.2015.03.034

Cunningham, W. A., & Brosch, T. (2012). Motivational salience: Amygdala tuning from traits, needs, values, and goals. *Current Directions in Psychological Science, 21*(1), 54-59.

Dauer, W., & Przedborski, S. (2003). Parkinson's disease: mechanisms and models. *Neuron, 39*(6), 889-909. doi:10.1016/s0896-6273(03)00568-3

Davis, K. D., Taylor, K. S., Hutchison, W. D., Dostrovsky, J. O., McAndrews, M. P., Richter, E. O., & Lozano, A. M. (2005). Human anterior cingulate cortex neurons encode cognitive and emotional demands. *Journal of Neuroscience, 25*(37), 8402-8406.

de la Rosa, S., Schillinger, F. L., Bulthoff, H. H., Schultz, J., & Uludag, K. (2016). fMRI Adaptation between Action Observation and Action Execution Reveals Cortical Areas with Mirror Neuron Properties in Human BA 44/45. *Front Hum Neurosci, 10*, 78. doi:10.3389/fnhum.2016.00078

Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., & Laland, K. N. (2012). Identification of the social and cognitive processes underlying human cumulative culture. *Science, 335*(6072), 1114-1118. doi:10.1126/science.1213969

Decety, J., & Jackson, P. L. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science, 15*(2), 54-58. doi:DOI 10.1111/j.0963-7214.2006.00406.x

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp Brain Res, 91*(1), 176-180. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1301372

Dudley, R., Taylor, P., Wickham, S., & Hutton, P. (2016). Psychosis, Delusions and the "Jumping to Conclusions" Reasoning Bias: A Systematic Review and Meta-analysis. *Schizophr Bull, 42*(3), 652-665. doi:10.1093/schbul/sbv150

Dunbar, R. I. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews, 6*(5), 178-190.

Dunbar, R. I. (2009). The social brain hypothesis and its implications for social evolution. *Ann Hum Biol, 36*(5), 562-572. doi:10.1080/03014460902960289

Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E., . . . Weinberger, D. R. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proc Natl Acad Sci U S A, 98*(12), 6917-6922. doi:10.1073/pnas.111134598

Eisenberger, N. I., & Lieberman, M. D. (2004). Why rejection hurts: a common neural alarm system for physical and social pain. *Trends Cogn Sci, 8*(7), 294-300. doi:10.1016/j.tics.2004.05.010

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition & Emotion, 6*(3-4), 169-200. doi:Doi 10.1080/02699939208411068

Ekman, P., & Cordaro, D. (2011). What is Meant by Calling Emotions Basic. *Emotion Review, 3*(4), 364-370. doi:10.1177/1754073911410740

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the Human Face: Guide-lines for Research and an Integration of Findings: Guidelines for Research and an Integration of Findings*: Pergamon.

Ekman, P., Friesen, W. V., Osullivan, M., Chan, A., Diacoyannitarlatzis, I., Heider, K., . . . Tzavaras, A. (1987). Universals and Cultural-Differences in the Judgments of Facial Expressions of Emotion. *Journal of Personality and Social Psychology, 53*(4), 712-717. doi:Doi 10.1037/0022-3514.53.4.712

El-Sourani, N., Wurm, M. F., Trempler, I., Fink, G. R., & Schubotz, R. I. (2018). Making sense of objects lying around: How contextual objects shape brain activity during action observation. *Neuroimage, 167*, 429-437. doi:10.1016/j.neuroimage.2017.11.047

Elsherif, M. M., Sahan, M. I., & Rotshtein, P. (2017). The perceptual saliency of fearful eyes and smiles: A signal detection study. *PLoS One, 12*(3), e0173199. doi:10.1371/journal.pone.0173199

Enticott, P. G., Arnold, S. L., Fitzgibbon, B. M., Hoy, K. E., Susilo, D. A., & Fitzgerald, P. B. (2012). Transcranial direct current stimulation (tDCS) of the inferior frontal gyrus disrupts interpersonal motor resonance. *Neuropsychologia, 50*(7), 1628-1631. doi:10.1016/j.neuropsychologia.2012.03.016

Enticott, P. G., Johnston, P. J., Herring, S. E., Hoy, K. E., & Fitzgerald, P. B. (2008). Mirror neuron activation is associated with facial emotion processing. *Neuropsychologia, 46*(11), 2851-2854. doi:10.1016/j.neuropsychologia.2008.04.022

Enticott, P. G., Kennedy, H. A., Bradshaw, J. L., Rinehart, N. J., & Fitzgerald, P. B. (2010). Understanding mirror neurons: evidence for enhanced corticospinal excitability during the observation of transitive but not intransitive hand gestures. *Neuropsychologia, 48*(9), 2675-2680. doi:10.1016/j.neuropsychologia.2010.05.014

Esslinger, C., Braun, U., Schirmbeck, F., Santos, A., Meyer-Lindenberg, A., Zink, M., & Kirsch, P. (2013). Activation of midbrain and ventral striatal regions implicates salience processing during a modified beads task. *PLoS One, 8*(3), e58536. doi:10.1371/journal.pone.0058536

Esslinger, C., Englisch, S., Inta, D., Rausch, F., Schirmbeck, F., Mier, D., . . . Zink, M. (2012). Ventral striatal activation during attribution of stimulus saliency and reward anticipation is correlated in unmedicated first episode schizophrenia patients. *Schizophr Res, 140*(1-3), 114-121. doi:10.1016/j.schres.2012.06.025

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron, 51*(6), 871-882. doi:10.1016/j.neuron.2006.07.029

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving Emotional Conflict: A Role for the Rostral Anterior Cingulate Cortex in Modulating Activity in the Amygdala. *Neuron, 51*(6), 871-882. doi:http://dx.doi.org/10.1016/j.neuron.2006.07.029

Ettinger, U., Mohr, C., Gooding, D. C., Cohen, A. S., Rapp, A., Haenschel, C., & Park, S. (2015). Cognition and brain function in schizotypy: a selective review. *Schizophr Bull, 41 Suppl 2*, S417-426. doi:10.1093/schbul/sbu190

Evans, C. E., Kemish, K., & Turnbull, O. H. (2004). Paradoxical effects of education on the Iowa Gambling Task. *Brain Cogn, 54*(3), 240-244. doi:10.1016/j.bandc.2004.02.022

Evans, K. L., & Hampson, E. (2014). Does risk-taking mediate the relationship between testosterone and decision-making on the Iowa Gambling Task? *Personality and Individual Differences, 61*, 57-62.

Field, T. M., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expression by neonates. *Science, 218*(4568), 179-181. doi:10.1126/science.7123230

Fitzgerald, D. A., Angstadt, M., Jelsone, L. M., Nathan, P. J., & Phan, K. L. (2006). Beyond threat: amygdala reactivity across multiple expressions of facial affect. *Neuroimage, 30*(4), 1441-1448. doi:10.1016/j.neuroimage.2005.11.003

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science, 308*(5722), 662-667. doi:10.1126/science.1106138

Fogassi, L., & Luppino, G. (2005). Motor functions of the parietal lobe. *Curr Opin Neurobiol, 15*(6), 626-631. doi:10.1016/j.conb.2005.10.015

Frank, D. W., Costa, V. D., Averbeck, B. B., & Sabatinelli, D. (2019). Directional interconnectivity of the human amygdala, fusiform gyrus, and orbitofrontal cortex in emotional scene perception. *J Neurophysiol, 122*(4), 1530-1537. doi:10.1152/jn.00780.2018

Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., & Bebbington, P. E. (2002). A cognitive model of persecutory delusions. *Br J Clin Psychol, 41*(Pt 4), 331-347. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12437789

Friston, K. J., & Frith, C. D. (1995). Schizophrenia: a disconnection syndrome? *Clin Neurosci, 3*(2), 89-97. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7583624

Frith, C. D., & Corcoran, R. (1996). Exploring 'theory of mind' in people with schizophrenia. *Psychol Med, 26*(3), 521-530. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8733211

Frith, C. D., & Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron, 60*(3), 503-510. doi:10.1016/j.neuron.2008.10.032

Frith, U. (2004). Emanuel Miller lecture: confusions and controversies about Asperger syndrome. *J Child Psychol Psychiatry, 45*(4), 672-686. doi:10.1111/j.1469-7610.2004.00262.x

Frith, U., & Frith, C. (2001). The biological basis of social interaction. *Current Directions in Psychological Science, 10*(5), 151-155. doi:Doi 10.1111/1467-8721.00137

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philos Trans R Soc Lond B Biol Sci, 358*(1431), 459-473. doi:10.1098/rstb.2002.1218

Fuelscher, I., Caeyenberghs, K., Enticott, P. G., Kirkovski, M., Farquharson, S., Lum, J., & Hyde, C. (2019). Does fMRI repetition suppression reveal mirror neuron activity in the human brain? Insights from univariate and multivariate analysis. *Eur J Neurosci*. doi:10.1111/ejn.14370

Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., . . . Politi, P. (2009). Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J Psychiatry Neurosci, 34*(6), 418-432. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19949718

Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and cognition, 17*(2), 535-543.
Gallese, V. (2003a). The manifold nature of interpersonal relations: the quest for a common mechanism. *Philos Trans R Soc Lond B Biol Sci, 358*(1431), 517-528. doi:10.1098/rstb.2002.1234
Gallese, V. (2003b). The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology, 36*(4), 171-180. doi:10.1159/000072786
Gallese, V. (2007a). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philos Trans R Soc Lond B Biol Sci, 362*(1480), 659-669. doi:10.1098/rstb.2006.2002
Gallese, V. (2007b). Embodied simulation: from mirror neuron systems to interpersonal relations. *Novartis Found Symp, 278*, 3-12; discussion 12-19, 89-96, 216-221. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/17214307
Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119 (Pt 2)*, 593-609. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8800951
Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends Cogn Sci, 2*(12), 493-501. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/21227300
Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends Cogn Sci, 8*(9), 396-403. doi:10.1016/j.tics.2004.07.002
Gayer-Anderson, C., & Morgan, C. (2013). Social networks, support and early psychosis: a systematic review. *Epidemiol Psychiatr Sci, 22*(2), 131-146. doi:10.1017/S2045796012000406
Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cereb Cortex, 19*(6), 1239-1255. doi:10.1093/cercor/bhn181
Geday, J., Gjedde, A., Boldsen, A. S., & Kupers, R. (2003). Emotional valence modulates activity in the posterior fusiform gyrus and inferior medial prefrontal cortex in social perception. *Neuroimage, 18*(3), 675-684. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12667845
Gläscher, J., Tüscher, O., Weiller, C., & Büchel, C. (2004). Elevated responses to constant facial emotions in different faces in the human amygdala: an fMRI study of facial identity and expression. *BMC Neurosci, 5*, 45. doi:10.1186/1471-2202-5-45
Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., & Haxby, J. V. (2007). Two takes on the social brain: a comparison of theory of mind tasks. *J Cogn Neurosci, 19*(11), 1803-1814. doi:10.1162/jocn.2007.19.11.1803
Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience, 41*(1), 1-24. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1676137
Green, M. F., Horan, W. P., Barch, D. M., & Gold, J. M. (2015). Effort-based decision making: a novel approach for assessing motivation in schizophrenia. *Schizophrenia bulletin, 41*(5), 1035-1044.
Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc Natl Acad Sci U S A, 100*(1), 253-258. doi:10.1073/pnas.0135058100
Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci, 10*(1), 14-23. doi:10.1016/j.tics.2005.11.006
Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci, 7*(5), 555-562. doi:10.1038/nn1224
Gur, R. C., Schroeder, L., Turner, T., McGrath, C., Chan, R. M., Turetsky, B. I., . . . Gur, R. E. (2002). Brain activation during facial emotion processing. *Neuroimage, 16*(3 Pt 1), 651-662. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12169250

Habel, U., Windischberger, C., Derntl, B., Robinson, S., Kryspin-Exner, I., Gur, R. C., & Moser, E. (2007). Amygdala activation and facial expressions: explicit emotion discrimination versus implicit emotion processing. *Neuropsychologia, 45*(10), 2369-2377. doi:10.1016/j.neuropsychologia.2007.01.023

Hahn, A. C., & Perrett, D. I. (2014). Neural and behavioral responses to attractiveness in adult and infant faces. *Neurosci Biobehav Rev, 46 Pt 4*, 591-603. doi:10.1016/j.neubiorev.2014.08.015

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., . . . Berke, J. D. (2016). Mesolimbic dopamine signals the value of work. *Nat Neurosci, 19*(1), 117-126. doi:10.1038/nn.4173

Hamilton, A. F. d. C. (2016). Gazing at me: the importance of social meaning in understanding direct-gaze cues. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1686), 20150080.

Harari, H., Shamay-Tsoory, S. G., Ravid, M., & Levkovitz, Y. (2010). Double dissociation between cognitive and affective empathy in borderline personality disorder. *Psychiatry Res, 175*(3), 277-279. doi:10.1016/j.psychres.2009.03.002

Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F., & Weinberger, D. R. (2002). The amygdala response to emotional stimuli: a comparison of faces and scenes. *Neuroimage, 17*(1), 317-323. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12482086

Hass, J., Hertag, L., & Durstewitz, D. (2016). A Detailed Data-Driven Network Model of Prefrontal Cortex Reproduces Key Features of In Vivo Activity. *PLoS Comput Biol, 12*(5), e1004930. doi:10.1371/journal.pcbi.1004930

Haxby, J. V., & Gobbini, M. I. (2011). *Distributed neural systems for face perception*: The Oxford Handbook of Face Perception.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in cognitive sciences, 4*(6), 223-233.

Heerey, E. A., Bell-Warren, K. R., & Gold, J. M. (2008). Decision-making impairments in the context of intact reward sensitivity in schizophrenia. *Biol Psychiatry, 64*(1), 62-69. doi:10.1016/j.biopsych.2008.02.015

Heinz, A., & Schlagenhauf, F. (2010). Dopaminergic dysfunction in schizophrenia: salience attribution revisited. *Schizophr Bull, 36*(3), 472-485. doi:10.1093/schbul/sbq031

Heinzel, A., Bermpohl, F., Niese, R., Pfennig, A., Pascual-Leone, A., Schlaug, G., & Northoff, G. (2005). How do we modulate our emotions? Parametric fMRI reveals cortical midline structures as regions specifically involved in the processing of emotional valences. *Brain Res Cogn Brain Res, 25*(1), 348-358. doi:10.1016/j.cogbrainres.2005.06.009

Hennenlotter, A., Schroeder, U., Erhard, P., Castrop, F., Haslinger, B., Stoecker, D., . . . Ceballos-Baumann, A. O. (2005). A common neural basis for receptive and expressive communication of pleasant facial affect. *Neuroimage, 26*(2), 581-591. doi:10.1016/j.neuroimage.2005.01.057

Herrington, J. D., Taylor, J. M., Grupe, D. W., Curby, K. M., & Schultz, R. T. (2011). Bidirectional communication between amygdala and fusiform gyrus during facial recognition. *Neuroimage, 56*(4), 2348-2355. doi:10.1016/j.neuroimage.2011.03.072

Herweg, N. A., Apitz, T., Leicht, G., Mulert, C., Fuentemilla, L., & Bunzeck, N. (2016). Theta-Alpha Oscillations Bind the Hippocampus, Prefrontal Cortex, and Striatum during Recollection: Evidence from Simultaneous EEG-fMRI. *J Neurosci, 36*(12), 3579-3587. doi:10.1523/JNEUROSCI.3629-15.2016

Herzmann, G., Schweinberger, S. R., Sommer, W., & Jentzsch, I. (2004). What's special about personally familiar faces? A multimodal approach. *Psychophysiology, 41*(5), 688-701. doi:10.1111/j.1469-8986.2004.00196.x

Howes, O. D., & Kapur, S. (2009). The dopamine hypothesis of schizophrenia: version III--the final common pathway. *Schizophr Bull, 35*(3), 549-562. doi:10.1093/schbul/sbp006

Huq, S. F., Garety, P. A., & Hemsley, D. R. (1988). Probabilistic judgements in deluded and non-deluded subjects. *Q J Exp Psychol A, 40*(4), 801-812. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/3212213

Iacoboni, M. (2009). Neurobiology of imitation. *Curr Opin Neurobiol, 19*(6), 661-665. doi:10.1016/j.conb.2009.09.008

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nat Rev Neurosci, 7*(12), 942-951. doi:10.1038/nrn2024

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol, 3*(3), e79. doi:10.1371/journal.pbio.0030079

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286*(5449), 2526-2528. doi:10.1126/science.286.5449.2526

Ishai, A., Pessoa, L., Bikle, P. C., & Ungerleider, L. G. (2004). Repetition suppression of faces is modulated by emotion. *Proceedings of the National Academy of Sciences, 101*(26), 9827-9832.

Izard, C. E. (1994). Innate and Universal Facial Expressions - Evidence from Developmental and Cross-Cultural Research. *Psychological Bulletin, 115*(2), 288-299. doi:Doi 10.1037/0033-2909.115.2.288

Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron, 58*(2), 284-294.

Jabbi, M., & Keysers, C. (2008). Inferior frontal gyrus activity triggers anterior insula response to emotional facial expressions. *Emotion, 8*(6), 775-780. doi:10.1037/a0014194

Jackson, M. E., & Moghaddam, B. (2001). Amygdala regulation of nucleus accumbens dopamine output is governed by the prefrontal cortex. *J Neurosci, 21*(2), 676-681. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11160446

Johnson-Frey, S. H., Maloof, F. R., Newman-Norlund, R., Farrer, C., Inati, S., & Grafton, S. T. (2003). Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron, 39*(6), 1053-1058. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12971903

Juckel, G., Schlagenhauf, F., Koslowski, M., Wüstenberg, T., Villringer, A., Knutson, B., . . . Heinz, A. (2006). Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage, 29*(2), 409-416. doi:10.1016/j.neuroimage.2005.07.051

Kaiser, D., Walther, C., Schweinberger, S. R., & Kovacs, G. (2013). Dissociating the neural bases of repetition-priming and adaptation in the human brain for faces. *J Neurophysiol, 110*(12), 2727-2738. doi:10.1152/jn.00277.2013

Kalbe, E., Schlegel, M., Sack, A. T., Nowak, D. A., Dafotakis, M., Bangard, C., . . . Kessler, J. (2010). Dissociating cognitive from affective theory of mind: a TMS study. *Cortex, 46*(6), 769-780. doi:10.1016/j.cortex.2009.07.010

Kanai, R., Bahrami, B., Roylance, R., & Rees, G. (2012). Online social network size is reflected in human brain structure. *Proc Biol Sci, 279*(1732), 1327-1334. doi:10.1098/rspb.2011.1959

Kanske, P., Böckler, A., Trautwein, F. M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and Theory of Mind. *Neuroimage, 122*, 6-19. doi:10.1016/j.neuroimage.2015.07.082

Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry, 160*(1), 13-23. doi:10.1176/appi.ajp.160.1.13

Kapur, S., Mizrahi, R., & Li, M. (2005). From dopamine to salience to psychosis--linking biology, pharmacology and phenomenology of psychosis. *Schizophr Res, 79*(1), 59-68. doi:10.1016/j.schres.2005.01.003

Kawasaki, H., Tsuchiya, N., Kovach, C. K., Nourski, K. V., Oya, H., Howard, M. A., & Adolphs, R. (2012). Processing of Facial Emotion in the Human Fusiform Gyrus. *Journal of Cognitive Neuroscience, 24*(6), 1358-1370. doi:DOI 10.1162/jocn_a_00175

Kennerley, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J., & Rushworth, M. F. (2006). Optimal decision making and the anterior cingulate cortex. *Nat Neurosci, 9*(7), 940-947. doi:10.1038/nn1724

Keysers, C., & Fadiga, L. (2008). The mirror neuron system: new frontiers. *Soc Neurosci, 3*(3-4), 193-198. doi:10.1080/17470910802408513

Keysers, C., & Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Curr Opin Neurobiol, 19*(6), 666-671. doi:10.1016/j.conb.2009.10.006

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *J Neurosci, 29*(32), 10153-10159. doi:10.1523/JNEUROSCI.2668-09.2009

Kirsch, P., Ronshausen, S., Mier, D., & Gallhofer, B. (2007). The influence of antipsychotic treatment on brain reward system reactivity in schizophrenia patients. *Pharmacopsychiatry*.

Kirsch, P., Schienle, A., Stark, R., Sammer, G., Blecker, C., Walter, B., . . . Vaitl, D. (2003). Anticipation of reward in a nonaversive differential conditioning paradigm and the brain reward system: an event-related fMRI study. *Neuroimage, 20*(2), 1086-1095. doi:10.1016/S1053-8119(03)00381-1

Kish, S. J., Shannak, K., & Hornykiewicz, O. (1988). Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. Pathophysiologic and clinical implications. *N Engl J Med, 318*(14), 876-880. doi:10.1056/NEJM198804073181402

Knafo, A., Zahn-Waxler, C., Davidov, M., Van Hulle, C., Robinson, J. L., & Rhee, S. H. (2009). Empathy in early childhood: genetic, environmental, and affective contributions. *Ann N Y Acad Sci, 1167*, 103-114. doi:10.1111/j.1749-6632.2009.04540.x

Knafo, A., Zahn-Waxler, C., Van Hulle, C., Robinson, J. L., & Rhee, S. H. (2008). The developmental origins of a disposition toward empathy: Genetic and environmental contributions. *Emotion, 8*(6), 737-752. doi:10.1037/a0014179

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci, 21*(16), RC159. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11459880

Kohler, C. G., Turner, T. H., Bilker, W. B., Brensinger, C. M., Siegel, S. J., Kanes, S. J., . . . Gur, R. C. (2003). Facial emotion recognition in schizophrenia: intensity effects and error pattern. *Am J Psychiatry, 160*(10), 1768-1774. doi:10.1176/appi.ajp.160.10.1768

Kohler, C. G., Walker, J. B., Martin, E. A., Healey, K. M., & Moberg, P. J. (2010). Facial emotion perception in schizophrenia: a meta-analytic review. *Schizophr Bull, 36*(5), 1009-1019. doi:10.1093/schbul/sbn192

Koski, L., Wohlschläger, A., Bekkering, H., Woods, R. P., Dubeau, M. C., Mazziotta, J. C., & Iacoboni, M. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cereb Cortex, 12*(8), 847-855. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12122033

Krach, S., Paulus, F. M., Bodden, M., & Kircher, T. (2010). The rewarding nature of social interactions. *Front Behav Neurosci, 4*, 22. doi:10.3389/fnbeh.2010.00022

Kraskov, A., Dancause, N., Quallo, M. M., Shepherd, S., & Lemon, R. N. (2009). Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for action suppression? *Neuron, 64*(6), 922-930. doi:10.1016/j.neuron.2009.12.010

Kret, M. E., Sinke, C. B., & de Gelder, B. (2011). Emotion perception and health. In *Emotion regulation and well-being* (pp. 261-280): Springer.

Kringelbach, M. L., & Berridge, K. C. (2010). The functional neuroanatomy of pleasure and happiness. *Discov Med, 9*(49), 579-587. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20587348

Kucharska-Pietura, K., & Mortimer, A. (2013). Can antipsychotics improve social cognition in patients with schizophrenia? *CNS Drugs, 27*(5), 335-343. doi:10.1007/s40263-013-0047-0

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

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind". *Trends Cogn Sci, 8*(12), 528-533. doi:10.1016/j.tics.2004.10.001

Leslie, K. R., Johnson-Frey, S. H., & Grafton, S. T. (2004). Functional imaging of face and hand imitation: towards a motor theory of empathy. *Neuroimage, 21*(2), 601-607. doi:10.1016/j.neuroimage.2003.09.038

Liberzon, I., Phan, K. L., Decker, L. R., & Taylor, S. F. (2003). Extended amygdala and emotional salience: a PET activation study of positive and negative affect. *Neuropsychopharmacology, 28*(4), 726-733. doi:10.1038/sj.npp.1300113

Lingnau, A., Gesierich, B., & Caramazza, A. (2009). Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proc Natl Acad Sci U S A, 106*(24), 9925-9930. doi:10.1073/pnas.0902262106

Liu, J., Harris, A., & Kanwisher, N. (2010). Perception of face parts and face configurations: an FMRI study. *J Cogn Neurosci, 22*(1), 203-211. doi:10.1162/jocn.2009.21203

LoBue, V. (2009). More than just another face in the crowd: Superior detection of threatening facial expressions in children and adults. *Developmental science, 12*(2), 305-313.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF. Retrieved from http://www.emotionlab.se/resources/kdef

Luyten, P., & Fonagy, P. (2015). The neurobiology of mentalizing. *Personal Disord, 6*(4), 366-379. doi:10.1037/per0000117

Maia, T. V., & Frank, M. J. (2017). An integrative perspective on the role of dopamine in schizophrenia. *Biological Psychiatry, 81*(1), 52-66.

Maldjian, J. A., Laurienti, P. J., & Burdette, J. H. (2004). Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *Neuroimage, 21*(1), 450-455. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/14741682

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage, 19*(3), 1233-1239. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12880848

Mars, R. B., Neubert, F.-X., Noonan, M. P., Sallet, J., Toni, I., & Rushworth, M. F. (2012). On the relationship between the “default mode network” and the “social brain”. *Frontiers in human neuroscience, 6*, 189.

Matthews, S. C., Simmons, A. N., Lane, S. D., & Paulus, M. P. (2004). Selective activation of the nucleus accumbens during risk-taking decision making. *Neuroreport, 15*(13), 2123-2127. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15486494

Matzke, B., Herpertz, S. C., Berger, C., Fleischer, M., & Domes, G. (2014). Facial reactions during emotion recognition in borderline personality disorder: a facial electromyography study. *Psychopathology, 47*(2), 101-110. doi:10.1159/000351122

McDougle, S. D., Bond, K. M., & Taylor, J. A. (2015). Explicit and Implicit Processes Constitute the Fast and Slow Processes of Sensorimotor Learning. *J Neurosci, 35*(26), 9568-9579. doi:10.1523/JNEUROSCI.5061-14.2015

McLean, B. F., Mattiske, J. K., & Balzan, R. P. (2017). Association of the Jumping to Conclusions and Evidence Integration Biases With Delusions in Psychosis: A Detailed Meta-analysis. *Schizophr Bull, 43*(2), 344-354. doi:10.1093/schbul/sbw056

Mehu, M., & Scherer, K. R. (2015). Emotion categories and dimensions in the facial communication of affect: An integrated approach. *Emotion, 15*(6), 798-811. doi:10.1037/a0039416

Meltzoff, A. N. (2002). Imitation as a mechanism of social cognition: Origins of empathy, theory of mind, and the representation of action. *Blackwell handbook of childhood cognitive development*, 6-25.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*(4312), 75-78. doi:10.1126/science.198.4312.75

Meyer-Lindenberg, A., Nichols, T., Callicott, J. H., Ding, J., Kolachana, B., Buckholtz, J., . . . Weinberger, D. R. (2006). Impact of complex genetic variation in COMT on human brain function. *Mol Psychiatry, 11*(9), 867-877, 797. doi:10.1038/sj.mp.4001860

Mier, D., Eisenacher, S., Rausch, F., Englisch, S., Gerchen, M. F., Zamoscik, V., . . . Kirsch, P. (2016). Aberrant activity and connectivity of the posterior superior temporal sulcus

during social cognition in schizophrenia. *Eur Arch Psychiatry Clin Neurosci*, 1-14. doi:10.1007/s00406-016-0737-y

Mier, D., Haddad, L., Diers, K., Dressing, H., Meyer-Lindenberg, A., & Kirsch, P. (2014). Reduced embodied simulation in psychopathy. *World J Biol Psychiatry, 15*(6), 479-487. doi:10.3109/15622975.2014.902541

Mier, D., & Kirsch, P. (2016). Social-Cognitive Deficits in Schizophrenia. *Curr Top Behav Neurosci*. doi:10.1007/7854_2015_427

Mier, D., & Kirsch, P. (2017). Social-Cognitive Deficits in Schizophrenia. *Curr Top Behav Neurosci, 30*, 397-409. doi:10.1007/7854_2015_427

Mier, D., Kirsch, P., & Meyer-Lindenberg, A. (2010). Neural substrates of pleiotropic action of genetic variation in COMT: a meta-analysis. *Mol Psychiatry, 15*(9), 918-927. doi:10.1038/mp.2009.36

Mier, D., Lis, S., Neuthe, K., Sauer, C., Esslinger, C., Gallhofer, B., & Kirsch, P. (2010). The involvement of emotion recognition in affective theory of mind. *Psychophysiology, 47*(6), 1028-1039. doi:10.1111/j.1469-8986.2010.01031.x

Mier, D., Lis, S., Zygrodnik, K., Sauer, C., Ulferts, J., Gallhofer, B., & Kirsch, P. (2014). Evidence for altered amygdala activation in schizophrenia in an adaptive emotion recognition task. *Psychiatry Res, 221*(3), 195-203. doi:10.1016/j.pscychresns.2013.12.001

Mier, D., Sauer, C., Lis, S., Esslinger, C., Wilhelm, J., Gallhofer, B., & Kirsch, P. (2010). Neuronal correlates of affective theory of mind in schizophrenia out-patients: evidence for a baseline deficit. *Psychol Med*, 1-11. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20056024

Mohamed, T. N., Neumann, M. F., & Schweinberger, S. R. (2011). Combined effects of attention and inversion on event-related potentials to human bodies and faces. *Cogn Neurosci, 2*(3-4), 138-146. doi:10.1080/17588928.2011.597848

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2009). Is the mirror neuron system involved in imitation? A short review and meta-analysis. *Neurosci Biobehav Rev, 33*(7), 975-980. doi:10.1016/j.neubiorev.2009.03.010

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2012). Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. *Neurosci Biobehav Rev, 36*(1), 341-349. doi:10.1016/j.neubiorev.2011.07.004

Molnar-Szakacs, I., Iacoboni, M., Koski, L., & Mazziotta, J. C. (2005). Functional segregation within pars opercularis of the inferior frontal gyrus: evidence from fMRI studies of imitation and action observation. *Cereb Cortex, 15*(7), 986-994. doi:10.1093/cercor/bhh199

Monroe, J. F., Griffin, M., Pinkham, A., Loughead, J., Gur, R. C., Roberts, T. P., & Christopher Edgar, J. (2013). The fusiform response to faces: explicit versus implicit processing of emotion. *Hum Brain Mapp, 34*(1), 1-11. doi:10.1002/hbm.21406

Montagne, B., Kessels, R. P., Frigerio, E., de Haan, E. H., & Perrett, D. I. (2005). Sex differences in the perception of affective facial expressions: do men really lack emotional sensitivity? *Cogn Process, 6*(2), 136-141. doi:10.1007/s10339-005-0050-6

Montgomery, K. J., Isenberg, N., & Haxby, J. V. (2007). Communicative hand gestures and object-directed hand movements activated the mirror neuron system. *Soc Cogn Affect Neurosci, 2*(2), 114-122. doi:10.1093/scan/nsm004

Moritz, S., & Woodward, T. S. (2005). Jumping to conclusions in delusional and non-delusional schizophrenic patients. *Br J Clin Psychol, 44*(Pt 2), 193-207. doi:10.1348/014466505X35678

Morris, R. W., Vercammen, A., Lenroot, R., Moore, L., Langton, J. M., Short, B., . . . Weickert, T. W. (2012). Disambiguating ventral striatum fMRI-related BOLD signal during reward prediction in schizophrenia. *Mol Psychiatry, 17*(3), 235, 280-239. doi:10.1038/mp.2011.75

Morrison, I., & Downing, P. E. (2007). Organization of felt and seen pain responses in anterior cingulate cortex. *Neuroimage, 37*(2), 642-651. doi:10.1016/j.neuroimage.2007.03.079

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Curr Biol, 20*(8), 750-756. doi:10.1016/j.cub.2010.02.045

Mukherjee, P., Whalley, H. C., McKirdy, J. W., Sprengelmeyer, R., Young, A. W., McIntosh, A. M., . . . Hall, J. (2014). Altered amygdala connectivity within the social brain in schizophrenia. *Schizophr Bull, 40*(1), 152-160. doi:10.1093/schbul/sbt086

Newman-Norlund, R., van Schie, H. T., van Hoek, M. E., Cuijpers, R. H., & Bekkering, H. (2010). The role of inferior frontal and parietal areas in differentiating meaningful and meaningless object-directed actions. *Brain Res, 1315*, 63-74. doi:10.1016/j.brainres.2009.11.065

Nicholson, T., Roser, M., & Bach, P. (2017). Understanding the Goals of Everyday Instrumental Actions Is Primarily Linked to Object, Not Motor-Kinematic, Information: Evidence from fMRI. *PLoS One, 12*(1), e0169700. doi:10.1371/journal.pone.0169700

Nielsen, M. Ø., Rostrup, E., Wulff, S., Bak, N., Lublin, H., Kapur, S., & Glenthøj, B. (2012). Alterations of the brain reward system in antipsychotic naive schizophrenia patients. *Biological Psychiatry, 71*(10), 898-905.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci, 10*(9), 424-430. doi:10.1016/j.tics.2006.07.005

Nummenmaa, L., Hirvonen, J., Parkkola, R., & Hietanen, J. K. (2008). Is emotional contagion special? An fMRI study on neural systems for affective and cognitive empathy. *Neuroimage, 43*(3), 571-580. doi:10.1016/j.neuroimage.2008.08.014

Ocampo, B., & Kritikos, A. (2011). Interpreting actions: the goal behind mirror neuron function. *Brain Res Rev, 67*(1-2), 260-267. doi:10.1016/j.brainresrev.2011.03.001

Öhman, A. (2005). The role of the amygdala in human fear: automatic detection of threat. *Psychoneuroendocrinology, 30*(10), 953-958. doi:10.1016/j.psyneuen.2005.03.019

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255-258. doi:10.1126/science.1107621

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychol Rev, 97*(3), 315-331. doi:10.1037/0033-295x.97.3.315

Owen, A. M. (2004). Cognitive dysfunction in Parkinson's disease: the role of frontostriatal circuitry. *Neuroscientist, 10*(6), 525-537. doi:10.1177/1073858404266776

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *J Neurosci, 30*(30), 10127-10134. doi:10.1523/JNEUROSCI.2161-10.2010

Pereira, C. (2000). *Dimensions of emotional meaning in speech.* Paper presented at the ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion.

Perner, J., & Leekam, S. (2008). The curious incident of the photo that was accused of being false: issues of domain specificity in development, autism, and brain imaging. *Q J Exp Psychol (Hove), 61*(1), 76-89. doi:10.1080/17470210701508756

Perner, J., & Roessler, J. (2010). Teleology and causal understanding in childrens' theory of mind. *Causing human action: New perspectives on the causal theory of action*, 199-228.

Perry, D., Walder, K., Hendler, T., & Shamay-Tsoory, S. G. (2013). The gender you are and the gender you like: sexual preference and empathic neural responses. *Brain Res, 1534*, 66-75. doi:10.1016/j.brainres.2013.08.040

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage, 16*(2), 331-348. doi:10.1006/nimg.2002.1087

Philiastides, M. G., Auksztulewicz, R., Heekeren, H. R., & Blankenburg, F. (2011). Causal role of dorsolateral prefrontal cortex in human perceptual decision making. *Curr Biol, 21*(11), 980-983. doi:10.1016/j.cub.2011.04.034

Pinkham, A. E., Griffin, M., Baron, R., Sasson, N. J., & Gur, R. C. (2010). The face in the crowd effect: anger superiority when using real faces and multiple identities. *Emotion, 10*(1), 141-146. doi:10.1037/a0017387

Pitcher, D., Walsh, V., & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Exp Brain Res, 209*(4), 481-493. doi:10.1007/s00221-011-2579-1

Poletti, S., Radaelli, D., Cavallaro, R., Bosia, M., Lorenzi, C., Pirovano, A., . . . Benedetti, F. (2013). Catechol-O-methyltransferase (COMT) genotype biases neural correlates of empathy and perceived personal distress in schizophrenia. *Compr Psychiatry, 54*(2), 181-186. doi:10.1016/j.comppsych.2012.06.008

Prather, J. F., Peters, S., Nowicki, S., & Mooney, R. (2008). Precise auditory–vocal mirroring in neurons for learned vocal communication. *Nature, 451*(7176), 305.

Preston, S. D., & de Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behav Brain Sci, 25*(1), 1-20; discussion 20-71. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12625087

Raine, A. (1991). The SPQ: a scale for the assessment of schizotypal personality based on DSM-III-R criteria. *Schizophr Bull, 17*(4), 555-564. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1805349

Rausch, F., Mier, D., Eifler, S., Esslinger, C., Schilling, C., Schirmbeck, F., . . . Zink, M. (2014). Reduced activation in ventral striatum and ventral tegmental area during probabilistic decision-making in schizophrenia. *Schizophr Res, 156*(2-3), 143-149. doi:10.1016/j.schres.2014.04.020

Rausch, F., Mier, D., Eifler, S., Fenske, S., Schirmbeck, F., Englisch, S., . . . Zink, M. (2015). Reduced activation in the ventral striatum during probabilistic decision-making in patients in an at-risk mental state. *J Psychiatry Neurosci, 40*(3), 163-173. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/25622039

Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nat Rev Neurosci, 20*(8), 495-505. doi:10.1038/s41583-019-0179-4

Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., & Rozzi, S. (2014). Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol Rev, 94*(2), 655-706. doi:10.1152/physrev.00009.2013

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu Rev Neurosci, 27*, 169-192. doi:10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron, 31*(6), 889-901. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11580891

Rizzolatti, G., Luppino, G., & Matelli, M. (1998). The organization of the cortical motor system: new concepts. *Electroencephalogr Clin Neurophysiol, 106*(4), 283-296. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9741757

Roca, M., Torralva, T., Gleichgerrcht, E., Chade, A., Arevalo, G. G., Gershanik, O., & Manes, F. (2010). Impairments in social cognition in early medicated and unmedicated Parkinson disease. *Cogn Behav Neurol, 23*(3), 152-158. doi:10.1097/WNN.0b013e3181e078de

Rogers, R. D., Ramnani, N., Mackay, C., Wilson, J. L., Jezzard, P., Carter, C. S., & Smith, S. M. (2004). Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. *Biological Psychiatry, 55*(6), 594-602. doi:http://dx.doi.org/10.1016/j.biopsych.2003.11.012

Russell, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178. doi:DOI 10.1037/h0077714

Sabatinelli, D., Bradley, M. M., Lang, P. J., Costa, V. D., & Versace, F. (2007). Pleasure rather than salience activates human nucleus accumbens and medial prefrontal cortex. *J Neurophysiol, 98*(3), 1374-1379. doi:10.1152/jn.00230.2007

Salgado, S., & Kaplitt, M. G. (2015). The Nucleus Accumbens: A Comprehensive Review. *Stereotact Funct Neurosurg, 93*(2), 75-93. doi:10.1159/000368279

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2012). Enhancing social ability by stimulating right temporoparietal junction. *Curr Biol, 22*(23), 2274-2277. doi:10.1016/j.cub.2012.10.018

Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: from imitation to Theory of Mind. *Cognition, 122*(2), 228-235. doi:10.1016/j.cognition.2011.11.004

Santos, A., Mier, D., Kirsch, P., & Meyer-Lindenberg, A. (2011). Evidence for a general face salience signal in human amygdala. *Neuroimage, 54*(4), 3111-3116. doi:10.1016/j.neuroimage.2010.11.024

Satpute, A. B., & Lieberman, M. D. (2006). Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain Res, 1079*(1), 86-97. doi:10.1016/j.brainres.2006.01.005

Savla, G. N., Vella, L., Armstrong, C. C., Penn, D. L., & Twamley, E. W. (2012). Deficits in domains of social cognition in schizophrenia: a meta-analysis of the empirical evidence. *Schizophrenia bulletin, 39*(5), 979-992.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annu Rev Psychol, 55*, 87-124. doi:10.1146/annurev.psych.55.090902.142044

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends Cogn Sci, 19*(2), 65-72. doi:10.1016/j.tics.2014.11.007

Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., & Pessiglione, M. (2012). Neural mechanisms underlying motivation of mental versus physical effort. *PLoS biology, 10*(2), e1001266.

Schulte-Rüther, M., Markowitsch, H. J., Fink, G. R., & Piefke, M. (2007). Mirror neuron and theory of mind mechanisms involved in face-to-face interactions: a functional magnetic resonance imaging approach to empathy. *J Cogn Neurosci, 19*(8), 1354-1372. doi:10.1162/jocn.2007.19.8.1354

Schurz, M., & Perner, J. (2015). An evaluation of neurocognitive models of theory of mind. *Front Psychol, 6*, 1610. doi:10.3389/fpsyg.2015.01610

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, 42*, 9-34. doi:10.1016/j.neubiorev.2014.01.009

Sergerie, K., Chochol, C., & Armony, J. L. (2008). The role of the amygdala in emotional processing: a quantitative meta-analysis of functional neuroimaging studies. *Neurosci Biobehav Rev, 32*(4), 811-830. doi:10.1016/j.neubiorev.2007.12.002

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain, 132*(Pt 3), 617-627. doi:10.1093/brain/awn279

Shifman, S., Bronstein, M., Sternfeld, M., Pisante-Shalom, A., Lev-Lehman, E., Weizman, A., . . . Darvasi, A. (2002). A highly significant association between a COMT haplotype and schizophrenia. *Am J Hum Genet, 71*(6), 1296-1302. doi:10.1086/344514

Speechley, W. J., Whitman, J. C., & Woodward, T. S. (2010). The contribution of hypersalience to the "jumping to conclusions" bias associated with delusions in schizophrenia. *J Psychiatry Neurosci, 35*(1), 7-17. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20040242

Spreckelmeyer, K. N., Krach, S., Kohls, G., Rademacher, L., Irmak, A., Konrad, K., . . . Gründer, G. (2009). Anticipation of monetary and social reward differently activates mesolimbic brain structures in men and women. *Soc Cogn Affect Neurosci, 4*(2), 158-165. doi:10.1093/scan/nsn051

St Onge, J. R., Ahn, S., Phillips, A. G., & Floresco, S. B. (2012). Dynamic fluctuations in dopamine efflux in the prefrontal cortex and nucleus accumbens during risk-based decision making. *J Neurosci, 32*(47), 16880-16891. doi:10.1523/JNEUROSCI.3807-12.2012

Stefanis, N. C., Smyrnis, N., Avramopoulos, D., Evdokimidis, I., Ntzoufras, I., & Stefanis, C. N. (2004). Factorial composition of self-rated schizotypal traits among young males undergoing military training. *Schizophrenia bulletin, 30*(2), 335-350.

Supekar, K., Uddin, L. Q., Khouzam, A., Phillips, J., Gaillard, W. D., Kenworthy, L. E., . . . Menon, V. (2013). Brain hyperconnectivity in children with autism and its links to social deficits. *Cell Rep, 5*(3), 738-747. doi:10.1016/j.celrep.2013.10.001

Taylor, M. J., Edmonds, G. E., McCarthy, G., & Allison, T. (2001). Eyes first! Eye processing develops before face processing in children. *Neuroreport, 12*(8), 1671-1676. doi:10.1097/00001756-200106130-00031

Thomas, R. M., De Sanctis, T., Gazzola, V., & Keysers, C. (2018). Where and how our brain represents the temporal structure of observed action. *Neuroimage, 183*, 677-697.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res, 168*(3), 242-249. doi:10.1016/j.psychres.2008.05.006

Tracy, J. L., & Randles, D. (2011). Four Models of Basic Emotions: A Review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emotion Review, 3*(4), 397-405. doi:10.1177/1754073911410747

Uddin, L. Q., Molnar-Szakacs, I., Zaidel, E., & Iacoboni, M. (2006). rTMS to the right inferior parietal lobule disrupts self-other discrimination. *Soc Cogn Affect Neurosci, 1*(1), 65-71. doi:10.1093/scan/nsl003

van der Gaag, C., Minderaa, R. B., & Keysers, C. (2007). Facial expressions: what the mirror neuron system can and cannot tell us. *Soc Neurosci, 2*(3-4), 179-222. doi:10.1080/17470910701376878

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Hum Brain Mapp, 30*(3), 829-858. doi:10.1002/hbm.20547

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage, 48*(3), 564-584. doi:10.1016/j.neuroimage.2009.06.009

Viinikainen, M., Jaaskelainen, I. P., Alexandrov, Y., Balk, M. H., Autti, T., & Sams, M. (2010). Nonlinear relationship between emotional valence and brain activity: evidence of separate negative and positive valence dimensions. *Hum Brain Mapp, 31*(7), 1030-1040. doi:10.1002/hbm.20915

Viviani, R., Dommes, L., Bosch, J., Stingl, J. C., & Beschoner, P. (2018). The Neural Correlates of Decisions About Sadness in Facial Expressions. *Journal of Neuroscience Psychology and Economics, 11*(2), 93-105. doi:10.1037/npe0000081

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron, 30*(3), 829-841. doi:10.1016/s0896-6273(01)00328-2

Wacker, J., Dillon, D. G., & Pizzagalli, D. A. (2009). The role of the nucleus accumbens and rostral anterior cingulate cortex in anhedonia: integration of resting EEG, fMRI, and volumetric techniques. *Neuroimage, 46*(1), 327-337. doi:10.1016/j.neuroimage.2009.01.058

Wai, M., & Tiliopoulos, N. (2012). The affective and cognitive empathic nature of the dark triad of personality. *Personality and Individual Differences, 52*(7), 794-799. doi:10.1016/j.paid.2012.01.008

Walter, H. (2012). Social Cognitive Neuroscience of Empathy: Concepts, Circuits, and Genes. *Emotion Review, 4*(1), 9-17. doi:10.1177/1754073911421379

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behav Res Methods, 42*(3), 671-684. doi:10.3758/BRM.42.3.671

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103-128. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/6681741

Winston, J. S., Henson, R. N., Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *J Neurophysiol, 92*(3), 1830-1839. doi:10.1152/jn.00155.2004

Woodward, T. S., Munz, M., LeClerc, C., & Lecomte, T. (2009). Change in delusions is associated with change in "jumping to conclusions". *Psychiatry Res, 170*(2-3), 124-127. doi:10.1016/j.psychres.2008.10.020

Wurm, M. F., Hrkac, M., Morikawa, Y., & Schubotz, R. I. (2014). Predicting goals in action episodes attenuates BOLD response in inferior frontal and occipitotemporal cortex. *Behav Brain Res, 274*, 108-117. doi:10.1016/j.bbr.2014.07.053

Zalocusky, K. A., Ramakrishnan, C., Lerner, T. N., Davidson, T. J., Knutson, B., & Deisseroth, K. (2016). Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature, 531*(7596), 642-646. doi:10.1038/nature17400

Zheng, J., Anderson, K. L., Leal, S. L., Shestyuk, A., Gulsen, G., Mnatsakanyan, L., . . . Lin, J. J. (2017). Amygdala-hippocampal dynamics during salient information processing. *Nat Commun, 8*, 14413. doi:10.1038/ncomms14413

8 OWN PUBLICATIONS

Erkic, M., Bailer, J., Fenske, S. C., Schmidt, S. N. L., Trojan, J., Schröder, A., . . . Mier, D. (2018). Impaired emotion processing and a reduction in trust in patients with somatic symptom disorder. *Clin Psychol Psychother, 25*(1), 163-172. doi:10.1002/cpp.2151

Hinterberger, T., Auer, J., Schmidt, S., & Loew, T. (2013). Evaluation of a salutogenetic concept for inpatient psychosomatic treatment. *Evid Based Complement Alternat Med, 2013*, 735731. doi:10.1155/2013/735731

Hinterberger, T., Schmidt, S., Kamei, T., & Walach, H. (2014). Decreased electrophysiological activity represents the conscious state of emptiness in meditation. *Front Psychol, 5*, 99. doi:10.3389/fpsyg.2014.00099

Schmidt, S. N. L., Fenske, S. C., Kirsch, P., & Mier, D. (2018). Nucleus accumbens activation is linked to salience in social decision making. *Eur Arch Psychiatry Clin Neurosci, 269*(6), 701-712. doi:10.1007/s00406-018-0947-6

Walter, C., Schmidt, S., Rosenstiel, W., Bogdan, M., & Gerjets, P. (2013). *Alpha-and theta frequencies as indicators for optimal cognitive load during learning.* Paper presented at the 6th International Cognitive Load Theory Conference.

Walter, C., Schmidt, S., Rosenstiel, W., Gerjets, P., & Bogdan, M. (2013). *Using cross-task classification for classifying workload levels in complex learning tasks.* Paper presented at the Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on.

Zamoscik, V., Mier, D., Schmidt, S. N., & Kirsch, P. (2016). Early Memories of Individuals on the Autism Spectrum Assessed Using Online Self-Reports. *Front Psychiatry, 7*, 79. doi:10.3389/fpsyt.2016.00079

Zamoscik, V. E., Schmidt, S. N. L., Gerchen, M. F., Samsouris, C., Timm, C., Kuehner, C., & Kirsch, P. (2018). Respiration pattern variability and related default mode network connectivity are altered in remitted depression. *Psychol Med, 48*(14), 2364-2374. doi:10.1017/S0033291717003890

9 CURRICULUM VITAE

PERSONALIEN

Vornamen und Nachname:	Stephanie Nicole Lyn Schmidt
Geburtsdatum:	17.02.1987
Geburtsort:	Bad Soden am Taunus
Familienstand:	ledig
Mutter:	Inge Sonja Schmidt, geb. Greiner
Vater:	Stephan Schmidt

SCHULISCHER WERDEGANG

1997-2006	Rabanus-Maurus Gymnasium, Mainz
17.03.2006	Abitur

UNIVERSITÄRER WERDEGANG

WS 2006/07	Beginn des Studiums Cognitive Science an der Universität Osnabrück
WS 2008/09	Forschungspraktikum Institute of Psychiatry, King‘s College, London, UK
08.12.2010	Bachelor of Science
WS 2010/11	Beginn des Studiums Bioinformatik an der Universität Tübingen *Masterarbeit: Development, Implementation and Application of a Learning Environment*
14.1.2013	Master of Science, Note: 2,2

10 ACKNOWLEDGEMENTS

Thank you, dear Prof. Dr. Peter Kirsch, for giving me the freedom to pursue these projects while still being there to offer a helping hand and critical thought. I highly appreciate that you cultivate best scientific standards in your department.

Thank you, dear Prof. Dr. Daniela Mier, for your guidance and mentorship, the value of which can hardly be expressed by words. Thank you, for making me the heart of your research ideas and inspiring me to learn such a variety of methods. Thank you, for challenging me to give it my all, and thank you for always supporting me!

I also thank the Heidelberger Akademie der Wissenschaften who financed the grant that allowed me to conduct these studies.

Thank you, dear Ellen Schmucker, for your valuable assistance with the measurements and for spreading the spirit of positivity and caring in the department.

Many thanks to my dear project assistants Manuel Vietze, Vera Eymann, Christian Sojer and Julian Schlierkamp and Zhimin Yan, for your help and companionship. Without you, the over 200 participants and almost 400 measurements would not have been half as joyful.

Thank you to my dear family and friends. Words cannot possibly express how grateful I am to have you. Thank you for being part of my wonderful life. I love you.

11 EIDESSTATTLICHE ERKLÄRUNG

zum Antrag auf Zulassung zur Promotion gemäß PromO „Dr.sc.hum.“

1. Ich habe an keiner anderen Stelle einen Antrag auf Zulassung zur Promotion gestellt oder bereits einen Doktortitel auf der Grundlage des vorgelegten Studienabschlusses erworben und mich auch nicht einer Doktorprüfung erfolglos unterzogen (dies schließt äquivalente Verfahren bzw. Titel ausserhalb Deutschlands ein).
2. Die an der Medizinischen Fakultät der Universität Heidelberg zur Promotion eingereichte Arbeit mit dem Titel:

Neural mechanisms of social cognition – the mirror neuron system and beyond

am

Zentralinstitut für Seelische Gesundheit

unter Anleitung von (Doktorvater/ -mutter)

Prof. Dr. phil. Peter Kirsch

habe ich selbst verfasst und bei der Abfassung der Arbeit keine anderen als die in der Abhandlung aufgeführten Hilfsmittel benutzt.
3. Die Arbeit oder Teile davon habe ich bislang an keiner Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.
4. Die Dissertation wurde ohne Hinzuziehung einer kommerziellen Promotionsberatung erstellt.
5. Mit der Veröffentlichung meines Lebenslaufes im Rahmen des Promotionsverfahrens (Dissertation) bin ich einverstanden.
6. Ich komme der Veröffentlichungspflicht gemäß § 13 PromO nach und stimme der Veröffentlichung der Zusammenfassung meiner Dissertation im Internet unter Angabe meines Namens und des Studienabschlusses zu.

7. Die Bedeutung der eidesstattlichen Erklärung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Erklärung sind mir bekannt.
Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Reichenau, 18.10.2019
Ort und Datum Unterschrift

The German text is legally binding.