

***Common and Distinct Neural Networks for Theory
of Mind Reasoning and Inhibitory Control***

Inaugural-Dissertation

zur Erlangung des Doktorgrades der Philosophie

an der Ludwig-Maximilians-Universität München



vorgelegt von

Christoph Rothmayr

Erstgutachter: *Prof. Dr. Beate Sodian*

Zweitgutachter: *Prof. Dr. Göran Hajak*

Tag der mündlichen Prüfung: *20. Juli 2009*

Contents

Acknowledgements.....	5
Abstract	6
Abbreviations	7
1. General Introduction.....	8
2. Theory of Mind (ToM) Reasoning	10
2.1 Development of ToM Reasoning.....	10
2.1.1 Cognitive Developments Preceding Belief-Reasoning.....	10
2.1.2 The Development of Belief-Reasoning and Its Investigation.....	11
2.2 Neural Correlates of ToM Reasoning.....	14
2.2.1 Results from Lesion Studies.....	15
2.2.2 Results from Neuroimaging Studies	19
2.2.2.1 PET Studies.....	20
2.2.2.2 FMRI Studies.....	21
2.3 The Mirror Neuron System, Theory Theory and Simulation Theory.....	27
3. Inhibitory Control (IC).....	31
3.1 Development and Measurement of Executive Functioning and IC.....	31
3.2 Neural Correlates of IC	33
3.2.1 Results from Lesion Studies.....	33
3.2.2 Results from Neuroimaging Studies	34
4. ToM and Inhibitory Control	37
4.1 Behavioral Results Concerning the ToM / IC Connection in Childhood	37
.....	37
4.1.1 Correlational Data	37
4.1.2 Training Effects	39
4.2 Adults and the ToM / IC Connection	40
4.3 Hypotheses Regarding the ToM / IC Relationship.....	41
4.4 Autism and ADHD: a Case in Point for the ToM / IC Connection?.....	43
4.4.1 Autism.....	43
4.4.2 ADHD.....	45
4.5 Neural Correlates of IC and ToM.....	46
4.5.1 Brain Lesions and Their Effects on ToM and IC	46
4.5.2 Neuroimaging of IC and ToM	47
4.6 Evidence for an Independent ToM Module.....	50
4.7 Investigating IC and Belief-Reasoning in a Single Study.....	51
4.8 Aim of the Present Study.....	57
5. Methods	60
5.1 fMRI Imaging	60
5.2 Pilot Experiment.....	62
5.3 Subjects.....	63
5.4 Experimental Design.....	63
5.4.1 Belief-Reasoning Experiment.....	64
5.4.2 Inhibitory Control Experiment	65
5.5 fMRI Set Up and Design.....	67
5.6 Data Analysis	67
5.6.1 Behavioral Data	67
5.6.2 Functional Data	67

5.6.2.1 Preprocessing	68
5.6.2.2 Statistical Analysis.....	69
6. Results.....	72
6.1 Behavioral Results.....	72
6.2 Functional Imaging Results	74
6.2.1 Belief-Reasoning Experiment.....	74
6.2.1.1 <i>True-Belief > Baseline</i>	74
6.2.1.2 <i>False-Belief > Baseline</i>	77
6.2.1.3 <i>False-Belief > True-Belief</i>	81
6.2.2 Inhibitory Control Experiment	85
6.2.2.1 <i>No-Go > Go</i>	85
6.2.3 Common Activity in the Belief-Reasoning and the Inhibitory Control Experiment	89
7. Discussion.....	92
7.1 Belief-Reasoning Networks.....	93
7.1.1 <i>True-Belief > Baseline</i> : Does True-Belief Reasoning Require Belief- Reasoning?.....	93
7.1.1.1 True-Belief Reasoning: Comparing Present Data to Previous Findings.....	96
7.1.2 <i>False-Belief > Baseline</i> : Decoupling and General Task Demands.....	97
7.1.2.1 Did the False-Belief Task Really Measure False-Belief Reasoning?	99
7.1.3 <i>False-Belief > True-Belief</i> : Neural Networks of a Decoupling Mechanism.....	100
7.1.3.1 Replication of Previous Studies	106
7.2 Brain Areas Related to IC	110
7.2.1 <i>No-go > Go</i> : A Neural Network for Response Inhibition.....	110
7.3 Common and Distinct Areas for Belief-Reasoning and IC.....	114
7.3.1 Distinct Areas Related to a Decoupling between Mentality and Reality ...	114
7.3.2 Distinct Areas Related to Inhibition.....	117
7.3.3 Common Neural Networks and Processes Involved in False-Belief Reasoning and IC	119
7.4 Integration of Current and Existing Data towards a Unified Theory for the Belief-Reasoning / IC Relation	124
7.5 Impact of the Current Findings on Modularity Accounts of Belief- Reasoning	129
7.6 Implications for the Understanding and Treatment of Compromised Belief-Reasoning.....	132
7.6.1 Autism.....	132
7.6.2 ADHD.....	134
7.6.3 The Papilio Project	136
7.7 Outlook	139
8. References.....	142
9. Index of Figures.....	164
10. Index of Tables	169
11. Appendices.....	170
12. Curriculum Vitae.....	181

Acknowledgements

First of all I would like to thank Professor Dr. Beate Sodian at the Department of Developmental Psychology at the Ludwig-Maximilians University for her constant support throughout the various stages in the making of this doctoral thesis. I greatly appreciate her valuable advice in all aspects of the project presented in this thesis.

Furthermore, I would like to thank Professor Dr. Göran Hajak at the Clinic and Policlinic for Psychiatry, Psychosomatics, and Psychotherapy at the University Medical Center Regensburg. I deeply value his constant encouragement to research freely and the opportunity he granted me to work with state-of-the-art functional magnetic resonance imaging.

Also, I am very grateful to Dr. Monika Sommer at the Clinical Neuroscience Center for Emotions and Social Cognition at the University of Regensburg who supervised all aspects of my thesis from the beginning. I appreciate her valuable and immediate help and support in the planning, measurement, analysis, and discussion of this project. Thank you.

Moreover, I would also like to thank all of my colleagues at my research group for their great help with data acquisition and analysis as well as for their valuable suggestions related to the discussion of the obtained results.

I also greatly appreciate the love and support given to me by my fiancée Mary and my family. Without them, the making of this thesis would not have been possible.

Last but not least, I would like to thank all subjects for participating in the study.

Abstract

Ample behavioral evidence has shown that the ability to attribute false beliefs as part of Theory of Mind (ToM) and the ability to inhibit a prepotent response are correlated in both children and adults. Lesion and functional imaging studies have indicated that both cognitive processes might even be supported by common areas of the brain. Some controversy also exists over whether there are areas in the brain that are specifically dedicated to the attribution of beliefs or whether these areas may also be engaged in other processes such as inhibitory control. Among these candidate regions are the right temporo-parietal junction (TPJ) and the medial prefrontal cortex (PFC). To date however, no imaging study has investigated both false-belief reasoning and inhibitory control in the same set of subjects using an equal set of picture stimuli.

This is what the present study attempted to do. A classical false belief task was used to study false-belief reasoning, whereas a Go / No-go paradigm was employed to probe inhibitory control. After an initial pilot experiment, a total of 12 subjects took part in the main study that was conducted in a high-field 3-Tesla functional magnetic resonance imaging (fMRI) scanner.

A subsequent random-effects analysis of the group data revealed common activity for both false-belief reasoning and inhibitory control in the bilateral TPJ, dorsal medial PFC, and right middle temporal gyrus. This result indicates that inhibitory control and false-belief reasoning may both rely on common underlying processes such as attention reorienting and conflict detection and inhibition. Additionally, the results show that false-belief reasoning may also require self-referential processes mediated by ventral medial PFC. Furthermore, this result yields the assumption that neither the right TPJ nor dorsal medial PFC may constitute a specific ToM module.

The current results are also discussed with respect to existing and future ToM and executive functioning training methods for children. Also, developmental disorders such as autism and attention deficit hyperactivity disorder (ADHD) are discussed in light of the novel results presented.

Abbreviations

ACC	<i>Anterior Cingulate Cortex</i>
ADHD	<i>Attention-Deficit Hyperactivity Disorder</i>
ALE	<i>Activation-Likelihood-Estimation</i>
ASD	<i>Autism Spectrum Disorders</i>
BA	<i>Brodmann Area</i>
BOLD	<i>Blood-Oxygen-Level Dependent</i>
DICOM	<i>Digital Imaging and Communications in Medicine</i>
df	<i>degrees of freedom</i>
DLPFC	<i>Dorso-Lateral Prefrontal Cortex</i>
EPI	<i>Echo-Planar Imaging</i>
EF	<i>Executive Functioning</i>
FB	<i>False-Belief</i>
fMRI	<i>functional Magnetic Resonance Imaging</i>
FoV	<i>Field-of-View</i>
FWHM	<i>Full-Width Half Maximum</i>
IC	<i>Inhibitory Control</i>
<i>M</i>	<i>Mean</i>
N/A	<i>Not Applicable</i>
NIfTI	<i>Neuroimaging Informatics Technology Initiative</i>
MNI	<i>Montreal Neurological Institute</i>
MNS	<i>Mirror Neuron System</i>
n.s.	<i>not significant</i>
PET	<i>Positron-Emission Tomography</i>
PFC	<i>Prefrontal Cortex</i>
RT	<i>Reaction Time</i>
SD	<i>Standard Deviation</i>
SNR	<i>Signal-to-Noise Ratio</i>
SPM	<i>Statistical Parametric Mapping</i>
STS	<i>Superior Temporal Sulcus</i>
TB	<i>True-Belief</i>
TE	<i>Time-to-Echo</i>
TMS	<i>Transcranial Magnetic Stimulation</i>
ToM	<i>Theory of Mind</i>
TPJ	<i>Temporo-Parietal Junction</i>
TR	<i>Time-to-Repeat</i>

1. General Introduction

“If there is any one secret of success, it lies in the ability to get the other person's point of view and see things from that person's angle as well as from your own.”

Henry Ford

More than 30 years ago, two behavioral scientists opened a new chapter in research on social cognition with their study of primates' cognitive abilities (Premack & Woodruff, 1978). Their study investigated for the first time the chimpanzee's ability to impute mental states to itself and others. Starting with Premack and Woodruff, this ability has subsequently been referred to as “Theory of Mind” (ToM).

Since then, researchers have tried to clarify the cognitive underpinnings of human ToM and its development across the life span. The investigation of the neural basis of ToM reasoning was sparked some 20 years ago when functional magnetic resonance imaging (fMRI) emerged as an innovative technology to identify the underlying neural networks of behavioral phenomena.

Imaging studies using fMRI have subsequently identified a number of areas in the brain that seem to play a prominent role in ToM reasoning. These are, among other candidate regions, the medial part of the prefrontal cortex and a region at the junction of the temporal and the parietal cortices.

At this point, conflicting accounts exist over whether any of these regions may constitute the neural correlate of a specific ToM module. Studies with patients suffering from lesions in these regions resulted in inconsistent findings. Some imaging studies hint that the ToM areas specified above may also subserve other cognitive abilities, among these especially inhibitory control. Behavioral studies seem to support this view. In 3 year olds, performance in inhibitory tasks predicts ToM performance and vice-versa (Carlson & Moses, 2001). Some studies show that people with autism, a lifelong developmental disorder with a severe impairment in social functioning, are impaired in both inhibitory control and ToM reasoning. By contrast, children suffering from ADHD show impaired executive functioning while possessing relatively intact ToM reasoning abilities.

However, no imaging study has directly and properly compared inhibitory control and belief-reasoning. This is what the present study has attempted to achieve. Using a Go / No-go task, we investigated inhibitory control. Within the same experimental session with

similar stimulus material and within the same subjects, we presented short cartoon stories that yielded false- or true-belief reasoning. The results from the study presented here may help clarify the relationship between inhibitory control and ToM, reveal common and distinct neural networks, and approve or reject accounts of specific ToM modules in the brain.

The following introductory part of this thesis will be separated into four larger sections. First, the development of ToM reasoning and its neural correlates will be delineated. This is followed by a section about the properties of inhibitory control and associated brain regions. Following this, the relationship between belief-reasoning and inhibitory control will be described on the behavioral as well as the neural level. The introductory part is concluded with a description of the present study's aim.

2. Theory of Mind (ToM) Reasoning

2.1 Development of ToM Reasoning

2.1.1 Cognitive Developments Preceding Belief-Reasoning

The term theory of mind (ToM) refers to the ability to attribute mental states such as intentions, desires and beliefs to oneself and to others and allows us to predict and explain the behavior of ourselves and others. Ample research has shown that the understanding of intentions and desires precedes the understanding of beliefs that emerges at roughly 3 to 4 years of age (Frith & Frith, 2003; Wellman et al., 2001; Wellman & Cross, 2001; Wellman & Liu, 2004).

Before the acquisition of actual belief understanding, infants achieve a variety of cognitive abilities that can be regarded as a prerequisite for attributing beliefs. The following paragraph delineates some of the major steps in human ontogenetic development and their relevance to the emergence of a ToM.

One of the first cognitive achievements that is innate to newborns is the preference for human faces as opposed to other objects (Johnson, 2003). Also, newborns are able to imitate facial expressions of others (Legerstee, 1991). Interestingly, these expressions are only imitated when the expressions are conducted by a human and not when they are simulated by two objects. At about 6 months of age, infants are able to differentiate between mechanical and biological motion. This finding is derived from observations showing that infants at that age direct significantly more attention to a point light display that mimics human motion as opposed to a display depicting non-biological motion (Moore et al., 2007). Also at around the age of 6 months, infants react with surprise when an object moves on its own compared to when the object is moved by a human (Spelke et al., 1995). At the end of the first year of life, dyadic interactions give way to triadic interactions which involve the infant, another person and an object to which the common attention is directed ("joint attention"; Butterworth & Jarrett, 1991). In socially ambiguous situations, infants now tend to use social cues such as an adult's facial expression in order to evaluate the situation. This cognitive achievement is also called "social referencing" (Carpenter et al., 1998). In an ingenious experiment, Gergely, Nadasdy, Csirba, & Biro (1995) were able to show that infants between the ages of 9 and 12 months are able to identify an agent's goal and interpret its actions based on this goal. In the experiment, infants reacted with surprise when an agent moving towards another object jumped over an invisible hurdle. A study conducted by Onishi, Baillargeon, & Leslie (2007) even gave rise to the assumption that infants as young as 15 months may predict actions based on

another person's belief. However, Sodian & Thoermer (2008) were able to show that this finding may be due to situational and behavioral clues. They argue that knowledge formation which results in later belief-reasoning capacities may form gradually in the second year of life based on the integration of situational clues.

A milestone in the development of social cognition takes place at the age of 18 months, when toddlers are able to understand the concept of pretence, which has been regarded by some as a first sign of mentalizing (Frith & Frith, 2003). This concept requires the decoupling of representations of real events from representations of thoughts. A toddler who has successfully adapted the concept of pretence will now appreciate her mother picking up a banana and pretending to use a phone (Leslie, 1994). Other researchers, however, argue that pretence play may not be a sign of early mentalizing. Rather, children may understand pretence on a mere operational level (Sodian & Thoermer, 2006).

First signs of empathic behavior in kids emerge at the end of the second year of life when children try to help others who experience a mishap. Some researchers tend to interpret this behavior as an early sign of the representation of others' mental states (Perner & Davies, 1991). At about 15 to 18 months of age, children are able to imitate an action that a protagonist has started but stopped soon thereafter. This finding suggests that children of that age develop an understanding of others' intentions (Bellagamba et al., 2006).

The understanding that other people have desires that differ from their own emerges in the second half of the third year of life. Children now start to understand the relationship between desires, the results of an action and emotional reactions. As such they understand that a protagonist will not be happy if a person catches a ball although the protagonist intended to throw it to another person (Wellman & Woolley, 1990).

2.1.2 The Development of Belief-Reasoning and Its Investigation

First accounts of understanding other individuals' beliefs have been reported for children at the age of about 3 years. Children's cognitive achievements prior the emergence of belief understanding seem to be a crucial prerequisite for belief-reasoning (Frith & Frith, 2003). Some of these achievements have been mentioned in the section above in more detail. Among these are the ability to discriminate biological from mechanical agents, focussing attention towards an object that another person has focused attention to (joint-attention), a preference for social stimuli, an understanding of pretence, and the ability to represent others' intentions and desires.

At about 3 to 4 years of age then, some children are able to pass tasks that require a discrimination between reality and belief. One of the first systematic studies of the development of this ability to attribute beliefs was conducted by Wimmer & Perner (1983). Their experimental task dubbed “Maxi and the chocolate” has been used in similar versions in many experiments conducted in the years thereafter. A variant of this task will also be used in the present study as a test of false-belief understanding. In the classic paradigm, the puppet or cartoon character Maxi puts a chocolate into a cupboard and leaves the room. After Maxi has left, his mother puts the chocolate from one cupboard into another. Maxi returns to the room and starts looking for the chocolate. The children in the experiment are then asked where Maxi will look for his chocolate. In order to successfully master this task, the child is required to attend to Maxi’s belief and not to pay attention to the actual location of the chocolate. Another task in which an object is not transferred to a new location is referred to as the true-belief condition. In order to answer this type of task correctly, no separation of reality and belief has to be conducted. As such, true-belief tasks can be answered based on the actual state of affairs.

A variant of the “Maxi and the chocolate” task is the so called “Sally-Anne paradigm” (Baron-Cohen et al., 1985). In this classic paradigm, Sally puts an object into a box and leaves the room. Anne then moves the object to a different location. When Sally enters the room again, the subject is asked where the object is in reality (reality question) and where Sally will look for the object (target question). The authors assume that a subject possesses the ability to attribute false beliefs if both test questions are answered correctly.

In a different experimental paradigm (“crayon task”) by Gopnik & Astington (1988), children at ages three to five years are shown deceptive objects (e.g., a Smarties box). Then the true content of the object is revealed to them (crayons, for instance). The children are then asked what they thought the object had contained and what another child would have thought the object had contained. In order to answer these questions correctly, children have to understand that others possess beliefs that differ from their own (false-belief understanding). Results show that most children at age three are unable to pass the test questions while most children at age five answer the questions correctly.

Sodian & Frith (1992) used a different type of task to investigate belief understanding in children. In their “sabotage-deception task”, children are presented with two puppets that are introduced as either friend or foe. A piece of candy is put into a box that can be locked and children are instructed to help the friend and never the foe. In the sabotage condition this can be achieved by either locking the box or keeping it unlocked. In the deception condition, however, this can only be achieved by telling a lie to the foe. Thus, the deception condition can only be solved when possessing the ability to understand and manipulate beliefs. Results in this experiment showed that autistic

children compared to healthy controls scored significantly worse in the deception condition, indicating a compromised ability to attribute beliefs. Scores in the sabotage task were equally high for both autistic and healthy children, indicating that both groups had understood the task.

Call & Tomasello (1999) have developed yet another task to assess false-belief understanding in both humans and apes. Their task does not require language. In a first set of control trials, the subject learns that another person (“communicator”) will point to a container that contains an object. In their crucial false-belief task, someone hides an object in one of two identical containers with the communicator watching. The communicator leaves the scene while the hider switches the location of the containers. The person returns and now points to the container that he falsely believes contains the object. The subject is then given the opportunity to find the hidden object. In order to solve this task, the subject has to realize that the person pointing possesses a false belief about the true state of affairs. This non-verbal task is mastered by most three year old humans but not by apes.

To date, dozens of studies have investigated the emergence of belief understanding in children. In an extensive meta-analysis, Wellman et al. (2001) included 178 studies from 77 reports that had used false-belief tasks. Their results suggest that age and not other factors such as task demands or cultural influences impact false-belief understanding. Wellman et al.’s (2001) analysis shows that children at about 44 months of age score approximately 50 % correct in false-belief tasks. Wellman and colleagues point out that task demands, particularly executive functions, seem only to play a role in the expression of false-belief understanding at an intermediate level of false-belief understanding.

As mentioned above, children between the ages of roughly 3 and 5 years acquire the ability to impute beliefs to oneself and others. This ability, however, is limited to so called first-order false-belief tasks that require a false-belief attribution to another person. Second-order tasks on the other hand require the attribution of a belief about another person’s belief. These types of false-belief tasks are typically mastered by children at the age of some 5 or 6 years (Sullivan et al., 1994).

Relatively little research has been dedicated to the further course of ToM development. This may be due to the fact that virtually all healthy subjects older than some 7 or 8 years pass standard ToM tasks such as the Sally-Anne paradigm. Therefore, some researchers have developed ToM tasks that may be better suited to tap more advanced ToM reasoning and to avoid the ceiling effect of simple ToM tasks in adults. In a study that investigated the ability to attribute intentions in young adults and older adults, Happe, Winner, & Brownell (1998) used short stories including double bluffs, mistakes, persuasions and white lies. These stories were then followed by several questions that

required the subjects to infer about a protagonist's intentions. Interestingly, the group of older adults with a mean age of 73 years performed significantly better in ToM tasks than the group of young adults with a mean age of 22 and a half years. The authors argue that this finding may indicate that social and non-social reasoning rely on distinct mechanisms because non-social reasoning usually declines with age (Zelazo et al., 2004).

An opposite pattern of results was found in a study by McKinnon & Moscovitch (2007). The authors compared a group of older adults (mean age= 78 years) to a group of younger adults (mean age= 20 years) by using first- and second-order belief stories about complex social situations such as social faux pas. Their results showed that older subjects performed significantly worse than younger subjects in second-order ToM tasks. In first-order tasks, however, younger and older subjects performed at an equal level. McKinnon and Moscovitch argue that declining executive control may account for older adults' poor performance in second order belief tasks.

In sum, various cognitive achievements precede the ability to reason about beliefs which emerges at about 3 to 4 years of age. Among these possible prerequisites is the ability to distinguish biological from mechanical movements, showing a preference for social stimuli, engaging in joint attention, understanding pretence, and the ability to reason about others' intentions and desires. In order to capture the developmental course of belief-reasoning researchers have come up with numerous different tasks such as Wimmer & Perner's (1983) "Maxi and the chocolate" task. Converging evidence now suggests that the vast majority of children are able to reason about beliefs at about age 6. To date, accounts of belief-reasoning abilities through adulthood and old age remain scarce and at times contradictory.

The neural correlates of ToM reasoning are described in the following section.

2.2 Neural Correlates of ToM Reasoning

An abundance of behavioral studies, some of which have been delineated in the above section, have investigated the properties of ToM reasoning and its precursors in humans and other species. Since the early 1990's, the emergence of fMRI has evoked yet another field of scientific investigation: social neuroscience. Next to other phenomena from the social realm, social neuroscience is dedicated to the investigation of ToM and related concepts by means of functional imaging. By identifying neural correlates of behavioral phenomena, researchers attempt to identify, among others, the processes that enable us to attribute mental states. Besides imaging techniques, social neuroscience also uses results from lesion studies. Several studies with patients suffering from brain lesions have

helped in identifying brain regions related to ToM reasoning. A selection of these studies will be discussed next, followed by a paragraph on the imaging of mental state attribution.

2.2.1 ToM: Results from Lesion Studies

Results from lesion studies provide an elegant way to speculate about the nature of ToM reasoning and underlying brain regions. Lesions in distinct brain areas that cause an impairment in ToM reasoning but not in other cognitive abilities could theoretically indicate that this very region may constitute a specific region for ToM reasoning. A distinct lesion that results in an impairment in ToM reasoning as well as in other cognitive impairments may on the other hand indicate that this region could theoretically constitute the neural correlate of a cognitive mechanism underlying both ToM reasoning and other specific cognitive mechanisms.

The first study of a person suffering from cognitive deficits resulting from brain injury dates back to the famous case of railroad worker Phineas Gage. In 1848, Gage experienced massive damage of the bilateral medial frontal cortex caused by an iron rod that had been propelled through his head (Damasio et al., 1994; Stone, 1999). Miraculously surviving the incident, Gage subsequently showed impulsive behaviour, a lack of ability to plan future actions ahead of time, and little consideration for the people around him. It can only be speculated here whether next to the historically described personality changes, Gage may also have suffered from an impairment in his ability to infer others' mental states. As we will see later, an area in the medial frontal cortex that is also likely to have been damaged in Gage seems to play a crucial role in ToM reasoning. A digital reconstruction of Gage's lesion is depicted in figure 2.1.

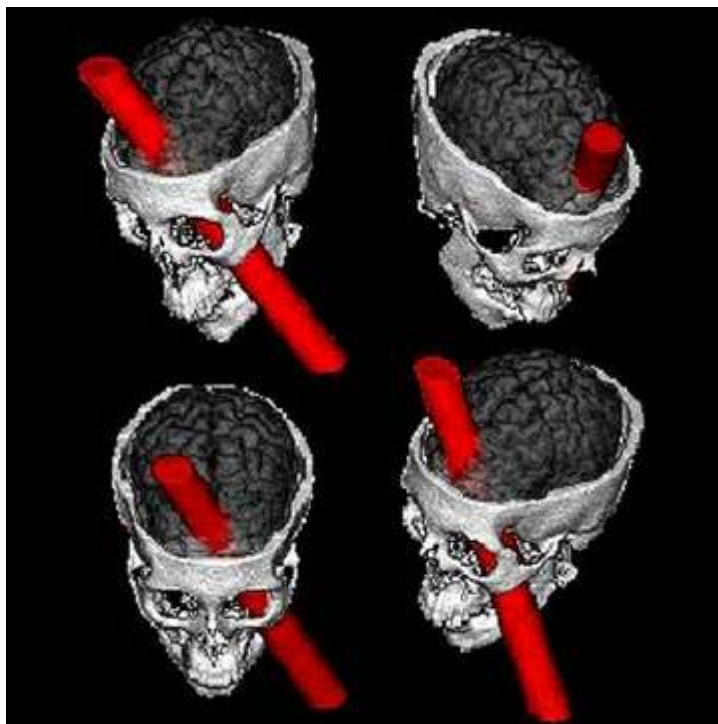


Figure 2.1. Digital reconstruction of brain damage to railroad worker Phineas Gage due to an iron rod propelled through his head. This picture shows that the medial part of the frontal cortex, an area associated with ToM reasoning, had also been affected. Historical accounts whether Gage had subsequently suffered from an inability to impute mental states, however, are vague. Picture taken from Damasio et al. (1994). Reprinted with permission from AAAS.

So far, only a small number of controlled lesion studies on ToM reasoning have been conducted. However, these studies point to a prominent role of a small number of candidate regions involved in ToM reasoning.

Fine, Lumsden, & Blair (2001) report the case of B.M., a patient suffering from congenital left amygdala damage who was diagnosed as an adult with schizophrenia and Asperger's syndrome. In the study, five measures of false-belief understanding were administered to the patient, two of which required first-order false-belief understanding, three of which required second-order false-belief understanding. The false-belief measures utilized were commonly used tasks such as Baron-Cohen's Sally-Anne paradigm (Baron-Cohen et al., 1985). Interestingly, the subject B.M. passed the two first-order false-belief tasks but failed in all second-order false-belief tasks. However, the patient managed to correctly answer the control questions contained in all five stories. This indicates that the poor performance shown was not due to general task demands such as story comprehension or memory. In an additional test of advanced ToM reasoning (Happe et al., 1998), the patient also scored below the score that would have been expected from a comparison group. Due to these results, the authors argue for a prominent role of the

amygdala either as a prerequisite for the development of belief-reasoning (B.M.'s amygdala lesion was caused by a congenital or early lesion to the amygdala) or as a crucial component of a neural circuitry for ToM reasoning. In this regard the authors note that the amygdala has extensive interconnections with the medial prefrontal cortex and the superior temporal sulcus. As the following lesion studies suggest, both of these structures also seem to play an important role in mentalizing.

In one of these studies, 31 patients with unilateral damage to various parts of the frontal lobes were investigated and compared to a control group of 31 healthy subjects (Rowe et al., 2001). Subjects were presented with newly constructed first and second-order false-belief stories that treated topics such as going to a restaurant or going grocery shopping. These stories were followed by a false-belief question and an inference question to assess the ability to draw inferences without a belief-reasoning component. The task also contained a fact question that tested if the subject had understood the events leading to a false belief, and a memory question testing whether the subject had an intact memory for story details. Compared to the control group, the patients with damage to the frontal lobes showed a significant impairment in both first- and second-order false-belief questions. The group of patients with right-frontal lesions showed no impairment in any of the control questions. Patients with left-frontal lesions, however, showed a slight impairment in the inference questions. However, covariate analyses showed that this impairment was independent of ToM performance. The results obtained in this study indicate that the prefrontal cortex plays a crucial role in the expression of ToM.

Another study compared a group of 19 patients with frontal lesions to a group of 13 patients with non-frontal lesions and a group of 14 controls (Stuss et al., 2001b). The group of frontal patients differed in the location of their lesion with 7 patients suffering from bi-frontal medial lesions, 4 people having right frontal lesions and 8 patients suffering from left frontal regions. The subjects' task in all conditions was to point to a cup underneath which an object had been hidden. One of the conditions in the experiment, the deception task, required the subject not to pick a cup that a "deceiving" assistant was pointing to, thus requiring the subjects to attribute a false belief. In the deception task administered, only the patients with damage to the medial part of the frontal cortex were significantly impaired when compared to the group of non-frontal patients and healthy controls. Once again, these results indicate a dominant role of the frontal cortex as part of a neural false-belief reasoning network. Moreover, the study's authors were able to show that the medial areas of the frontal cortex in particular seem to be of special importance for the ability to infer others' beliefs.

Another study, however, stemming from the data of one patient only, found no impairment of belief-reasoning after extensive damage to the bilateral medial frontal cortex

(Bird et al., 2004). The patient presented in this study suffered from a dysexecutive syndrome including problems in planning and memory. Yet when presented with a battery of false-belief tasks such as the Sally-Anne paradigm or Happe's advanced ToM tasks (Happe et al., 1996), the patient achieved a level of performance that was well in the range of healthy subjects. Based on the results of this single-subject study, the authors argue that the medial frontal gyrus may not necessarily constitute a part of the neural ToM network. This finding is in clear contradiction to most lesion studies and neuroimaging studies. The authors then also note that the stroke patient's lesion may not have involved all areas of the medial frontal lobe implicated in ToM reasoning and suggest further studies.

Another candidate region as part of a network for ToM is the temporo-parietal junction (TPJ), an area at the border of the temporal and the parietal cortices, located mostly in Brodmann Areas (BA) 39, 40 and 41 and partly comprising the angular gyrus, supramarginal gyrus and superior temporal gyrus. Some researchers also refer to this region as the posterior superior temporal sulcus (pSTS; Frith & Frith, 2003). In a study by Samson, Apperly, Chiavarino, & Humphreys (2004) three patients with damage to the left TPJ participated in a story-based and a video-based version of a false-belief reasoning task. In the video-based task (unlike in the story-based task), none of the three subjects had any difficulties with the control questions included in the task. However, none of the subjects scored above chance level in either the story-based or the video-based belief task. These results indicate that the left TPJ may be a necessary component of the ability to infer others' mental states and that the observed compromised ToM reasoning performance may not have been due to general task demands.

A study by Apperly, Samson, Chiavarino, & Humphreys (2004a) with 12 brain-damaged patients with circumscribed lesions in either the left TPJ or the prefrontal cortex used the same stimulus material as in the study presented above (Samson et al., 2004). The authors found that 3 patients with left TPJ lesions as well as 4 patients with prefrontal lesions performed only at chance level in both false-belief tasks. However, only the patients with the prefrontal lesions also showed an impairment in a working memory control task. The authors argue that prefrontal patients may fail in false-belief tasks due to the executive demands of the task, while left TPJ patients fail not because of the processing demands of the task but because the left TPJ constitutes a crucial part of a neural belief-reasoning network. The authors also explicitly note that this finding does not mean that the right TPJ is not an essential part of a ToM network. To this point however, no lesion study has investigated belief-reasoning performance in subjects with circumscribed lesions to the right TPJ. Nevertheless, most neuroimaging studies investigating belief-reasoning have consistently found significant activity in this anatomic

region when comparing false-belief to control tasks. This finding will be discussed in more detail in the following section.

A novel method of studying social cognition is transcranial magnetic stimulation (TMS). This non-invasive method elicits activity in neuron populations by applying a rapidly changing magnetic field. This technique is related to lesion studies as it can also “turn off” circumscribed areas of the brain. As such, it can demonstrate causal relationships between affected brain regions and related behavioral phenomena. Therefore, lesion studies and TMS share the same methodological background. In one of the first TMS studies regarding ToM reasoning, Costa, Torriero, Oliveri, & Caltagirone (2008) applied magnetic currents onto 11 healthy subjects. The magnetic pulses were applied over the right/left dorsolateral prefrontal cortex as well as over the right TPJ. Subjects were then administered written false-belief stories as well as faux-pas stories as a measure of advanced false-belief reasoning. Compared to control tasks during which no magnetic pulses were sent from the head coil (sham condition), subjects showed longer reaction times in both false-belief tasks. In the faux-pas task only, subjects also showed significantly poorer accuracy compared to the sham condition. In concordance with previous lesion studies, the results of this first TMS study point to a prominent role of the prefrontal cortex and the TPJ in belief-reasoning. This view is corroborated by neuroimaging studies that will be delineated in the following section.

To conclude, most lesion studies and one TMS study have investigated patients with damage to a small number of circumscribed areas. The studies presented here were able to show that especially the prefrontal medial cortex and the TPJ seem to play a crucial role in ToM reasoning. The role of the amygdala is less clear; only one single-subject study with a patient suffering from congenital amygdala damage has explicitly investigated belief-reasoning. It could be that the amygdala plays a role in the emergence of belief-reasoning but not in its expression in adulthood. Also, a number of important structures such as the right TPJ or more circumscribed areas of the frontal cortex (e.g., the anterior cingulate cortex) have not yet been investigated in lesion studies but were commonly found in neuroimaging studies and, as far as the right TPJ is concerned, in the only TMS study so far.

2.2.2 ToM: Results from Neuroimaging Studies

Neuroimaging studies of ToM reasoning and related concepts emerged with the invention of modern functional imaging techniques. These methods have helped identify areas associated with the attribution of mental states. The identification of brain regions involved in ToM reasoning can help to shed light on issues such as the development of mentalizing,

the question over whether a specific ToM module exists and on the debate over cognitive processes supporting the attribution of mental states. Also, neuroimaging techniques provide further insight on disorders such as autism, a severe developmental disorder associated with a frequent impairment in ToM reasoning.

2.2.2.1 PET Studies

Positron emission tomography (PET) is one of the methods used in the functional imaging of ToM. This technology is able to measure activity of the brain by detecting gamma rays that are emitted whenever an uptake of glucose takes place in the nerve cell. In order to obtain this indirect measure of neural activity, radioactive tracer isotopes have to be introduced into the subject's body. PET has a spatial resolution of some 4 to 8 mm and a temporal resolution of approximately 1 second. Although PET requires the injection of possibly harmful radioactive substances, it yields better images compared to fMRI in areas that are susceptible to scanner artefacts such as the amygdala or the orbitofrontal cortex (D'Esposito, 2000).

In one of the first PET studies investigating belief-reasoning, short stories followed by questions were presented to 6 healthy subjects (Fletcher et al., 1995). The belief-reasoning condition required the attribution of false beliefs. Two other conditions served as control tasks. One of these, the physical stories condition, was similar to the belief-reasoning condition in terms of content and complexity. However, it did not require the attribution of mental states. The second control condition consisted of unlinked sentences that were not connected to each other in terms of content. Results revealed significantly more activity for the false-belief condition as well as for the physical stories condition in the temporal poles, the left superior temporal gyrus and the posterior cingulate cortex when compared to the unlinked sentences condition. Since the condition "physical stories" did not require the attribution of mental states, the areas mentioned above did not constitute a neural correlate of belief attribution. Therefore the conditions belief-reasoning and physical stories were compared as well in order to isolate a specific mentalizing component. Significantly higher activity for the belief-reasoning condition was found only in the left medial frontal gyrus (BA 8). These results point to a prominent role of the medial prefrontal cortex for the attribution of false beliefs.

A different type of task in an attempt to study belief-reasoning was used by Goel, Grafman, Sadato, & Hallett (1995). In their PET study, subjects had to attend to visual and semantic properties of commonly used objects presented on a screen. These tasks served as control conditions and did not require the attribution of mental states. In the ToM task, subjects had to infer whether a person living in the 15th century such as Christopher

Columbus would have already known these objects. When comparing this ToM condition to the control conditions, significantly increased activity for the ToM condition was found in left medial prefrontal lobe (BA 9), the left TPJ and in the temporal poles. Although the brain areas related to belief-reasoning reported here are in line with previous findings, the authors may have failed to actually measure belief attribution. It may possibly be sufficient to recall historic knowledge about life in the 15th century without having to infer another person's mental state.

A more recent PET study investigated ToM reasoning with yet another fairly unusual mentalizing task (Calarge et al., 2003). In this experiment, subjects were asked to make up a story about an unknown woman sitting on a bench crying. After 30 seconds of reading the instructions and planning their narrative, subjects were given 100 seconds to speak and tell their made up story. This "ToM" story was then compared to a control condition during which subjects had to read out loud a story presented on a screen. When subtracting the activation in the control task from activation in the ToM task, the authors found activity in the medial frontal cortex (BA 10, 32), the superior frontal cortex (BA 6, 8, 9, 32), the right anterior cingulate (BA 24, 32) and the right cerebellum. According to the authors, these areas seem to be related to ToM reasoning and show once again the importance of the prefrontal cortex for ToM reasoning. However, it is unclear which component of ToM has been investigated here. No information is given as to whether the task presented requires the attribution of intentions, emotions, desires, or beliefs. The possibility exists that the task presented, making up a story, may not even have tapped ToM reasoning at all but rather only have required the engagement of other cognitive abilities such as episodic memory recall, working memory or language production.

While PET studies on ToM and especially belief-reasoning are relatively scarce, a larger amount of fMRI studies investigating the inferring of mental states has been conducted. A selection of fMRI studies will be presented next.

2.2.2.2 FMRI Studies

fMRI measures the so called hemodynamic response. The hemodynamic response is a measure of neural activity (Logothetis et al., 2001) and describes the process during which oxygen is released to active neurons. Oxygenated (diamagnetic) and deoxygenated (paramagnetic) hemoglobin have differing magnetic properties. The resulting difference in magnetic susceptibility can then be detected in an fMRI scanner by applying a strong magnetic field to the human cortex. This procedure does not rely on radiation which is inherent in other procedures such as computer tomography or PET and is therefore harmless to human tissue. fMRI scanning has become the most commonly used imaging

technique due to its wide availability and its excellent spatial resolution of some 2 to 6 millimeters.

One of the first fMRI studies investigating ToM reasoning was conducted by Gallagher et al. (2000). In the first of two ToM conditions subjects were presented with written stories followed by questions that required the attribution of false beliefs. This condition was compared to two control stories (physical stories and unlinked sentences) that did not require any mentalizing. In the second ToM condition subjects were presented with cartoons depicting events that required the attribution of mental states. Unfortunately it is unclear whether the cartoons of this second ToM condition actually required the attribution of false beliefs or simply the attribution of desires or intentions. The ToM cartoons were compared to cartoons with no mentalizing component or to cartoons that consisted of jumbled pictures. When comparing the first ToM condition to the control stories, significantly increased activity for the ToM condition was found in the medial prefrontal cortex (BA 8, 9), the temporal poles (BA 38) and bilateral TPJ (BA 39, 40). The comparison of ToM cartoons against non-ToM cartoons revealed increased activity in the medial prefrontal cortex (BA 8), the right middle frontal gyrus (BA 6), the right TPJ (BA 40), and precuneus (BA 7, 31). When comparing both ToM conditions as well as the control conditions against baseline, only the medial prefrontal cortex (BA 9) was uniquely activated by the ToM task. Although not all of the tasks used by Gallagher et al. may have tapped belief-reasoning, these results point once again to a strong involvement of the medial prefrontal cortex in the attribution of beliefs.

The same set of belief stories and control stories as in Gallagher et al.'s study (physical stories and unlinked sentences) were also used to investigate belief-reasoning in a study by Gobbini, Koralek, Bryan, Montgomery, & Haxby (2007). In addition to using these stories, the authors also compared activity related to animations of intentional movements of geometric shapes to movements of geometrical shapes with random movements. Significantly increased activity in the false belief stories compared to the physical stories was revealed in medial prefrontal cortex (PFC; BAs 9 and 10), bilateral TPJ, left precuneus, and the bilateral temporal poles. Possibly, these areas might constitute components of a hypothesized neural network supporting the attribution of beliefs. A differing network for attributing intentions and goals based on actions as executed by geometrical shapes was revealed by subtracting activity in the random movement condition from activity in the intentional movement condition. This comparison revealed increased activity in the bilateral ventral lateral PFC (BA 45), bilateral superior temporal sulcus, bilateral temporal poles, precuneus, bilateral inferior parietal lobule and the medial PFC (BA 9). Taken together, these findings point to a belief-reasoning network consisting of the temporal poles, precuneus and the ventral medial PFC. The data also

suggest that the attribution of intentions may be supported by a network that additionally includes the inferior parietal lobule and superior temporal sulcus and that this network does not recruit the bilateral TPJ.

Stories requiring the attribution of a false belief were also used in a study conducted by Saxe & Kanwisher (2003; experiment 2). The false-belief task contained in this study was subsequently compared to a false photograph task. The false photograph stories utilized in the study were logically similar to false-belief stories but did not include any social content. Significantly increased activity in the false-belief condition across all 21 subjects was revealed in the bilateral TPJ, precuneus, superior medial PFC, the frontal poles and the right superior temporal sulcus. A subsequent ROI analysis showed that the bilateral TPJ was not activated by a condition that merely described the physical appearance of a person. Taken together, the authors argue that their findings provide strong support that the bilateral TPJ plays a specific role in the attribution of beliefs.

A subsequent fMRI study by Saxe & Wexler (2005) presented short stories describing a protagonist's social background. This was followed by a description of the protagonist's desire and a story outcome. Saxe & Wexler's study focused on four regions of the brain that have frequently been found in other studies investigating ToM reasoning. These were the left and right TPJ, the medial prefrontal cortex and the posterior cingulate cortex. All of these regions were introduced into a ROI-analysis. A significant increase in brain activity for the description of a protagonist's desire compared to the description of the protagonist's social background was only found in the right TPJ and not in any of the other areas. Also, only the right TPJ showed no differing activation when the protagonist's social background was changed to a background that was unfamiliar to the subject. Due to these observations, the authors claim that the right TPJ may constitute a specific region for the attribution of mental states. However, results from this study only apply to the attribution of intentions; the attribution of beliefs was not required at any given point.

Another study conducted by the same research group underscored the important role of the right TPJ for mental state attribution (Saxe & Powell, 2006). However, this time the posterior cingulate also responded selectively for mental state attribution. Once again, the subjects were presented with short stories that either described a protagonist's thought (ToM condition), his physical appearance (control condition) or a bodily sensation (control condition) experienced by the protagonist. The bilateral TPJ, posterior cingulate and ventral, lateral and middle medial prefrontal cortex were introduced into a subsequent ROI analysis. Compared to both control conditions, the ToM condition was associated with significantly increased activity in the bilateral TPJ and the posterior cingulate. The medial prefrontal cortex showed no significant activation related to the attribution of mental states. Based on these results, the authors claim that the medial prefrontal cortex may not be

selective to the attribution of mental states but rather that it may be involved in a broader representation of socially or emotionally relevant information about other individuals. This study also raises some doubt over whether the ToM condition presented here may actually have tapped belief-reasoning. The subjects merely read stories without having to answer any questions concerning the story. Thus, the authors were unable to probe story comprehension or the actual attribution of mental states. Further, some of the stories did not even include beliefs but referred to concepts that the authors vaguely describe as “thoughts” or “reasoning”.

A commonly used task in behavioral studies assessing belief-reasoning is the Sally-Anne paradigm that requires the attribution of false beliefs (Baron-Cohen et al., 1985). False-belief tasks are regarded as the crucial test for the ability to attribute beliefs. A variant of the latter task was used for the first time in an fMRI experiment by Sommer et al. (2007). A total of 16 healthy subjects took part in the experiment. In their false-belief condition, subjects were shown cartoon stories consisting of 7 pictures each that depicted the transfer of an object without one of the protagonists watching. In the last picture the protagonist was shown looking for the object in one of the containers. Subjects were then asked whether they had expected the cartoon character looking into that container. Thus, this condition required the attribution of a false belief to the story character. In the true-belief condition, however, no decoupling between reality and a false belief was required. In this condition, the cartoon character looking for the object was also watching when the object was transferred from one container to the other. Thus, both conditions were virtually identical in story content and visual stimulation except for the fact that the false-belief condition required the attribution of a false belief. By subtracting the activation in the true-belief condition from activation in the false-belief condition, the authors identified areas related to a process that separates mental states from the real state of affairs. This decoupling of mentality and reality is a crucial component in the attribution of false beliefs. Activity for this decoupling mechanism was revealed in the dorsal part of the anterior cingulate cortex (ACC; BA 32), right dorsolateral prefrontal cortex (BA 9), right middle frontal gyrus (BA 6), right lateral ventral frontal cortex (BA 10), right TPJ (BA 39), right middle temporal gyrus (BA 21), and precuneus (BA 7). This result emphasizes once again the importance of the prefrontal cortex and the right TPJ for belief-reasoning. The authors in this study also investigated common neural networks for false-belief and true-belief reasoning. The authors claim that brain areas dedicated specifically to belief-reasoning should theoretically show increased activity in both conditions when compared to baseline activity. Therefore a conjunction analysis was conducted that revealed common activity in superior, middle and inferior frontal gyrus but not in areas previously identified in mentalizing such as the TPJ or the anterior ventral medial frontal cortex. However, this

finding is based on the assumption that the true-belief condition used in this study would also tap the attribution of beliefs, albeit in this case the attribution of a true-belief. Nevertheless, it may also be argued that this condition did not require the attribution of belief at all. Subjects could also answer the task by simply memorizing the true location of the object and comparing it to the location where the protagonist is looking for the object. This strategy may not require the attribution of a mental state. The subtraction of activity in the true-belief condition from activity in the false-belief condition, as done in this study, may therefore be sufficient to isolate activation dedicated the attribution of beliefs. Imaging data from the Sommer et al. (2007) study are depicted in figure 2.2.

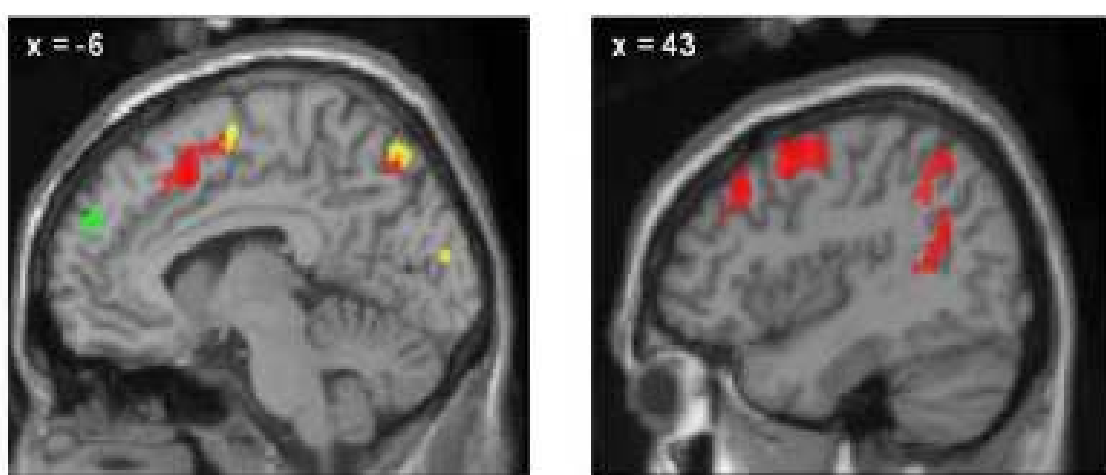


Figure 2.2. Two lateral views at MNI coordinates $x = -6$ and $x = 43$ of significantly increased activity in three comparisons conducted in a study by Sommer et al. (2007). Significantly increased activity in the contrast *False-Belief* > *True-Belief* is depicted in red shading, significantly increased activity in the contrast *True-Belief* > *False-Belief* is shown in green shading. Areas in yellow color depict common activation for the two contrasts *False-Belief* > *Baseline* and *True-Belief* > *Baseline*. Picture reprinted from Sommer et al. (2007) with permission from Elsevier.

Unlike Sommer et al.'s non-verbal cartoon study, an fMRI study conducted by Perner, Aichhorn, Kronbichler, Staffen, & Ladurner (2006) used verbal stories to investigate the neural correlates of belief-reasoning. Therefore the authors presented written false belief stories to the subjects. These were contrasted to three control conditions: a false photo task, a false sign task and a temporal change control task. A false sign task was utilized as an additional control task because according to Perner et al., the false photograph task may not capture an understanding of perspective differences that is part of a false-belief task. However, this understanding of perspective differences may also be included in the false-sign task. False sign tasks, unlike false photograph tasks, may be genuinely false while still including non-mental objects. When comparing the false-belief task to the false

photo task, significantly increased activity in the false-belief task was revealed in the bilateral TPJ, the right middle and superior temporal gyrus, precuneus and medial PFC (BA 9). Compared to the false sign task, however, significantly increased activity for the false-belief task was only revealed in the right TPJ. Due to these results, the authors argue that the right TPJ may indeed constitute a specific area for the processing of mental states.

To summarize, several imaging studies have investigated the neural correlates of ToM reasoning. Although a relatively circumscribed array of brain regions have emerged as possible components of a neural ToM network, it remains unclear what some of the studies presented here have truly investigated. First of all, the experimental tasks used here ranged widely from viewing cartoons, to making up stories to merely reading stories. It is likely that although ToM reasoning may have been tapped in most studies presented here, the brain activation measured may also have included other processes such as working memory, language and so forth due to insufficiently controlled comparison tasks. Secondly, and more importantly, it remains unclear whether some of the studies actually investigated the attribution of beliefs or rather the attribution of intentions, desires or emotions. Some studies did not even state which kind of attribution they wanted to study. The attribution of desires, intentions and emotions differs from the attribution of beliefs. In the development of ToM, the attribution of desires, intentions and emotions precedes the ability to attribute beliefs (Sodian & Thoermer, 2006). It is likely that these two types of mental state attribution even depend on distinct neural networks.

Most likely, only five studies have so far managed to actually isolate the neural correlate of belief-reasoning in carefully controlled paradigms. These are the studies conducted by Fletcher et al. (1995), Gallagher et al. (2000; story condition only), Perner et al. (2006), Gobbini et al. (2007), and Sommer et al. (2007). Their results point to a number of candidate regions for the attribution of beliefs. These are the bilateral TPJ (BAs 39, 40), the medial prefrontal cortex (BAs 9, 10, 32) and perhaps the temporal poles (BA 38). Interestingly, these results are corroborated by the results stemming from lesion studies as well as from the only TMS study conducted so far. These studies have also, with only few exceptions, underscored the importance of the left TPJ and the medial prefrontal cortex for ToM and possibly belief-reasoning. No lesion study has yet investigated the role of the right TPJ.

Areas in proximity of the candidate regions mentioned above for a possible neural network for belief-reasoning have also been discussed as possible components of a mirror neuron network. The mirror neuron system, its relation to the ToM reasoning system, and theories on the underlying mechanisms of belief-reasoning are discussed in more detail in the following section.

2.3 The Mirror Neuron System, Simulation Theory and Theory Theory

In 1992, di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti (1992) discovered a class of neurons in the macaque brain that responded not only when performing a goal directed action but also when the primates watched an experimenter perform that action. These neurons located in area F5 within the ventral premotor cortex of the primate's brain have been dubbed "mirror neurons". Next to areas in inferior frontal cortex, mirror neurons in the primate brain have also been identified in the anterior part of the anterior frontal cortex (Iacoboni & Dapretto, 2006). In a unique series of experiments, Umiltà et al. (2001) were able to show that mirror neurons in the macaque brain fire not only when the monkey observes and executes an action, but also when the actions observed are partially hidden, thus requiring the monkey to anticipate the intended goal of the action observed. Also, the same class of neurons in the premotor cortex fires when the primates hear a sound that is associated with a previously observed or executed action (Kohler et al., 2002). These findings seem to indicate that at least macaques infer the outcome of an intended action by simulating the action in the very brain areas supporting the actual execution of that action.

In the human brain, recruitment of the posterior inferior frontal gyrus and the rostral inferior parietal lobe has frequently been observed during motor action observation and execution (Rizzolatti et al., 2001; Rizzolatti & Craighero, 2004). These areas in the human brain may thus represent the human homologue of the macaque's mirror neuron system (MNS). The posterior STS on the other hand may provide the main visual input to the MNS (Allison et al., 2000) and receive back motor efference copies from the MNS in order to match predicted motor plans to the visually observed action (Iacoboni et al., 2001; Iacoboni & Dapretto, 2006). This network between the MNS and the STS may support the neural basis of imitation.

It has been suggested that the human MNS may also represent the basis for the attribution of mental states (Lieberman, 2007). However, none of the areas identified as part of the human MNS have so far been identified in any functional imaging study investigating mental state attribution. Although there is some evidence that the macaques' MNS is involved when predicting intended motor actions (Umiltà et al., 2001), there is no evidence to this point that this may also be true for the attribution of beliefs. Besides that, several studies suggest that primates may not even be able to attribute beliefs (Call & Tomasello, 2008; Herrmann et al., 2007).

As studies with human subjects have suggested, there is no empirical evidence to date that humans might predict the intentions, desires and beliefs of other individuals by means of simulation. This is, however, assumed in simulation theory (ST). ST basically

assumes that attributing mental states to other individuals is achieved by using the own mind as a model (Apperly, 2008). After having worked out the other person's initial set of mental states (which differs at least in part from our own mindset), these mental states could then be modelled in our own mind using mechanisms used for guiding our own behaviour. Based on this model, the behavior of others could then be predicted. Although empirical data from imaging experiments are unable to confirm the assumptions made by advocates of ST, some behavioral results seem to be in favor of ST. For instance, it has been reported that even infants who are only a few days old spontaneously imitate the facial expressions displayed by an adult (Meltzoff & Moore, 1977; Meltzoff & Moore, 1983). According to the authors, this finding cannot be explained by innate mechanisms or instrumental learning but only by a matching mechanism requiring the simulation of others' behaviors.

Theory theory (TT) on the other hand makes different assumptions about the underlying mechanisms of mental state attribution. According to TT, a set of concepts or theories exists for the various aspects of ToM (desires, beliefs, etc.). These concepts are connected to each other by means of rules about how these concepts influence each other (e.g., Apperly, 2008; Gopnik & Wellman, 1994; Saxe, 2005). However, it remains unclear whether these concepts or representations could be represented in the brain as symbols, algorithms or even language of thought. One of the advocates of TT is Perner (1993), who proposes a representational theory of mind that is required for mental state attribution. According to his model, representations represent something as "being a certain way". In the case of the "Maxi and the chocolate" task which is almost identical to the task used in the present study, Maxi forms a representation about the chocolate being in a certain location. This representation is kept up although the chocolate is transferred to a different location. The subject subsequently has to understand that by looking for the chocolate in its original location, Maxi is guided by a representation of the real situation and not the real situation per se.

In an ingenious experiment, Perner & Howes (1992) attempted to test the assumptions implied in ST and TT. In one part of this experiment, 32 children between the ages of 4,10 and 6,4 years were shown dolls enacting a short story. In this story the two characters John and Mary return home with a box of chocolates. Since Mary has to leave John tells her that he will put the chocolate either in the top or the bottom drawer. After Mary has left he puts the chocolate in the top drawer and leaves. In the meantime their mother transfers the chocolate from the top to the bottom drawer. In the following test questions the subject is asked where John thinks the chocolate is (think question), whether John says he knows where the chocolate is when he is asked (self-reflection question) and what Mary will say when she is asked whether John knows where the

chocolate is. The study's results showed that the self-reflection question was a lot more difficult to answer for the subjects than the think question. This result clearly contradicts the view that children might attempt to answer the task by simulation. If children simulated John's belief, they should in theory be able not only to think that John believes the chocolate is in its original location but also to think that he is convinced that it is there. This result therefore speaks against accounts of simulation as a means of attributing mental states and favors a representational theory of mind.

Another developmental experiment has come to a similar conclusion (Gopnik & Astington, 1988). This study used the so called "Crayon task" in which 3 to 5 year old children were shown a deceptive container (e.g., a Smarties box) and were asked what they thought it contained. The true content of the object was then revealed to them (e.g., crayons). Children were then asked what they thought the object was when they first saw it (representational change question). Also, the children were asked what other children thought the object contained (false belief question). Results showed that the 3 to 5 year children had as many difficulties in the false-belief question as they did in the representational change question. This result suggests that the children cannot form representations of objects. Furthermore, the observed deficit in the representational change question also suggests that this deficit is not due to a lack of understanding about how other people form beliefs. Therefore, Gopnik & Astington argue that the ability to represent a representation is at the core of understanding false beliefs.

Furthermore, TT also implies that a specified cognitive mechanism responsible for the forming and maintenance of representations may exist. ST on the other hand does not imply such a distinct mechanism. This mechanism assumed in TT may be related to a distinct area in the brain supporting this cognitive process. Various imaging studies in the past have thus attempted to identify such a specific mentalizing module in the brain. The right TPJ and parts of the medial PFC have been mentioned in some of these studies as candidate regions for a ToM reasoning module in the brain. The issue of a specific ToM module is discussed in more detail in chapter 4.6.

Besides exclusive accounts of TT and ST, other theories have proposed that ToM abilities may rely on assumptions made by both ST and TT. These hybrid accounts propose that people may rely in some instances on simulation, while others instances could require the use of concepts as proposed by TT (see also Stich & Nichols, 1997).

The imaging results of the study presented in this thesis may help to further support or refute assumptions made by either ST or TT. Significant activity for the attribution of false beliefs in areas associated with the human MNS could theoretically be seen as supporting evidence for ST.

The following section will review empirical data on executive functioning. As we shall see later, especially inhibitory control (IC) as one of the parts of executive functioning seems to be partially involved in the emergence and the expression of belief-reasoning. This is of special importance as the study presented here aims to shed light on the connection between IC and belief-reasoning in healthy adults.

3. Inhibitory Control

Generally speaking, executive functioning (EF) refers to the cognitive processes underlying goal-directed behavior (Sodian & Hülksen, 2005). EF encompasses a variety of different cognitive concepts such as inhibitory control (IC), planning, working memory, coordination and control of action sequences, and attention shifting (Carlson et al., 2005). Several studies investigating the relationship between EF and belief-reasoning have found that the strongest correlations for these concepts exist for IC (Kain & Perner, 2005). IC is defined as the ability to suppress actions that are unwanted and to select actions that are goal-directed (Simmonds et al., 2008). The following paragraph delineates the development of EF with special emphasis on IC. Also, experimental paradigms used to assess IC are reviewed. This is followed by a paragraph on the neural networks of IC.

3.1 Development and Measurement of Executive Functioning and IC

Young children are often described by their care-takers as present-oriented, impulsive, and stimulus-bound. As they grow older, children show improved planning skills, better error detection, and a behavior that is more goal-directed while representing multiple aspects of their goals (Zelazo et al., 2004). This observation by lay people is complemented by empirical data. A large body of research has found a U-shaped curve for the development of EF (Carlson, 2003; Carlson, 2005; Carlson et al., 2005; Rubia et al., 2006; Zelazo et al., 2004) with EF improving from childhood throughout adulthood and declining in old age. This finding is also true for IC (Bedard et al., 2002). Pronounced changes in the development of EF and IC take place especially during infancy.

Using an extensive battery of EF tasks, Carlson (2005) investigated the development of various aspects in children from ages 2 to 6 years. The results showed a significant increase in virtually all tasks administered with age. Further analyses were able to show that these changes were not related to the children's verbal ability. Carlson argues that these changes might reflect structural and functional maturation of the brain as well as children's increased exposure to social experiences independent of task-rule comprehension or verbal abilities.

These behavioral data are complemented by biological data showing maturational changes in the brain such as synaptic pruning and reorganization as well as changes in the gray matter / white matter ratio due to myelination processes (Huttenlocher & Dabholkar, 1997).

A variety of tasks have so far been used to study IC across the life span. In young children the so called Bear/Dragon task is widely used to assess IC. In this task, the child

is asked to do what a “nice” bear puppet tells them to do (e.g., “touch your nose”) and to refrain from doing what a “naughty” dragon puppet asks from them (Reed et al., 1984). In order to master this task, the child has to successfully suppress a response required by one of the characters.

A more advanced task that is related to but more difficult than the Bear/Dragon task is the “Simon says” task (Strommen, 1973). In this task children are required to execute a behavior such as touching the feet only if the experimenter starts the command with the words “Simon says”. For all other instances the child is required to remain still. Once again, this test requires children to inhibit a prepotent response and thus serves as a measure for response inhibition.

A frequently used test to assess IC abilities in children and also in adults is the Dimensional Change Card Sorting (DCCS) task (Zelazo et al., 2003). In this task the subject has to sort a stack of cards depicting one of two objects (a rabbit or a boat, for instance) presented in one of two colors (e.g., blue or red) according to shape. After several trials sorting by shape, the subject is then asked to start sorting the cards according to color. The rules are then changed back and forth several times. In order to solve this task, the children have to inhibit a previously valid behavioral rule and adopt a new behavioral strategy.

Another task similar to the DCCS task but more difficult is the Wisconsin Card Sorting Task (WCST). This test measures task switching and response suppression (Buchsbaum et al., 2005). In this test, the subject has to sort a stack of cards according to either color, shape, or number. The experimenter secretly determines the sorting rules and only tells the subject whether he or she has made an error or sorted correctly. This rule is then changed after 10 correct trials. Again, the experimenter does not state the new rule but keeps giving the subject feedback about his or her actions. Subjects lacking cognitive flexibility and IC will tend to keep sorting cards according to the original rule.

Another task that is predominantly used in adults is the Stroop task (Adleman et al., 2002; Stroop, 1935; Stuss et al., 2001a). In the original version of the task subjects are required to read the written meaning of a colored word or to name the color of a colored word. This task leads to increased reaction times and more errors compared to words that have a congruent color and meaning. To solve the task, a prepotent answer has to be overridden in favor of an alternative response. Thus, the Stroop task serves as a measure of response inhibition.

The most widely used task to study IC, however, is the so called Go / No-go paradigm (Simmonds et al., 2008). The original Go / No-go task involves two stimuli: A Go stimulus and a No-go stimulus. Subjects are instructed to press a button as quickly as possible whenever a Go stimulus is presented. Whenever a No-go stimulus appears,

subjects are asked to refrain from any response. In order to elicit a prepotent tendency to also falsely react to a No-go stimulus, there is usually a larger number of Go stimuli than No-go stimuli. The classic Go / No-go task measures IC with little demands on other cognitive processes such as working memory. Variants of the Go / No-go paradigm exist that also require other cognitive concepts to varying extents. For example, in one frequently used variant of the task, two stimuli are alternately presented on a screen (Garavan et al., 2006; Kelly et al., 2004). The subject is then required to press a button whenever the stimuli alternate, but to refrain from doing so if two identical stimuli are presented successively. This procedure measures inhibitory control but may also include, albeit to a lesser extent, working memory demands.

The next section deals with brain areas associated with IC. As we shall see in one of the subsequent sections of this paper, some substantial overlap exists between areas associated with IC and areas related to belief-reasoning.

3.2 Neural Correlates of IC

3.2.1 IC: Results from Lesion Studies

Lesion studies provide an elegant way to show causal relationships between circumscribed brain regions and corresponding cognitive concepts. Lesion studies have shown that particularly the prefrontal cortex plays an important role in IC.

Stuss et al. (2001a) compared the performance of 51 patients with frontal and non-frontal lesions to 26 healthy matched controls in the Stroop interference task, a measure of IC. No significant impairment in the task was found for the group of patients with non-frontal regions when compared to the control group. All patients with frontal lesions, however, showed a significant impairment in the task when compared to the controls. Among the group of patients with frontal lesions, damage to the non-cingulate superior medial frontal lobe was revealed to have the greatest effect on Stroop performance. These results point to a particular role of the superior medial cortex in IC.

This latter finding was corroborated by a study with 23 patients with focal lesions that were compared to a group of 19 healthy controls (Floden & Stuss, 2006). In order to measure IC performance, a Stop signal task was used. In this task, subjects had to press a button when a stimulus appeared (Go stimulus). Subjects had to withhold a response when the stimulus was accompanied by a tone (No-go stimulus). The results showed that among the patient group, those with lesions to the superior medial frontal gyrus of the right hemisphere showed the greatest impairment in the task. Results from both studies presented here indicate that the frontal cortex might be a key region supporting IC. The

right superior medial part of the frontal cortex in particular seems to be of special importance for response inhibition.

Another lesion study by Samson et al. (2004) investigated three patients with circumscribed lesions to the left TPJ. Of these three patients, only one patient was significantly impaired in measures of IC when compared to healthy controls. Unfortunately, no lesion study has so far investigated IC performance after circumscribed right TPJ lesions. Therefore, lesion studies (unlike imaging studies) cannot yield reliable evidence as to whether the TPJ could be implicated in IC.

Taken together, lesion studies suggest that the superior medial cortex may play an especially prominent role in IC.

Neuroimaging studies of IC corroborate this finding. However, imaging studies also point to other important areas related to IC, such as the TPJ. Results from imaging studies with respect to IC are presented in the following section.

3.2.2 Results from Neuroimaging Studies

A large number of studies using fMRI have investigated the neural correlates of IC (Blasi et al., 2006; Casey et al., 1997; Ciesielski et al., 2006; Garavan et al., 2006; Kelly et al., 2004; Liddle et al., 2001; Menon et al., 2001; Mostofsky et al., 2003; Stevens et al., 2007; Wager et al., 2005; Watanabe et al., 2002). The following section contains a selection of studies that are of importance to the study presented in this thesis. Also, the results of meta-analyses of response inhibition will be delineated.

An fMRI study that utilized a Go / No-go paradigm in order to investigate response inhibition was conducted by Kelly et al. (2004). In this study, subjects viewed the letters X and Y that were presented successively on a screen for 700 to 1100ms with a 100ms inter-stimulus interval. Subjects were instructed to press a button after every letter unless one of the letters was repeated. In that case subjects were instructed to withhold a response. A No-go stimulus was presented here approximately every 11s. In order to find brain areas associated with response inhibition, the events during which a successful response-inhibition for a No-go trial had occurred were contrasted with successful Go trials. Activity related to IC was found in a largely right-hemispheric network consisting of areas such as the dorso-lateral prefrontal cortex (DLPFC; BA 9, 46), middle frontal gyrus (BA 9, 10), medial frontal gyrus (BA 8, 24), bilateral TPJ (BA 40), precuneus (BA 7), and several subcortical regions.

Similar activity in a largely right-hemispheric network was found in other studies investigating the neural correlate of response inhibition. This result is also depicted in several meta-analyses combining previous studies on response inhibition.

In 2005, Buchsbaum et al. (2005) conducted a meta-analysis of studies that had used a Go / No-go paradigm to investigate IC. A total of 18 studies were introduced into an analysis based on activation-likelihood-estimation (ALE). In ALE, areas of brain activation concurring across several studies are estimated. In the present meta-analysis, highly lateralized activation related to IC was revealed in the right frontal cortex, encompassing the DLPFC (BA 9, 46), the inferior frontal gyrus (BA 44, 46) and the medial frontal gyrus (BA 6). Activity was also found in a small cluster in the left middle frontal gyrus (BA 46). Activity in non-frontal areas was revealed in the right TPJ (BA 40).

A more recent meta-analysis that also investigated brain activity related to IC as measured in Go/No-go tasks was able to confirm the above results (Simmonds et al., 2008). Common patterns of activation for 11 IC studies were revealed by also using an ALE method. Once again, activation related to IC was found in a largely right-hemispheric network consisting of superior frontal gyrus (BA 9), middle frontal gyrus (BA 6, 10), inferior frontal gyrus (BA 9, 44), superior medial frontal cortex (BA 6, 32) bilateral TPJ (BA 40), and subcortical structures. Results from the Simmonds et al. (2008) meta-analysis are depicted in figure 3.1.

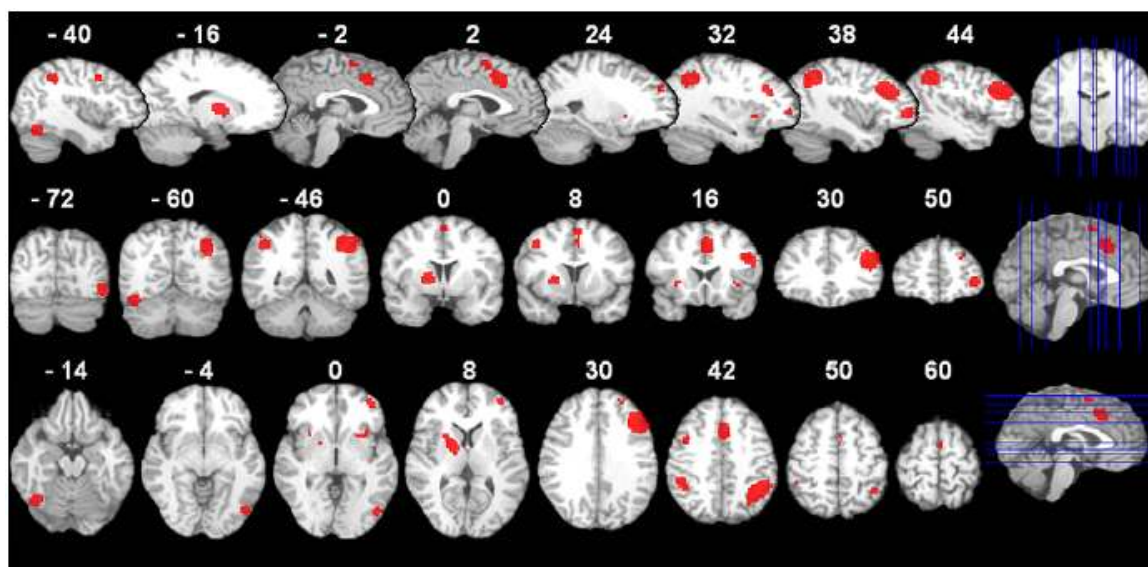


Figure 3.1. Results from the activation-likelihood-estimation (ALE) meta-analysis of 11 IC studies by Simmonds et al. (2008). Brain areas associated with IC are depicted in red. All of the studies included in the meta-analysis had used a Go / No-go paradigm to measure response inhibition. Numbers above each slice refer to corresponding Talairach coordinates. As can be seen here, IC-related activity is found in a largely right-hemispheric fronto-parietal network. This picture is taken from Simmonds et al. (2008). Reprinted with permission from Elsevier.

Taken together, a close examination of fMRI studies and meta-analyses dealing with the neural correlates of IC reveals surprisingly homogeneous patterns of activation. Firstly, virtually all studies report activation in a largely right-hemispheric network. Also, almost all studies find activation in quite similar fronto-parietal areas. Among these are the medial prefrontal cortex (BA 8, 9, 10, 32), the DLPFC (BA 9, 46) and areas either in or in the vicinity of the TPJ bilaterally (BA 39, 40). Lesion studies, probably due to their scarcity and focus on the frontal cortex, are only able to corroborate this view regarding the superior medial frontal cortex.

The role of the right TPJ in IC and other cognitive concepts has been scrutinized in more detail in a meta-analysis by Decety & Lamm (2007). This meta-analysis included 18 studies on IC (or attention reorienting, as it is referred to in the meta-analysis) and 24 studies on ToM reasoning. Interestingly, Decety & Lamm's meta-analysis found largely overlapping activity in the right TPJ region for both concepts.

This finding and a large body of other functional and behavioral data, especially from developmental studies, give rise to the assumption that IC and belief-reasoning may be closely related. Evidence and hypotheses regarding a possible connection between belief-reasoning on the one hand and IC on the other hand will be discussed in the next section. In addition, the study presented in this thesis is aimed at clarifying a possible connection between belief-reasoning and IC.

4. ToM and Inhibitory Control

At first glance, EF including IC and the ability to attribute mental states seem to describe two quite different concepts. However, developmental psychologists discovered that both concepts show surprisingly similar developmental courses with significant improvements between the ages of 3 and 6 years (Bedard et al., 2002; Wellman et al., 2001; Wellman & Liu, 2004; Zelazo et al., 2003; Zelazo et al., 2004). Subsequent behavioral studies were able to find strong correlations between belief-reasoning and EF. For instance, a meta-analysis by Perner & Lang (1999) found a mean effect size of 1.08 for the correlation between EF and general ToM reasoning in children.

Evidence for this connection comes from behavioral studies which will be discussed in more detail in this section. Furthermore, the following paragraph will examine evidence for a ToM / IC connection stemming from observations of children suffering from autism and ADHD. Neuroimaging results complement this finding and will also be delineated. These scientific data have also sparked hypotheses trying to explain this possible connection. Several of these hypotheses will thus be mentioned in the following section as well. Special emphasis is dedicated to IC, which among EF shows the strongest correlation with ToM reasoning.

4.1 Behavioral Results Concerning the ToM / IC Connection in Childhood

4.1.1 Correlational Data

The observation of a similar developmental timetable for IC and the ability to attribute mental states has led to several studies that investigated the relationship between the two latter concepts. A lot of this work has examined the connection in preschoolers at ages 3 to 6 years.

In an extensive study by Carlson et al. (2001), 107 children between the ages of 3 and 4 years were administered a battery of tasks including both measures of belief-reasoning and IC. Among the belief-reasoning task were two tests of false-belief understanding. IC was assessed by using tasks such as the Stroop task, the Bear / Dragon task or the DCCS task. The authors found a strong correlation for performance in the IC and the belief-reasoning tasks ($r = 0.66$; $p \leq .001$). This correlation remained significant even after controlling for factors such as age, verbal ability and gender.

Among the IC tasks administered, the tasks requiring the suppression of a prepotent response (conflict task; e.g., the Bear / Dragon task) were revealed to have a higher

predictive influence on ToM performance compared to tasks requiring the delay of a prepotent response (delay task; e.g., executing a motor response after a delay).

A subsequent study examined the connection between ToM and IC in 47 children ages 3 to 5 years (Carlson et al., 2002). This time, general intelligence as well as working memory capacity was assessed in addition to measures of verbal ability. Once again, a strong connection was found between performance in belief-reasoning tasks and performance in IC tasks. Interestingly, this connection remained strong once the factors working memory, verbal ability and even general intelligence were controlled for. This finding emphasizes the special relationship between IC and the ability to attribute beliefs. Among EF, IC seems to be the only sufficient factor to explain the relationship between belief-reasoning and EF.

A similar result was also found in a study that examined IC, false-belief reasoning and other cognitive concepts in 69 children between the ages of 37 and 65 months of age (Müller et al., 2005). As a measure of IC, the authors used the DCCS task. The “crayon task” as used in a study by Gopnik & Astington (1988) was used to assess false-belief understanding. The results showed that the performance in the IC task was correlated to performance in the false-belief task. This correlation remained robust even after the factors age and verbal ability had been cancelled out.

A strong correlation between measures of IC (among them the Bear / Dragon task and a simple Go / No-go task) and false-belief understanding (as measured with standard false-belief tasks such as the Sally-Anne paradigm) was also found in a study by Flynn (2007). The children between the ages of 3 and 4 years were tested monthly on all measures over a period of 6 months. This longitudinal approach revealed that early IC competence predicted later false-belief performance. Interestingly, the same was not true for the opposite direction: the mastery of early false-belief understanding did not predict later IC performance. This finding may provide strong support for the hypothesis that IC is a necessary prerequisite for the emergence of false-belief reasoning.

Taken together, correlational data provide strong support for a relationship between IC and belief-reasoning in childhood. Although only a few studies were exemplarily delineated here, this finding has been replicated in several other studies as well (Carlson et al., 2004; Hughes & Ensor, 2007; Perner et al., 2002a; Perner & Lang, 1999).

Interestingly, this strong relationship between IC and belief-reasoning has also been found in children from various cultural backgrounds. Sabbagh, Xu, Carlson, Moses, & Lee (2006) compared U.S. preschoolers with an age-matched group of Chinese preschoolers. For the U.S. preschoolers, the same sample was used as in the previously described study by Carlson et al. (2001). Performance in IC tasks proved to be a strong predictor for false-belief performance in both samples. The Chinese sample, however, had significantly

higher scores in the IC battery. Nevertheless, the Chinese preschoolers had false-belief reasoning scores that were comparable to the American sample. The authors argue that this result may indicate that other factors along with IC abilities may also play a role in the emergence of belief-reasoning. Since most Chinese children do not have any siblings due to China's "one child policy", they may be less exposed to social interactions and thus learn less about others' mental states. It could therefore be that the number of siblings constitutes a factor for ToM development.

Another study investigated the belief-reasoning / IC connection in preschoolers from Germany, Costa Rica and Cameroon (Chasiotis et al., 2007). A correlation between conflict inhibition and belief-reasoning was found across all three cultures. Delay inhibition tasks, however, were not related to false-belief performance, as shown in a previous study by Carlson et al. (2001).

4.1.2 Training Effects

An ingenious study corroborating findings of a strong relation between IC and belief understanding comes from Kloo & Perner (2003). Instead of taking on a correlational approach, the authors investigated the effects of training in either a measure of IC or a false-belief task. Seventy-four children between the ages of 3,0 and 4,7 years took part in the experiment. The children were assigned to one of three groups. Each of these groups then received training in false-belief stories, the DCCS task (a measure of IC), or in a number conservation task (control group). The training sessions started one week after a pre-test session. The training consisted of two training sessions within a week and were followed by a post-test session one week after the last training. Post- and pre-test sessions consisted of an assessment of performance in all three task batteries (false-belief, IC, number conservation). A comparison of post- and pre-test results revealed that training in the DCCS task significantly improved performance both in the DCCS task itself as well as in the false-belief tasks. Also, the children who received training in the false-belief tasks showed significantly improved performance in the DCCS task. The children who were trained in the control task did not show any significant improvement in either task. This exciting finding has shown for the first time a causal relationship between IC and belief-reasoning that cannot be attributed to mere maturational processes or to general training effects in both tasks. Nevertheless, this finding cannot explain the processes mediating this connection. However, the authors tentatively propose that both the IC and the false-belief tasks might as basic processes require the re-description of objects as different things.

As has been specified in this paragraph, an abundance of correlational data as well as one training study point to a strong relationship between IC and belief-reasoning in early childhood. Fewer studies have investigated this connection in adulthood. These studies will be discussed in the following section.

4.2 Adults and the ToM / IC Connection

One of the first studies that investigated the relationship between inhibition and ToM reasoning in adults was conducted by Chasiotis & Kiessling (2004). In this study, ToM reasoning in adults between the ages of 17 and 59 years was assessed with auditorily presented stories that contained social interactions. At the end of each story subjects had to answer questions that required the attribution of desires, emotions and beliefs to the story's protagonists. In order to measure IC abilities, subjects performed a computer game that required the suppression of a prepotent answer. The results revealed a significant correlation between IC and ToM reasoning in the adult group. This correlation was similar to results from studies with preschoolers. However, according to the authors these correlations were somewhat smaller compared to previous children's studies.

Another study compared a group of older individuals between 62 and 90 years of age to a group of younger subjects between the ages of 18 and 26 years (German & Hehman, 2006). Both groups were assessed in their EF abilities by using a battery of different tests, including measures of IC such as the Stroop task and the Day / Night task. In the Day / Night task, subjects are required to say the opposite of what is shown on a card. If a picture of the moon is depicted, subjects have to say "day". A picture of the sun on the other hand requires the answer "night". In order to measure false-belief reasoning abilities, subjects were prompted to answer questions regarding written stories. Only the false-belief condition required the subjects to attribute mental states, not so the true-belief condition. The difficulty of the belief stories was altered by varying the protagonist's desire to either approach or avoid a certain goal. By doing this, the authors claim to have varied the amount of executive demands posed in the belief story. The results revealed that both the younger and the older group of subjects showed poorer performance in the belief task with increasing executive demands. Moreover, this effect was significantly increased in the group of older subjects. Also, regression analyses including the results from the EF battery showed that especially measures of IC and processing speed were able to explain the most variance in accuracy and processing speed in the belief-reasoning task. These results indicate that compromised IC in particular could explain poor belief-reasoning abilities in old age. A decline in belief-reasoning abilities in old age has also been described in previous studies (e.g., Maylor et al., 2002; McKinnon & Moscovitch, 2007).

Taken together, behavioral studies indicate that IC and belief-reasoning are strongly connected. This connection has been confirmed in young children and in young and old adults. Theories trying to explain a hypothesized connection between IC and belief-reasoning are presented next.

4.3 Hypotheses Regarding the ToM / IC Relationship

Various theories have been suggested trying to explain the strong empirical evidence indicating a connection between IC and belief-reasoning.

One prominent theory is referred to as “emergence theory” (Carlson et al., 2002; Perner & Lang, 1999; Sabbagh et al., 2006). According to this theory, IC is a necessary prerequisite for the development of belief attribution. The theory suggests that it is necessary for a child to inhibit his or her impulses and to be able to attend to new situations in order to pay attention to other people’s mental states. It also assumes that IC is not necessary anymore for attributing beliefs once a child has acquired a fully efficient ToM. Emergence theory states that from this point on, ToM reasoning modules may work independently of EF abilities such as IC.

Richer versions of emergence theory suggest that IC may be necessary in order to become exposed to social situations during which the children can then learn about discrepancies between mental states and reality. This view is backed by a study that found advanced IC abilities in Chinese preschoolers when compared to U.S. preschoolers (Sabbagh et al., 2006). Although IC abilities predicted belief-reasoning performance in both samples, Chinese preschoolers showed no superior belief-reasoning abilities. The authors argue that this finding may be due to the fact that U.S. preschoolers grow up with more siblings and are thus exposed to more everyday social experiences in which they can learn about others’ mental states. Further evidence comes from a longitudinal study showing that early IC performance predicts later belief-reasoning performance in children between 3 and 4 years of age (Flynn, 2007). Early belief-reasoning abilities on the other hand did not emerge as a predictor for IC performance. Emergence theory can thus account for some of the behavioral results that show a relationship between IC and belief-reasoning in 3 to 5 year old children. However, it cannot account for findings that IC performance and belief-reasoning abilities are also correlated in adulthood and old age (German & Hehman, 2006; McKinnon & Moscovitch, 2007)

Another theory, dubbed “expression” or “performance” theory attempts to explain the connection between IC and belief-reasoning from childhood to old age (Kloo & Perner, 2003; Perner & Lang, 1999; Siegler & Varley, 2002). Expression theory proposes that belief-reasoning tasks require the suppression of the true state of affairs in favor of one’s

own or others' mental states by means of IC (Hughes & Russell, 1993; Russell et al., 1991). As such it is argued that young children may possess a concept of belief long before they actually are able to use it. Only their poor IC abilities may hamper the expression of their ability to attribute mental states.

Expression theory therefore implies that decreasing inhibitory demands in belief-reasoning tasks could actually improve children's belief-reasoning performance. However, this prediction of expression theory is not true. Explanation versions of false-belief tasks with little to no inhibitory demands are as hard for children as the original task versions (Hughes, 1998; Perner et al., 2002a; Perner & Lang, 1999). It could be that only children at an intermediate belief-reasoning level with intermediate IC abilities are aided by decreasing inhibitory task demands (Wellman et al., 2001). This hypothesis has not been tested empirically, though. Further, according to the expression theory, a lack of inhibitory control and thus belief-reasoning would imply that young children will typically answer a false-belief question with the actual state of reality. This, however, has not been observed either. As a matter of fact, children who do not pass false-beliefs tasks yet answer questions about false-beliefs at chance level, suggesting that they do not possess any concept of mental states at all (Wellman et al., 2001). Due to several arguments against expression theory, this theory has little explanatory power regarding the relation between IC and the attribution of beliefs.

Another theory suggests that belief-reasoning may be a prerequisite for IC (Perner, 1998; Perner & Lang, 1999). According to this theory, children first have to understand that mental states have causal power, meaning that these states are the driving force behind people's actions. IC on the other hand requires the subject to realize that a tendency exists towards executing a wrong action and that this action needs to be inhibited. Hence, this tendency within exhibits causal power as well. The understanding that mental states possess causal power may therefore also foster IC, which requires an understanding of causal power as well. Evidence for this theory comes from the finding that hyperactive children suffering from an impairment in IC show relatively unimpaired belief-reasoning abilities (Sodian et al., 2003).

More advanced theories concerning the IC / belief-reasoning relation propose a bi-directional relationship between the two latter concepts. The acquisition of a ToM in this theory may lead to improved IC which then facilitates mental state attribution, especially in situations with high inhibitory demands (Sodian & Hülken, 2005).

Yet another class of more elaborate theories has focused on various components underlying the attribution of mental states. According to these theories, mental state attribution is supposed to require a core ToM system ("ToM Mechanism; ToMM;) as well as a co-opted system, also referred to as a "selection processor" (Leslie et al., 2004;

Leslie et al., 2005). This selection processor is related to IC. According to this theory, the maturation of belief-reasoning abilities may rely on an improvement in IC skills with the core belief-reasoning system remaining relatively stable. However, accounts that belief-reasoning performance in early childhood is not facilitated by lowering executive task demands oppose this view (Leslie et al., 2004; Wellman et al., 2001). The power of this approach lies nevertheless in its ability to explain instances of compromised belief-reasoning capacities while showing intact IC as well as of relatively intact belief-reasoning with compromised IC, as found in children with ADHD (Sodian et al., 2003).

In sum, a series of hypotheses has attempted to explain results suggesting a strong relation between belief-reasoning and IC. So far no hypothesis has been able to sufficiently account for all existing behavioral results. More elaborate hypothesis may be needed to explain and predict phenomena related to the connection between IC and belief-reasoning. This matter will be discussed in more detail with respect to the present study's results in the discussion section of this paper.

As briefly mentioned in the previous section, individuals with disorders such as autism or ADHD show deficits to varying extents in either belief-reasoning or IC, corroborating findings of an IC / belief-reasoning connection. Systematic investigations of impaired IC and belief-reasoning functioning in autism and ADHD can provide insights about the relation between the two latter concepts. Findings from patients with compromised IC / belief-reasoning abilities will be discussed next.

4.4 Autism and ADHD: a Case in Point for the ToM / IC Connection?

4.4.1 Autism

Autism is a severe developmental disorder with a prevalence between 0.16 to 0.22 % (Sodian & Thoermer, 2006). The main symptom in autism is a dysfunction in verbal and non-verbal communication. Patients suffering from autism are unable to understand and cope with the requirements of the social world. Another hallmark of autism is a frequently observed engagement in stereotyped behaviors. A subgroup of autistic spectrum disorders (ASD) is Asperger's syndrome. Patients with Asperger's also show impaired social interactions but are not impaired in their verbal and cognitive development. This disorder has also been referred to by some as "high-functioning autism".

It has been found that patients suffering from ASD show a severe impairment in their ability to attribute beliefs. In one of the first studies involving autistic children, Baron-Cohen et al. (1985) presented the Sally-Anne paradigm to healthy children, to children with Down's syndrome and to autistic children. Compared to the children with Down's

syndrome and to healthy controls, children with autism were significantly impaired in their ability to attribute false beliefs.

A similar result was found in a study by Sodian & Frith (1992). In this study, performance in the sabotage-deception task (a measure of false-belief reasoning) was compared between healthy, mentally retarded and autistic children. Compared to healthy controls and mentally retarded children, autistic children performed significantly worse when required to tell a lie in order to deceive another person. Since autistic children were equally able to physically manipulate the protagonist's behavior, this impairment is not likely to be due to general task demands.

Several studies have also shown that patients suffering from ASD are significantly impaired in measures of EF (Hill & Bird, 2006; McEvoy et al., 1993). Evidence for this finding is now considered quite robust (Hughes, 2002). For instance, a study conducted by Ozonoff, Pennington, & Rogers (1991) was able to show that, compared to healthy children, autistic children with an IQ in the normal range were significantly impaired in the WCST, a measure of IC. A review article by Pennington & Ozonoff (1996) found that 13 out of 14 studies assessing EF in autistic children showed a significant EF dysfunction. These behavioral findings suggest that the observed impairment in autistic individuals' belief-reasoning abilities could be related to their poor EF abilities.

While several behavioral studies have investigated EF and belief-reasoning in ASD patients, there are only a small number of functional imaging studies that have investigated ASD patients. One such imaging study investigated the attribution of mental states. Compared to controls, autistic patients showed significantly less activity in medial prefrontal cortex (BA 9) and bilateral TPJ in a task requiring the attribution of mental states (Castelli et al., 2002). Another belief-reasoning study by Nieminen-von Wendt et al. (2003) found less activation that was less wide-spread compared to controls in the medial prefrontal cortex.

Some EF tasks in autistic patients show a similar pattern. In an EF task requiring the subject to switch between stimulus-oriented and stimulus-dependent phases, autistic subjects showed abnormal patterns of activation in the medial PFC when compared to healthy controls (Gilbert et al., 2008).

Hence, preliminary empirical findings suggest that ToM deficits in autistic patients may be mediated by the medial PFC. This abnormal functioning is also likely to be reflected in an impairment in EF.

4.4.2 ADHD

An impairment in EF is the cardinal symptom of yet another developmental disorder, ADHD. This disorder has its onset in childhood and a world-wide prevalence of some 5%. It is estimated that about 60% of children diagnosed with ADHD maintain this disorder through adulthood (Polanczyk et al., 2007). ADHD symptoms include EF deficits such as inattention, impulsive behavior, hyperactivity, distractability and an impairment in inhibitory control (Sodian et al., 2003). This finding has been tested empirically and has been confirmed in a large number of studies (Hughes et al., 1998; Hughes, 2002; Seidman et al., 1997). Naturally, this finding has led to the speculation that children diagnosed with ADHD may also be impaired in their ability to attribute beliefs. This assumption is derived from the finding in healthy subjects that the performance in IC tasks and belief-reasoning are highly correlated. Hence, several studies have investigated this hypothesis by testing subjects with ADHD and healthy controls in measures of belief-reasoning and IC.

In a study by Perner, Kain, & Barchfeld (2002b), 24 children at risk of ADHD were compared to a group of 22 non-ADHD children. The children's ages ranged between 4 ½ and 6 ½ years. The children were assessed in their ability to attribute beliefs with a variety of second-order ToM tasks, including measures of false belief. Also, they were administered several tests of EF, including tests for IC. As expected, the group of ADHD children showed a significant impairment in several EF tasks. However, they showed no impairment at all in second-order ToM reasoning performance. The authors argue that this finding speaks against accounts that EF may serve as a prerequisite for later belief-reasoning. This finding of unimpaired mental state attribution despite significant EF impairment is accounted for by the theory that the ability to attribute mental states could lead to improved IC (Sodian et al., 2003).

A study conducted by Sodian & Hülken (2005) found similar results. In this study, 32 children with ADHD were compared to 101 normally-developing controls. Children were administered a battery of ToM tasks, including a second-order false-belief task. Also, children were assessed with several measures of IC. With the exception of a task that required the understanding of epistemic states, children with ADHD showed no different performance compared to healthy controls in measures of second-order ToM reasoning. However, children with ADHD were significantly impaired to controls in IC tasks. The authors argue that poor performance in the epistemic state attribution task as a measure of ToM reasoning may have been due to the relatively high inhibitory demands of the task. It is suggested then that children diagnosed with ADHD may be able to attribute mental states but that they are unable express this ability in some tasks due to high inhibitory task demands. This finding is best accounted for by the theory that the acquisition of a ToM

leads to better inhibitory skills. These IC abilities could then in turn help to apply mental state knowledge in the social environment.

Imaging experiments have also investigated the neural correlate of the observed EF deficit in ADHD. According to one of these studies, the frequently observed EF impairment in ADHD seems to be associated with dysfunctional patterns of activation in anterior cingulate cortex, DLPFC and inferior PFC (Dickstein et al., 2006).

Summing up, results of compromised IC and/or belief-reasoning in ADHD and autism can help collect evidence for or against some of the previously discussed theories concerning the connection between IC and the attribution of beliefs. As has been delineated here, ToM reasoning deficits in autism could be due to poor IC abilities that are mediated by the medial PFC or vice versa. Children with ADHD, however, show EF deficits in spite of relatively intact mental state attribution. It seems as if children with ADHD are only impaired in ToM tasks that require relatively high inhibitory demands.

In addition to behavioral results stemming from both healthy subjects and impaired individuals, results from neuroimaging and lesion studies may help shed further light on the IC / belief-reasoning dispute. This evidence is given in the following section.

4.5 Neural Correlates of IC and ToM

Functional imaging and lesion studies have attempted to clarify the strong behavioral connection between IC and belief-reasoning. Lesion studies provide a powerful tool to draw causal conclusions about the impact of circumscribed brain regions for IC and belief-reasoning. Functional imaging studies on the other hand can give helpful hints about the nature of the belief-reasoning / IC relation by investigating the degree of overlapping brain activity for both concepts.

4.5.1 Brain Lesions and Their Effects on ToM and IC

A small number of studies have tested patients with brain lesions both in their ability to attribute mental states as well as in their ability to inhibit prepotent responses. The focus in these studies has been primarily on the PFC.

One of the first such study assessed belief-reasoning and IC in a group of 31 patients with unilateral frontal brain damage (Rowe et al., 2001). Performance in a false-belief test and measures of executive control (e.g., Stroop task, WCST) was compared to a group of healthy controls. Patients with right-sided as well as left-sided lesions were significantly impaired in their ability to attribute beliefs. However, both the patient and the control group were able to answer control questions that did not require the attribution of

mental states. Thus, the observed belief-reasoning deficit seems not to have been due to general task demands. Also, the patient group showed a significant impairment in nearly all measures of EF, including measures of IC, when compared to the healthy controls. Although the patients with PFC lesions showed an impairment in both EF and belief-reasoning, these two may not be causally related. This was shown in a subsequent analysis of covariance. According to this analysis, EF and belief-reasoning deficits were independent of each other. The authors argue that this result indicates that a specialized ToM module that is not related to EF may exist in the frontal lobes.

Another study also investigated belief-reasoning and IC in 12 patients with brain lesions (Apperly et al., 2004). Among those, 4 patients had widespread lesions in the bilateral PFC. Compared to controls, these patients showed a significant impairment in their ability to attribute beliefs (as assessed with a story- and video based false-belief task) that was accompanied by a significant EF impairment (including a measure of IC). Thus, results from this lesion study indicate that the PFC supports both EF as well as belief attribution. However, the data of this study are not able to clarify whether one function possibly underlies the other or whether a part of the PFC might serve as a modularized area for belief attribution.

The same study also included three subjects with lesions to the left TPJ. Interestingly, these subjects were significantly impaired in their ability to attribute beliefs. Nevertheless, two of the three patients showed relatively unimpaired EF abilities.

Viewed as a whole, these results indicate that the PFC is likely to be one of the regions responsible for IC and belief-reasoning. Unfortunately, differing lesion sites throughout the PFC as well as only vague descriptions of the lesion sites do not yield a more exact localisation for either function. Also, the results are not clear enough to make an assumption about whether there might be a module for belief-reasoning in the PFC. More solid evidence exists for the left TPJ. As a dissociation for this region has been found for EF and belief-reasoning, it could be that this region is indeed specifically activated in belief-reasoning tasks (Apperly et al., 2004).

Further data concerning common networks for belief-reasoning and IC come from imaging studies. These data and their relevance to modular accounts of belief-reasoning are outlined in the following paragraph.

4.5.2 Neuroimaging of IC and ToM

Neuroimaging results concerning activity related to IC and belief-reasoning have been described in detail in two previous sections of this thesis. Surprisingly consistent activation

within each cognitive concept has been reported in single studies and meta-analyses alike.

As far as IC is concerned, several regions have emerged as possible key regions for response inhibition. A close examination of available literature reveals that most single studies found activity in a largely right-hemispheric network consisting of the medial PFC (BA 6, 8, 9, 32), the DLPFC (BA 9, 46) and the TPJ (BA 39, 40). The TPJ seems to be one of the few regions that is activated bilaterally for IC. Two recent meta-analyses have come to a similar conclusion (Buchsbaum et al., 2005; Simmonds et al., 2008). According to these analyses, commonly observed activity for IC is found in the medial PFC (BA 6, 8, 32), the bilateral TPJ (BA 39, 40), right DLPFC (BA 9) and right inferior frontal gyrus (BA 44). Generally speaking, there is converging evidence that three key regions seem to play a crucial role for response inhibition. These are the medial PFC (BA 6, 8, 32), the right DLPFC (BA 9) and the TPJ bilaterally with emphasis on the right hemisphere (BA 39, 40).

Relatively consistent activation patterns have been identified for belief-reasoning as well, although these studies have used a tremendous variety of different tasks. Across virtually all studies, only two key regions have been consistently activated in the attribution of mental states (Kain & Perner, 2005; Perner & Aichhorn, 2008; Saxe et al., 2004). These are the medial PFC (encompassing BAs 9, 10, 32) and the bilateral TPJ (BA 39, 40). Although some researchers have claimed that the temporal poles might also constitute a key region for belief-reasoning (Frith & Frith, 2003; Gallagher & Frith, 2003), more recent studies have not reported any activity in this area (e.g., Sommer et al., 2007). No meta-analysis has so far been dedicated to belief-reasoning at the whole brain level.

A comparison between areas associated with belief-reasoning and areas implicated in IC shows quite clearly that both the medial PFC as well as the TPJ bilaterally are implicated in the attribution of beliefs and IC alike. Only the right DLPFC as a key region for IC is not as commonly activated in studies investigating belief-reasoning.

A thorough meta-analysis has investigated activity for both IC and ToM reasoning in the right TPJ (Decety & Lamm, 2007). This was done in order to compare the exact localisation of TPJ activity in both cognitive concepts. For their meta-analysis, the authors included 18 studies related to IC (“attention reorienting”) and 24 studies related to ToM reasoning. Results showed a substantial overlap for both concepts in the right TPJ region (see figure 4.1).

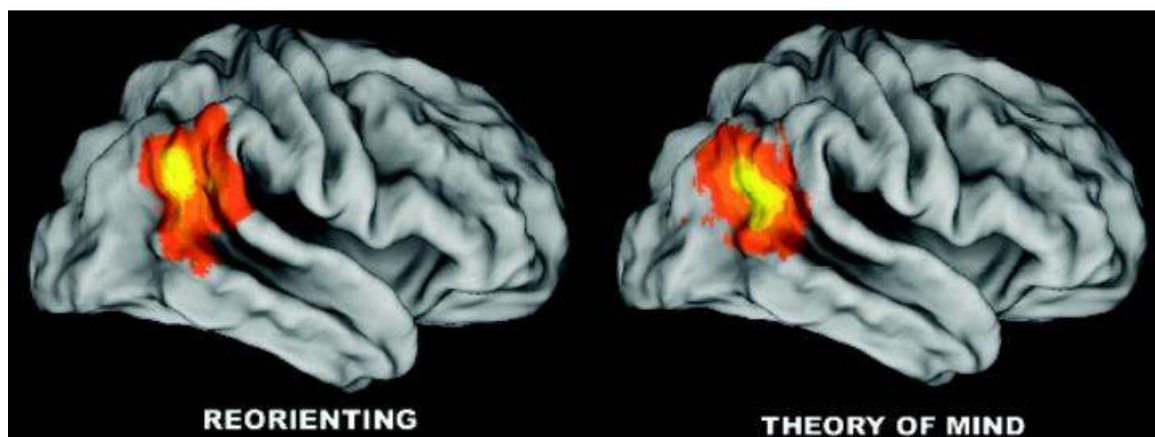


Figure 4.1. Activity in the right TPJ region stemming from a meta-analysis including 18 studies related to attention reorienting (left image) and 24 studies related to ToM reasoning (right image). Image taken from Decety et al. (2007). Reprinted with permission from SAGE publications.

So far, no meta-analysis has investigated activity associated with belief-reasoning and IC in the medial PFC. However, a careful examination of relevant studies reveals that medial PFC activity related to IC seems to be slightly posterior to areas commonly activated by belief-reasoning. While most studies investigating IC report activity in and around BAs 6, 8, 32, activity related to belief-reasoning seems to be restricted to the more anterior BAs 8, 9, 10 and 32. However, this interpretation especially concerning the medial PFC has to be treated carefully. Also it seems impossible to draw any conclusions about the nature of the IC / belief-reasoning connection based on the results available. All imaging data used to compare the two concepts stem from largely differing studies. Firstly, results are compared based on differing sample groups. Also, even the modalities used to compare activations in both concepts vary. While most studies have used fMRI, some of the data originate from PET experiments. Even though most data stem from fMRI experiments, scanners of differing field strengths have been used to provide data. This may result in differing activation. Furthermore, neither belief-reasoning studies nor IC studies have used similar paradigms. Belief-reasoning studies have used paradigms that sometimes may not have tapped belief-reasoning at all. Instead, the attribution of intentions or desires may have been measured. The same is true for IC studies. While some studies have used Go / No-go studies, other studies have used Stroop tasks or card sorting tasks, thus requiring differing extents of other concepts such as working memory or verbal ability. Last but not least, differing analyses of imaging data may have yielded differing activations.

Based on data stemming from largely differing paradigms, some researchers have nevertheless attempted to make assumptions about the connection between belief-reasoning and IC. For instance, Saxe et al. (2004) have claimed that “[...] belief attribution – even of false beliefs - appears to rely on distinct neural mechanisms from those

responsible for response conflict, selection, and inhibitory control.”(Saxe et al., 2004; p. 101). Others have interpreted existing results differently. Concerning the right TPJ, Decety et al. (2007; p. 583) have claimed that “[...] the TPJ implements a component that is not domain-specific in social cognition but is rather a more general and powerful computational process that operates in many other contexts besides theory of mind.”

Clearly, these differing interpretations of the same data call for studies that investigate both belief-reasoning and IC in one single study. Valid inferences about the belief-reasoning / IC connection on the neural level can only be drawn from such a methodological approach.

4.6 Evidence for an Independent ToM Module

Results from neuroimaging studies have also been used to clarify whether regions in the brain exist that are specific for belief-reasoning, as has frequently been suggested in some studies.

In order for a region to qualify as a specialized region for attributing beliefs (“belief-reasoning module”), two criteria should be met: generality and specificity (Saxe et al., 2004). According to the generality criterion, a candidate region should always show increased activity to all experimental paradigms requiring the attribution of beliefs. The specificity criterion on the other hand asks that such a region is only active during belief-reasoning and not during other mental operations. For instance, a specialized belief-reasoning area should not be activated during reasoning about non-mental states or because of the paradigm containing social stimuli. In order to meet the specificity criterion, control conditions in an experiment should only differ to the respective belief-reasoning condition with regard to the attribution of beliefs. All other parameters need to be held constant. The so called subtraction logic may only be applied under these circumstances. According to this procedure, subtracting the control condition from the belief-reasoning condition will result in an area dedicated to the attribution of beliefs. This region would then qualify as a belief-reasoning module if it showed increased activity in all other tasks requiring the attribution of beliefs (generality) and if no other cognitive process were to be associated with activity in this region (specificity).

One research group has frequently reported to have found a region in the brain that meets both criteria to be a belief-reasoning module (Saxe & Kanwisher, 2003; Saxe & Powell, 2006; Saxe & Wexler, 2005). The belief-reasoning tasks used by this group frequently revealed activity in the right TPJ. According to these results, the generality criterion seems to be met by the right TPJ. Saxe and colleagues have also claimed that no other cognitive concept would be associated with activity in a similar area. Therefore they

claim that the right TPJ also meets the specificity criterion and therefore constitutes a highly-specialized belief-reasoning module. This view, however, is not shared by other researchers who report activity in a similar region in the right TPJ for other cognitive concepts independent of mental state attribution (Corbetta et al., 2008; Decety & Lamm, 2007).

The dispute over whether the right TPJ or another region qualify as a belief-reasoning module is exacerbated by a methodological problem. Areas identified as neural correlates of belief-reasoning and areas identified as correlates of other cognitive concepts stem from different studies. As has been stated above, this methodological flaw can only yield vague assumptions about a possible ToM module. Due to several methodological considerations, this approach cannot answer the question of whether activity in regions related to concepts other than belief-reasoning might overlap with a possible belief-reasoning region. Data stemming from different studies are therefore unable to serve as evidence for meeting the specificity criterion. However, this criterion has to be met in order for a brain region to qualify as a belief-reasoning module.

Clearly, this methodological dilemma can only be solved by investigating belief-reasoning and another related cognitive concept within a single study. This procedure can account for differing sample groups, differing paradigms, differing analyses and differing scanning properties. One of the cognitive concepts related to belief-reasoning with possible overlapping brain activity in the vividly disputed right TPJ is IC. Thus, single studies investigating both IC and belief-reasoning may help shed light on the nature of the IC / belief-reasoning connection and on the ToM modularity dispute.

Two studies have so far attempted to investigate both concepts in a single study. These studies will be presented next.

4.7 Investigating IC and Belief-Reasoning in a Single Study

The first fMRI study that attempted to investigate belief-reasoning and EF in one single study using a within-subjects design was conducted by Saxe, Schulz, & Jiang (2006). Their main experiment was preceded by a reference experiment, which consisted of a response selection task and a ToM task.

The response selection task required subjects to press a button corresponding to one line that differed in length from three other lines presented on a screen. In a compatible condition of this task, subjects had to press a button that corresponded to the actual target location. In the incompatible condition of the task, subjects were required to press a button that was shifted two locations from the actual target position. A schematic depiction of the response selection task of the reference experiment is shown in figure 4.2.

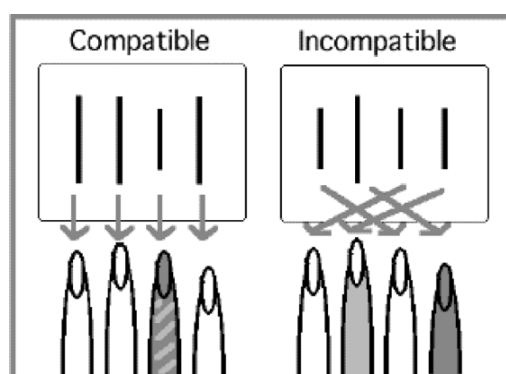


Figure 4.2. The response selection task used in a study by Saxe et al. (2006). In the compatible condition, subjects had to press a button that referred to a line differing in length from the other presented lines. In the incompatible condition, subjects had to press a button that was shifted two locations from the location showing the line that differed in length. This figure is taken from Saxe et al. (2006). Reprinted with permission from Psychology Press.

Conditions were presented in a block design. A ToM reasoning task was also included in the reference experiment. This ToM task consisted of written false photograph stories and belief stories that were followed by questions that the subject had to answer. An outline of Saxe et al.'s ToM task is presented in figure 4.3.

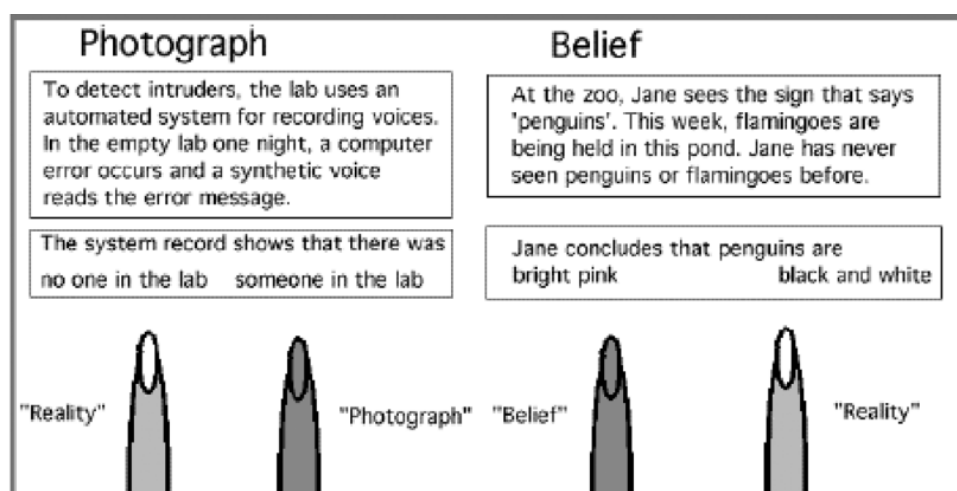


Figure 4.3. The ToM reasoning task in Saxe et al.'s (2006) reference experiment is depicted here. In both conditions, subjects were required to complete sentences by pressing one of two buttons. In the false-belief condition subjects had to attribute a mental state to a protagonist. The false-photograph condition on the other hand required no such decoupling between a mental state and reality. Reproduced from Saxe et al. (2006) with permission from Psychology Press.

In Saxe et al.'s reference experiment, activation in the incompatible response selection condition was greater than in the compatible response selection condition in regions such as bilateral intraparietal sulcus, pre-supplementary motor area (pre-SMA), bilateral frontal gyrus, and right middle temporal gyrus. Activity in the belief condition was greater than in the false photograph condition in bilateral TPJ, medial PFC, posterior cingulate, bilateral amygdala, and right superior temporal gyrus. The authors claim not to have found any overlapping brain regions. Unfortunately, their paper does not show significant activation from the group analyses rendered on a template brain. Instead, activation patterns of four individual subjects were depicted on a template brain. Also, the authors fail to specify respective Brodmann areas.

Results from the reference experiment as described above were then used to tailor individual ROIs for the main experiment. During the main experiment, subjects watched a false-belief cartoon during which a chocolate bar moved from one box into another box or back into its original location. A girl was either watching this or turning away as the bar moved again to one of the two boxes. In one condition ("ToM condition") subjects were then asked where they thought the girl thinks that the chocolate was. Another condition ("response selection") asked subjects to apply one of two algorithms to solve the task (algorithm 1: "girl facing, then pick last box"; algorithm 2: "girl looking away, pick first box" and vice versa). A schematic outline of both tasks in the main experiment is depicted in figure 4.4.

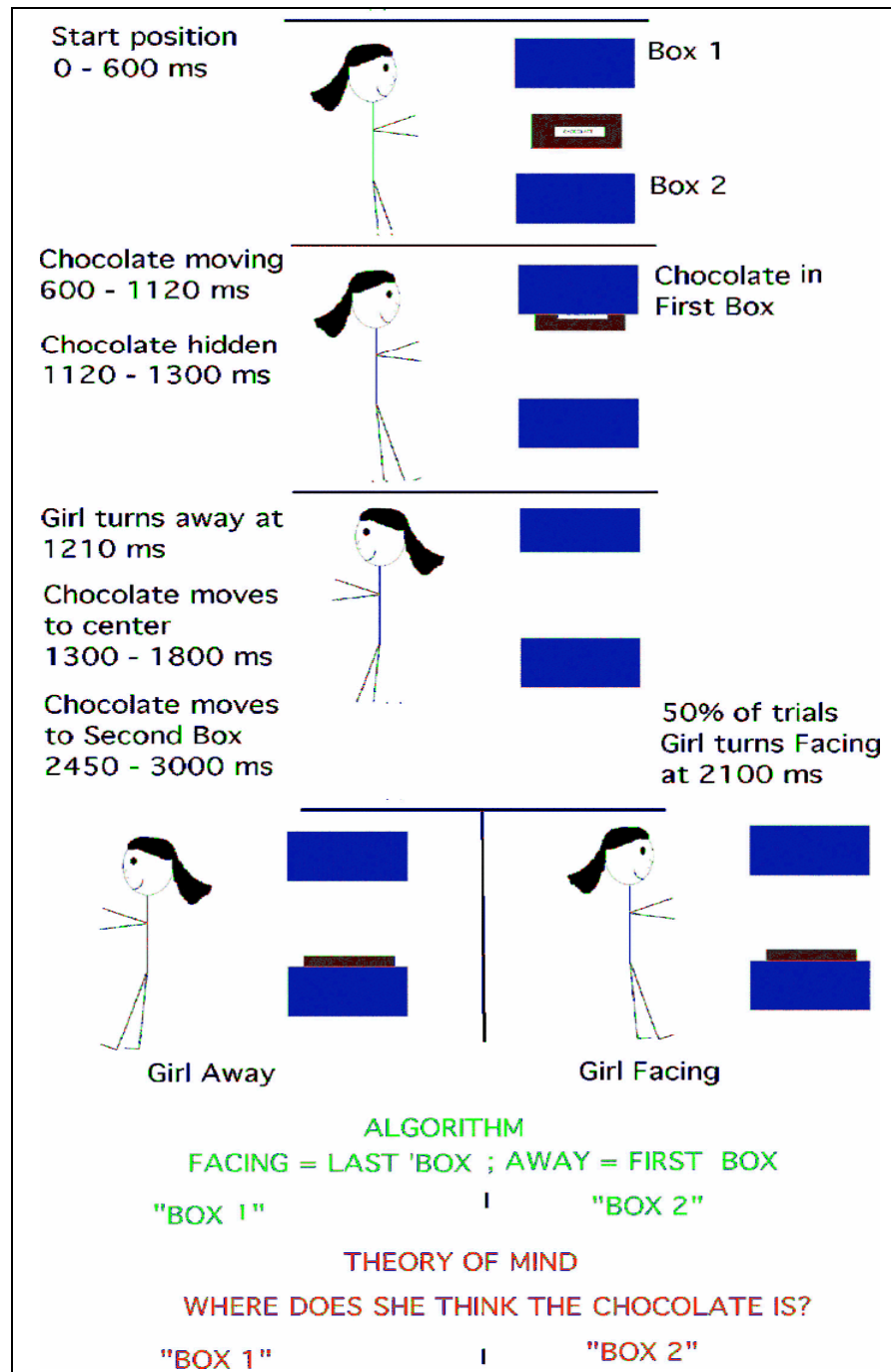


Figure 4.4. Outline of the response selection task and the ToM task in Saxe et al.'s (2006) main experiment. Both tasks used the exact same visual stimuli. The stimuli presented showed a girl either facing or looking away as a chocolate bar was moved from one box (known to the girl) back to the same location or to the other box. In the ToM task subjects were instructed to press a button corresponding to the box where the girl would think the chocolate bar was. In the response selection task subjects were instructed to press buttons according to one of two rules. Rule one, for example, required subjects to press a button corresponding to the location of the last box containing the chocolate bar, but to do so only if the girl was facing the two boxes. If the girl was not facing the boxes subjects were instructed to press a button corresponding to the first box containing the chocolate bar. Reproduced from Saxe et al. (2006) with permission from Psychology Press.

Saxe et al.'s ROI analysis revealed activity in a common neural network for both conditions consisting of regions such as medial PFC, bilateral parietal sulcus, the anterior cingulate cortex (ACC) and the left TPJ. Only the right TPJ showed significantly higher activation in the ToM condition compared to the algorithm task. According to the authors, this finding may support the claim that belief-reasoning relies on executive control but that it also requires the recruitment of a specific region, namely the right TPJ. This region seems to support the representation of others' beliefs.

With their main experiment, Saxe et al. investigated the recruitment of one aspect of EF in belief-reasoning: response selection. However, response selection is quite different than IC. Unlike IC, response selection does not require the suppression of a prepotent response. Therefore, the main experiment does not yield any valid evidence about possible contributions of inhibitory control to belief-reasoning. Their reference experiment, however, may have tapped IC by requiring subjects to inhibit a prepotent response in the incompatible condition of the reference EF experiment. Unfortunately, the belief tasks and the EF tasks varied largely which makes it hard to deduct hypotheses about the connections between ToM and IC based on these results. Differing results in this experiment, for example, may have been due to the largely differing stimuli used. Also, the use of a lengthy blocked design may not be suited to build up a prepotent response in subjects.

Furthermore, the authors claimed that this result shows that the right TPJ may constitute a highly specialized belief-reasoning module. However, this was done under the assumption that no other cognitive process elicits activity in the same brain region. Although this was the case in their study, other processes that were not investigated here might yield similar patterns of activation.

A study that has shown just that in a single experiment has recently been conducted (Mitchell, 2008). The study attempted to investigate whether the right TPJ is really specific for belief-reasoning. As other studies have suggested, the same region seems to be engaged in attention reorienting as well (Corbetta et al., 2000; Corbetta & Shulman, 2002; Serences et al., 2005). An fMRI study was therefore conducted that investigated both attentional reorienting and belief-reasoning in one single study. An attentional cueing task was used to study attentional reorienting. In this task, subjects had to press a button indicating the location of a visual object. This location was cued in 75% and miscued in 25% of all cases. To study belief-reasoning, the same task as in the reference experiment in Saxe et al. (2006; false-belief versus false photograph) was used. The results revealed activity related to belief-reasoning in the right TPJ, medial PFC, and the precuneus. Attentional reorienting was associated with activity only in the right TPJ region. This activation, however, was virtually identical to right TPJ activity associated with belief-

reasoning. Results from Mitchell's study in the right TPJ superimposed on a standard brain are depicted in figure 4.5.

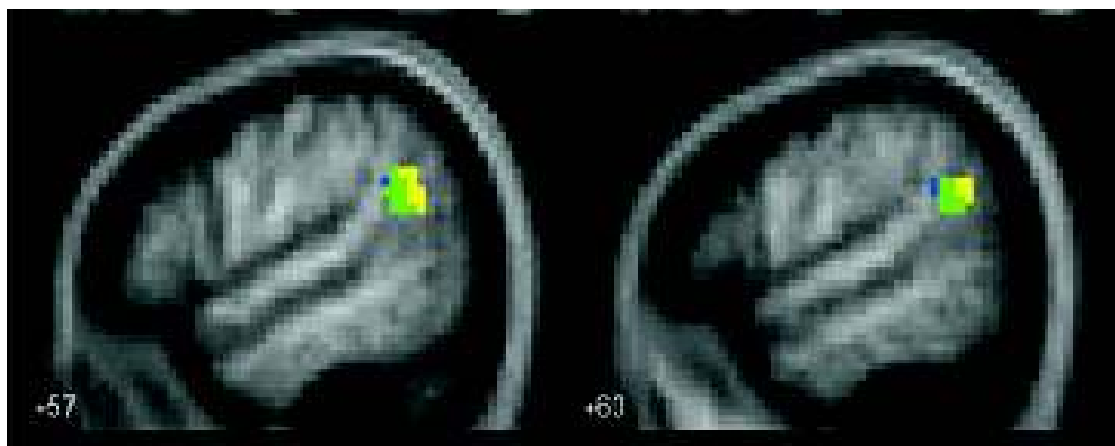


Figure 4.5. Activation in the right TPJ region in a study conducted by Mitchell (2008). Two lateral views at MNI coordinates $x = 57$ and $x = 60$ are shown. Depicted in the color yellow is activity related to belief-reasoning. Blue areas refer to significantly increased activity associated with attentional reorienting as revealed in an attentional cueing task. The green area refers to activity related to both attentional reorienting and belief-reasoning. Reproduced from Mitchell (2008) with permission from Oxford journals.

Based on these results, Mitchell argues that the right TPJ is not specific for ToM reasoning. As another cognitive process shows activity in the same region, the specificity assumption for the right TPJ cannot be upheld. The author then proposes that both attentional reorienting and belief-reasoning rely on common processes that need to be further investigated.

For the reasons stated above, Saxe et al.'s (2006) attempt to prove that the right TPJ constitutes a specific ToM module was clearly flawed. Mitchell (2008) on the other hand has presented solid evidence that other cognitive processes may also engage a region in the right TPJ that is identical to activity elicited by the attribution of beliefs. However, Mitchell's study was unable to find activity in the medial PFC, a region that is frequently activated in response inhibition tasks. Last but not least, visual stimulation between his tasks varied largely, which may also have caused the observed differences in activation.

The study presented in this thesis attempts to account for some of the flaws contained in previous studies attempting to investigate IC and belief-reasoning in a single imaging study. This attempt is outlined in the following paragraph.

4.8 Aim of the Present Study

The ability to attribute beliefs as well as the ability to inhibit a prepotent response are closely related. First hints towards this assumption stemmed from observations that both cognitive abilities share a common developmental timetable in childhood. Behavioral studies in children have then provided solid evidence that both concepts are indeed strongly correlated. IC is the strongest predictor for later false-belief reasoning, far above several other factors such as working memory or verbal ability (Perner & Lang, 1999). Also, training in IC abilities significantly improves false-belief performance in young children and vice versa (Kloo & Perner, 2003). A close behavioral relation between belief-reasoning and IC has been found across many different cultures (Chasiotis et al., 2007; Sabbagh et al., 2006). Furthermore, there is evidence that this close relationship persists throughout adulthood and old age. Compromised belief-reasoning in old age seems to be related to weakening IC functions (German & Hehman, 2006; McKinnon & Moscovitch, 2007). Further evidence indicating a strong connection between the attribution of beliefs and IC comes from developmental disorders with an impairment in one or both concepts. In autism, for example, a frequently observed disability to infer others' mental states is accompanied with compromised EF abilities (Pennington & Ozonoff, 1996). Compromised IC is also one of the cardinal symptoms in ADHD. However, belief-reasoning abilities in this disorder are widely intact, suggesting a complex interaction between belief-reasoning and IC (Sodian et al., 2003).

Functional imaging and lesion studies have corroborated the view of a close relationship between IC and belief-reasoning. Lesions in the medial PFC result in an impairment of both belief attribution and IC (Rowe et al., 2001). Lesions in the left TPJ, however, seem only to have an effect on belief-reasoning (Samson et al., 2004). Functional imaging of both processes indicates an overlap of associated brain regions in medial PFC and the bilateral TPJ. This observation, however, is based on largely differing studies that have for the most part failed to scrutinize both cognitive processes in a single study. Only the results of such combined studies are suitable to test the various hypotheses regarding the nature of the belief-reasoning / IC connection. Such studies would also help to clarify whether some brain regions may qualify as specialized belief-reasoning modules. Although two previous studies have attempted such a methodological approach, only one of them (Mitchell, 2008) may have actually tapped IC. Results of this study indicate that the right TPJ, a former candidate region for a belief-reasoning module, is also engaged in other processes independent of mental state attribution. Nevertheless, this study used stimuli that differed largely across the different conditions. It is undeniable that the activation found in this study may have been due to differing stimulus properties.

Also, Mitchell's study was unable to show IC-related activity in the medial PFC, a region frequently reported as a neural correlate of response inhibition.

The study presented in this thesis will therefore try to further clarify the connection between belief-reasoning and IC. This is done by investigating both concepts in a single study by means of a within-subjects design. Also, this study attempts to account for methodological flaws that have become apparent in previous studies. Stimuli will be used this time that only differ marginally between conditions in terms of visual properties. Also, the most commonly used task for response inhibition, a Go / No-go task, is taken to assess IC. In order to measure neural activity related to belief-reasoning, a false-belief task will be used. The false-belief task is the crucial task to measure the ability to attribute beliefs. Results from this study can help clarify the nature of the belief-reasoning / IC relationship in adulthood. However, the results stemming from this study are unable to make any assumption about this relationship in childhood or how either of these concepts influences the other during their emergence in childhood. A possible overlap between brain activity related to the two cognitive processes can indicate how closely these are related to each other. Largely overlapping areas of activation, for instance, could indicate that the ability to inhibit a response is possibly an essential component of belief-reasoning or even vice versa. Regions that are explicitly activated by belief-reasoning on the other hand could qualify as candidate regions for a belief-reasoning module. This of course would also entail clarifying the role of the right TPJ as a hypothesized belief-reasoning module. While some researchers claim that this region does indeed constitute a specific area for belief-reasoning, others have claimed that this exact region also supports other cognitive processes. Finding concurring activity for both IC and belief-reasoning in the right TPJ would contradict accounts of the right TPJ serving a specialized area for ToM reasoning.

Last but not least, this study's results could also help replicate findings of a previous belief-reasoning study that had used a false-belief task and compared it to a true-belief task (Sommer et al., 2007). Both of Sommer et al.'s tasks were modelled according to the Sally-Anne paradigm. An almost identical approach was used in this study.

Based on previous studies investigating belief-reasoning and IC, I suggest the following hypotheses for the present study:

- Increased brain activity in the false-belief condition compared to the true-belief condition is likely to be found in the following areas: bilateral TPJ, medial PFC. This hypothesis is based on a large array of imaging studies that have found concurring activity in these areas.

- Neural correlates of IC (as reflected in the contrast *No-go* > *Go*) are likely to be located in a largely right-lateralized network consisting of the DLPFC, the bilateral TPJ and the medial PFC. Based on previous results it is suggested that medial PFC activity related to false-belief reasoning is slightly anterior to IC-related medial PFC activity.

- Resulting overlapping activity for IC and belief-reasoning is expected in areas such as the right TPJ and possibly parts of the medial PFC. This would indicate that basic processes exist that support both belief-reasoning and IC.

- In addition, it seems unlikely that the right TPJ could only be implicated in this study's belief-reasoning condition. Based on results from previous studies, the right TPJ is likely to be implicated in this study's IC condition as well. This could refute accounts of the right TPJ as a specialized ToM module.

- No hypotheses can be made at this point about the role of the left TPJ in both belief-reasoning and IC. Although lesion studies point to a crucial role of the left TPJ in belief-reasoning, only few imaging studies have reported activity in this area. Also, accounts of left TPJ activity during IC tasks are inconsistent across the studies reviewed here. Thus, the results from this thesis may clarify the role of the left TPJ in both IC and the attribution of beliefs.

5. Methods

The following methods section begins with a brief description of fMRI, the imaging technique used in this study. The subject characteristics and the experimental design will be specified next. This is followed by a description of the statistical analysis of behavioral and functional imaging data.

5.1 FMRI imaging

Functional Magnetic Resonance Imaging (fMRI) as used in the present study is based on the measurement of an electro-magnetic field which can be measured in all atomic particles with an odd number of protons (Goebel & Kriegeskorte, 2004). This magnetic field is induced by protons revolving around their axis, a property that has also been dubbed “spin”. A common element in biological tissue that possesses only one proton is hydrogen.

By applying a strong static magnetic field to the tissue in the scanner, these spins are aligned on a common axis. In the present study a magnetic field of 3 Teslas was used, which is equal to about the 60,000 fold amount of the earth’s magnetic field.

The application of another temporary electro-magnetic pulse at radio-frequency causes a slight tilting of the proton’s circular movement. This type of movement is referred to as “precession”. All protons now circulate around a common axis. Then the previously applied short electro-magnetic pulse is halted. This causes a decay of the former magnetic field which was previously aligned transversally towards the static magnetic field. This mechanism is referred to as transverse (T2) relaxation. The time for the decay of the transverse magnetic field is shorter in biological tissue compared to most non-biological tissue. Therefore, the time for transverse relaxation in biological tissue has also been called T2*. The time for transverse magnetic field decay due to the on- and offset of an electro-magnetic stimulus is influenced by the type of tissue or molecules contained in a certain area. The same is true for the brain. Hemoglobin as a crucial component of the blood that is able to bind to oxygen and release it to the nerve cells as neural activity demands energy in the form of oxygen and glucose.

Interestingly, neural activity in nerve cells is related to an increase in oxygenated hemoglobin that takes place several seconds after the neural activity. This may be regarded as an anticipatory mechanism to prepare the affected nerve cell for further activity. Thus, active nerve cells in the brain show an increase of oxygenated hemoglobin that is also depicted in a characteristic T2* time that differs significantly from the T2* time of de-oxygenated hemoglobin. This mechanism leads to a specific signal that can be

detected after the offset of the previously described radio-frequency pulse and yields information about the underlying ratio of de-oxygenated and oxygenated hemoglobin in the brain tissue investigated. This effect is referred to as the Blood-Oxygen-Level-Dependant (BOLD) effect and was first discovered by Ogawa, Lee, Kay, & Tank (1990). A combined study including local field potential measurements, single- and multi unit spiking activity measurements and BOLD-dependant fMRI showed that the BOLD signal serves as a reliable estimate of underlying neural activity elicited by a stimulus (Logothetis et al., 2001). The BOLD response shows a specific time course that is also depicted in figure 5.1. After a stimulus has been presented to a subject, this will cause a so called “initial dip” of the measured magnetic signal in affected brain regions after some 2 seconds. Unfortunately, this dip cannot be measured properly in fMRI as of yet. However, the following positive BOLD response that is related to an increase of oxygenated hemoglobin can be measured. The signal obtained reaches its maximum some 6 to 8 seconds after stimulus onset. After about 12 seconds the BOLD response will go into a so-called “undershoot” and retreat back to its baseline.

In scientific fMRI experiments, the BOLD response measured is then associated with its respective experimental condition (also referred to as “regressor”) which is specified as a series of time points with specific durations. Comparisons in terms of neural activity between conditions can now be conducted statistically by comparing regressor-related estimates of the BOLD response.

The statistic analysis of the current study’s imaging data is described in more detail in chapter 5.6.2.2.

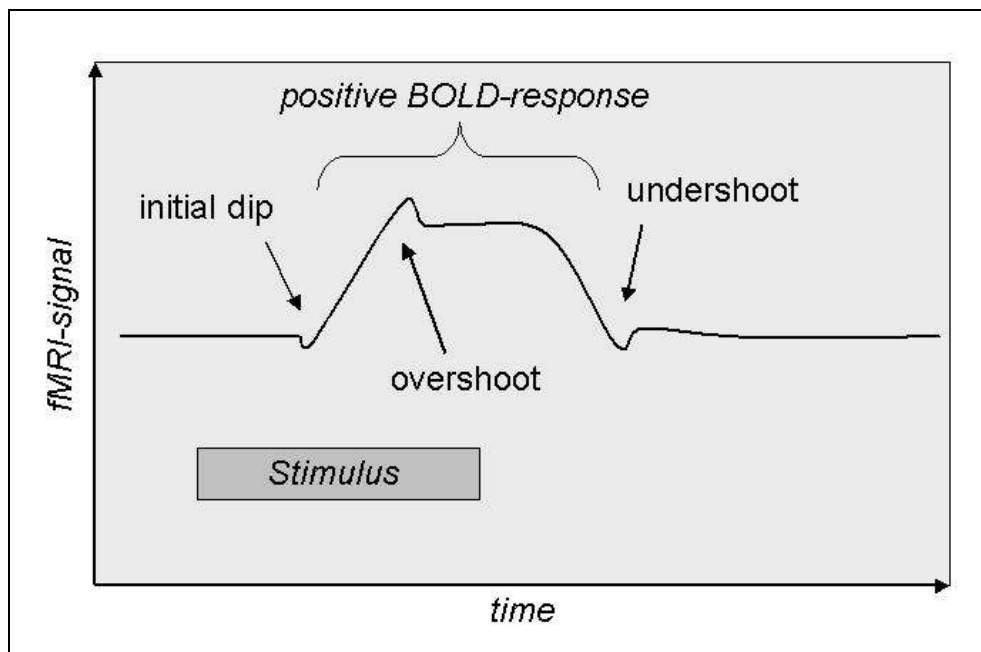


Figure 5.1. Blood-Oxygen-Level-Dependant (BOLD) response as measured in fMRI. After an “initial dip” some 2 seconds after stimulus onset, the BOLD response reaches its peak after about 6 to 8 seconds. This overshoot related to an increase of oxygenated hemoglobin is called “positive BOLD-response”. Before retreating back to its baseline after approximately 12 seconds, the BOLD response briefly dips below baseline (“undershoot”).

The following section will describe the subject sample and properties of the experimental design, followed by detailed specifications about the analysis of obtained data.

5.2 Pilot Experiment

Prior to the actual experiment, a pilot experiment was conducted. In this pilot experiment 9 subjects ($M = 23.8$ years; range 23-24; 3 male) conducted several experiments in a 3-Tesla Siemens Allegra Head Scanner (Siemens Inc., Erlangen, Germany). The same scanner was also used in the subsequent main experiment. The purpose of the pilot experiment was to test whether the paradigms used to measure IC-related and belief reasoning-related would result in a BOLD signal that would be sufficient for further comparisons. During the pilot experiment in the fMRI scanner, different picture presentation durations, different inter-trial interval lengths as well as differing numbers and arrays of pictures were presented in order to find an optimal paradigm measuring the underlying BOLD-response. In the subsequent analysis of the data from the pilot experiment, different analysis approaches of data analysis were used in order to find the

optimal analysis for rendering statistical estimates of the BOLD signal attributable to observed behavioral phenomena.

The different approaches testing for both optimal stimuli presentation as well as optimal statistical analysis thereof resulted in an experimental design and analysis that was subsequently used in the main experiment. The properties of this experimental design as well as the statistical analysis used in the main experiment are presented in the following chapters of this methods section. None of the subject data from the pilot experiment were used in the analysis presented in the following sections.

5.3 Subjects

Twelve subjects ($M = 23.7$ years, range 23-24; 5 male) with no reported history of neurological or psychiatric impairment participated in the study. All subjects had either normal or corrected to normal vision. All participants were right-handed. Subjects were recruited from the University of Regensburg student body. A rough estimate of general intelligence was assessed with the German version of the culture fair intelligence test (CFT-1; Cattell et al., 1997). Mean intelligence as measured by the CFT-1 in the subject group was 117.73 ($SD = 10.67$). All participants gave their written informed consent for participation (appendix A).

5.4 Experimental Design

Every subject conducted both a belief-reasoning experiment and an IC experiment. Both experiments were administered within a single session in the fMRI scanner.

Prior to the actual experiments in the scanner, subjects were shown a standardized paper instruction about the upcoming tasks (see appendix B). Also, all subjects received a training session on a standard personal computer outside the scanner that consisted of 12 belief-reasoning trials and 20 IC trials.

Half of the subjects were then randomly assigned to start with the belief-reasoning experiment, the other half started with the IC experiment first. fMRI scanning was not interrupted between the first and the second experiment in order to prevent inhomogeneities of the scanner's magnetic field. Between the two experimental tasks the subjects were shown a fixation cross at the center of the screen for 40 seconds. This was followed by a 12-second instruction for the subsequent task. After completion of functional T2* image acquisition, a structural image of the subject's brain was obtained.

Entire scanning time including functional and structural image acquisition for every subject was approximately 32 minutes (belief-reasoning experiment: 10 min., IC experiment: 15 min., structural image acquisition: 7 min.).

After completing the experiment, subjects were administered the CFT-1 paper-pencil test as a rough measure of general intelligence abilities.

The Belief-Reasoning experiment and the IC experiment are described in more detail in the following two paragraphs.

5.4.1 Belief-Reasoning Experiment

In each trial, we presented a non-verbal cartoon story consisting of 3 pictures each. Each story showed a protagonist acting on the basis of a correct (true-belief) or incorrect (false-belief) representation of reality.

A total of 10 different story plots (consisting of differing protagonists, objects, and hiding places) were used in this experiment.

Stories were modelled according to the Sally-Anne paradigm (Baron-Cohen et al., 1985).

The first picture showed two children in a room with one of the children putting an object into an empty container such as box or a basket. The second picture showed that one of the children had either left the room (false-belief condition) or had remained in the room (true-belief condition) with the other child transferring the previous object from one hiding place to the other. In the third picture the child that had either left (false-belief condition) or stayed in the room (true-belief condition) was shown searching for the object in one of the two hiding places present in the room. Subjects were instructed to press a button with the left index finger of their right hand if they had expected the behavior of the child or to press a different button with the right middle finger if they had not expected the child's behavior. Cartoon characters showed an expected behavior in 50% of all trials. We instructed the subjects to press the button as accurately and as quickly as possible after the onset of the third picture. Subjects were instructed to do so within the presentation time of the third picture.

Every picture appeared on the screen for 2000ms. The fixation period between trials varied between 2000 and 3000ms in order to jitter the trial onset time, thus rendering an average trial length of 8500ms. The stimulus shown during the fixation period was made up out of a scrambled story picture in order to prevent changes in brain activity that could be attributed to differing visual properties of the fixation stimulus compared to the visual stimuli of the actual story (e.g., contrast, luminance).

A schematic outline of both belief-reasoning conditions is depicted in figure 5.2.

Thirty-five true-belief (TB) trials and 35 false-belief (FB) trials were randomly presented within the belief task. Thus, the belief-reasoning experiment lasted approximately 10 minutes.

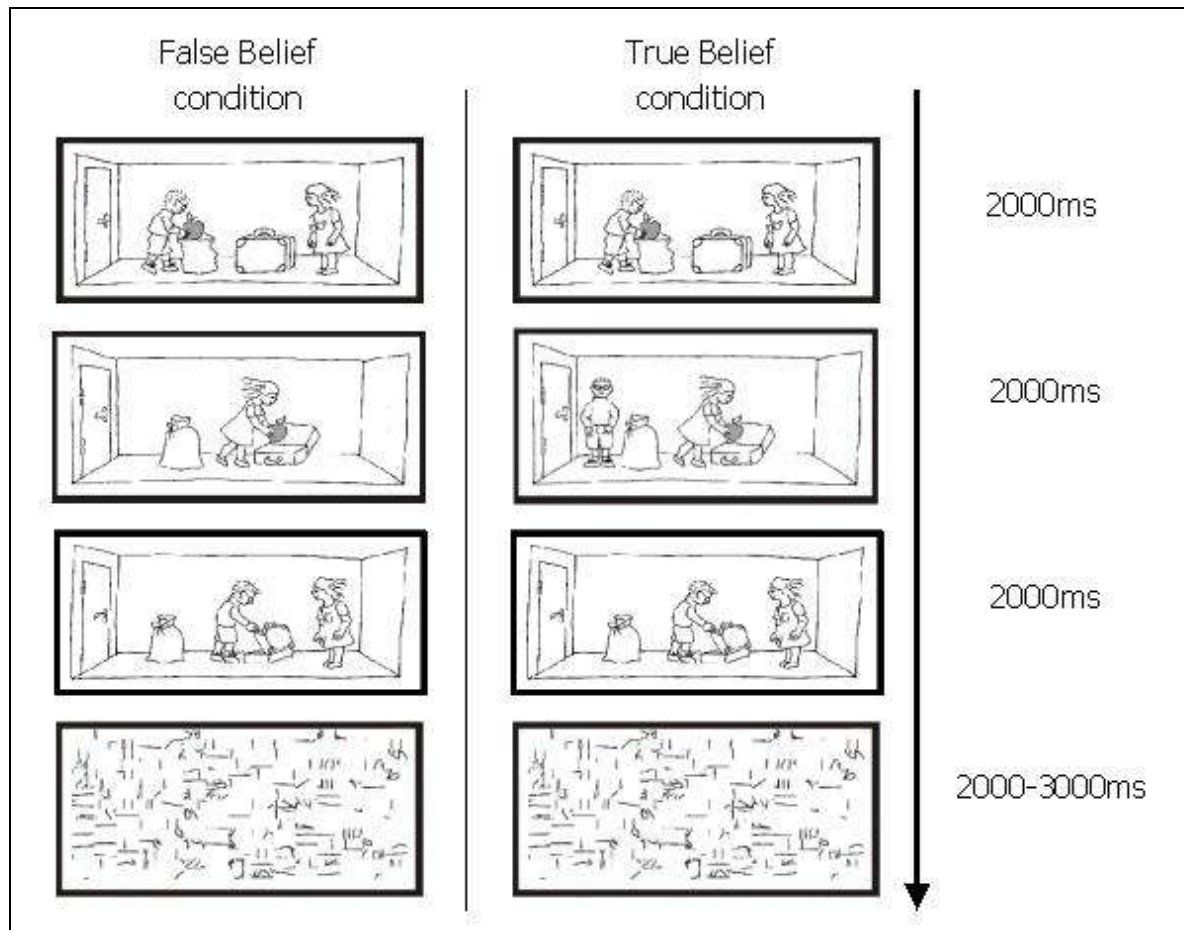


Figure 5.2. Course of the false-belief and the true-belief condition of the belief-reasoning experiment with respective durations. Subjects were instructed to indicate with the press of a button whether they had expected the person who either left or stayed in the room to perform the behavior shown in the third picture.

5.4.2 Inhibitory Control Experiment

As an IC task, a classical Go / No-go paradigm was used. Subjects were shown the identical pictures that were also used as the second picture in the belief-reasoning task (thus showing one or two children in a room). Pictures were randomly drawn from 10 different story plots, thus differing in terms of story characters, objects, and hiding places. Pictures were separated by a fixation period (consisting of a scrambled story picture). Subjects were instructed to press a button with the index finger of their right hand

whenever the number of children in the present picture differed from the number of children in the previous picture (Go condition). Subjects were told to refrain from pressing a button if the number of children shown in the picture was the same as in the previous picture (No-go condition). Stimuli were presented for 1000ms each with a fixation period between pictures that varied between 3000 and 4000ms in order to jitter the trial onset time. Thus, average trial length was 4500ms. A schematic outline of the IC experiment is shown in figure 5.3.

A total of 160 Go trials and 40 No-go trials were randomly presented. Hence, the ratio between Go and No-go trials was 4:1. This was done in order to build up a prepotent tendency to respond with a button press to No-go trials as well. The IC experiment lasted about 15 minutes.

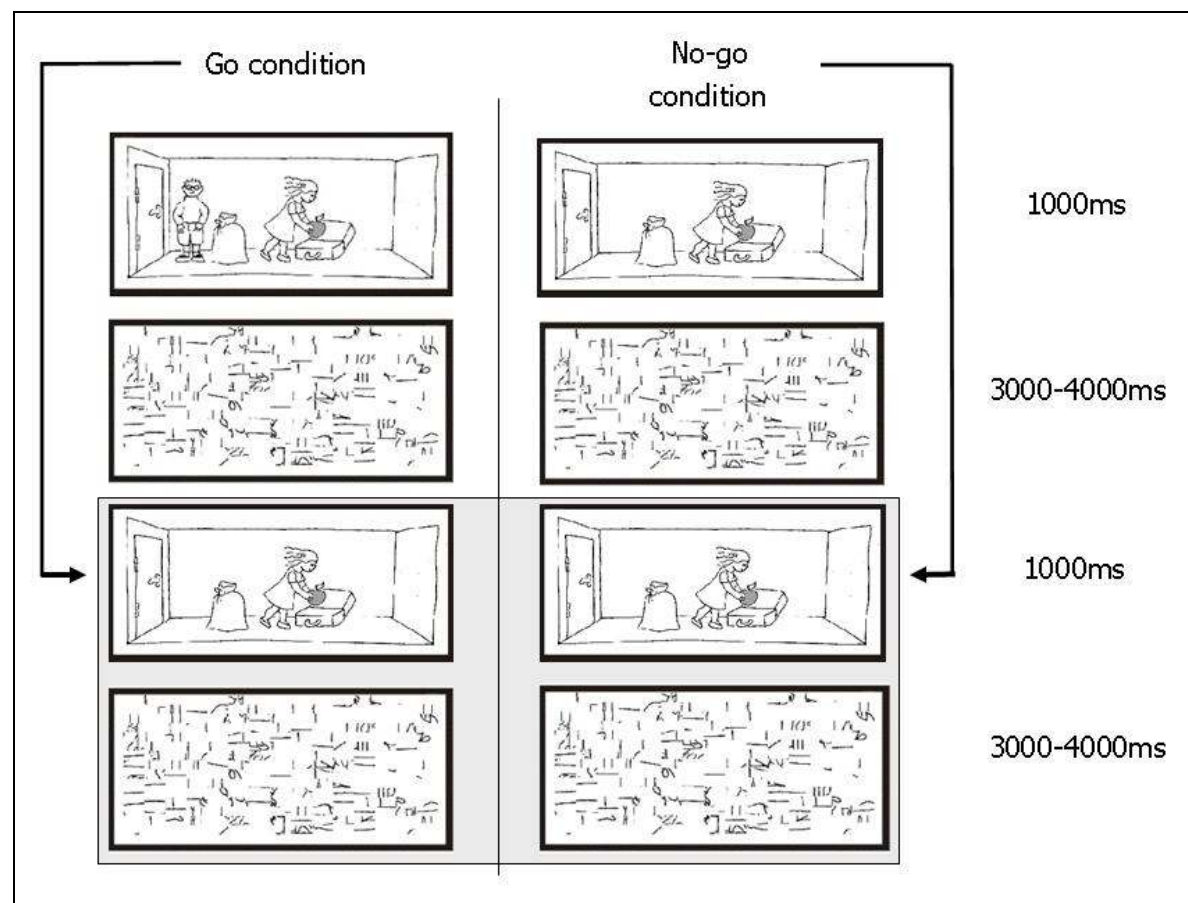


Figure 5.3. The two conditions of the inhibitory control experiment (inside the grey frame) and respective stimulus durations. The Go condition showed a number of children different from the previous picture and required subjects to press a button with their right index finger. In the No-go condition the number of children in the present picture was the same as in the previous picture. Subjects were instructed to withhold a response in this condition.

Both the belief-reasoning and the IC experiment were generated with Presentation 11.3 software (Neurobehavioral Systems Inc., Albany, CA). Behavioral data such as reaction times and accuracy were also recorded with Presentation 11.3. Inside the scanner, responses were retrieved with a LUMItouch optical response device (Photon Control, Burnaby, Canada). Stimuli were back-projected onto a mirror inside the scanner with a D-ILA LCD Projector (JVC Corp., Yokohama, Japan).

5.5 FMRI Set Up and Design

A 3-Tesla Siemens Allegra Head Scanner (Siemens Inc., Erlangen, Germany) was used to record imaging data. The scanner acquired echo-planar-imaging (EPI) sequences using fast gradients. During T2* data acquisition, 32 slices (whole brain; slice thickness = 3mm, no skip) were recorded in interleaved order with a Time-to-Repeat (TR) of 2000ms, a Time-to-Echo (TE) of 30ms, a flip angle of 90°, a Field-of-View (FoV) of 192 x 192 mm and a voxel size of 3 x 3 x 3 mm. A total of 790 images were recorded in the entire experiment.

A structural image was recorded from every subject at the end of functional data acquisition. These T1-weighted images were obtained using an MPRAGE (*Magnetization Prepared Rapid Acquisition Gradient Echo*) pulse sequence (TR = 2250ms, TE = 3.93ms, flip angle = 9°, FoV = 256 x 256 mm) scanning 160 slices with isotropic voxels of 1 x 1 x 1 mm.

5.6 Data Analysis

5.6.1 Behavioral Data

Accuracies and response times for performance in the belief-reasoning experiment were compared statistically between conditions by using paired t-tests at a statistical significance level of $p = .05$. SPSS 16 (SPSS Corp., Chicago, IL) software was used for the latter analyses. As the No-go condition of the IC experiment did not require a button press, no statistical comparisons between the Go and the No-go condition were conducted.

5.6.2 Functional Data

All images obtained in the scanner were preprocessed and analyzed statistically with SPM5 software (<http://www.fil.ion.ucl.ac.uk/spm/software/spm5/>), a free program that is based on Matlab 7.0 software (The MathWorks, Natick, MA).

Preprocessing and analyzing steps will be described next.

5.6.2.1 Preprocessing

a) DICOM to NIFTI Transformation

Functional and structural imaging data in the scanner were recorded in DICOM (*Digital Imaging and Communications in Medicine*) format. In order to preprocess and analyze these data in a format recognizable by SPM5, these data were transformed to NIfTI (*Neuroimaging Informatics Technology Initiative*) format.

b) Slice Timing

Next, all functional data were slice timed. This routine accounts for the fact that an entire scan of the brain takes several seconds to record (time-to-repeat (TR); 2 seconds in the present study). This leads to the problem that all picture slices recorded within this time-frame are measured at differing time points. As slices were recorded in an interleaved fashion (meaning that slices with an even number were recorded first), adjacent slices were measured with a time difference of $TR / 2$, which equals 1000ms in the present case. However, each slice needs to correspond to the same point in time for further analysis. Thus, the slice timing routine slightly shifts the original data. This results in new data estimates simulating that all slices were recorded at the same time.

c) Realignment

This routine accounts for motion artefacts. Functional images were realigned to one functional image. To achieve this, a least squares approach was used and images were spatially transformed using 3 parameters for translation and 3 parameters for rotation. These parameters were acquired during the scanning session. A functional mean image was written in this routine for further use in the next preprocessing routine.

d) Coregistration

In this preprocessing step, parameters were estimated that are needed to realign the functional and the structural images. To facilitate this procedure, the necessary parameters were estimated for realigning the structural image to the previously estimated functional mean image.

e) Normalization

Structural brain properties differ largely across subjects. In order to introduce single-subject data into a group analysis, the individual imaging data have to be altered for comparisons across subjects. In the present study, functional images were normalized into a standard space as defined by a template structural brain image. Here, a standard template brain ("T1.mni") contained in SPM5 was used. This template represents the average of 152 structural brain images from different subjects as recorded at the Montreal Neurological Institute (MNI). Also in this step, functional images were re-written to voxels measuring 2 x 2 x 2 mm.

f) Smoothing

In this routine, voxels were spatially smoothed by applying a Gaussian kernel of a specified width. In the present study a full-width half maximum (FWHM) of 8 mm was used. Smoothing increases the signal-to-noise ratio (SNR) and accounts for inter-individual functional and anatomical differences.

5.6.2.2 Statistical Analysis

Events of Interest / Design Matrix

In this first step after preprocessing, time intervals (regressors) of particular interest to the study were specified first. This was conducted on single-subject level. A so-called event-related design was used in this study. Unlike block-designed studies, this approach investigates BOLD-signal changes related to relatively short time-periods. In the present study, events of interest lasted for 1000ms. So called events of no interest were also included in the analysis. These events within the experiment were of no particular interest for later comparisons but improve the signal to noise ratio by accounting for variance related to the underlying BOLD-signal.

Events of interest as well as events of no interest were then synchronized with corresponding functional images. Also, 6 motion parameters for translation and rotation were included in the statistical analysis for further explanation of variance. All of the above information made up the so-called design matrix.

In the belief-reasoning experiment, trials for the false-belief and the true-belief condition were separated. In order to capture false- and true-belief reasoning, a regressor

lasting 1000ms was defined as starting 500ms after the onset of the second picture. In the false-belief condition, this second picture showed a child transferring an object with the other child having left the room. In the true-belief condition, the second picture showed a child transferring an object with the other child watching. All 70 false-belief and true-belief trials were introduced into the subsequent statistical analysis.

For the IC experiment, Go and No-go trials were modelled separately. In addition, only correct trials in both conditions were introduced into the analysis for further comparisons. However, missing Go trials and No-go false alarms were also included separately in the design matrix as events of no interest. In both the Go and the No-go condition, the 1000ms presentation time of the picture stimuli was specified as a regressor.

Regressors from the belief-reasoning experiment and the IC experiment were all introduced into the same statistical model.

General Linear Model-Estimation

Statistical significance on single subject level for the previously compiled model was subsequently estimated. This was done by means of multiple t-tests based on the assumptions of the General Linear Model (GLM). In the GLM the equation $Y = X * \beta + e$ is assumed for every voxel. Y stands for the imaging data obtained, X stands for the specifications of the design matrix and β stands for to be determined parameters (thus the BOLD signal change attributed to the experimental paradigm). The variable e stands for errors that cannot be accounted for by the other variables X and β .

Subsequent results were then computed based on the β s estimated according to the GLM.

Results

After the estimation of the β parameters on single subject level, these results were introduced into a group analysis. For this purpose a random-effects analysis was conducted. This type of group-analysis accounts for inter-individual anatomical and functional variability. As such, a random effects analysis is more conservative than the analysis of group data in a fixed-effects analysis. In fixed-effects analyses, variability between subjects is ignored, therefore no inferences beyond the specific subjects at the specific time of data acquisition may be drawn. Because a random-effects analysis was conducted here, it allows for inferences beyond the specific subject sample investigated in this study.

In the results routine, time intervals (regressors) that were previously specified in the design matrix were now contrasted with other regressors of interest. One-tailed t-tests were utilized for comparisons between conditions. A cluster defining threshold of $T > 5$ (equal to $p \leq .0001$) was set for the comparisons *True-Belief > Baseline* and *False-Belief > Baseline*. This conservative threshold was chosen because of the expected strong activations typically revealed in comparisons against baseline. For the comparisons *False-Belief > True-Belief* and *No-Go > Go* a cluster-defining threshold of $T > 4$ (equals $p \leq .001$) was defined. Significant activation of a cluster was reported if the corrected p-value of a cluster did not surpass .05.

Significant results were ultimately displayed on template images that are part of the SPM5 software package.

Results from the group analysis will be described in detail in the following section.

6. Results

6.1 Behavioral Results

In both the false-belief and the true-belief condition of the belief-reasoning experiment subjects were asked if they had expected the protagonist's behavior. In the false-belief condition, subjects answered this question correctly in 95.5% (SD = 3.3%) of all trials (33.4 (SD = 1.16) out of 35 trials). In the true-belief condition, average accuracy was 96.4% (SD = 2.1%) or 33.7 (SD = .75) trials. A comparison of accuracy in both conditions by means of a paired t-test revealed no statistically significant difference between the two conditions (t (df=11) = 0.979; p = .349, n.s.).

Average accuracy in the IC experiment was assessed for the conditions Go and No-go. Mean accuracy in the Go condition was 96.3% (SD = 2.1%), equalling an average of 154.1 (SD = 3.3) correct trials out of 160 Go trials presented. In the No-go condition subjects showed a mean accuracy (i.e. refraining from button press) of 89.8% (SD = 12.0%) or 35.9 (SD = 4.8) correct trials out of 40 No-go trials presented in the experiment. There was a statistical tendency for higher accuracy in the Go condition when compared to the No-go condition (t (df=11) = 1.997; p = .071). Accuracy in the belief-reasoning and the IC experiment is shown in figure 6.1.

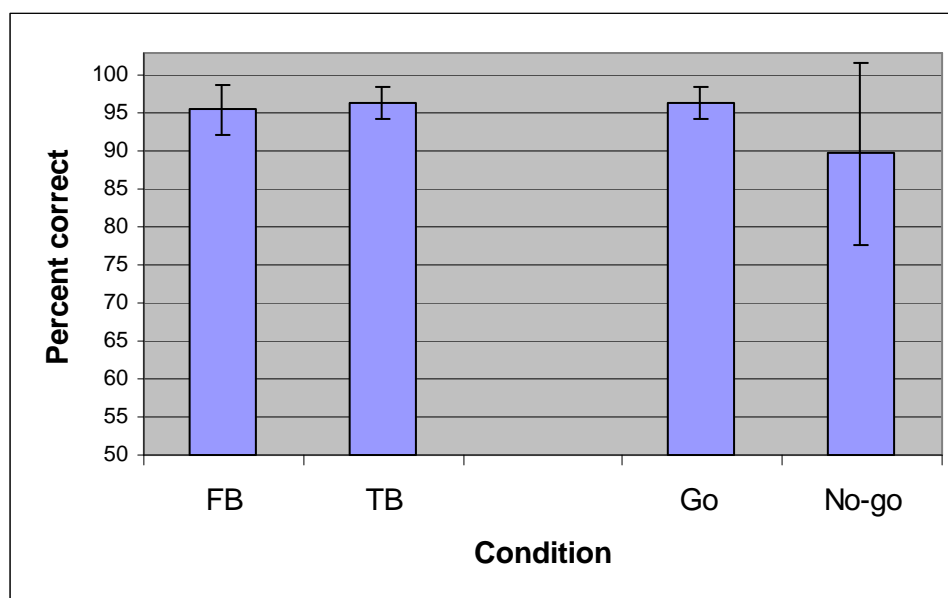


Figure 6.1. Accuracy and corresponding standard deviations in percent of correct trials for the false-belief (FB), the true-belief (TB), the Go, and the No-go condition. Paired t-tests revealed no significant difference in accuracy between the FB and the TB condition. However, there was a statistical tendency towards higher accuracies in the Go condition when compared to the No-go condition (t (df=11) = 1.997; p = .071).

Mean reaction time in the false-belief condition of the belief-reasoning experiment was 754.9ms (SD = 148.9ms). In the true-belief condition, average reaction time across all subjects was 744.8ms (SD = 150.1ms). This difference was not significant as revealed by a paired t-test ($t(df=11) = 1.144$; $p = .277$, n.s.).

In the IC experiment, mean reaction time in correct Go trials was 534.0ms (SD = 64.3ms). No response times were recorded for the No-go trials as they did not require any buttons to be pressed.

Reaction times in the belief-reasoning and the IC experiment are displayed in figure 6.2.

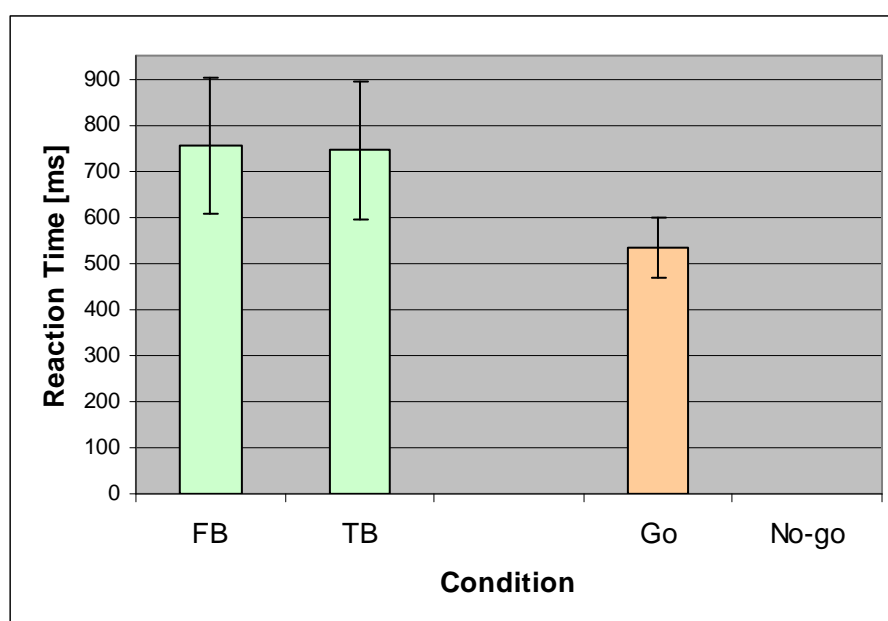


Figure 6.2. This figure depicts reaction times (RTs) and standard deviations for the false-belief (FB) and the true-belief (TB) condition in the belief-reasoning experiment as well as for correct Go trials in the IC experiment. There was no significant difference in RTs between the FB and the TB condition. No RTs were assessed for correct No-go trials since the No-go condition required the withholding of a response.

6.2 Functional Imaging Results

Functional imaging results in the belief-reasoning and the IC experiment are described in more detail in the following section. Also, a graphical comparison of brain areas associated with false-belief reasoning and brain areas involved in IC will be presented. All results presented in this section stem from the random effects group analysis based on the individual results of the 12 subjects participating in this study.

6.2.1 Belief-Reasoning Experiment

6.2.1.1 *True-belief > Baseline*

Baseline activity was subtracted from activity in the true-belief condition in order to reveal brain areas associated with cognitive processes required for solving the true-belief task. Mastery of this task does not require a separation between reality and the protagonist's belief (decoupling). Therefore, this contrast should in theory not contain brain activity specifically dedicated to the decoupling between mentality and reality. Significant activity in this comparison was found in the following areas: bilateral TPJ, bilateral medial frontal gyrus, bilateral superior frontal gyrus, bilateral inferior frontal gyrus, right middle temporal gyrus, bilateral postcentral gyrus, left precentral gyrus, left thalamus and bilateral precuneus. Significant activity in this comparison is reported at a voxel-wise threshold of $T = 5$ and a corrected p-value of $p \leq .05$ (cluster-level). Table 6.1 contains corresponding BAs, cluster sizes, peak MNI coordinates and T-values. Figure 6.3 shows significant brain activity for this contrast on a template brain.

This activity is also depicted in figure 6.5 (green and yellow shading) together with activity in the *False-Belief > Baseline* comparison (red and yellow shading).

Table 6.1. Significantly increased brain activity related to the contrast *True-belief > Baseline*. This table includes information on the name and hemisphere of activated brain regions as well as associated Brodmann areas, cluster sizes in voxels, MNI coordinates, and T-values for each cluster's peak voxel.

True-Belief > Baseline

Hemisphere & region	Brodmann areas	Cluster size in voxels ^a	MNI coordinates			T-values ^b
			x	y	z	
L precentral gyrus, L postcentral gyrus, L supramarginal gyrus, L angular gyrus, L inferior parietal lobule (includes TPJ region)	1, 2, 3, 4, 39, 40	5298	-32	-20	72	15.68
R middle temporal gyrus, R superior temporal gyrus, R middle occipital gyrus (includes TPJ region)	37, 39, 40	2146	60	-40	6	11.27
R postcentral gyrus	1, 2	721	62	-18	44	9.67
R inferior frontal gyrus	44, 45	936	52	8	46	9.20
L thalamus	N/A	233	-12	-14	4	9.16
L inferior frontal gyrus, L precentral gyrus	9, 44, 45	679	-56	0	42	9.13
Bil. superior frontal gyrus, bil. medial frontal gyrus	6, 8	357	0	20	60	8.47
Bil. precuneus	7	586	6	-66	58	7.04

Notes. ^a Voxel sizes amounted to 2 x 2 x 2 mm.

^b T-values are given at a corrected p-value of $p \leq .05$.

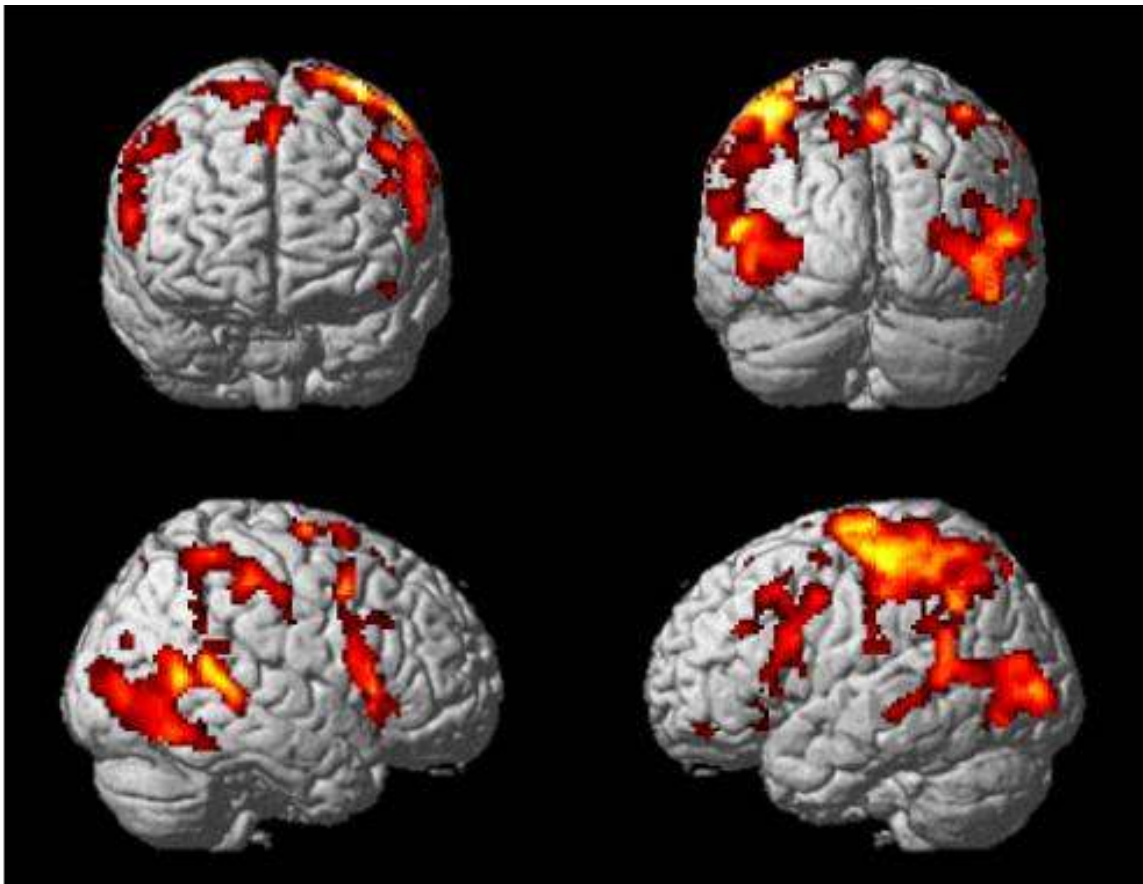


Figure 6.3. Significantly increased brain activity in the true-belief condition when compared to baseline. Activity is shown on a template brain at a voxel-wise threshold of $T = 5$ and a corrected p-value of $p \leq .05$ on cluster-level. A detailed description of activated brain areas can be found in table 6.1.

6.2.1.2 *False-Belief > Baseline*

In this comparison activity related to the baseline of the entire experiment (i.e. average brain activity in the experiment minus activity related to the regressor “false-belief”) was subtracted from activity in the false-belief condition. This was done in order to identify areas related to false-belief reasoning. This includes activity related to a decoupling between reality and a protagonist’s false belief as well as activity related to more general task demands. These general task demands are also likely to be present when solving a true-belief task.

Significant activity in the comparison *False-Belief > Baseline* at a voxel-wise threshold of $T = 5$ and a corrected p-value of $p \leq .05$ on cluster-level was found in the following prominent areas: bilateral TPJ, left superior parietal lobule, bilateral middle occipital gyrus, left precentral gyrus, left postcentral gyrus, right middle and superior temporal gyrus, bilateral middle frontal gyrus, left inferior frontal gyrus, right superior frontal gyrus, and right precuneus.

Corresponding BAs, cluster sizes (in number of voxels), peak MNI coordinates and T-values are specified in table 6.2. Significant brain activity superimposed on a standardized brain template is depicted in figure 6.4. Figure 6.5 shows the same activation in red and yellow shading on a template brain together with significant activity in the comparison *True-Belief > Baseline* (green and yellow shading).

Table 6.2. Brain activity as revealed in the contrast *False-Belief > Baseline*. Given in this chart are the names and hemisphere information for every active brain region as well as corresponding Brodmann areas, cluster sizes, MNI coordinates, and T-values of each cluster's peak voxel.

False-Belief > Baseline

Hemisphere & region	Brodmann areas	Cluster size in voxels ^a	MNI coordinates			T-values ^b
			x	y	z	
L precentral gyrus, L postcentral gyrus, L middle occipital gyrus, L supramarginal gyrus, L angular gyrus, L inferior parietal lobule, L superior parietal lobule (includes TPJ region)	1, 2, 3, 4, 5, 6, 7, 8, 39, 40	8935	-46	-80	16	13.05
R middle temporal gyrus, R superior temporal gyrus, R middle occipital gyrus	22, 37, 39	2599	64	-36	0	13.04
R inferior frontal gyrus, R middle frontal gyrus	9, 44, 45	2095	18	-4	2	12.38
L inferior frontal gyrus, L middle frontal gyrus, L thalamus	9, 44, 45	2730	-44	22	44	9.78
R supramarginal gyrus, R inferior parietal lobule (includes TPJ region)	40	1433	50	-50	32	9.44
R superior frontal gyrus, R middle frontal gyrus	6	136	14	-10	70	8.84
L middle frontal gyrus	10	127	-48	46	-14	8.01

Notes. ^a Voxel sizes amounted to 2 x 2 x 2 mm.

^b T-values are given at a corrected p-value of $p \leq .05$.

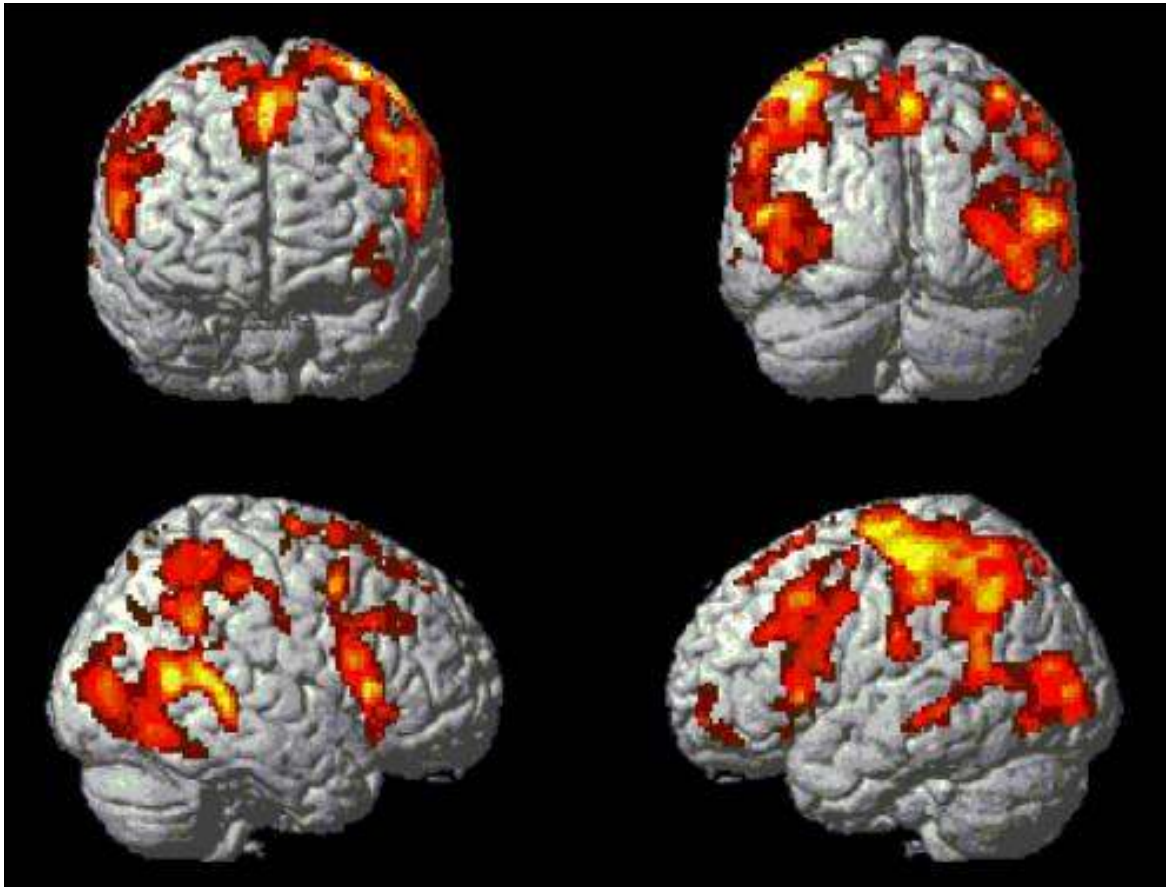


Figure 6.4. Significantly increased brain activity in the false-belief condition when compared to baseline. Activity is shown on a template brain at a voxel-wise threshold of $T = 5$ and a corrected p-value of $p \leq .05$ (cluster-level). For a description of activated brain areas, please refer to table 6.2.

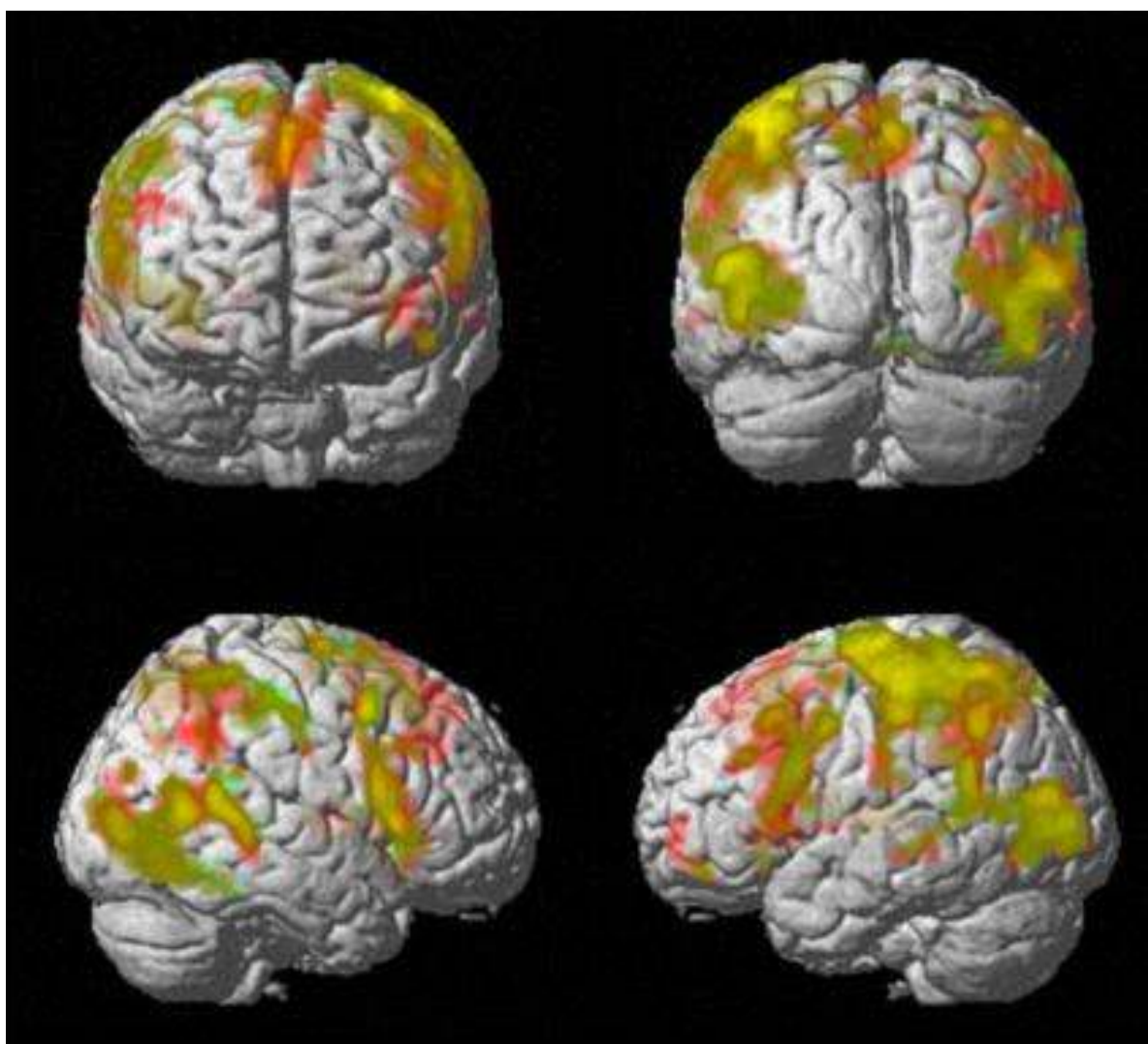


Figure 6.5. Brain activity in the contrast *False-Belief > Baseline* (red shading) and *True-Belief > Baseline* (green shading) at a voxel-wise threshold of $T = 5$ and a corrected p-value of $p \leq .05$ (cluster-level). Areas in yellow shading show activity in both comparisons. A detailed description of active brain areas in the contrasts *False-Belief > Baseline* and *True-Belief > Baseline* can be found in tables 6.1 and 6.2, respectively.

6.2.1.3 *False-Belief > True-Belief*

A direct statistical comparison was conducted in order to identify brain areas related to a mere decoupling between an individual's false belief and the true state of reality. This was achieved by subtracting brain activity associated with the true-belief condition from brain activity related to the false-belief condition. For this comparison a voxel-wise threshold of $T = 4$ and a p-value of $p \leq .05$ on cluster level was used. Significant activity in this comparison was revealed in the bilateral TPJ, bilateral middle frontal gyrus, left precentral gyrus, bilateral medial frontal gyrus, bilateral middle temporal gyrus, left thalamus, bilateral precuneus, and bilateral superior frontal gyrus. BAs, cluster sizes, peak MNI coordinates and T-values corresponding to the above areas are delineated in table 6.3. Figure 6.6 displays significant brain activity for the latter contrast superimposed on a standard brain. Sectional views of significant clusters of activity are shown in the subsequent figure 6.7.

Table 6.3. Brain regions showing significantly increased activity in the contrast *False-Belief > True-Belief*. The table contains the names and hemisphere of every activated cluster, associated Brodmann areas, cluster sizes, MNI coordinates, and T-values for each cluster's peak voxel.

False-Belief > True-Belief

Hemisphere & region	Brodmann areas	Cluster size in voxels ^a	MNI coordinates			T-values ^b
			x	y	z	
L middle frontal gyrus, L precentral gyrus	8, 9	523	-42	22	42	10.04
R middle frontal gyrus	6, 8	129	32	4	52	9.12
R supramarginal gyrus, R inferior parietal lobule (incl. TPJ region)	40	505	52	-50	38	8.92
Bil. medial frontal gyrus	9, 10	269	6	66	22	7.06
R middle temporal gyrus	21	161	66	-36	-2	6.89
L supramarginal gyrus, L inferior parietal lobule (incl. TPJ region)	39, 40	543	-52	-50	32	6.89
L middle frontal gyrus	10	182	-40	52	-4	6.86
L middle temporal gyrus	21	89	-64	-38	-10	6.68
L thalamus	N/A	162	-18	4	14	6.36
Bil. precuneus	7	179	-2	-62	44	6.24
Bil. medial frontal gyrus, bil. superior frontal gyrus	6, 8	352	10	34	46	6.19
R middle frontal gyrus	9, 46	84	52	28	30	5.98
R middle frontal gyrus, R superior frontal gyrus	10	110	30	58	10	5.72

Notes. ^a Voxel sizes amounted to 2 x 2 x 2 mm.

^b T-values are given at a corrected p-value of $p \leq .05$.

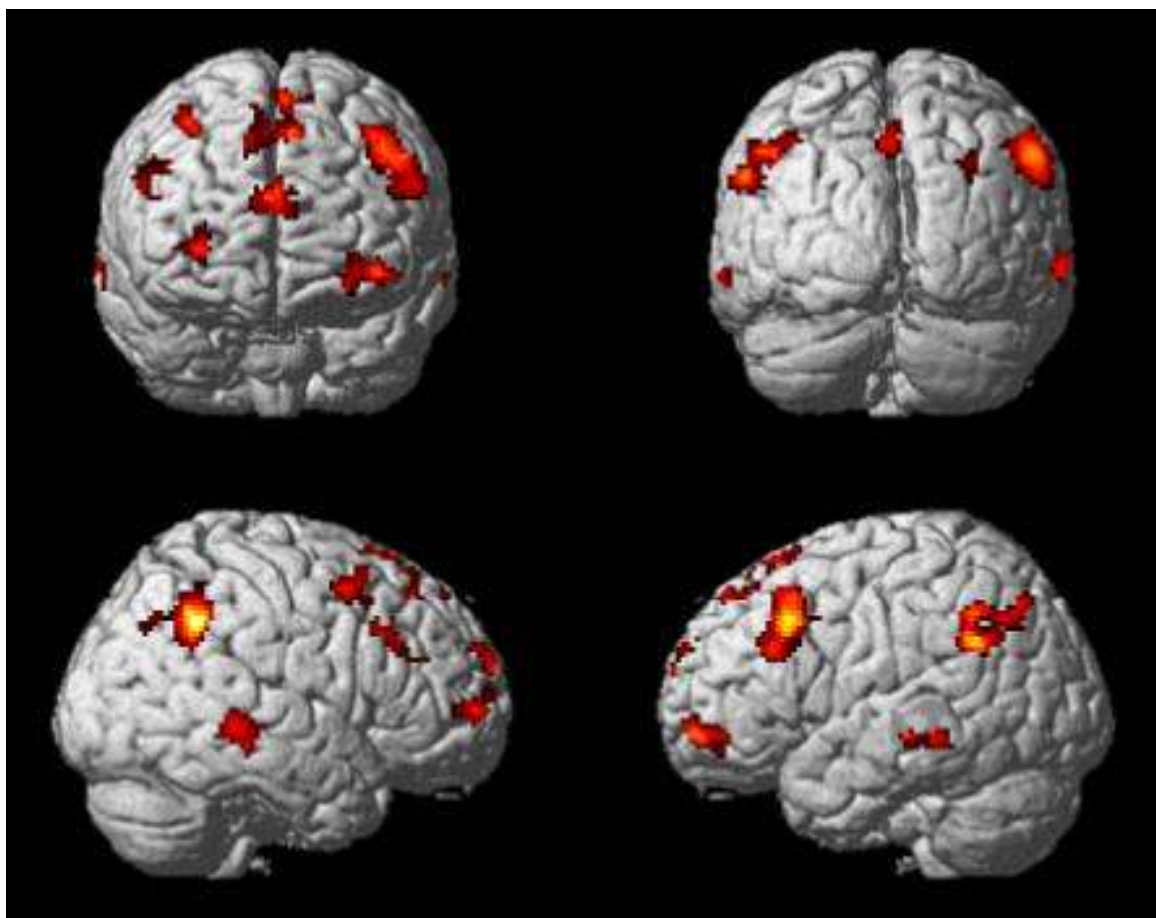


Figure 6.6. Significantly increased brain activity in the false-belief condition when compared to the true-belief condition. Significant activity is shown on a template brain at a voxel-wise threshold of $T = 4$ and a corrected p-value of $p \leq .05$ (cluster-level). Please refer to table 6.3 for a detailed description of activated brain areas.

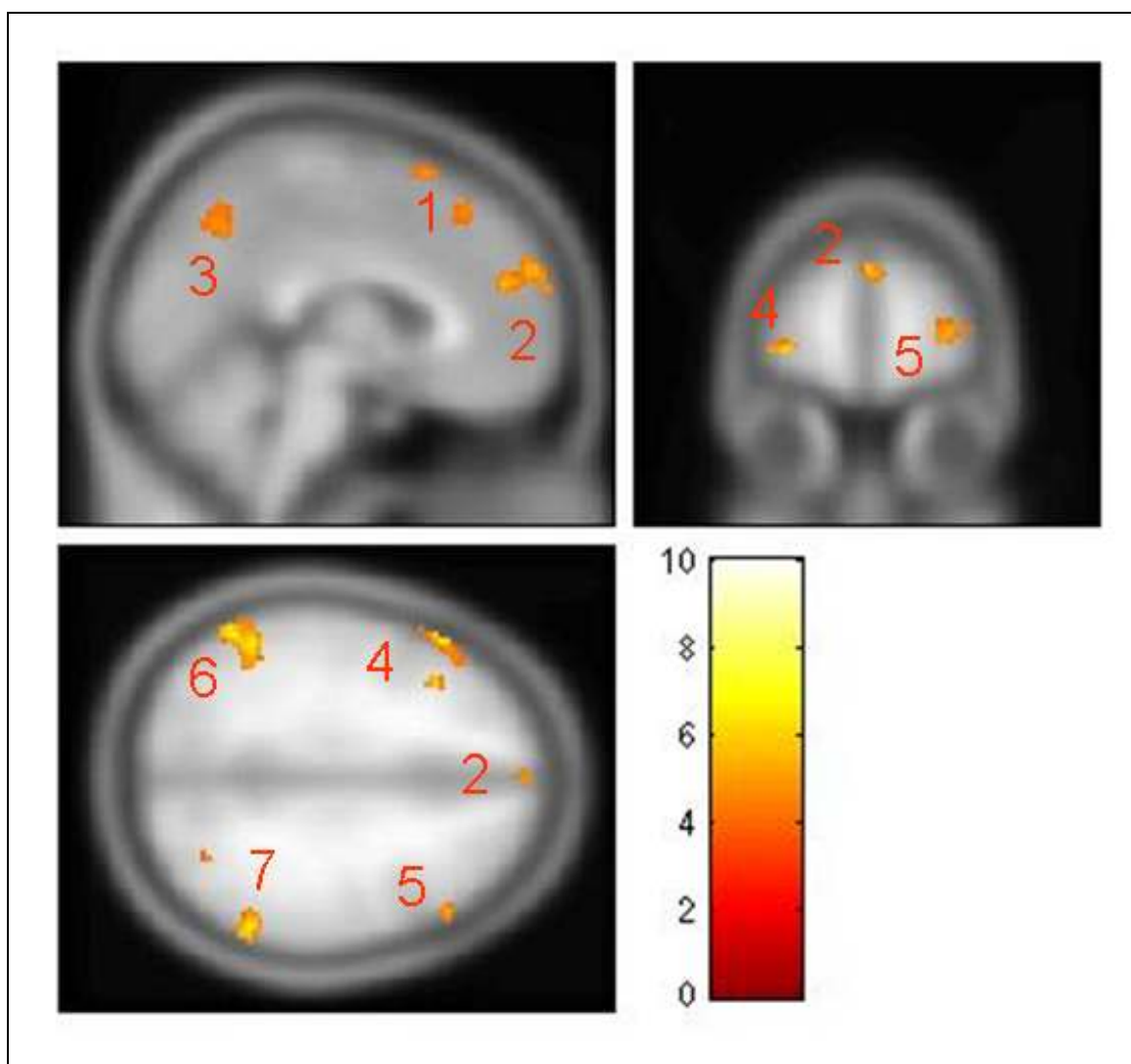


Figure 6.7. A sectional view of significant brain activity in the contrast *False-Belief > True-Belief* at MNI coordinates $x = 3$, $y = 60$ and $z = 35$. Brain activity in this contrast is shown at a voxel-wise threshold of $T = 4$ and a corrected p -value of $p \leq .05$ on cluster-level. Numbers correspond to significantly active brain regions (1 = bilateral medial frontal gyrus and superior frontal gyrus (BA 6, 8), 2 = bilateral medial frontal gyrus (BA 9, 10), 3 = bilateral precuneus, 4 = left middle frontal gyrus, 5 = right middle frontal gyrus, 6 = left TPJ, 7 = right TPJ). The color bar depicts corresponding T-values.

6.2.2 Inhibitory Control Experiment

6.2.2.1 *No-go > Go*

Activity in the condition *Go* was subtracted from activity related to the *No-go* condition in order to isolate brain areas associated with response inhibition. A voxel-wise threshold of $T = 4$ and a p-value of $p \leq .05$ on cluster level was used for this comparison. Brain activity related to IC was found in a largely right-hemispheric network consisting of right inferior and middle frontal gyrus, right middle and superior temporal gyrus, right medial frontal gyrus, right superior frontal gyrus, and the right TPJ. The only significantly activated area in the left hemisphere was the left TPJ. Brain areas as revealed in the contrast *No-go > Go* are described in more detail in table 6.4. Figure 6.8 depicts this IC-related pattern of activity superimposed on a template brain. Significant brain activity in areas below the surface of the brain is displayed in a sectional view in figure 6.9.

Table 6.4. Significantly increased brain activity in the contrast *No-go > Go*. Depicted here are the names and hemisphere of activated brain regions, associated Brodmann areas, cluster sizes in voxels, MNI coordinates, and T-values corresponding to every cluster peak voxel.

No-go > Go

Hemisphere & region	Brodmann areas	Cluster size in voxels ^a	MNI coordinates			T-values ^b
			x	y	z	
R inferior frontal gyrus, R middle frontal gyrus	8, 9, 44, 45, 47	2390	48	20	-4	9.93
L supramarginal gyrus, L inferior parietal lobule (includes TPJ region)	40	138	-62	-56	36	8.78
R middle temporal gyrus, R superior temporal gyrus, R angular gyrus, R supramarginal gyrus, R inferior parietal lobule (includes TPJ region)	21, 22, 40	2664	64	-46	30	8.20
R medial frontal gyrus, R superior frontal gyrus	6, 8	559	6	24	62	6.87

Notes. ^a Voxel sizes amounted to 2 x 2 x 2 mm.

^b T-values are given at a corrected p-value of $p \leq .05$.

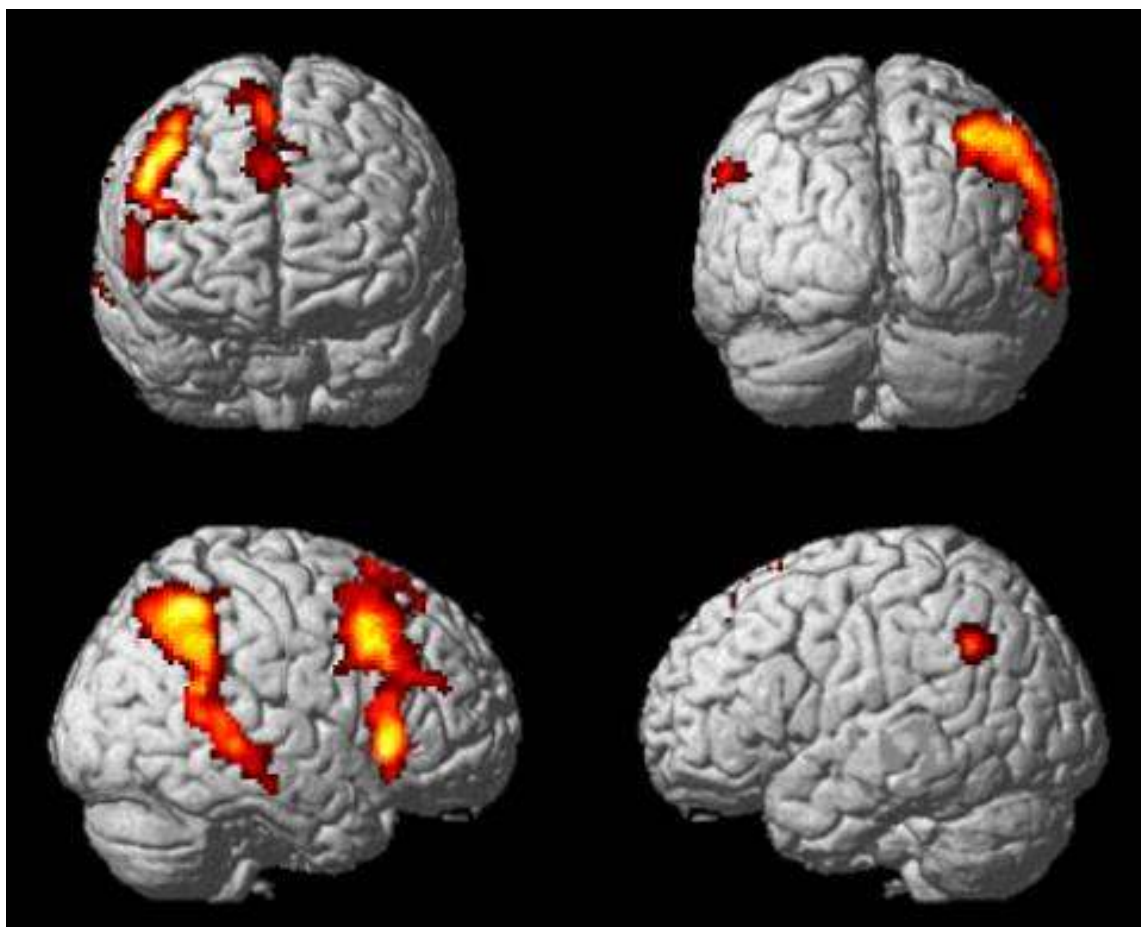


Figure 6.8. Significantly increased brain activity in the No-go condition when compared to the Go condition is depicted here. Activity is shown on a template brain at a voxel-wise threshold of $T = 4$ and a corrected p -value of $p \leq .05$ on cluster-level. For a detailed description of active brain regions please refer to table 6.4.

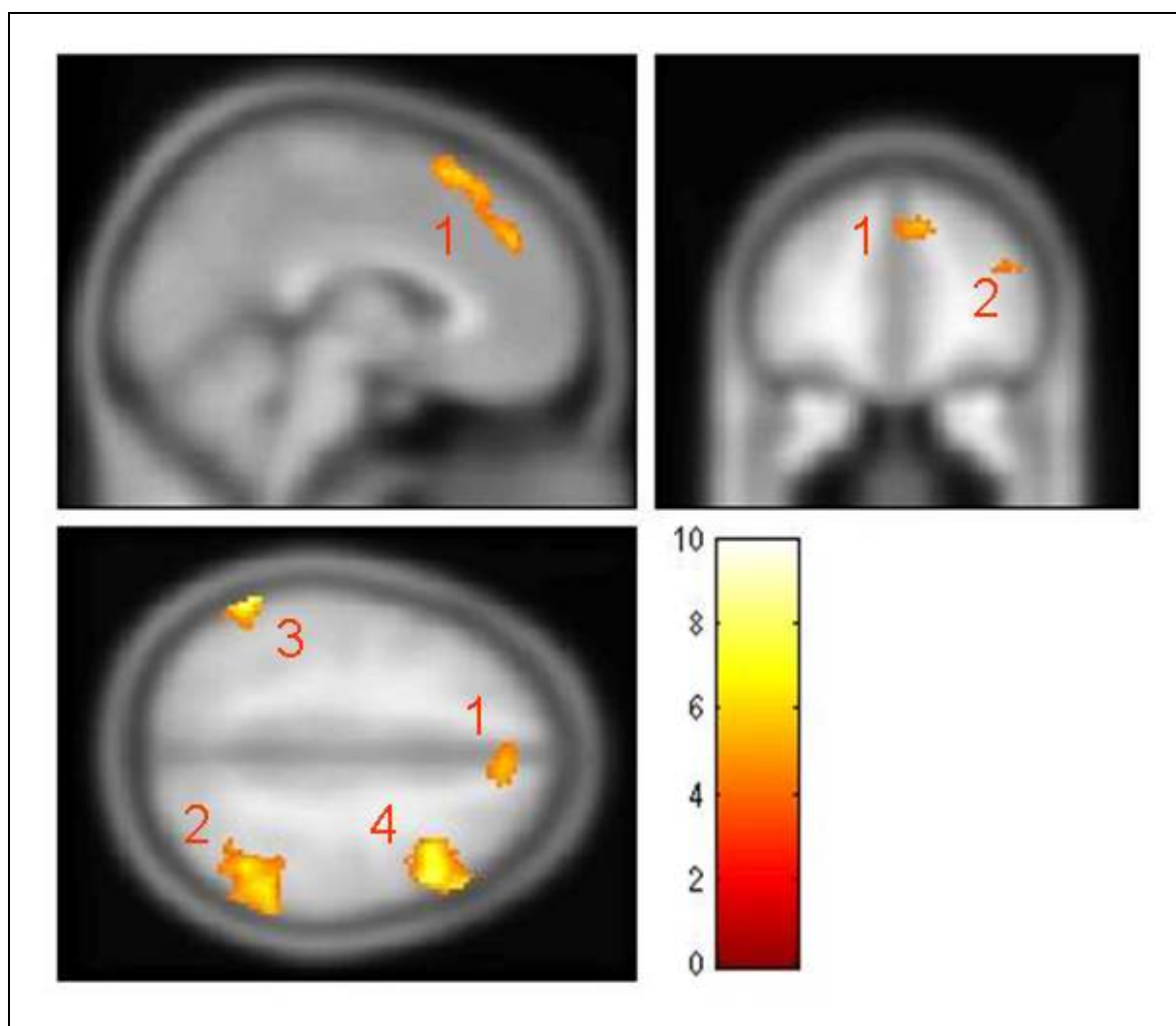


Figure 6.9. This image depicts a sectional view of significant brain activity in the contrast *No-go > Go* at MNI coordinates $x = 3$, $y = 39$ and $z = 38$. Brain activity in this contrast is depicted at a voxel-wise threshold of $T = 4$ and a corrected p -value of $p \leq .05$ on cluster-level. Numbers correspond to significantly active brain regions (1 = right medial and superior frontal gyrus, 2 = right TPJ region and middle and superior temporal gyrus, 3 = left TPJ region, 4 = right inferior and middle frontal gyrus). The color bar depicts corresponding T-values.

6.2.3 Common Activity in the Belief-Reasoning and the Inhibitory Control Experiment

Activity in the comparison *False-Belief* > *True-Belief* and *No-go* > *Go* was overlaid on a template brain in order to show areas that are implicated in both comparisons. These common areas are thus associated with IC processes and a decoupling mechanism between reality and someone else's false belief alike. Overlapping activity was found in the bilateral TPJ, right middle temporal gyrus, medial PFC and right middle and superior frontal gyrus. Respective areas are shown in yellow shading on a template brain in figure 6.10. Areas solely active in the contrast *False-Belief* > *True-Belief* are depicted in red shading. Areas exclusively activated in the contrast *No-go* > *Go* are shown in green shading. Both contrasts show significant clusters only at a voxel-wise threshold of $T = 4$ and a p-value of $p \leq .05$ (cluster level). This finding is also depicted in a sectional view of the brain in figure 6.11.

It has to be noted that the presented comparison of the two latter contrasts is not the result of a statistical comparison. Images with corresponding brain activity were merely superimposed on a template brain. The two contrasts *False-Belief* > *True-Belief* and *No-go* > *Go* could not be compared statistically as the two experiments differed in terms of parameters such as stimulus duration, sequence of stimulus presentation, or differing pre- and succeeding visual stimuli. However, the visual stimulus underlying both contrasts was identical for all regressors making up the latter contrasts.

A sectional view depicting common and differing activity for the latter contrasts in medial and right lateral areas of the brain can also be found in the following discussion section of this thesis (figure 7.4).

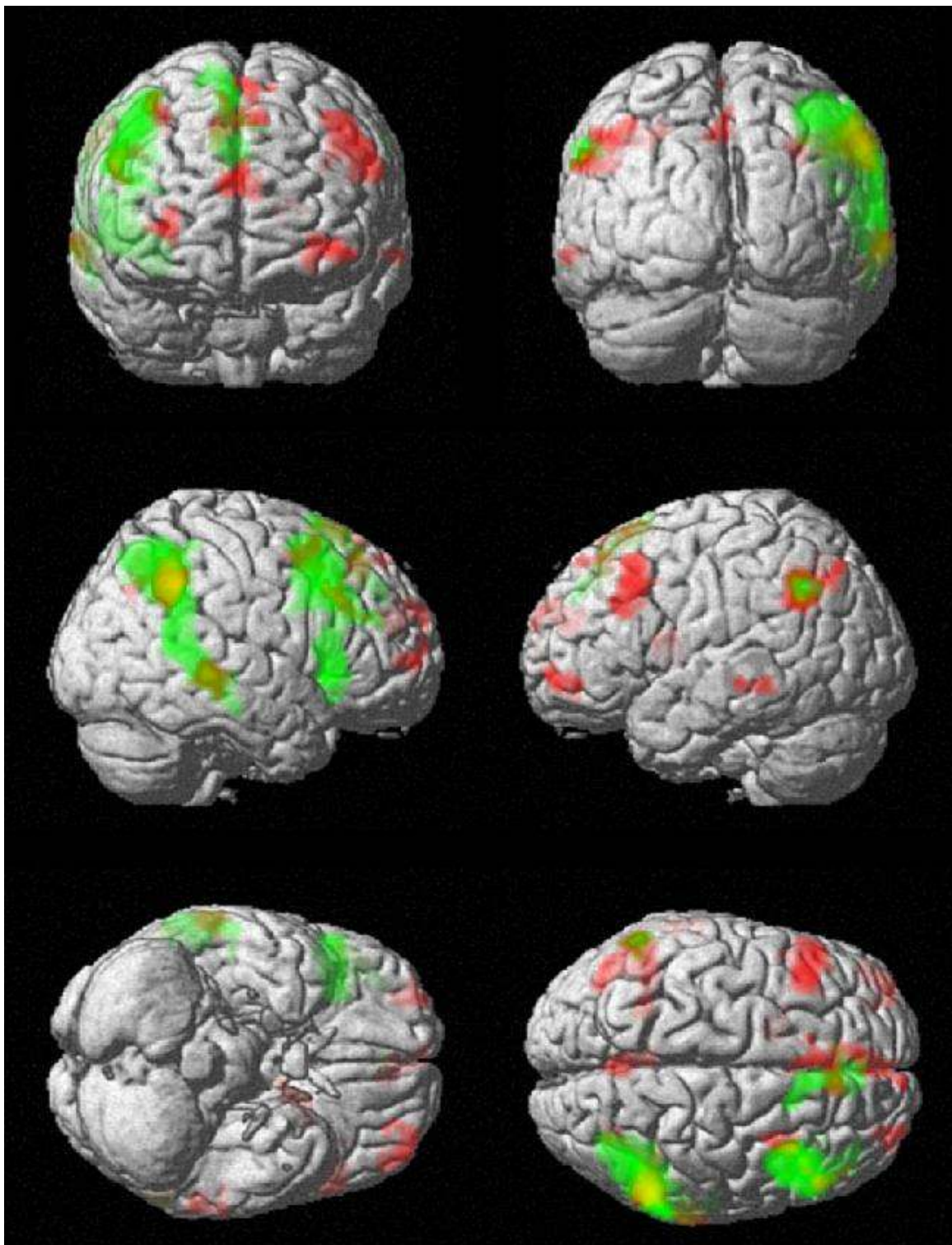


Figure 6.10. Significant brain activity in the contrasts *False-Belief* > *True-Belief* (red shading) and *No-go* > *Go* (green shading) superimposed on a template brain. Areas in yellow shading depict significantly increased activity in both contrasts. A voxel-wise threshold of $T = 4$ and a p-value of $p \leq .05$ on cluster level was set for both comparisons.

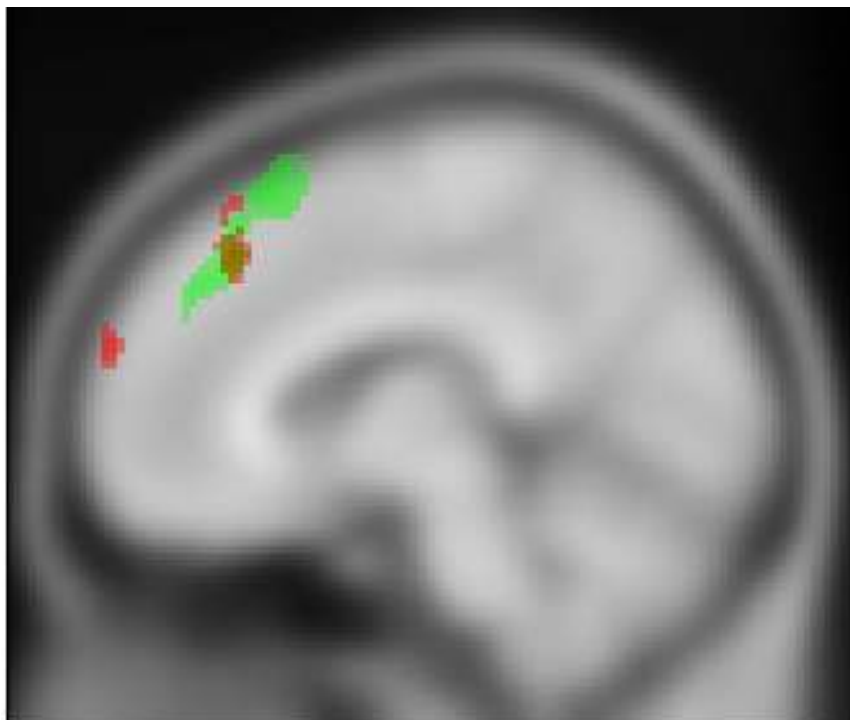


Figure 6.11. Activity related to response inhibition as revealed in the contrast *No-go* > *Go* (green shading) and to a decoupling between reality and a false belief as revealed in the contrast *False-Belief* > *True Belief* (red shading) in the medial prefrontal cortex. Significantly increased activity from the group-analysis is depicted on a sagittal view of the brain at a voxel-wise threshold of $T > 4$ and a cluster-defining threshold of $p \leq .05$.

7. Discussion

The study presented in this thesis attempted to identify neural correlates of belief-reasoning and IC within a single study, using a within-subjects approach and stimuli that differed only marginally between conditions. This was done in order to further clarify a frequently described connection between response inhibition and belief attribution that has become apparent in countless behavioral studies, independent imaging studies, and accounts of compromised belief-reasoning and IC abilities in ASD. Also, this study attempted to probe whether a specialized ToM reasoning module might exist in previously disputed candidate regions such as the TPJ or medial PFC. Ultimately, results from the present study may also be suitable for replicating previous findings of brain regions constituting the neural network for belief-reasoning and IC.

In order to measure neural activity related to the attribution of beliefs, all subjects conducted a belief-reasoning experiment that showed a cartoon character acting on the basis of a true or false belief. Also, a Go / No-go experiment was employed in order to assess neural activity related to IC.

Before discussing the obtained results in more detail, the most important findings from this study will briefly be delineated again.

In the belief-reasoning experiment baseline activity was subtracted from the true-belief condition in order to identify areas associated with general task demands independent of a decoupling mechanism between reality and a false belief. This comparison revealed activity in bilateral lateral and medial frontal areas as well as in precuneus, right middle temporal gyrus, bilateral occipital gyrus, post- and precentral gyrus, and the bilateral TPJ (figure 6.3).

Significantly increased activity in the false-belief condition when compared to baseline was found in almost identical areas as in the contrast *True-Belief > Baseline* (figure 6.4). However, a graphical comparison of both contrasts (figure 6.5) shows that activity in the contrast *False-Belief > Baseline* was more widespread than in the *True-Belief > Baseline* comparison in areas such as bilateral middle frontal gyrus, left TPJ, and lateral and medial BA 10. The *False-Belief > Baseline* contrast was conducted in order to reveal areas related to both general task demands as well as to a decoupling between a false belief and reality.

Also, activity in the true-belief condition was subtracted from the false-belief condition in order to reveal areas specifically dedicated to the attribution of beliefs independent of general task demands. This contrast revealed, among others, significantly increased activity in the bilateral TPJ, bilateral medial frontal gyrus, bilateral middle and superior frontal gyrus, bilateral middle temporal gyrus and bilateral precuneus (figures 6.6 and 6.7).

In the IC experiment, activity associated with response inhibition was revealed by subtracting activity in the Go condition from activity in the No-go condition. This comparison showed significant increases in brain activity in a right-hemispheric network consisting of inferior, middle and superior frontal gyrus, medial frontal gyrus, superior temporal gyrus and TPJ. Significantly increased activity in the left hemisphere was only revealed in the left TPJ (figures 6.8 and 6.9).

A comparison between IC-related areas (*No-go > Go*) and areas dedicated to a decoupling between a false belief and reality (*False-Belief > True Belief*) showed overlapping activity in areas such as the bilateral TPJ, parts of the medial PFC, right middle temporal gyrus, and right middle and superior frontal gyrus (figures 6.10 and 6.11).

Behaviorally, there was no significant difference among conditions in the belief-reasoning experiment in terms of accuracy or reaction times. In the IC experiment, there was a statistical tendency towards higher accuracy in the Go condition when compared to the No-go condition.

The following discussion section is divided into several sections. First, significant activity in all of the contrasts mentioned above is discussed in more detail. This study's findings of common areas related to both IC and belief-reasoning are then discussed in terms of possible common underlying cognitive mechanisms. This section is followed by an attempt to integrate the present findings into a unified hypothesis about the relationship between IC and belief-reasoning. The following section then discusses the impact of this study's results on accounts of hypothesized modular brain regions for the attribution of beliefs. The implications of the current findings for compromised belief-reasoning (such as in ASD) are discussed next, followed by a discussion of the "Papilio" project, an ongoing project that offers kindergarten children training in both social skills and EF abilities. Ultimately, an outlook for future research is given.

7.1 Belief-Reasoning Networks

7.1.1 *True-Belief > Baseline*: Does True-Belief Reasoning Require Belief-Reasoning?

In the belief-reasoning experiment, baseline activity throughout the experiment was subtracted from activity in the true-belief condition in order to identify areas related to general task demands independent of a decoupling mechanism between reality and a protagonist's false belief, which is the crucial component in false-belief tasks.

In the true-belief condition the story protagonist watches the other child transfer the object of interest. The last picture then shows this character looking for the previous object in one of the two containers. When the subject is prompted as to whether he or she

expected the protagonist's behavior, the subject can answer this question by merely comparing the location searched by the protagonist to the actual state of reality. Clearly, this does not require attributing a mental state to the protagonist. Thus, answering a question related to a protagonist's true belief does not require the attribution of a belief, as it is sufficient and more parsimonious to respond to such a question by merely comparing someone else's action to the true state of reality. Therefore, the contrast *True-Belief* > *Baseline* should in theory result in activation that is independent of any sort of mental state attribution. It may therefore be argued that the term "true-belief reasoning" is somewhat misleading as it does not necessarily require the attribution of a belief. Nevertheless, this contrast should depict activity that is related to general task demands that occur in the second story picture, as the statistical analysis of imaging data specifically focused on this period. General task demands in the second story picture involve cognitive processes such as viewing a complex visual scene, shifting attention towards the number of people present in the room (i.e. determining whether the trial depicts a true-belief or a false-belief condition), paying attention to the location that the object is transferred to, and storing both of these scenes in working memory for immediate recall in the third picture. Also, a motor response (i.e. pressing a button with the right index or middle finger) that is required thereafter may already be prepared at this point.

Activity in the contrast *True-Belief* > *Baseline* was revealed in bilateral lateral and medial frontal areas as well as in precuneus, right middle temporal gyrus, bilateral occipital gyrus, post- and precentral gyrus, and the bilateral TPJ (figure 6.3).

Some of the brain regions showing significantly increased activity in this contrast have long been known to be related to some of the general task demands assumed for the present true-belief condition.

Activity in the bilateral occipital cortex in the present comparison may represent the story picture's higher degree of visual complexity in comparison to the overall experiment's baseline as well as a possible storing process of visual items into visual working memory. The role of the occipital cortex for visual perception but also for visual working memory has widely been acknowledged (e.g., Cabeza & Nyberg, 2000; Pasternak & Greenlee, 2005; Rothmayr et al., 2007).

Activity in lateral frontal areas (such as superior, medial and inferior frontal cortex) is also likely to reflect the storage of visually presented items into working memory. In order to successfully respond as to whether the cartoon character's behavior was expected or not in the third story picture, the subject has to memorize several aspects of the second story picture. These include facts such as whether there are one or two children present in the room (thus determining whether a true-belief or a false-belief task is presented) and remembering the location of the container that now holds the transferred object. Lateral

frontal regions active in the contrast *True-Belief > Baseline*, among them BAs 6, 8, 9, 44, and 45, are known to play a prominent role in working memory processes. This view has also been confirmed in review studies including a large number of relevant studies on working memory (e.g., Cabeza & Nyberg, 2000; D'Esposito, 2007).

Activity in the precuneus area (BA 7) was also revealed in the present contrast. Activity in the precuneus has frequently been reported in studies investigating working memory and attentional processes (Cabeza & Nyberg, 2000). Thus, the precuneus activity observed may be associated with working memory demands present in the second story picture. Also, this activity may reflect an increased level of selective attention necessary to respond to the experimental task. Also, activity in the precuneus has been associated with an internal switching process between a first-person and a third-person perspective and with interpreting actions as being controlled by oneself or by someone else (Cavanna, 2007; Cavanna & Trimble, 2006; Vogele & Fink, 2003). Precuneus activity as revealed in the present contrast may thus be associated with attending to a third person (i.e. the story character transferring the object) who is in control of the actions depicted in the cartoon story presented. Clearly, this interpretation is highly speculative and needs further investigation. The data obtained in the present experiment are unable to clarify whether increased precuneus activity could be related to attentional or memory-dependent processes or even to processes related to a sense of agency.

Significantly increased activity was also revealed in the pre- and postcentral gyrus with emphasis on the left hemisphere. This area represents primary and secondary sensory motor areas of the contra-lateral fingers and the hand (Kleinschmidt et al., 1997). Thus, the left-hemispheric activity observed is very likely to relate to the preparation of an imminent motor response as well as to a hypothesized preparatory increase in sensitivity in the right hand. In the experiment, subjects were instructed to use their right middle and index fingers in order to press one of two buttons of a response pad in the following third story picture.

Increased activity in the true-belief condition compared to baseline was also revealed in the medial PFC (BA 6, 8) and the bilateral TPJ (BA 39, 40). As delineated in chapter 2.2.2.2, activity in these areas has frequently been reported in studies investigating the attribution of beliefs (e.g., Fletcher et al., 1995; Gallagher et al., 2000; Saxe et al., 2006; Sommer et al., 2007). At first this seems surprising as this experiment's true-belief condition did not require a decoupling between reality and another person's belief. A mere comparison of the protagonist's action to the real state of affairs was sufficient to answer the present true-belief task. However, activity in areas previously related to belief-reasoning does not necessarily mean that this specific task also required the attribution of beliefs. It is true, however, that areas such as the TPJ and the medial PFC have been

related by some to mental state attribution only, claiming that either one of these specific regions may even constitute the neural substrate of a belief-reasoning module (Frith & Frith, 2003; Saxe, 2006). This view, however, seems outdated. As more recent studies have shown, the dorsal medial PFC and the TPJ are also implicated in other cognitive processes independent of mental state attribution (Corbetta et al., 2008; Mitchell, 2008). Among these processes are the ability to detect and inhibit conflicting stimuli (dorsal medial PFC) as well as the ability to reorient attention to behaviourally relevant stimuli in the environment (right TPJ; Aron, 2007; Corbetta et al., 2008; Gilbert et al., 2008). It can be argued that the true-belief condition does also require subjects to reorient their attention towards the events shown in the second story picture. Clearly these events presented in the subject's environment are behaviourally relevant, as some of the information presented in this picture is crucial for being able to respond to the story character's action in the following third story picture. Thus, activity in areas previously identified in areas related to mental state attribution may rather represent reorienting processes, as this study's true-belief condition did not require the attribution of beliefs. This view is also corroborated by findings from the study presented here. IC, which is related to attentional reorienting, seems to engage areas such as the TPJ and the medial PFC, which have also been found to be dedicated to mental state attribution. This important finding will be discussed in more detail in one of the following sections that deal with the common neural networks for belief-reasoning and IC as revealed in the present study.

7.1.1.1 True-Belief Reasoning: Comparing Present Data to Previous Findings

Another aim of this study was to replicate previous findings on the neural networks for true and false-belief reasoning. In the present case of true-belief reasoning, however, it is hardly possible to determine whether this has been achieved. Of the belief-reasoning studies reviewed, only one study implemented a true-belief condition (Sommer et al., 2007). In this specific study, neural correlates related to mere true-belief reasoning were identified by contrasting two pictures depicting an observed object transfer to four pictures depicting the events leading to this object transfer. Also, Sommer and colleagues used seven pictures for their story compared to the three pictures used in the present study. In the present study the second story picture (depicting the object transfer) could not be contrasted to the first story picture, which would have been analogous to Sommer et al.'s procedure. This is due to the fact that the second story picture in the present study follows immediately after the first picture. Due to this circumstance no statistical comparison of imaging data can be conducted. This is caused by the large covariance of the BOLD-

response elicited by both pictures, thus leading to a blurring of two BOLD-responses that cannot be separated reliably in order to attribute them to the specific pictures. Thus, the present study compared the true-belief condition to baseline activity. Unfortunately, the comparison presented here and Sommer et al.'s contrast can hardly be compared, as the regressors that the true-belief conditions were contrasted to differed largely. The baseline activity, as used in the present investigation, thus reflects other cognitive processes than Sommer et al.'s regressor consisting of four story pictures.

The comparison of Sommer et al.'s direct contrasts between the false and true-belief conditions to the same contrast in the current study, however, is statistically meaningful and will be discussed in one of the following sections.

Next, the present study's results in the comparison between the false-belief condition and the baseline are discussed.

7.1.2 *False-Belief > Baseline*: Decoupling and General Task Demands

When subtracting activity in the experiment's baseline from activity in the false-belief condition, significantly increased activation was revealed in the bilateral TPJ, bilateral lateral and medial frontal areas, bilateral precuneus, right middle temporal gyrus, bilateral occipital gyrus, and post- and precentral gyrus (figure 6.4). Interestingly, activity in this contrast was found in almost identical areas as in the comparison *True-Belief > Baseline*. However, a graphical comparison between both contrasts shows that activity in the false-belief condition is more widespread in areas such as the right TPJ, left lateral frontal areas and the medial PFC (figure 6.5). This finding is also reflected in a statistical comparison of both conditions that is discussed in the following chapter.

Compared to the true-belief condition, the false-belief condition additionally requires subjects to decouple the true state of affairs from another person's false belief. However, the false-belief task also includes general task demands that are also reflected in the *True-Belief > Baseline* contrast. Among these are cognitive processes such as the viewing of a complex visual scene, the focusing of attention towards the number of people present in the room (i.e. determining whether a true-belief or a false-belief trial is shown), paying attention to the location that the object is transferred to, and storing both of these scenes in working memory for immediate recall in the third picture. Also, the contrast *False-Belief > Baseline* is likely to reflect contra-lateral neural activity related to the preparation of a motor response. Activity related to these more general task demands has been discussed in more detail in the previous chapter discussing the contrast *True-Belief > Baseline* and also applies largely to the contrast *False-Belief > Baseline*. As such, activity in the occipital cortex may be dedicated to focusing attention on a behaviorally relevant visual scene and

to storage of crucial visual information in visual working memory. Increased activity in lateral frontal areas is likely to represent working memory-related storage processes for immediate recall in the third story picture. This activity may be more widespread in the false-belief condition as compared to the true-belief condition due to increased working memory demands which additionally require subjects to store the representation of a false belief in working memory. Precuneus activity as observed in the present contrast may be related to either selective attentional processes or processes associated with the taking over of a third-person perspective.

As previously mentioned, activity in the medial PFC (BAs 6, 8, 10) and the bilateral TPJ was more widespread when graphically comparing true and false-belief reasoning by superimposing both contrasts on a template brain (figure 6.5). This observation that also becomes apparent in the statistical comparison can only be explained by an additional cognitive process required by the false-belief task. Obviously, this process constitutes the decoupling of another person's mental state from reality. This is a crucial precondition in order to respond to the third story picture. However, a mere graphical comparison is unable to state whether this decoupling process relies on specialized belief-reasoning mechanisms or whether this activation can be explained by a process of attention reorientation that is more elaborate and demanding in the false-belief condition when compared to the true-belief condition. As this question constitutes one of the main focuses of the present study, an IC experiment was conducted here in order to test whether these specific areas may also be engaged in attentional reorienting independent of mental state attribution. These findings suggest that attentional reorienting and other processes do indeed play a major role in belief-reasoning. Therefore, it is likely that more widespread activity especially in the TPJ in the false-belief condition as opposed to the true-belief condition largely reflects increased demands on the reorienting of attention to behaviorally relevant stimuli in the environment and other processes. This assumption is discussed in more detail in one of the following sections which discusses the implication of the observed common networks for IC and belief-reasoning.

Furthermore, a graphical comparison between the contrasts *False-Belief > Baseline* and *True-Belief > Baseline* may give hints about an underlying decoupling mechanism between reality and a false belief. Nevertheless, this graphical comparison cannot constitute sufficient evidence in terms of a possible replication of previous studies. Previous studies investigating belief-reasoning have for the most part contrasted a false-belief condition against a control condition and not against baseline as conducted in the present contrast. Whether the present study may thus have succeeded in replicating previous accounts of the neural basis of false-belief reasoning will therefore be discussed

in one of the following chapters comparing true-belief and false-belief reasoning directly and statistically.

7.1.2.1 Did the False-Belief Task Really Measure False-Belief Reasoning?

The interpretation of current data presented above is based on the assumption that the false-belief condition actually measured the attribution of a false belief to one of the cartoon characters (experimental validity). However, theoretically it could also be argued that the false-belief task used in the current paradigm could be solved by applying a cognitive strategy without having to impute a mental state to the story's protagonist. In theory, subjects are also able to successfully master both belief tasks by merely applying a set of rules. For instance, if only one person was present (false-belief) in the second story picture and if in the third picture the story protagonist searched the container that contained the object at last, then the subject could use the following rule in order to find out that this observed behavior was unexpected: One person – search in last container → unexpected. Three additional rules would then be sufficient to answer all true and false-belief tasks correctly: one person – no search in last container → expected; two people – search in last container → expected; two people – no search in last container → unexpected. Although applying a set of four three-tiered rules would in theory be sufficient to correctly respond to all belief-reasoning conditions, it seems very unlikely that this strategy was adopted by any of the subjects. It can be argued that it is far more parsimonious in terms of mental effort to merely observe the false-belief trials and respond to the third picture by relying on taking over the protagonist's perspective in order to decouple his belief from the true state of reality. The attribution of beliefs is a highly-specialized and automated process that is probably used by humans countless times every day in order to predict other people's intentions, desires and beliefs (Frith & Frith, 2007). It is very likely that this process is faster, more reliable and more economical in terms of mental effort than applying a rather complicated set of rules that the subject would have had to discover first. Evidence for this assumption comes from a study that required subjects to use specified rules similar to the ones stated above in order to solve a false-belief task. In a different experimental condition, the same set of subjects was then also instructed to solve the same task by merely attributing beliefs to the story character (Saxe et al., 2006). When later asked which "rule" seemed easier to them, 11 out of 12 subjects stated that the attribution of beliefs as opposed to the application of the rules felt easier to them.

Due to these considerations it can be argued that the false-belief task used in the current experiment is a valid measure of the attribution of false-beliefs. It seems highly

unlikely that subjects may have discovered a possible though rather difficult set of rules in order to solve the tasks presented. Nevertheless, it may be that false-belief reasoning in general does rely on covert, unconscious and automated rules that may be structured in a similar fashion as the rules stated above. Underlying basic processes that could possibly support false-belief reasoning will be discussed in more detail in one of the following sections.

Taken together, results in the contrast *False-Belief > Baseline* reflect activity in brain areas related to more general task demands such as working memory, selective attention, or a preparation of motor areas. In addition, this contrast also reflects activity of a decoupling mechanism between reality and mental states, which may in part rely on attention reorienting and other processes. Furthermore, present and previous data suggest that subjects solved the false-belief task used here by imputing a mental state to the protagonist and not by using a rather complicated set of rules. The false-belief task used in the present condition is thus a valid measure of belief attribution.

Next, results from the statistical comparison between the false-belief and the true-belief condition are discussed.

7.1.3 *False-Belief > True Belief*: Neural Networks of a Decoupling Mechanism

Activity in the true-belief condition was subtracted from activity in the false-belief condition to specifically identify the neural correlates of a decoupling mechanism between the true state of reality and another person's false belief. Activity in this comparison was revealed in the bilateral TPJ, medial frontal gyrus, bilateral middle and superior frontal gyrus, bilateral middle temporal gyrus, left thalamus, and bilateral precuneus. As activity in the true-belief condition was subtracted from the false-belief condition, activity in the above areas is not likely to reflect task demands that are also present in the true-belief condition. Thus, activity in these latter areas is specifically dedicated to cognitive functions supporting the attribution of false beliefs. Nevertheless, the areas revealed in the present contrast may also in part or entirely support other basic cognitive processes that may also be present in other cognitive tasks. The contrast *False-Belief > True-Belief* can thus make no inferences as to whether any of the areas revealed constitute the neural correlate of a highly-specialized belief-reasoning module.

Next, the areas revealed in the statistical comparison *False-Belief > True-Belief* are discussed in more detail. This is followed by an assessment of whether the current results are in accordance with previous studies employing similar experimental paradigms.

Middle Temporal Gyrus

As mentioned above, one of the areas with significantly increased activity in the decoupling between reality and another person's false-belief is bilateral middle temporal gyrus. This area is equivalent to BA 21. A review of relevant imaging literature reveals that this area is predominantly found in tasks requiring the encoding of semantic material into working memory (Cabeza & Nyberg, 2000; Hickok & Poeppel, 2004). Furthermore, this region is among other areas that are active during the processing of sentence and text material (Vigneau et al., 2006). It may be that subjects decode the visually presented scene in the second picture by covertly recounting the depicted events. Such a procedure may help subjects with storing the depicted event in working memory for recall in the subsequent third picture. This covert retelling of the depicted situation may happen in the fashion "girl puts teddy in box, other girl is outside the room, thinks it's in the drawer though". This verbal recounting of the events depicted may help subjects to process the viewed scene more deeply and thereby enhance storage into working memory (Craik & Lockhart, 1972). Possibly, such a retelling of the story events in the true-belief task may either not exist or be less extensive since this task can merely be solved by comparing the location searched by the character with the object's real location, i.e. comparing two locations in the third and final story picture. Although speculative, bilateral middle temporal gyrus activity in the false-belief condition may thus reflect more elaborate semantic processing of the visually presented story events than is the case in the true-belief condition. This could reflect the subject's effort to successfully store the false-belief task's more complex content into working memory.

Interestingly, middle temporal and superior temporal gyrus are also activated in the analysis of facial and bodily features (Allison et al., 2000). Putatively, the subjects may pay more attention to the events presented in the false-belief condition, as this condition additionally requires the attribution of a false belief. This increased focusing of attention may also lead to a more elaborate yet automatic analysis of the bodily properties of the story characters.

However, it can only be speculated here whether the observed increased middle temporal gyrus activity in the false-belief condition as opposed to the true-belief condition is due to increased operations of working memory, an analysis of bodily features, or a combination of both.

Middle Frontal Gyrus

Activity in the false-belief condition when compared to the true-belief condition was also significantly increased in the bilateral middle frontal gyrus. In the left hemisphere there

were three clusters of activation throughout the middle frontal gyrus. One cluster was situated within the boundaries of BAs 6 and 8, another one in BA 9, and another one in lateral BA 10. In the right hemisphere, middle frontal gyrus activity was confined to two clusters, one including BAs 8 and 9 and the other one in lateral BA 10. Activity in the dorsal middle frontal gyrus region (BAs 6, 8, 9) has typically been associated in the past with storage and retrieval processes of working memory (Cabeza & Nyberg, 2000; D'Esposito, 2007). Presumably, the false-belief condition may pose higher working memory demands on the subjects than the true-belief condition. While it is sufficient in the true-belief task to compare the location that is searched by the story character to the actual object location, the false-belief task poses higher demands on the subjects. Next to remembering the new location of the object the subject also has to remember the story character's false-belief about the object location.

However, dorsal middle frontal gyrus activity when attributing beliefs may also be explained differently. Recent converging evidence suggests that the middle frontal gyrus may constitute part of a so-called ventral attention network (Corbetta et al., 2008; Corbetta & Shulman, 2002). Next to the middle frontal gyrus, this network also comprises the TPJ and the inferior frontal gyrus. This attentional network is dedicated to detecting and reorienting attention towards behaviorally relevant objects. Activity in the ventral attentional network is for instance observed in paradigms requiring subjects to attend to a novel unattended and behaviorally relevant stimulus (e.g., Bledowski et al., 2004). Although accounts of a ventral attention network are based on data from visual experiments, more recent evidence points to a supramodal function beyond visual stimuli (e.g., Stevens et al., 2005). The false-belief task in the present study may require activity of the ventral attention system, possibly more so than the true-belief task does. Additional attention has to be dedicated to the fact that the story's protagonist is not present in the room at the time of object transfer, thus requiring the subject to additionally attribute a false belief to this person. This mental process is required due to the situation depicted in the second story picture. This picture may thus trigger activity of the ventral attention network by requiring the subject to reorient attention to this behaviorally relevant stimulus which in turn leads to the attribution of a false belief. Although the subject has to focus on the events presented in the true-belief task as well, it does not require a reorienting of attention to the protagonist's false belief. The true-belief task can be solved by merely comparing the true state of reality (i.e. the object location) with the location where the subject looks in the third picture. Activity of the ventral attentional system may constitute one of the basic processes underlying belief-reasoning. Also, it may be one of the basic processes supporting IC. This possible link between belief-reasoning and IC will be

discussed more thoroughly in one of the following sections on the common networks of belief-reasoning and IC (chapter 7.3.3).

Activity in the more ventral bilateral lateral PFC (BA 10) as detected in the current comparison has been revealed in previous studies as a neural correlate of working memory and episodic memory retrieval (Cabeza & Nyberg, 2000). This view has been corroborated in a recent meta-analysis on the subdivisions of BA 10 (Gilbert et al., 2006). As delineated above, the false-belief task may pose higher working memory demands on the subject than the true-belief task. Speculatively, the false-belief task may require subjects to additionally form and store a representation of another person's false belief in working memory. This encoding and storing process of the newly formed representation of a false belief may be reflected in increased activity in ventral lateral PFC comprising lateral BA 10. Clearly, this highly speculative assumption needs further empirical testing.

Precuneus

Next to the lateral frontal activations discussed above, significantly increased activity for the false belief condition was also found in the bilateral precuneus. As mentioned earlier, the precuneus as part of the parietal cortex is a region that is associated with a variety of cognitive concepts ranging from visuo-spatial imagery, attentional processes, sense of agency and perspective taking to recall from episodic memory (Blakemore et al., 2007; Cavanna, 2007; Cavanna & Trimble, 2006). In the current false-belief task, the observed precuneus activity may be associated with taking over another person's perspective, a crucial component of belief-reasoning (Vogeley et al., 2001), or an increased focusing of attention towards an external stimulus (i.e. the scene depicted in the second story picture). However, the precuneus is not a component of the recently proposed ventral attentional system (Corbetta et al., 2008) dedicated to the reorienting of attention towards behaviorally relevant stimuli in the environment. Thus, the precuneus activity that was observed in the present study may play a more specific role in the attribution of beliefs. It is likely that the precuneus may support perspective taking as well as self-processing operations, which in the present case would entail switching from a self perspective to a third person perspective (i.e. the cartoon character's perspective).

Bilateral TPJ

Activity in the bilateral TPJ region was also significantly increased in the false-belief condition when compared to the true-belief condition. The right TPJ is among the most vividly discussed candidate regions for a possible specific belief-reasoning module (e.g.,

Mitchell, 2008; Saxe et al., 2006). Although this region was significantly active in the contrast presented, this does not necessarily imply that either the left or the right TPJ constitute a belief-reasoning module that meets the specificity and the generality criterion for such a region (Saxe et al., 2004). Previous research has shown that the TPJ is also implemented in a variety of other tasks, ranging from IC, the processing of semantic material and attentional reorienting to the analysis of biological motion (Allison et al., 2000; Buchsbaum et al., 2005; Constable et al., 2004; Serences et al., 2005). An area that has been investigated less extensively is the left TPJ, which also seems to play an important role for the attribution of beliefs. This view is especially corroborated by a lesion study that found three patients with damage to the left TPJ significantly impaired in two measures of false-belief reasoning (Samson et al., 2004). Furthermore, the left TPJ is not considered a part of the ventral attentional network that supports the reorienting of attention to behaviorally relevant stimuli (Corbetta et al., 2008). This evidence suggests that the left TPJ may play a specific role in the attribution of beliefs that may not necessarily be explained by attentional processes. The possibly differing roles of the left and the right TPJ in belief-reasoning and also IC are discussed more extensively in one of the following chapters.

Medial PFC

In the present study, the medial PFC emerged as another significantly active region for the attribution of false beliefs. This activation extended over two neighboring clusters. One more ventral cluster encompassed BAs 9 and 10. The other more dorsal cluster was restricted to BAs 6 and 8. Next to the TPJ, the medial PFC is one of the most frequently observed regions associated with the attribution of beliefs (Fletcher et al., 1995; Frith, 2008; Gallagher et al., 2000).

Interestingly, most studies investigating the attribution of mental states revealed medial PFC activity in more ventral prefrontal areas (BAs 9 and 10). These results unfortunately stem from studies that largely differed in terms of experimental paradigms, contrasted conditions and the actual concepts investigated. Many studies also fail to specifically state the location of the observed activations. Nevertheless, a review of respective studies indicates that especially BAs 9 and 10 are associated with belief-reasoning. This view is also supported by a meta-analysis of 104 functional imaging studies conducted by Gilbert et al. (2006) who found that the medial areas of BA 10 are implicated in the attribution of mental states. Interestingly, medial BA 10 activity is also significantly increased in tasks requiring a switching between stimulus-oriented and stimulus-independent thought (Gilbert et al., 2008). It can be argued that in the present false-belief tasks subjects are required to reorient their attention and thought towards an

outside stimulus, namely the protagonist's false belief. This orientation towards a stimulus may be more pronounced than it is the case in the true-belief task, which does not require an additional belief attribution in order to solve the task.

Response inhibition has also been associated with medial PFC activity (see also chapter 3.2). However, activity in experiments investigating IC is usually confined to more dorsal areas of the medial PFC (BAs 6, 8) as well as to the dorsal part of the ACC (BA 32). This view is also corroborated by two lesion studies showing that among patients with frontal brain lesions, dorsal medial frontal lesions have the largest impact on the ability to inhibit a response (Floden & Stuss, 2006; Stuss et al., 2001a). Possibly, activity in this study's dorsal medial PFC cluster (BAs 6, 8) for belief-reasoning could be related to inhibitory demands required for the attribution of a false belief. Although not a part of the ventral attentional network, the dorsal medial PFC could also be dedicated to low-level processes supporting both IC and belief-reasoning. A discussion of the possible nature of such low-level processes is presented in one of the subsequent sections.

Summing up, medial PFC activity associated with a decoupling mechanism was revealed in two distinct clusters, a dorsal cluster consisting of BAs 6 and 8 and a more ventral cluster consisting of BAs 9 and 10. Activity in the dorsal cluster may be related to inhibitory processes supporting belief-reasoning and IC alike. Previous research suggests that the ventral medial PFC on the other hand may be related to a more specific function possibly exclusively implicated in the attribution of beliefs.

Thalamus

A part of the thalamus was another, yet unexpected cluster of activity in the present contrast *False-Belief > True-Belief*. Up to now no accounts of thalamus activity in the attribution of beliefs have either been reported or discussed. The thalamus as part of the diencephalon serves multiple functions in the brain and maintains nerve tracts to various parts of the cortex, including the PFC. One of the thalamus's main functions is to relay sensory signals for further processing in other parts of the brain. However, the thalamus also supports attentional processes underlying arousal, sleep, wakefulness and selective attention (Schiff, 2008). The thalamus activity observed in the present study is likely to be associated with increased attentional demands in the false-belief task when compared to the true-belief task. Next to general task demands also present in the true-belief task, the false-belief task requires an additional attribution of a false-belief, which may be associated with an increased demand for attentional processes. This increased level of selective attention is likely to be reflected in thalamic activity as found in the present comparison. Possibly, the thalamus initiates an attentional arousal of PFC regions that seem to be of special importance in belief-reasoning. Activity of the thalamus as part of an

attention system in the brain may not have been revealed in previous belief-reasoning studies due to the imaging modalities used to detect brain activity. The current study may have succeeded in measuring thalamic activity due to the use of 3-Tesla high-field functional magnetic imaging, whereas most previous studies have relied on field strengths of 1.5 Tesla or even on PET scanning.

7.1.3.1 Replication of Previous Studies

Replication of Lesion and TMS Studies

One of the aims of the present study was to replicate previous findings of neural networks associated with the attribution of beliefs. Lesion studies for instance were able to underscore the significant influence of a number of regions directly related to the attribution of beliefs. Among these areas were the left TPJ, the PFC (with emphasis on the medial part) and the amygdala (Apperly et al., 2004b; Fine et al., 2001; Rowe et al., 2001; Samson et al., 2004; Stuss et al., 2001a; Stuss et al., 2001b). The only TMS study conducted so far found a significant impairment in the attribution of beliefs when “knocking out” the bilateral dorso-lateral PFC and the right TPJ (Costa et al., 2008).

The data obtained in the current study were thus able to replicate the existing results from lesion studies almost entirely. Furthermore, areas not investigated in previous lesion and TMS studies were revealed in the present study as additional components of a neural network for the attribution of beliefs. Taken together, the current study found false-belief reasoning-related activity in, among other areas, the right TPJ, the medial PFC and the bilateral dorso-lateral PFC. All of these areas were also identified in existing lesion and TMS studies as part of a neural belief-reasoning network.

However, the present study did not find activity in the amygdala region, which may also play a role in belief-reasoning, as one lesion study suggests (Fine et al., 2001). This lesion study’s findings are nevertheless based on the data of only one subject with congenital amygdala damage and may therefore be considered as only vague evidence for a role of the amygdala in belief-reasoning. Also, not a single fMRI study has so far reported amygdala activity in belief-reasoning. Nevertheless, it may be that the present study did not capture amygdala activity due to the properties of the functional scanning method used here. Difficulties in capturing amygdala activity in high-field magnetic imaging due to susceptibility artefacts have widely been acknowledged and may also have resulted in the present study’s inability to detect amygdala activity (Morawetz et al., 2008). In addition to this methodological challenge, previous studies suggest that the amygdala is implicated in other aspects of social cognition, including the processing of social visual stimuli (Adolphs & Spezio, 2006). This finding and the lack of reported amygdala activity in

the current and previous studies suggest that the amygdala may not be specifically implicated in the attribution of beliefs. Improved scanning procedures in the future may result in clarifying a possible contribution of the amygdala in ToM reasoning. Taken together, the present study was by and large able to replicate lesion-based findings of a belief-reasoning network consisting, among others, of the right TPJ, medial PFC and the dorso-lateral PFC.

Replication of Functional Imaging Studies

The belief-reasoning areas identified in the current study are also in accordance with most of the previous functional imaging studies that have investigated belief-reasoning by means of PET and fMRI. Furthermore, the current study also identified areas of the brain that have previously not as frequently been associated with the attribution of beliefs.

A detailed overview of imaging studies investigating belief-reasoning is found in chapter 2.2.2.2. This review suggests that the bilateral TPJ, the medial PFC (with preponderance in ventral areas such as BAs 9 and 10) and lateral prefrontal areas constitute the core components of a neural belief-reasoning network. It was also argued previously that probably only five imaging studies may have yielded valid results in terms of identifying a neural belief-reasoning network (Fletcher et al., 1995; Gallagher et al., 2000; Gobbini et al., 2007; Perner et al., 2006; Sommer et al., 2007). Although other studies attempted to measure the attribution of beliefs as well, these studies may in reality have tapped other concepts such as the attribution of desires or intentions or may even have failed to investigate any sort of ToM reasoning at all.

The only PET study among these three investigated the neural correlate of belief-reasoning by comparing belief-reasoning stories to physical control stories that did not require any reasoning about beliefs (Fletcher et al., 1995). This comparison resulted in increased activity in the belief-reasoning condition in the medial PFC (BA 8). The current study also found activity related to belief attribution in BA 8. This replication of previous findings thus underscores the prominent role of medial PFC in the attribution of beliefs. The role of the dorsal medial PFC will be further discussed in one of the subsequent sections of this thesis.

One of two ToM conditions utilized in an fMRI study by Gallagher et al. (2000) may also have captured brain activity related to the attribution of beliefs by using stories requiring belief-attribution as well as control stories that did not require the attribution of beliefs. Gallagher and colleagues found increased activity related to belief-reasoning in the medial PFC (BAs 8, 9), bilateral TPJ, precuneus, and the temporal poles. With the exception of the temporal poles, the current study was able to replicate findings of the bilateral TPJ, medial PFC and precuneus as parts of a neural network underlying the

attribution of beliefs. A recent review suggests that the temporal poles (BA 38) may play a role in processes such as semantic memory, face recognition or emotional processing (Olson et al., 2007). The present study may have been unable to capture temporal pole activity due to the use of a head-only high-field fMRI scanner resulting in a relatively small field-of-view (192 X 192 mm). A review of individual data from this study suggests that most subjects' temporal poles were outside this field-of-view. Thus, the present study is unable to determine the role of the temporal poles in either belief-reasoning or inhibitory control.

Another study conducted by Gobbini et al. (2007) used the identical stimuli to investigate belief-reasoning as in the Gallagher et al. study discussed above. Gobbini et al.'s results point to a network dedicated to belief-reasoning consisting of medial PFC (BAs 9 and 10), bilateral TPJ, left precuneus, and the bilateral temporal poles. As discussed above, the current study was thus also able to replicate Gobbini et al.'s findings almost entirely, with the exception of the temporal poles. As recently mentioned, this may have been due to the methodological issues presented above or due to an unspecific role of the temporal poles in belief-reasoning.

A study also using an appropriate methodology to capture belief-reasoning was conducted by Perner et al. (2006). Their comparison of false-belief stories to false photograph stories resulted in a belief-reasoning-related activity in the bilateral TPJ, the right middle and superior temporal gyrus, precuneus, and medial PFC (BA 9). The current study revealed activity over and above these areas. Perner and colleagues also compared activity in the false-belief task to a false sign task which may only differ to the false-belief tasks in the belief attribution component. This comparison resulted in increased activity for belief-reasoning in the right TPJ which was also found in the current study. Although Perner et al. argue that this points to a specific role of the right TPJ in belief-reasoning, this result could also be explained by an additional attentional reorienting process as described earlier in this section. The issue of the right TPJ as a possible belief-reasoning module will be discussed in more detail in one of the following chapters.

The current study also attempted to replicate findings of a study conducted by Sommer et al. (2007) that used a similar approach to identify the neural correlates of a belief-reasoning network. In Sommer et al.'s study, correlates of a belief-reasoning network were identified by subtracting activity in a true-belief condition from activity in a false-belief task. The same cartoon stories as in the present study were used, however containing seven instead of the three pictures used in this story. Sommer et al. revealed a neural network for a decoupling between reality and a false belief in the right DLPFC (BA 9), right middle frontal cortex (BA 6), right lateral PFC (BA 10), right middle temporal gyrus, right TPJ, and ACC (BA 32). With exception of the ACC the current study was able

to replicate the findings from Sommer et al.'s study and even reveal further components of a belief-reasoning network. The hypothesized roles of the brain areas commonly identified in both studies have been discussed above. The observed ACC activity in Sommer et al.'s study has been explained by increased action monitoring demands for the false-belief task. Action monitoring in this sense refers to situations involving response conflicts including IC. As has been delineated further above, IC is also related to a more dorsal yet neighboring area encompassing BA 8. The peak activity voxel for the ACC as revealed in Sommer et al.'s study was located at MNI coordinates $-18, 16, 46$ and is thus 1.8 centimeters lower than the possibly corresponding medial PFC cluster at MNI coordinates $10, 34, 46$ that has been revealed in the present study. Although the peak voxels of the clusters differ across the two studies, a graphical comparison of the two clusters reveals that some of the active voxels revealed in both studies may actually overlap. Although the respective dorsal medial cluster activated in the current study may be located primarily in BAs 8 and 6, some minor parts of it also seem to be located in the ACC. The ACC cluster revealed in Sommer et al.'s study on the other hand seems to have voxels located in BAs 8 and 6 as well. Thus, it is likely that the two clusters revealed in the two differing studies may actually represent a functionally homogenous region. This specific region may however have shown up in slightly differing locations due to differences between subjects in the anatomic allocation of this specific functional area. This differing peak voxel allocation ultimately becomes apparent in the group analysis representing an average of individual data. Furthermore, the current and Sommer et al.'s study have utilized slightly different statistical regressors in order to isolate a neural network dedicated to the attribution of beliefs. While Sommer et al.'s study included two pictures in their statistical analysis, the current study focused on one story picture during which belief-reasoning was assumed to take place. Thus, both the methodological issues of varying regressors as well as inter-individual differences may have contributed to the fact of finding a possibly functionally homogenous yet not perfectly overlapping region which seems to be associated with belief-reasoning. Taken together, Sommer et al.'s and the current study accordingly point to a neural belief-reasoning network consisting of the right TPJ, right lateral PFC, right middle temporal gyrus, right dorso-lateral cortex, precuneus, and dorsal medial PFC. Furthermore, the current study revealed that also the left TPJ, left lateral PFC, ventral medial PFC, left thalamus, and left middle temporal gyrus are implicated in the attribution of beliefs.

A review of several studies investigating belief-reasoning thus suggests that the current study succeeded in replicating previously identified components of a neural belief-reasoning network. In accordance with previous studies, the study presented here found activity related to belief-reasoning in the right TPJ, medial PFC, lateral PFC, precuneus,

and the middle temporal gyrus. Also, activity related to belief-reasoning was found in areas previously not related to the attribution of beliefs, such as the thalamus for instance.

Summing up, the current study succeeded in identifying a neural network associated with belief-reasoning by subtracting activity in the true-belief condition from activity in the false-belief condition. This hereby identified neural network consists of the bilateral TPJ, ventral and dorsal medial PFC, bilateral lateral PFC, bilateral middle temporal gyrus, and bilateral precuneus. Converging evidence including existing studies suggests that the right TPJ, bilateral middle temporal gyrus, dorsal medial PFC (BAs 8, 6), and lateral PFC seem to be implicated in the attribution of beliefs. However, these areas also support cognitive functions such as working memory, inhibition and attentional processes. These functions may thus constitute basic processes of belief attribution. Previous data further indicate that the left TPJ and ventral medial PFC (BAs 10 and 9) may play a role that could be more specific to the attribution of beliefs. Furthermore, the current study was able to replicate previous findings from lesion and functional imaging studies and identify areas that have previously not been associated with belief-reasoning.

Concurrent evidence points to a close relation between belief-reasoning and IC both behaviorally and on the neural level. This is why in addition to belief-reasoning the current study also investigated the neural correlates of IC. Brain regions revealed in this study as neural correlates for response inhibition are thus discussed in the following chapter.

7.2 Brain Areas Related to IC

7.2.1 *No-Go* > *Go*: A Neural Network for Response Inhibition

In an attempt to identify brain areas showing significantly increased activity associated with IC, a Go / No-go paradigm was utilized in the present study. This Go / No-go paradigm used the identical visual stimuli that were presented as the second story picture in the belief-reasoning experiment. IC-related activity was revealed by subtracting activity in the Go condition from activity in the No-go condition. This contrast revealed a largely right-hemispheric network dedicated to IC consisting of inferior, middle and superior frontal gyrus, medial frontal gyrus, superior temporal gyrus, and TPJ. The only significantly active region in the left hemisphere was the left TPJ.

A review of previous studies that employed a similar approach to investigate IC by means of Go / No-go or related tasks revealed activity in areas that were surprisingly similar to the areas identified in the current study. Interestingly, virtually all previous studies that focused on the neural correlate of response inhibition also revealed activity in

a largely right-hemispheric network (Blasi et al., 2006; Casey et al., 1997; Ciesielski et al., 2006; Garavan et al., 2006; Kelly et al., 2004; Liddle et al., 2001; Menon et al., 2001; Mostofsky et al., 2003; Stevens et al., 2007; Wager et al., 2005; Watanabe et al., 2002). This result which is depicted in various studies investigating IC is also depicted in two recent meta-analyses that comprised a total of 29 different studies (Buchsbaum et al., 2005; Simmonds et al., 2008). According to these data, a network consisting of the right dorso-lateral PFC (BAs 9, 46), the right inferior frontal cortex (BAs 44, 46), the superior medial frontal cortex (BAs 6, 8, 32), and the bilateral TPJ constitute the core components of the neural correlate for IC.

IC is possibly a process that is supported by various underlying cognitive functions that are in turn reflected in activity in distinct regions of the brain.

The right DLPFC (BAs 9 and 46), which has also been found for IC in the present study and virtually all previous IC studies, seems to support working memory which is also a necessary component for successful response inhibition (e.g., D'Esposito, 2000). In order to successfully inhibit a response in the present IC task, the subject has to remember the information given in the previous picture (i.e. number of children present) and compare it to the present picture. Subsequently, the subject also has to retrieve the previously learned rule for pressing a button (i.e. refrain from pressing the button if number of children is identical as in previous picture) from working memory in order to successfully master the task. In addition to forming these stimulus-response associations, the DLPFC also exerts top-down control on motor areas which need to refrain from performing a motor response that is only required in the Go condition (Curtis & D'Esposito, 2003; Simmonds et al., 2008). Although these memory and control processes are also necessary in the Go condition, these processes may be more automatic in the Go trials which outnumber the No-go trials by far. Therefore, the No-go condition may require more elaborate operations in working memory for both picture and task rule retrieval which are reflected in DLPFC activity for response inhibition.

Another significantly active region for IC as revealed in the current study is superior and middle temporal gyrus, which in the present study were part of a larger cluster also encompassing the TPJ region. As discussed in one of the previous chapters, this area has been found in previous studies investigating the processing of sentence and text material as well as the encoding and retrieval of semantic material into and from working memory (Cabeza & Nyberg, 2000; Hickok & Poeppel, 2004; Vigneau et al., 2006). It is likely that subjects use a verbal, albeit covert strategy to correctly respond to the requirements of the present IC task. The subjects may code the information presented to them (i.e. number of children present in the picture) by covertly rehearsing the number words "one" and "two". In the No-go condition the verbal switching between the words "one" and "two" is

interrupted, thus requiring the subjects to override an impulse to use the opposite number word and to instead decode the information given to the appropriate word. Thus, middle and superior temporal gyrus activity could reflect the verbal decoding of the picture stimuli presented as well as verbal working memory processes. Interestingly, the superior temporal gyrus, a region that was not active for decoupling mentality from reality in the belief-reasoning experiment, has also been associated with social perception (Allison et al., 2000). Several studies report superior temporal gyrus activity in subjects observing biological movements (e.g., Grezes et al., 2001). Furthermore, the same region is also activated by static images of face and body as well as by stimuli indicating an individual's action (e.g., Kourtzi & Kanwisher, 2000; Morris et al., 2006). Presumably, increased attention towards the behaviorally relevant No-go stimuli may also result in an automatic processing of the social information contained in the picture presented (i.e. processing faces, bodies, intended movements and actions). This more elaborate processing of the picture content may result in increased superior temporal gyrus activity for IC in the current experiment. Taken together, superior and middle temporal gyrus activity as part of an IC network may either reflect verbal working memory processes, social stimulus processing, or a combination of both.

Another significantly active region associated with IC in the present as well as previous studies is the right inferior frontal gyrus region encompassing BAs 44 and 45 (Buchsbaum et al., 2005; Kelly et al., 2004; Simmonds et al., 2008). While the DLPFC for example may play a more specific role in working memory, the inferior frontal gyrus region could be more related to an actual inhibitory process. This view is also shared by a review article dealing with the function of this specific region (Aron et al., 2004). According to the results reviewed in this paper, the inferior frontal cortex may constitute the only area in the brain dedicated to IC. Although the DLPFC and the dorsal medial cortex were also reviewed in this article as they are also part of a hypothesized IC network, the authors argue that these two areas support working memory and conflict detection, respectively. Further support for a distinct role of the right inferior frontal cortex in response inhibition also comes from a study showing that the extent of damage to this region (but not to other prefrontal regions) shows a significant negative correlation with performance in an IC task (Aron et al., 2003).

Activity related to IC was also revealed in the medial PFC in BAs 6 and 8. This result is in accordance with previous studies and meta-analyses reporting IC-related activity in a similar area (Buchsbaum et al., 2005; Kelly et al., 2004; Mostofsky et al., 2003; Simmonds et al., 2008). Both meta-analyses reviewed here consisting of a total of 29 IC-related studies revealed concurring activity in BA 6, an area that has also been referred to by some as pre-SMA (supplementary motor area). A large body of electrophysiological, lesion

and imaging studies point to a prominent role of the dorsal (or superior) medial PFC in IC. Two lesion studies, for example, were able to show that among patients with damage to the PFC, only the patients with damage to the right superior frontal gyrus were significantly impaired in tasks assessing response inhibition (Floden & Stuss, 2006; Stuss et al., 2001a). Also, patients with ADHD, a disorder that involves poor IC abilities, show a dysfunctional activity pattern in dorsal medial PFC, thus corroborating accounts that this region serves as a key area for response inhibition (Dickstein et al., 2006).

Due to these accounts it is argued here that the dorsal medial PFC activity and the right inferior frontal cortex activity as observed in the current experiment reflects activity of a region that constitutes an area specifically implicated in response inhibition.

The bilateral TPJ was another region that was significantly active in the contrast *No-go* > *Go*. Although the TPJ is associated primarily with ToM reasoning experiments, a close examination of relevant literature reveals that it is also frequently reported in studies assessing IC. A recent meta-analysis conducted by Simmonds et al. (2008) that included results from 11 IC studies reported concurrent bilateral TPJ activity. Another meta-analysis that included 18 studies related to IC (“attentional reorienting”) found activity in the right TPJ that was almost identical to activity elicited by ToM reasoning tasks (Decety & Lamm, 2007). Converging evidence suggests that at least the right TPJ may support processes that are necessary for both IC and belief-reasoning. Respective studies point to a role of the right TPJ in breaking the current attentional set to orient to task-relevant stimuli (Decety & Lamm, 2007; Mitchell, 2008). Furthermore, the right TPJ (and possibly the left TPJ, which however shows less frequent activations) seem to be part of a hypothesized ventral attention network which reorients attention to behaviorally relevant stimuli in the environment. In the present *Go* / *No-go* task the *No-go* stimulus presents such a novel and behaviorally relevant stimulus. The *No-go* stimulus requires subjects to reorient attention to this specific stimulus. Subsequently, the subject is required to inhibit a prepotent motor response which he or she has previously gotten used to due to its frequent occurrence. This process is likely to require additional attentional resources. Taken together, activity of the right and possibly the left TPJ may reflect activity of an attention reorienting system involved in IC that redirects attention to a behaviorally relevant stimulus, represented in the present case by the *No-go* stimulus. This reorienting response may also be a component of belief attribution and will be discussed in one of the subsequent chapters.

Summing up, the comparison *No-go* > *Go* revealed a predominantly right-hemispheric network consisting of the DLPFC, the middle and superior temporal cortex, the inferior frontal gyrus, dorsal medial PFC, and bilateral TPJ. The areas identified match previous findings of IC-related brain regions almost entirely. While the DLPFC and the middle temporal gyrus may support necessary and increased working memory demands,

the superior temporal gyrus may additionally and automatically process the social information contained in the picture presented. The inferior frontal cortex and the dorsal medial PFC on the other hand may constitute regions that are specific to response suppression. Ultimately, the bilateral TPJ may support processes necessary to reorient attention to a task-relevant stimulus.

The following section discusses common and distinct areas for belief-reasoning and IC as revealed in the present study. Furthermore, these findings will be discussed in terms of common underlying cognitive processes for both concepts as well as in light of modularity accounts for belief-reasoning

7.3 Common and Distinct Areas for Belief-Reasoning and IC

A graphical comparison of areas significantly active in the decoupling between reality and a false belief (as revealed in the comparison *False-Belief* > *True-Belief*) and areas dedicated to IC (*No-go* > *go*) was conducted in the present study. This comparison was conducted in an attempt to clarify the frequently reported behavioral and functional connections between the two latter concepts. The finding of common brain areas could indicate possible underlying cognitive commonalities that are present in both concepts. Finding distinct brain areas on the other hand might give hints about specific properties of the one or the other cognitive ability. Furthermore, the current comparison is able to give further evidence for or against a specific ToM module in the brain. For a region to qualify as a specific ToM module, this region must not be activated in any other cognitive process but the attribution of mental states (Saxe et al., 2004).

The following paragraph first delineates distinct brain areas activated in one or the other process, followed by a discussion of the possible role of these specific areas. This is followed by a description of areas that have been identified in the current study as being part of both a belief-reasoning as well as an IC network. Furthermore, the implications of overlapping neural activity in both concepts towards a revised theory of the IC / belief-reasoning connection are discussed.

7.3.1 Distinct Areas Related to a Decoupling Between Mentality and Reality

Activity related to a hypothesized decoupling mechanism between the true state of affairs and a false belief as revealed by the contrast *False-Belief* > *True-Belief* was found in the bilateral TPJ, medial frontal gyrus, bilateral middle and superior frontal gyrus, bilateral middle temporal gyrus, left thalamus, and bilateral precuneus. When comparing this pattern of activation to the corresponding IC-related comparison (*No-go* > *Go*), distinct

activity for the decoupling-related comparison became apparent in the bilateral precuneus, left thalamus, left middle temporal gyrus, left middle frontal gyrus (BAs 6, 8, 10), right middle frontal gyrus (BA 10), and ventral medial frontal gyrus (BAs 9, 10). This finding is also depicted in figures 7.1, 7.2, and 7.3.

The above areas, which do not seem to play a role in the current response inhibition task, have also been discussed in detail in a previous section of this paper and are reviewed here briefly.

The bilateral precuneus was engaged in the present belief-reasoning task but not for inhibitory control. This finding can be accounted for by viewing the precuneus as an area possibly implicated in perspective taking. Perspective taking may constitute a crucial component of false-belief reasoning but not play a role in inhibitory control (e.g., Blakemore et al., 2007; Vogeley et al., 2001).

The left thalamus as another structure implicated in belief-reasoning but not in IC may be engaged in the attentional arousal of medial prefrontal areas which in turn seem to be of special importance for mental state attribution. However, the thalamus does not seem to play a specific role for belief-reasoning but rather serve as a multi-functional component of the brain's attentional system that is also recruited during other cognitive processes (Schiff, 2008)

Left middle temporal gyrus activity during the attribution of beliefs may be related to a verbal rehearsal strategy that is more elaborate in the more demanding false-belief condition than in the true-belief condition (Hickok & Poeppel, 2004). As the left hemisphere is specific for language-related processes, this indicates that subjects might rely on verbal working memory processes especially for belief attribution but only to a lesser extent for response inhibition.

Left middle frontal gyrus (BAs 6, 8, 10) and right middle frontal gyrus (BA 10) may reflect activity of storage and retrieval processes in working memory that are distinctly active for false-belief reasoning but not for IC (Cabeza & Nyberg, 2000; D'Esposito, 2007). While the medial part of BA 10 may be more specific to the attribution of beliefs, lateral BA 10 activity may reflect working memory demands necessary for belief attribution (Gilbert et al., 2006). Putatively, lateral BA 10 may be dedicated to the storage of the mental representation of a false belief and thus not be engaged in working memory during response inhibition.

Furthermore, an area in the ventral medial PFC (BAs 9, 10) showed significantly increased activation in the belief-reasoning but not in the IC comparison. This finding is once again depicted in a sectional view in figure 7.1.

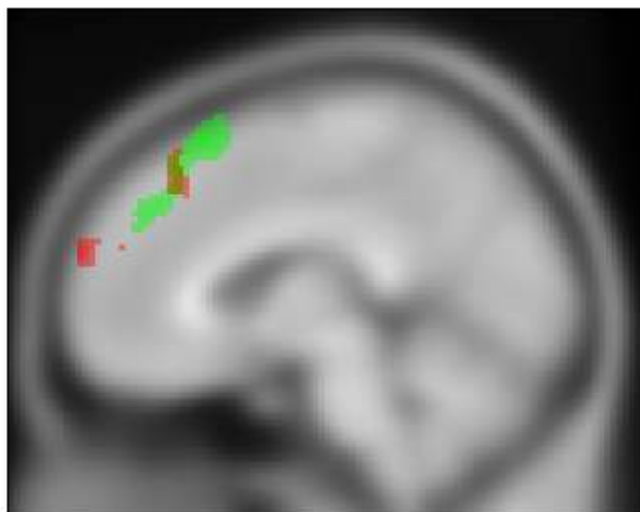


Figure 7.1. Sagittal view of a template brain depicting activity in the comparison *No-go > Go* (green and yellow shading) and *False-Belief > True-Belief* (red and yellow shading) at a voxel-wise threshold of $T > 4$ and a cluster-defining threshold of $p \leq .05$. As can be seen here, a region in the medial ventral PFC cortex was distinctly activated by the comparison *False-Belief > True-Belief*.

Ventral medial PFC (BA 10) seems to support self-referential processes that may be considered crucial for belief-reasoning (Northoff et al., 2006). More dorsal areas of the medial PFC (BAs 6, 8) on the other hand seem to play a role in the inhibition of unwanted responses as various imaging and lesion studies have suggested (e.g., Floden et al., 2006). This distinction in the medial PFC that becomes apparent when reviewing relevant data is also supported by a meta-analysis of 104 functional imaging studies (Gilbert et al., 2006). Furthermore, this is also what the current study has found. Activity related to IC was more prevalent in the dorsal medial PFC whereas belief-reasoning associated activity was more confined to ventral areas.

The left TPJ was activated by both a decoupling mechanism as well as by response inhibition. However, there was only a minor overlap between the two clusters of activation. As stated before, the exact role of the left TPJ for belief attribution and its functional distinction to its counterpart in the right hemisphere remains unclear. However, the left TPJ does not seem to constitute an area specific to belief attribution as it is also engaged in IC paradigms.

Summing up, distinct activations in the decoupling mechanism-related comparison but not in the IC-related comparison were revealed in bilateral precuneus, left thalamus, left middle temporal gyrus, left middle frontal gyrus (BAs 6, 8, 10), right middle frontal gyrus (BA 10), ventral medial frontal gyrus (BAs 9, 10), and left TPJ. All of the above areas seem to support processes related to the attribution of false beliefs that may not necessarily be important in response inhibition. A review of previous studies indicates that

distinct or more elaborate processes for false-belief reasoning when compared to response inhibition may include working memory (middle temporal and middle frontal gyrus), attentional arousal (thalamus), perspective taking (precuneus) and self-referential processes (ventral medial PFC). The role of the left TPJ for belief attribution, however, remains unclear.

7.3.2 Distinct Areas Related to Inhibition

In the IC condition, significantly increased activity related to response inhibition was revealed in a largely right-hemispheric network consisting of inferior, middle and superior frontal gyrus, medial frontal gyrus, superior and middle temporal gyrus, and TPJ. The only significantly active region in the left hemisphere was the left TPJ. When compared to activity related to a decoupling between reality and a false belief (*False-Belief > True-Belief*), response inhibition recruited distinct brain areas in the superior temporal gyrus and the right inferior, middle and superior frontal gyrus. In the middle temporal gyrus and the TPJ of the right hemisphere IC-related activity was more widespread than in areas related to a decoupling process specific for false-belief reasoning. This finding is once again shown in figure 7.2. IC-related activity is depicted in green and yellow color, activity related to a decoupling mechanism is shown in red and yellow color.

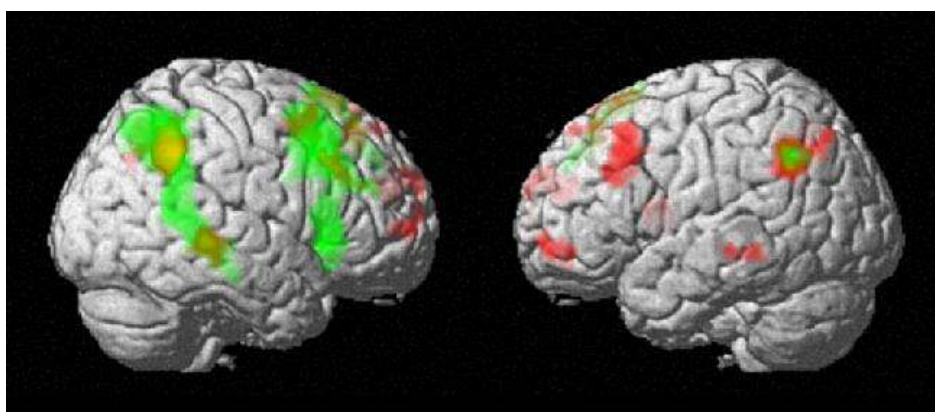


Figure 7.2. Activity in the right and the left hemisphere related to IC (green shading) and a decoupling process between reality and a false belief (red shading) as revealed in the current study. Common activity is depicted in yellow shading. Activity is depicted on a template brain at a voxel-wise threshold of $T > 4$ and a cluster-defining threshold of $p \leq .05$.

Further activity related to response inhibition was revealed in the left TPJ. Activity in the belief-reasoning task related to a decoupling mechanism was also revealed in the left TPJ. However, activity in the IC task in this region was confined to one rather small cluster of

activity (138 voxels) that was located towards the brain's surface. Activity in the corresponding belief-reasoning comparison in this region was more widespread (543 voxels) and further below the surface of the brain. There was only a small overlap of active voxels for response inhibition and a decoupling mechanism. Thus, although both contrasts of interest showed activity in the left TPJ region, the two foci of activation were situated in largely distinct areas. This finding is depicted on a coronal and a sagittal view of the brain in figure 7.3. Activity related to response inhibition is depicted in green shading. Activity related to a decoupling mechanism between reality and a false belief is shown in red. On the other hand this figure also shows that in the right TPJ activity in the false-belief related cluster was almost entirely immersed in the significant cluster related to response inhibition.

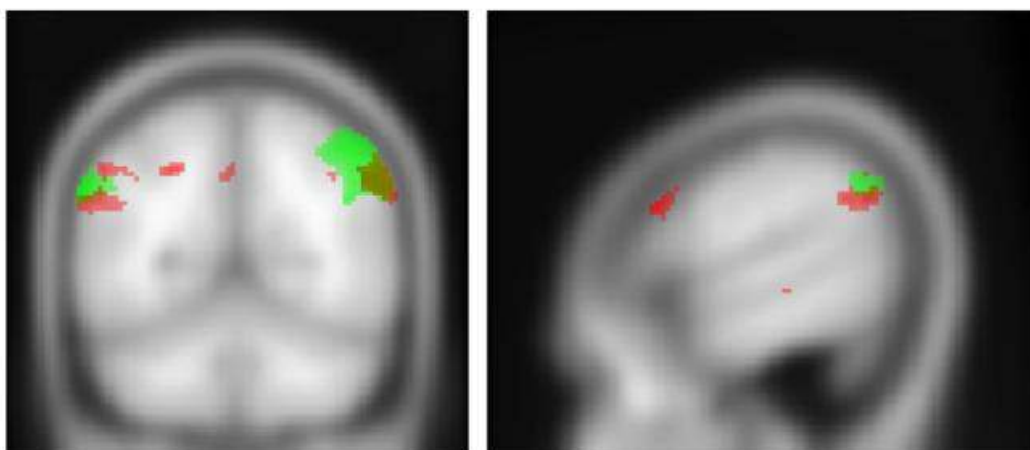


Figure 7.3. Activity related to response inhibition (green shading) and to a decoupling between reality and a false belief (red shading) in the left TPJ region on a coronal (left picture) and a sagittal view of the brain (right picture). Common activity is depicted in yellow shading. Although both cognitive concepts showed significant activations in the left TPJ, there was no substantial overlap between the two in this region. In the right TPJ, however, there was an almost complete overlap between the two respective clusters (only shown in the left picture here). Nevertheless, activity in the right TPJ was more widespread for IC.

The description given above of brain areas distinctly or at least predominantly active in response inhibition yields possible assumptions about underlying components included in response inhibition but not necessarily in the attribution of beliefs. Distinct activity for IC in the current study has been discussed in more detail in previous chapters and will only briefly be reviewed here. Taken together, previous studies indicate that all of the areas distinctly activated in IC but not for false-belief reasoning are indeed related to processes associated with various aspects of executive functioning (e.g., D'Esposito, 2000; D'Esposito, 2007).

Right superior temporal gyrus as an area distinctly activated in IC but not belief-reasoning has for instance been associated with working memory for verbal stimuli and may reflect increased verbal decoding of the picture content that may be less prevalent in the attribution of false-beliefs (Vigneau et al., 2006). Right inferior, middle and superior frontal gyrus on the other hand may reflect further activity of working memory processes or, as in the case of the inferior frontal cortex, of a relatively specialized inhibition area. These processes may constitute necessary components of IC but not of belief-reasoning. The exact nature of these processes implicated in IC but not in belief-reasoning, however, remains unclear and cannot be resolved by the present study. Although left TPJ activity was found in both the IC-related and the belief-reasoning-related comparison, there was only a small overlap between the two activated clusters. The exact role of the left TPJ in response inhibition remains unclear. While the right TPJ seems to constitute a part of a hypothesized ventral attention system, the left TPJ is to date not considered part of this system (Corbetta et al., 2008). Future research thus needs to explicitly focus on the role of the left TPJ in response inhibition.

Taken together, distinctly activated areas for IC such as the right superior temporal gyrus and the right lateral PFC seem to specifically support functions such as working memory (superior temporal and middle and superior frontal gyrus) or the inhibition of automated responses (inferior frontal gyrus). Areas with some minor overlap for activity related to belief-reasoning include the left TPJ. To date, the exact role of the left TPJ for response inhibition remains an enigma.

Next, common activity for both a decoupling mechanism between reality and a false belief as well as for response inhibition will be delineated and discussed with respect to common underlying processes involved in both cognitive concepts.

7.3.3 Common Neural Networks and Processes Involved in False-Belief Reasoning and IC

The assessment of possible common underlying processes for IC and false-belief reasoning is based on this study's finding of common activity in both concepts. As previously stated, an overlap of activation for the comparison *No-go* > *Go* and *False-Belief* > *True-Belief* was revealed in right middle temporal gyrus, right TPJ and dorsal medial PFC. The current findings and previous data show that the above areas support cognitive functions that are part of both the ability to inhibit responses as well as of the ability to attribute a false belief.

Right Middle Temporal Gyrus

Right middle temporal gyrus as a common region for both false-belief reasoning and response inhibition may reflect a process that is necessary in both concepts. As previously described in more detail, the middle temporal gyrus seems to be engaged in tasks requiring encoding and storage processes into verbal working memory. As subjects may try to improve their performance in both tasks by using a verbal rehearsal strategy, the common middle temporal gyrus activity may reflect operations of verbal working memory. Several studies also indicate that the superior temporal sulcus region encompassing the middle temporal gyrus is recruited when perceiving and analyzing biological motion as well as static images of face and body (Allison et al., 2000). All conditions presented in this study have used images containing at least outlines of human faces and bodies. Also, these images depicted still images of humans performing goal-directed motions. This was also true for the IC conditions which also showed images of humans in order to keep the visual input closely matched to its corresponding belief-reasoning condition. Although activity in the Go condition was subtracted from the No-go condition, No-go stimuli may have been analyzed more deeply due to their task relevance. In the belief-reasoning experiment, the false-belief stimuli may also have been processed more elaborately than the stimuli presented in the true-belief task. This may have been due to the fact that subjects dedicated more attentional resources towards this task as it additionally required the representation of a false belief. Taken together, middle temporal gyrus activity in both the IC and the belief-reasoning comparisons may also reflect processing of specific human facial and bodily properties as well as the intended biological movements contained therein. Possibly, such processing of facial and bodily features may be an automatic process and not necessarily be a crucial component of either mental state attribution or response inhibition. Middle temporal gyrus activity in these tasks may only be revealed if these tasks contain depictions or descriptions of faces and bodies. This question, however, remains open to further investigation. Also, it can only be speculated here whether using a response inhibition task without depicting humans may have resulted in a similar result. In the current study this was knowingly avoided. Differing brain activity could otherwise be attributed to largely differing stimulus properties and not to differences between the cognitive mechanisms investigated. Summing up, both inhibitory control and belief-reasoning seem to rely on processes of working memory mediated by the right middle temporal gyrus. Possibly, this activity might also reflect analyses of face and body features as well as of biological motion. This analysis may be automatic and not be a crucial component of either false-belief reasoning or IC.

Right TPJ

Common activity for both the belief-reasoning and the IC comparison was also revealed in the right TPJ (figure 7.4). Converging evidence suggests that the right TPJ may be part of an attention reorienting system of the human brain (Corbetta et al., 2008). Attentional reorienting also seems to be a crucial component of the ability to attribute false beliefs. Response inhibition seems to rely on this attention reorienting system as well.

Evidence that the right TPJ may support attentional reorienting comes from studies that required the subjects to reorient their attention to behaviorally relevant stimuli. In one of these studies, for example, the location of an upcoming target stimulus was cued in one condition whereas in another condition the target stimulus appeared at an unexpected location that had not been cued beforehand (Kincade et al., 2005). Compared to the cued target condition, the uncued target condition requiring attentional shifting to a behaviorally relevant stimulus resulted in significantly increased activity in the right TPJ. Interestingly, the right TPJ was not modulated in additional conditions that presented stimuli that were irrelevant to solving the task. This study and others thus suggest that the right TPJ plays a crucial role in the reorienting of attention to exogenous stimuli. However, the right TPJ is only engaged if these stimuli are behaviorally relevant (Corbetta et al., 2008). As previously discussed, attention reorienting seems to be required for successful false-belief reasoning. In the false-belief task additional attention has to be directed to the events presented in this study's second story picture. While in the true-belief task internal predictions based on the state of reality can be made about the story outcome in the third story picture, the false-belief task also requires that the subjects represent a protagonist's false-belief. The information that one of the protagonists may possess a false belief is given to the subjects by the information that at the time of object transfer only one character is present in the room. This information thus represents an exogenous stimulus that is clearly of behavioral relevance to the subject. While the current true-belief task may possibly also require attentional reorienting to some extent (as reflected in right TPJ activity in the comparison *True-Belief* > *Baseline*), the demands on the reorienting of attention towards a protagonist's false belief are probably more elaborate.

Furthermore, the reorienting of attention to a behaviorally relevant stimulus seems to constitute a crucial part of response inhibition. Due to the majority of Go trials subjects tend to emerge with an almost automated routine in responding by the press of a button to the stimuli presented. This routine is only interrupted by the appearance of a No-go stimulus. The No-go stimulus thus represents an exogenous stimulus of behavioral relevance as it requires the withholding of a motor response. Therefore, the No-go

condition can be viewed as a task that may rely in large parts on the reorienting of attention towards a behaviorally relevant stimulus.

Summing up, both false-belief reasoning and response inhibition require an attentional shifting to exogenous and behaviorally relevant stimuli. This common process is likely to be supported by the right TPJ.

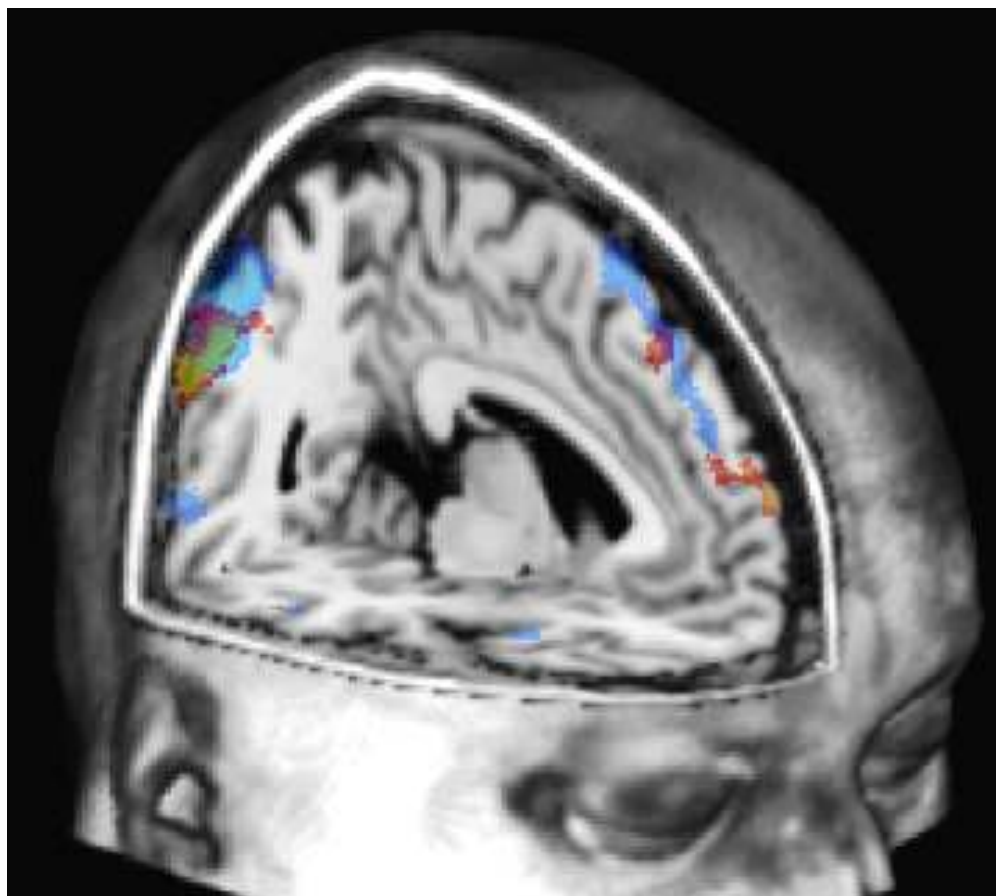


Figure 7.4. Activity related to response inhibition (blue shading) and a decoupling between reality and mentality (orange shading) in the right parietal and medial prefrontal cortices. Significantly increased activation as revealed in the present study is depicted on a template brain contained in MRICro software (www.mricro.com). This figure shows common activity for both mechanisms in the right TPJ (BA 39, 40) and dorsal medial PFC (BA 6, 8). Distinct activity for a decoupling mechanism was revealed in the ventral medial PFC (BA 10).

Dorsal Medial PFC

Last but not least, common activity for false-belief reasoning and IC was also revealed in dorsal medial PFC in BAs 6 and 8 (figure 7.4). Independent studies investigating belief-reasoning or IC have also found activity in this area (e.g., Gallagher et al., 2000; Garavan

et al., 2006; Sommer et al., 2007; Stevens et al., 2007). Activity of the dorsal part of the medial PFC has been related in numerous studies to cognitive processes such as error detection, motor response selection, conflict detection and inhibition (e.g., Mostofsky et al., 2003; O'Connell et al., 2007; Rubia et al., 2003; Simmonds et al., 2007). It is assumed that these processes are actually closely related (Simmonds et al., 2008). Taken together, it can be assumed that the dorsal medial PFC is involved in conflict detection and subsequent inhibition of a prepotent response (Aron, 2007).

Conflict detection and the inhibition of a prepotent answer are clearly the core mechanisms of response inhibition as captured in the present Go / No-go comparison. This task requires subject to first detect the conflicting No-go stimulus and to then inhibit a prepotent tendency to respond by pressing a button.

The false-belief reasoning task may also require some sort of conflict detection as well as a possible inhibitory component. As has been delineated in Leslie et al. (2004, p. 528), "...people's mundane beliefs are usually true, the best guess about another person's belief is that it is the same as one's own." Accordingly, subjects will have a prepotent tendency to respond even to the false-belief task according to the true state of reality. However, in order to respond to this task correctly, subjects have to detect the conflicting belief presented to them, namely that one of the story protagonists is not present in the room and that he or she thus possesses a false belief about reality. After having detected this circumstance subjects may subsequently need to inhibit their prepotent tendency to respond according to their own true belief equaling to the true state of reality. It is thus likely that the dorsal medial PFC activity as revealed in the present study may represent operations of an inhibitory control process.

Summing up, both belief-reasoning and IC at least in adulthood seem to rely on a common error detection and inhibition process which may be mediated by dorsal medial PFC.

In conclusion, the hereby presented discussion of common underlying processes for false-belief reasoning and response inhibition was based on the finding of common activity in three major brain regions. A close look at the possible roles of these prominent brain areas indicates that both false-belief reasoning and response inhibition may depend on basic processes such as working memory of possibly the verbal domain, attention reorienting to behaviorally relevant stimuli in the environment as well as conflict detection and subsequent inhibition of a prepotent response.

7.4 Integration of Current and Existing Data towards a Unified Theory for the Belief-Reasoning / IC Relation

In this following section this study's findings in light of previous data will be used to characterize a hypothesized connection between IC and belief-reasoning. As this study's results are based on the data of healthy adults, the current findings primarily apply to this relationship in adulthood.

Based on the current findings and previous studies from all fields of neuroscience, the components of both a false-belief reasoning and response inhibition network may have become clearer. These components and processes will be delineated in more detail in the following section and are depicted schematically in the brain's right and left hemisphere in figures 7.5 and 7.6, respectively. Components related to a decoupling between reality and a false belief are depicted in red. IC-related components are shown in green. Components shared by both networks are displayed in yellow ellipses.

Although IC as investigated in the presented Go / No-go paradigm may already constitute a rather basic cognitive process, this process may nevertheless be made up of several underlying components. As the involved brain areas suggest, IC may rely on working memory operations, the reorienting of attention, conflict detection and a specialized inhibitory mechanism.

The decoupling between mentality and reality as the crucial component of false-belief reasoning on the other hand may rely on increased attentional arousal, working memory operations (differing in part from those involved in IC), attentional reorienting, conflict detection, inhibitory mechanisms, and self-referential processes including perspective taking.

The present data also suggest that the neural networks for false-belief reasoning and IC rely on common components. These components may support processes such as attention reorienting towards stimuli in the environment, distinct working memory operations, and specific mechanisms involved in conflict detection and subsequent inhibition.

Based on these findings, the following paragraph represents an attempt to integrate these findings with existing results on the IC / belief-reasoning connection in adulthood.

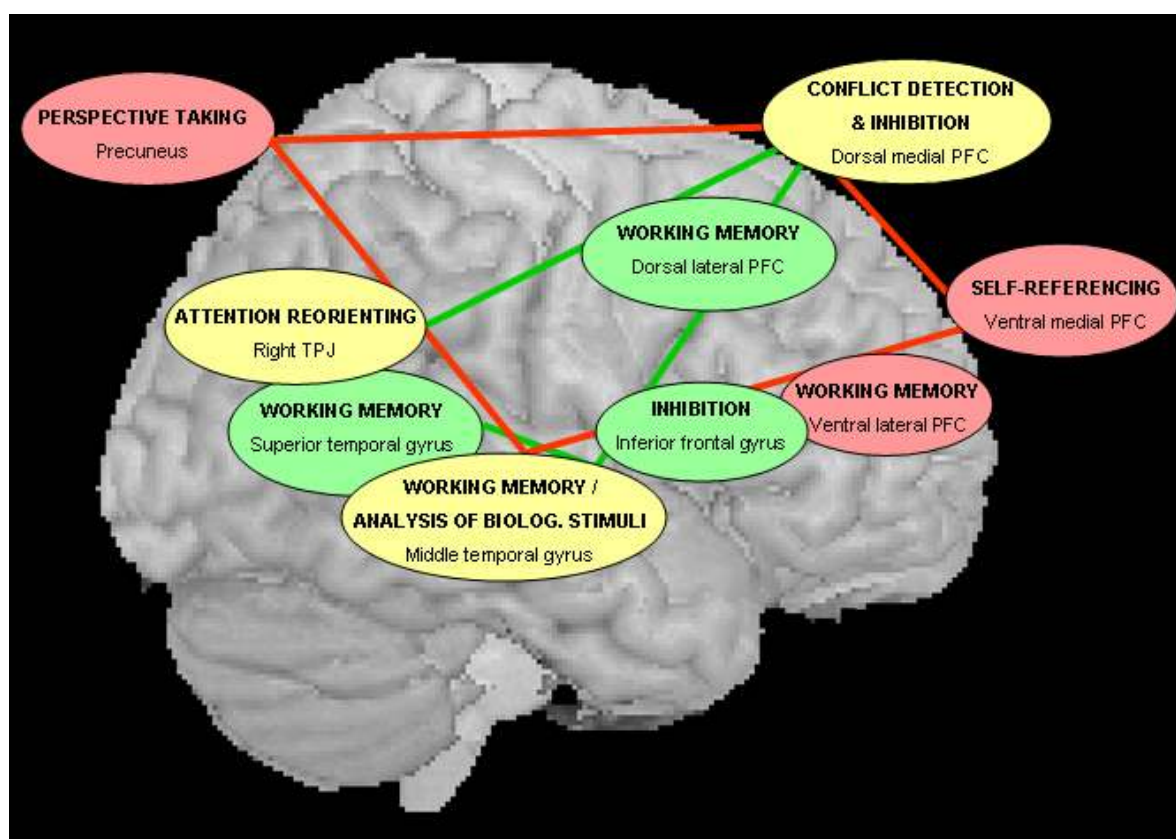


Figure 7.5. Schematic outline of the hypothesized neural networks for a decoupling between reality and a false belief (red color) and response inhibition (green color) in the right hemisphere based on the findings from the current study. Yellow ellipses depict common processes and structures possibly shared by both cognitive concepts.

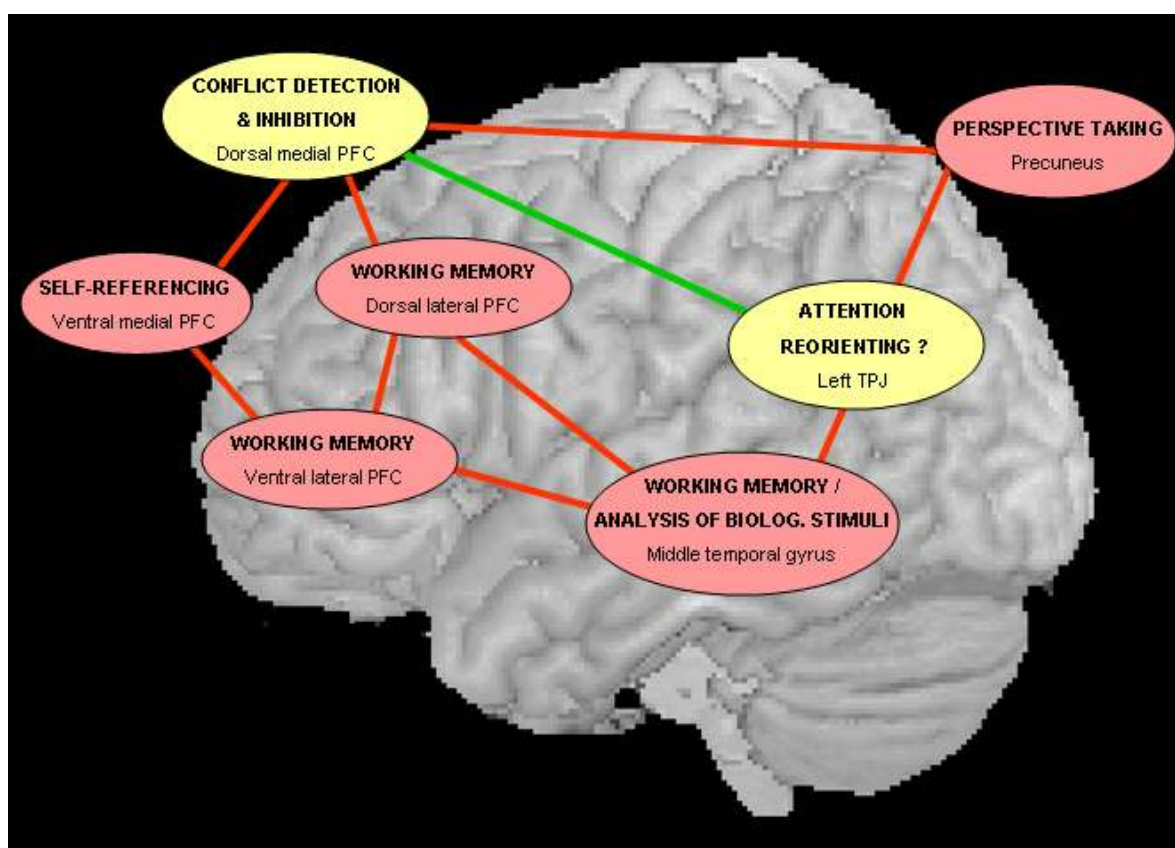


Figure 7.6. Components and hypothesized functions of the neural networks for inhibitory control and a decoupling between mentality and reality (red shading) in the brain's left hemisphere. Yellow ellipses represent components shared by both processes.

It has been assumed that working memory may be at the core of the observed IC / Belief-Reasoning relationship. Based on previous findings, however, it seems unlikely that working memory only may explain the correlation between belief-reasoning and inhibition. As several studies have revealed, working memory does not show any substantial correlations with belief-reasoning (Perner & Lang, 1999). Albeit in the present study an overlap for both concepts was found in the middle temporal gyrus that has previously been associated with working memory. This, however, may have been due to the middle temporal gyrus's also described role in the analysis of facial and bodily features. Since the present Go / No-go task contained depictions of people, this may have caused an automatic processing of the facial and bodily features contained therein. Middle temporal gyrus activity in the IC task may thus be due to the analysis of facial and bodily features and may not necessarily constitute a crucial part of the ability to inhibit responses.

Attention reorienting on the other hand clearly seems to constitute a component of both false-belief reasoning and IC. Both concepts require the switching of attention to a stimulus in the environment. As far as IC is concerned, this outside stimulus is represented by the No-go stimulus requiring the withholding of a response. The exogenous stimulus represented in false-belief reasoning on the other hand may be comprised of the events presented in the environment. This observation of outside events (e.g., object transfer with one of the protagonists not present) then subsequently leads to the forming of a representation of false belief. Thus, attention reorienting mediated by the TPJ may be a common component in both concepts and could explain the close relationship of both.

Another area implicated in both cognitive processes is dorsal medial PFC, which seems to support broader functions such as conflict detection and possibly a specialized inhibitory component. In the No-go condition, the No-go stimulus has to be detected and identified as a conflicting stimulus and a subsequent inhibition of a motor response has to be initiated. In the false-belief task on the other hand, it has to be detected that an event leads to false belief which also conflicts with the true belief (i.e. the representation of reality). The false belief may represent a conflicting stimulus as subjects may by default assume that beliefs are generally true. After detecting this conflict the representation of reality and a motor response according to this state has to be inhibited. Taken together, there is some indication that the relationship between IC and belief-reasoning is also mediated by a cognitive process involving conflict detection and possibly initiating subsequent inhibition.

The explanation attempts presented above are only partially included in so called expression theory which has also attempted to explain the connection between IC and belief-reasoning in adulthood (Kloo & Perner, 2003; Perner & Lang, 1999; Siegal & Varley, 2002; Sodian & Hülken, 2005). Due to conflicting results, expression theory

provides little explanatory power for behavioral findings in childhood (Hughes, 1998; Perner et al., 2002a; Perner & Lang, 1999). According to expression theory, belief-reasoning tasks require the suppression of the true state of affairs in favor of one's own or others' mental states by means of IC (Hughes & Russell, 1993; Russell et al., 1991). For adults, the results from this study however suggest that in addition to a possible role of a conflict detection and inhibition process, there may also be another process at the core of this relationship: attention reorienting. This breaking down of components included in both concepts may be able to explain phenomena that have in the past not sufficiently been accounted for by expression theory. For instance, it has been shown that ADHD patients with typically low IC abilities do not show a substantially poor performance in false-belief tasks. However, expression theory would predict a significant ToM impairment under the assumption that IC as a whole could be part of mental state attribution. By splitting up the components involved in both IC and belief-reasoning and thus finding common components involved in both concepts, behavioral phenomena as in the case of unimpaired belief-reasoning with compromised IC may be explained. Such an attempt in the case of ADHD is presented in one of the following sections.

As previously mentioned, conclusions concerning the nature of the belief-reasoning / IC connection in childhood are limited. As some researchers have reported stronger correlations between the two concepts in childhood than in adulthood (Chasiotis & Kiessling, 2004), it may be that the differing components implicated in both concepts may be of differing importance in childhood. Previous studies indicate that children under the age of about 3 years may not possess any concept of belief at all, thus maybe not sharing any common processes at all due to undeveloped networks for belief-reasoning. This view is backed by the observation that these children answer false-belief questions at chance level and not always with the true state of reality (Wellman et al., 2001). This contradicts accounts that the children's failure in false-belief tasks may be due to a lack of inhibition of the real state of affairs. In the further course of development children are aided in false-belief tasks when decreasing the tasks' inhibitory demands (Wellman et al., 2001). This indicates that at an intermediate level of false-belief performance only (and thus at the beginning stages of forming a concept of false belief) children may rely more heavily on conflict detection and inhibitory processes. In the further developmental course into adulthood false-belief reasoning may become more automatic and rely even less on conflict detection and inhibition. The remaining yet weaker correlation between IC and belief-reasoning in adulthood may be explained by the finding that both cognitive abilities may also rely on a common component responsible for reorienting attention towards stimuli of importance in the environment.

Summing up, two basic cognitive processes may be at the core of the frequently observed relationship between IC and belief-reasoning. One of these processes is attention reorienting towards a behaviorally relevant stimulus in the environment. This process seems to be supported by the right TPJ, possibly also involving the left TPJ. Another process that mediates this relationship may also be conflict detection and the initiation of inhibition. The neural correlate of this process seems to be situated in the dorsal medial PFC.

The subsequent section discusses the impact of the here presented findings on accounts assuming the existence of highly-specialized modularized areas for belief-reasoning in the brain.

7.5 Impact of the Current Findings on Modularity Accounts of Belief-Reasoning

An area in the brain needs to fulfil two crucial requirements in order to qualify as a specific module for belief attribution: generality and specificity. This implies that such an area is engaged in any task requiring the attribution of beliefs and that this area is not recruited in any other distinct cognitive process. Two areas have so far been discussed as possible ToM modules in the brain: the right TPJ and the medial PFC (Frith & Frith, 2003; Saxe et al., 2004).

The right TPJ has for quite some time been considered as specific for the attribution of mental states especially by a group of researchers around Rebecca Saxe at the Massachusetts Institute of Technology. This research group reported to have found in various studies that the right TPJ is recruited in mental state attribution but not during control tasks that do not require mental state attribution. Although other studies have also frequently reported right TPJ activity as part of a neural network for belief-reasoning, these studies have not argued that this region may necessarily be specific to the attribution of beliefs (e.g., Gallagher et al., 2000; Gobbi et al., 2007; Sommer et al., 2007). Taken together, the right TPJ does therefore indeed seem to meet the generality criterion which has to be considered indispensable for this region to qualify as a specific ToM module. However, ample converging evidence exists that the right TPJ does not meet the requirements of the specificity criterion in order to qualify as an independent ToM module. As Mitchell has pointed out, researchers investigating the neural correlates of mentalizing have "...neglected substantial evidence that this region may also subserve a set of attentional processes that are not specific to social contexts." (Mitchell, 2007; p. 262). A review of studies investigating response inhibition and other concepts related to attentional processes shows indeed that most of these report activity of the right TPJ region. This view is also corroborated by a meta-analysis that has focused specifically on the right TPJ

(Decety & Lamm, 2007). This extensive meta-analysis found virtually identical right TPJ activations for ToM and attention reorienting tasks (see figure 4.1). Also, a single study investigating the neural correlates of belief-reasoning and attentional reorienting found overlapping activity in the right TPJ, thus refuting accounts of this area to be specific for mental state attribution (Mitchell, 2007). Ultimately, the current study was also unable to confirm accounts of the right TPJ as a ToM module. In the present study the right TPJ was recruited by both IC and false-belief reasoning as reflected by largely overlapping activity (e.g., figure 7.3). Thus, the right TPJ does not meet the specificity criterion necessary for this region to qualify as a specific module for belief attribution. Overwhelming evidence corroborated by the current findings suggests that the right TPJ is not specific for the attribution of beliefs. As Mitchell (2007) has pointed out, the issue of the right TPJ as a possible specific area shows parallels to the fusiform face area, which was in the past assumed to be a specific area for the detection of faces. However, recent evidence suggests that the fusiform face area is also recruited when subjects are presented with non-face stimuli that the subjects have become accustomed to (e.g., Bukach et al., 2006). The right TPJ may be doomed to the same fate. As delineated above, the right TPJ may not be specific to the attribution of mental states but support a broader function in terms of attentional reorienting.

Next to the right TPJ, the medial PFC has also been dubbed a candidate region for a specialized ToM module. As Frith & Frith (2003) have suggested, the medial PFC is activated in tasks that require subjects to attend to mental states of the self or others. This view is confirmed by the fact that most imaging studies investigating mental state attribution have indeed found activity in the medial PFC. Unfortunately, Frith & Frith do not explicitly specify as to which part of the medial PFC their assessment is aimed at. The medial PFC is a large and heterogeneous anatomical area with various possible foci of activation. However, the rather unspecific reference to “the medial PFC” may also have been due to the largely differing locations of activity throughout the medial PFC in studies investigating mental state attribution. A thorough review of relevant literature however reveals that activity attributed to mental state attribution is usually confined to the ventral medial PFC. Activity in the dorsal medial PFC on the other hand has for the most part been associated with inhibitory processes (Simmonds et al., 2008). In the present study there was a significant overlap of activity in the dorsal medial PFC for IC and belief-reasoning. This finding and previous reports thus indicate that at least the dorsal medial PFC is not specific for mental state attribution. Although this dorsal part of the medial PFC fulfils the generality criterion required to qualify as a specific ToM module, it does not fulfil the specificity criterion due to its role in tasks with no mentalizing content such as response inhibition.

Hypotheses regarding the functions of the ventral medial PFC (BA 10) are much more vague. The majority of studies investigating mental state attribution report increased activity in this area for mental state attribution (Gilbert et al., 2006). Also in the present study, this area exhibited increased activity in the belief-reasoning comparison. This finding indicates that the ventral medial PFC may indeed show a general recruitment during the attribution of beliefs. However, is this area also specifically recruited for the attribution of beliefs? Indeed, the IC-related comparison of the study presented here did not show any increased activity in this area. Other accounts of activity in this area for cognitive processes other than mental state attribution are scarce. However, ventral medial PFC has previously been reported in cognitive mechanisms that could also be a crucial part of the ability to attribute beliefs. For instance, a meta-analysis conducted by Northoff et al. (2006) revealed that an area in the ventral medial PFC similar to the area found in the current and previous studies might support self-referential processes of the verbal domain. Results of this meta-analysis of self-referential tasks are depicted in figure 7.7.

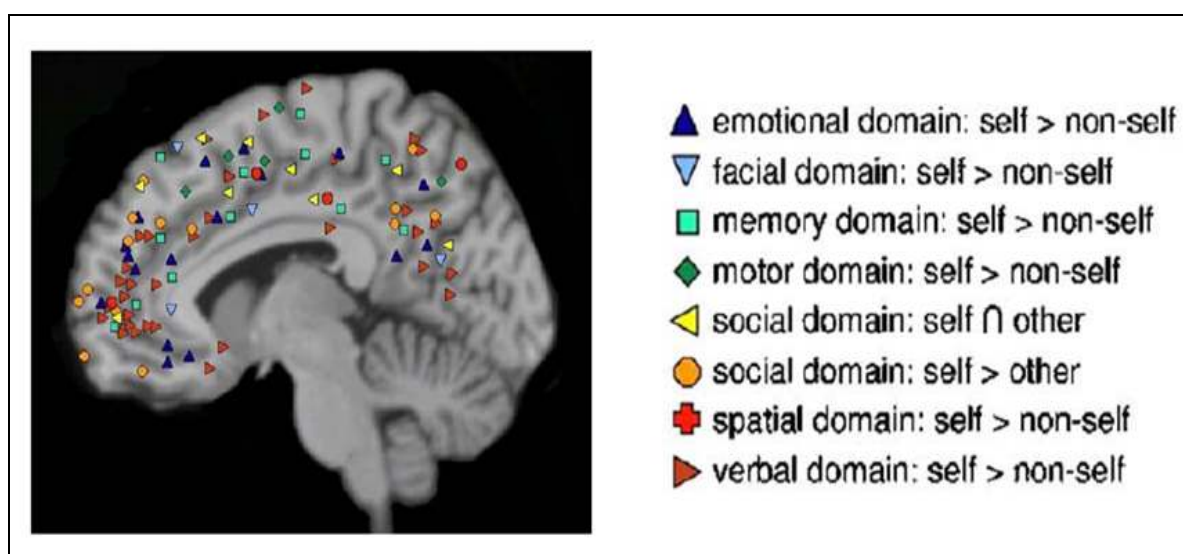


Figure 7.7. Activity as revealed in several imaging studies investigating self-referential processes. Picture taken from Northoff et al. (2006). Reprinted with permission from Elsevier.

However, most of the studies included in Northoff et al.'s meta-analysis may have also included mental state attribution of some sort. This would also imply that self-referential processes may be either largely equivalent to mental state attribution or at least be a crucial part of it. If one considered processes of self – non-self distinction a separate ability also necessary for other cognitive abilities independent of mental state attribution, then this would imply that the ventral medial PFC cannot be specific for mental state attribution.

If on the other hand self-referential processes were considered a necessary prerequisite for mental state attribution only (and not for other concepts), this would in turn allow the hypothesis that ventral medial PFC might actually constitute a specific ToM module. Other studies, however, have also reported activity in ventral medial PFC in tasks that did not require any mental state attribution. A study investigating executive functioning in autistic patients and healthy controls revealed that the ventral medial PFC may be implicated in the switching between stimulus-oriented versus stimulus-independent attention (Gilbert et al., 2008). This switching process related to self-referential processes hints that these basic processes may indeed also support other cognitive mechanisms independent of mental state attribution, thus refuting accounts of this area as possible module for mental state attribution.

Taken together, solid evidence corroborated by the data from the present study indicates that the right TPJ is not specific for mental state attribution. Furthermore, the dorsal medial PFC also does not seem to represent a neural module specific for the attribution of beliefs. This view is based on previous findings as well as on the current study's results. The ventral medial PFC on the other hand does not play a role in response inhibition and may support self-referential processes that are a crucial part of mental state attribution and possibly of other cognitive mechanisms. If not the neural correlate of a specific ToM module, this area could be specific to self-referential processes. The exact role of the ventral medial PFC with respect to belief attribution needs to be further investigated.

The following chapter focuses on the impact of the current findings for instances of compromised belief-reasoning as is the case in developmental disorders such as autism and ADHD.

7.6 Implications for the Understanding and Treatment of Compromised Belief-Reasoning

7.6.1 Autism

Autism is a developmental life long disorder with a pronounced impairment in social functioning. It has widely been acknowledged that autistic patients show compromised belief-reasoning abilities in experimental tasks as well as in daily life (e.g., Baron-Cohen et al., 1985; Ozonoff et al., 1991; Sodian & Frith, 1992). Interestingly, this deficit in ToM reasoning is accompanied by an impairment in executive functioning (e.g., Domes et al., 2008; Pennington & Ozonoff, 1996; Sinzig et al., 2008). Working memory as a part of

executive functioning, however, seems to be relatively unimpaired in autistic individuals (Ozonoff & Strayer, 2001).

Neuroimaging studies investigating ToM reasoning in autistic patients are relatively scarce. However, two studies have shown that autistic individuals compared to healthy controls show less extensive activations especially in the ventral medial PFC when solving ToM tasks (Happe et al., 1996; Nieminen-von Wendt et al., 2003). While Happe et al.'s study investigated the attribution of intentions, Nieminen-von Wendt et al.'s study focused on the attribution of beliefs. A subsequent regions-of-interest approach conducted in this study revealed less belief-reasoning related activity in the ventral medial PFC. This result is also depicted in figure 7.8.

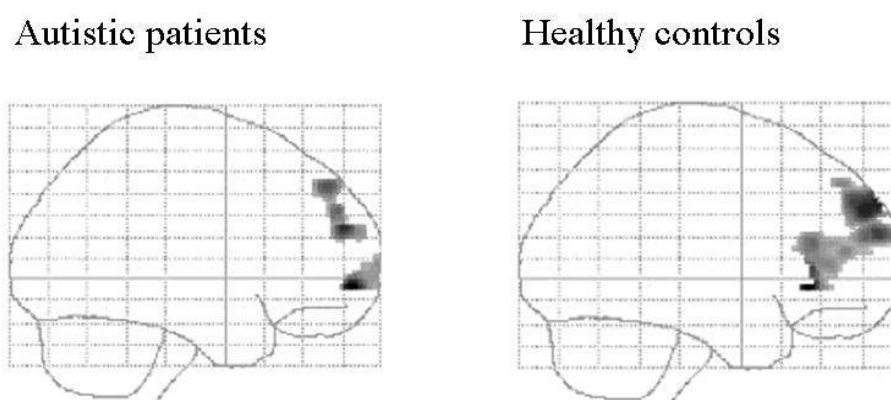


Figure 7.8. Activity in the ventral medial PFC in the comparison *ToM Stories > Physical Stories* in 8 autistic patients and 8 healthy controls in a study conducted by Nieminen-von Wendt et al. (2003). Image is taken from Nieminen-von Wendt et al. (2003). Reproduced with friendly permission from Springer.

Another study conducted by Kennedy et al. (2006) investigated resting state activations in autistic individuals. Compared to healthy controls, autistic controls showed pronounced dysfunctional activations particularly in the ventral medial PFC. Furthermore, measures of social impairment in autistic individuals were highly correlated with the degree of neurofunctional abnormality in the ventral medial PFC. Patients with higher social impairment scores showed higher levels of abnormal brain functioning in ventral medial PFC.

The above results are in line with the findings of the current study showing that ventral medial PFC may constitute a crucial component of a neural belief-reasoning network. The ventral medial PFC seems to support functions related to self-referential processes and may as such be considered a necessary prerequisite for false-belief attribution. The current findings thus indicate that autistic individuals may be impaired in

their ability to attribute beliefs due to a dysfunction of self-referential processes mediated by the ventral medial PFC. Although this may explain the observed ToM reasoning impairment in autism, it does not sufficiently explain why this impairment is accompanied by poor executive functioning abilities. This circumstance could possibly be accounted for by assuming that not only may the ventral medial PFC show dysfunctional activity in autism but that this dysfunction could also apply to more dorsal areas of the medial PFC. As has been widely acknowledged, the dorsal medial PFC is engaged in response inhibition (e.g., Simmonds et al., 2008). This is also what the present study has found. There is evidence that this region of the medial PFC may support conflict detection as well as an actual inhibitory response (Aron et al., 2004). The present study also revealed that both false-belief-reasoning and response inhibition recruit the dorsal medial PFC. It is thus argued here that the dysfunctional activation observed in autism may not be restricted to the ventral medial PFC only. Possibly more widespread areas of the medial PFC may be affected, thus resulting in the belief-reasoning and executive functioning impairment observed. While the executive dysfunctioning in autism may be explained by an impairment of a conflict detection and inhibition component located in dorsal medial PFC and maybe DLPFC, the observed belief-reasoning deficit in autism may be mediated by an impairment in the ventral medial PFC. Furthermore, an impairment in dorsal medial PFC may also contribute to the observed belief-reasoning problems in autism, as the present study was able to show that dorsal medial PFC constitutes a common component of the neural networks for belief-reasoning and response inhibition.

It is tentatively argued here that taken together, the observed impairment in belief-reasoning and response inhibition in autistic individuals may be mediated by the medial PFC. The present study revealed that both response inhibition and false-belief-reasoning recruit the medial PFC. This finding is in line with previous results indicating dysfunctional medial prefrontal patterns of activation in these cognitive abilities in autistic patients.

The next section discusses ADHD in light of the current results. Although patients with ADHD show pronounced difficulties in their overall executive functioning, they seem to be relatively unimpaired in their ability to attribute beliefs. This observation also indicates that belief-reasoning abilities in autism may not be due entirely to difficulties in executive functioning.

7.6.2 ADHD

ADHD is a disorder with an impairment in executive functioning abilities such as inattention, impulsive behavior, hyperactivity, distractibility, and an impairment in inhibitory control (Sodian et al., 2003). However, belief-reasoning abilities in individuals with this

disorder seem to be relatively unimpaired. There is evidence that children diagnosed with ADHD are only impaired in belief-reasoning tasks containing relatively high inhibitory demands (Sodian & Hülshen, 2005).

As the present and previous studies suggest, response inhibition as one of the executive functioning abilities affected in ADHD seems to recruit right-lateralized fronto-parietal regions of the brain. Furthermore, response inhibition recruits areas in the dorsal part of the medial PFC. A large body of studies with ADHD patients suggests that among other areas, the dorsal medial PFC exhibits less activation compared to healthy controls in response inhibition tasks (e.g., Sinzig et al., 2008). An ALE meta-analysis including 16 neuroimaging studies on executive functioning revealed significantly decreased activity for ADHD patients in dorsal medial PFC when compared to healthy controls (Dickstein et al., 2006). A review of relevant studies also suggests that more ventral medial PFC regions are largely unimpaired. These and the present results indicate that patients with ADHD may not exhibit pronounced belief-reasoning difficulties because the disorder may not affect areas in the brain that seem to be crucial for belief-reasoning. One of these areas possibly not affected in ADHD seems to be the ventral medial PFC mediating self-referential processes. While this process may be crucial for belief-reasoning, the current findings indicate that this process is not implicated in response inhibition.

Although dorsal medial PFC on the other hand is recruited by both belief-reasoning and response inhibition in healthy adults and although it shows dysfunctional activity in ADHD, it may not constitute a crucial component for belief-reasoning. As there are additional areas of the brain implicated in conflict detection and inhibition, the activation of these areas may be sufficient to provide the necessary conflict detection and inhibition resources also required for belief-reasoning in ADHD. It is likely that only belief-reasoning tasks with high inhibitory demands require the additional recruitment of the dorsal medial PFC, thus explaining the observed difficulties for ADHD patients in tasks with high inhibitory demands.

Summing up, areas such as the ventral medial PFC that are crucial for belief-reasoning may not be affected in ADHD and thus explain the relatively unimpaired belief-reasoning capacities despite an impairment in executive functioning. Although ADHD seems to affect areas implicated in belief-reasoning and IC, such as dorsal medial PFC, these areas do not seem to be a crucial component for belief-reasoning and may be compensated by other areas. Autism on the other hand seems to show a dysfunction in a crucial area for belief-reasoning, namely the ventral medial PFC. This area seems to support self-referential processes. As the dysfunctional engagement of the medial PFC in autism may also extend to the dorsal medial PFC, this could explain the frequently described

impairment of executive functions in this developmental disorder. Without doubt this assessment of possible neural underpinnings of ADHD and autism is highly speculative and needs further and more sophisticated research. Future research also needs to investigate the engagement of other areas of the brain and explore how these different areas work in common.

The following section describes a training program for young children in social skills that seems to have an effect on executive functioning and behavioral problems. This practical application of existing findings in light of previous and the current findings is critically assessed. Furthermore, the following section discusses possible applications of the current findings with respect to developmental disorders.

7.6.3 The Papilio Project

The Papilio project is a program aimed at reducing behavioral problems and fostering social and emotional abilities in kindergarten children. The program was developed in 2002 and subsequently applied in several German kindergartens. First results have previously been released and will be discussed in this section (Scheithauer et al., 2008).

The idea that training in social skills but also in executive functioning skills may possibly show benefits for future social functioning as well as towards reducing behavioral problems is based on a large body of research. As has been delineated in detail throughout this whole thesis, data stemming from behavioral and neuroimaging studies have revealed that ToM reasoning as a crucial component for social functioning and executive functioning as the key compromised set of functions in behavioral disorders such as ADHD are indeed closely related. Furthermore, it has also been shown in young children that training in an IC task improves performance in a ToM task and vice versa (Kloo & Perner, 2003). The present study also revealed that IC and belief-reasoning are indeed closely related on a neural basis. Both processes engage neural networks that show a surprising overlap in a number of key regions. As has been argued in the thesis presented here, these common areas seem to support basic functions such as attentional reorienting or conflict detection and subsequent inhibition.

The core training in the Papilio project involves three measures aimed at improving the three to six year olds' skills in attributing emotions, desires and beliefs as well as improving further social skills and fostering moral behavior. However, the training program does not include a specific training in executive functioning.

One of the measures included in the program involves stories that are enacted by dolls impersonating goblins. These goblins engage in social interactions that entail emotions such as anger, guilt, happiness, or sadness. Children are also taught about

facial and bodily gestures associated with certain emotions. Furthermore, the children are instructed about possible actions necessary whenever they encounter someone exhibiting these emotions. As the stories also depict the characters acting on the basis of intentions and beliefs, this measure is aimed at improving the children's abilities to attribute emotions, intentions and beliefs to themselves and others.

Another measure contained in the Papilio training program is the "toys are on vacation game". This training requires children to come up with creative ideas on how to play games with other children without having any toys available. Although this training is supposed to foster social relations with others, it may also require children to recruit executive functions such as the ability to plan ahead.

The last training measure contained in the program is the "good behavior game" that teaches children about social rules and ethical behavior. Such behavior is then reinforced throughout a period of two weeks by means of positive reinforcement. This training method is thus aimed at fostering ethical behavior but may also help children in understanding the effects of their actions on others. As ethical behavior seems to rely on processes also included in mental state attribution (e.g., Young & Saxe, 2008), this training may also improve ToM reasoning in the children.

In the Papilio program, the children's social and emotional skills as well as behavioral problems were assessed prior to the training phase and 4 months after the beginning of the training by means of standardized questionnaires given to the children's trainers. These results were compared to the data of children who were randomly assigned to a waiting control group. The subsequent analysis included the data of 716 three to six year old children from 25 kindergartens. The results revealed a significant decrease in the training group when compared to the control group in terms of behavioral problems and hyperactivity and attention deficit syndromes. Also, the children of the training group showed higher levels of ethical behavior after completion of the training program. The study revealed no difference between the groups after the training phase in terms of aggressive behavior or emotional skills (<http://www.papilio.de/download/papilio-ergebnisse.pdf>).

Taken together, the results of the Papilio program as an example of a training program in social and emotional skills did find a positive effect, particularly on behavioral problems. These results indicate that a standardized training program for children in social skills may actually also improve executive functioning as a key component in behavioral and attention deficit problems. Although the scientific evaluation of the program presented here was subject to a number of methodological flaws (e.g., children were assessed by the trainers, statistical analyses, etc.), it also hints that executive functioning and ToM reasoning may be closely related. Furthermore, it shows that training programs may

improve children's executive functioning and ToM reasoning alike. When taking the findings presented in this thesis into account, it seems that this effect may be mediated by common mechanisms shared by both functions. These include attention reorienting and conflict detection and inhibition. As the present study suggests, the latter mechanisms may be mediated by the bilateral TPJ and the dorsal medial PFC, respectively. Although the findings from the present study stem from an adult sample, there is ample evidence that in children the observed connection may even be stronger (Chasiotis & Kiessling, 2004) and that the current findings thus also apply to the Papilio program's results.

Future training programs with children may thus be improved by especially focusing on the training of tasks that require attention reorienting abilities as well as conflict detection and inhibition. Such training may prove beneficial for future executive functioning, a decreased likelihood of behavioral problems such as ADHD, and enhanced ToM abilities which could in turn result in improved ethical behavior.

While the Papilio program is aimed at protecting healthy children from behavioral problems and teaching social skills, there are also several training programs for children exhibiting existing behavioral problems as well as an impairment in social functioning (e.g., Ozonoff & Miller, 1995; Sinzig et al., 2008). These programs, which entail a number of different training techniques, comprise methods for fostering both executive functioning as well as social functioning. In autism, for instance, training programs such as Lovaas technique have proven to be effective in improving autistic individuals' social and executive functioning skills (e.g., Lovaas, 1987; Ospina et al., 2008).

The next and final section contains a brief review of the current findings, an outlook towards possible applications of the research presented here, and a discussion of possible future directions in ToM reasoning research.

7.7 Outlook

An abundance of behavioral, imaging and lesion studies have indicated in the past that response inhibition and belief-reasoning are closely related. It has been shown that training in either a belief-reasoning or an IC task improves performance in both tasks. Furthermore, an impairment in ToM reasoning as the key impairment in ASD is frequently accompanied by an impairment in IC. The nature of this relationship on the neural level and the debate over whether there is a module in the brain specific to the attribution of beliefs, however, has remained puzzling. The functional imaging study presented within this thesis was therefore aimed at clarifying these questions by investigating both IC and false-belief reasoning within a single study in the same set of subjects. The results revealed a largely right-hemispheric fronto-parietal network for IC consisting of brain regions supporting working memory, conflict detection and inhibition, and attention reorienting. A neural network for a decoupling between mentality and reality as a crucial component for false-belief reasoning revealed a neural network consisting of several components. These components seem to subservise basic cognitive processes such as perspective taking, self-referencing, working memory, conflict detection and inhibition, attentional arousal, and possibly an automated analysis of bodily and facial features.

These results indicate that the attribution of beliefs and IC may be related to each other via a recruitment of common underlying basic processes including attention reorienting, conflict detection and specific inhibition, and working memory. Furthermore, the results have shown that the right TPJ and medial PFC which had previously been discussed as possible ToM modules in the brain are not specific to the attribution of beliefs. The right TPJ was active in the present study for both false-belief reasoning as well as response inhibition, thus refuting accounts that this area is specific to mental state attribution. Medial PFC also does not seem to qualify as a specific module for the attribution of beliefs. While the dorsal part was active in the current study in both IC and belief-reasoning, previous studies have shown that the ventral part, although not active in the present IC comparison, is also frequently recruited in self-referential processes. These results indicate that future research needs to focus on the underlying processes of belief-reasoning and stop searching for a specific ToM module. It has widely been acknowledged that cognitive functions in the brain rely on networks consisting of components also used in other processes. Belief attribution is unlikely to be an exception to this parsimonious mode of operation in the brain. Future research thus needs to further clarify the roles of the differing components of the neural network for belief attribution that have been identified in the current study. There is also little known about how these components work in common to support the attribution of beliefs. Future imaging studies should therefore

analyze their data in terms of functional connectivity between the various components of these networks. Already existing statistical methods suitable for such an approach include psycho-physiological interaction analysis (PPI) or dynamic causal modelling (DCM).

Future research also needs to be dedicated in particular to the role of the left TPJ in false-belief reasoning and IC. Although both concepts recruited parts of the left TPJ in the present study, there was only a minor overlap between the two. It remains unclear at this point whether this area is part of a ventral attention network or whether this area may play a more specific role in the attribution of beliefs, as previous lesion studies suggests.

Based on existing and present findings, it has furthermore been hypothesized in this thesis that social and behavioral problems in autistic individuals may be due to a dysfunction of medial PFC. Compromised belief-reasoning in this disorder may be mediated by the ventral part of the medial PFC. This region seems to support self referencing as a crucial component of ToM reasoning. The also affected dorsal part of the medial PFC supports conflict detection and inhibition and may explain a frequently reported impairment in autistic executive functioning. In ADHD, however, medial prefrontal dysfunctions may be limited to the dorsal part only and spare the more ventral area of medial PFC which seems to be indispensable for ToM reasoning. This may be the reason why patients with ADHD show no impairment in ToM reasoning tasks unless they include high inhibitory demands.

Future applications of present and previous results should therefore be applied to improve already existing training programs for both healthy children and children with disorders such as ADHD and ASD. The findings indicate that training in attentional reorienting as well as conflict detection and inhibition may prove beneficial for executive functioning and ToM reasoning throughout adulthood.

As mentioned, previous and current results indicate that compromised belief-reasoning in ASD may be due to a dysfunction of ventral medial PFC. This finding may lead to the development of novel treatment approaches. Possibly, one such novel approach may be the application of deep brain stimulation techniques for autistic individuals. Deep brain stimulation is a method that has successfully been used in a variety of psychiatric and neurological disorders such as Parkinson's disease, chronic pain, major depression, and obsessive-compulsive disorder (e.g., Larson, 2008). In deep brain stimulation, a specific area in the brain is stimulated by an implant in the brain emitting electrical impulses. In ASD, a stimulation of ventral medial PFC may possibly have beneficial effects on the patient's ability to attend to stimuli in the environment and thereby improve their ability to attribute mental states. However, such an application would clearly need further research into which other structures may also contribute to this developmental disorder. Furthermore, it is not sufficient to focus exclusively on the

dysfunction of specific brain structures in ASD. In addition to focusing on how different areas work together in common networks, future research also needs to further investigate how this dysfunction is mediated on the neurotransmitter level. There is some indication, for example, that autistic patients exhibit increased levels of serotonin in certain areas of the brain (e.g., Kolevzon et al., 2006; Prasad et al., 2008). Furthermore, recent studies indicate that the neurohypophyseal peptide oxytocin may enhance affective mental state attribution in healthy adults (Domes et al., 2007). This and other novel findings indicate that oxytocin may also have a beneficial effect on compromised ToM reasoning in autistic patients (Bartz & Hollander, 2008; Heinrichs & Domes, 2008).

Taken together, the study presented here was able to provide clarifying insights into the nature of the connection between belief-reasoning and IC in adulthood. While the study has indicated that both concepts consist of a variety of differing components with an overlap in some substantial areas, it has also raised several new questions. Among these yet to be investigated topics is the nature of the connection between belief-reasoning and IC in childhood. Furthermore, it needs to be clarified how the neural networks for belief-reasoning and IC are altered in compromised ToM, as is the case in individuals with autism. In addition, future research in these complex topics needs to be open to other approaches and integrate existing findings from diverse fields such as genetics, neurophysiology, neuroinformatics, neuroimaging, psychology, psychiatry, and special education.

8. References

- Adleman, N. E., Menon, V., Blasey, C. M., White, C. D., Warsofsky, I. S., Glover, G. H. et al. (2002). A developmental fMRI study of the Stroop color-word task. *Neuroimage.*, 16, 61-75.
- Adolphs, R. & Spezio, M. (2006). Role of the amygdala in processing visual social stimuli. *Prog.Brain Res.*, 156, 363-378.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn Sci.*, 4, 267-278.
- Apperly, I. A. (2008). Beyond Simulation-Theory and Theory-Theory: why social cognitive neuroscience should use its own concepts to study "theory of mind". *Cognition*, 107, 266-283.
- Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004a). Frontal and temporo-parietal lobe contributions to theory of mind: neuropsychological evidence from a false-belief task with reduced language and executive demands. *J.Cogn Neurosci.*, 16, 1773-1784.
- Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004b). Frontal and temporo-parietal lobe contributions to theory of mind: neuropsychological evidence from a false-belief task with reduced language and executive demands. *J.Cogn Neurosci.*, 16, 1773-1784.
- Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *Neuroscientist.*, 13, 214-228.

- Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nat.Neurosci.*, 6, 115-116.
- Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn Sci.*, 8, 170-177.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21, 37-46.
- Bartz, J. A. & Hollander, E. (2008). Oxytocin and experimental therapeutics in autism spectrum disorders. *Prog.Brain Res.*, 170, 451-462.
- Bedard, A. C., Nichols, S., Barbosa, J. A., Schachar, R., Logan, G. D., & Tannock, R. (2002). The development of selective inhibitory control across the life span. *Dev.Neuropsychol.*, 21, 93-111.
- Bellagamba, F., Camaioni, L., & Colonnesi, C. (2006). Change in children's understanding of others' intentional actions. *Dev.Sci.*, 9, 182-188.
- Bird, C. M., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on 'Theory of Mind' and cognition. *Brain*, 127, 914-928.
- Blakemore, S. J., den Ouden, H., Choudhury, S., & Frith, C. (2007). Adolescent development of the neural circuitry for thinking about intentions. *Soc.Cogn Affect.Neurosci.*, 2, 130-139.
- Blasi, G., Goldberg, T. E., Weickert, T., Das, S., Kohn, P., Zolnick, B. et al. (2006). Brain regions underlying response inhibition and interference monitoring and suppression. *Eur.J.Neurosci.*, 23, 1658-1664.

- Bledowski, C., Prvulovic, D., Goebel, R., Zanella, F. E., & Linden, D. E. (2004). Attentional systems in target and distractor processing: a combined ERP and fMRI study. *Neuroimage.*, *22*, 530-540.
- Buchsbaum, B. R., Greer, S., Chang, W. L., & Berman, K. F. (2005). Meta-analysis of neuroimaging studies of the Wisconsin card-sorting task and component processes. *Hum.Brain Mapp.*, *25*, 35-45.
- Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: the power of an expertise framework. *Trends Cogn Sci.*, *10*, 159-166.
- Butterworth, G. & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, *9*, 55-72.
- Cabeza, R. & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *J.Cogn Neurosci.*, *12*, 1-47.
- Calarge, C., Andreasen, N. C., & O'Leary, D. S. (2003). Visualizing how one brain understands another: a PET study of theory of mind. *Am.J.Psychiatry*, *160*, 1954-1964.
- Call, J. & Tomasello, M. (1999). A nonverbal false belief task: the performance of children and great apes. *Child Dev.*, *70*, 381-395.
- Call, J. & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn Sci.*, *12*, 187-192.
- Carlson, S. M. (2003). Executive function in context: development, measurement, theory, and experience. *Monogr Soc.Res.Child Dev.*, *68*, 138-151.
- Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Dev.Neuropsychol.*, *28*, 595-616.

- Carlson, S. M., Davis, A. C., & Leach, J. G. (2005). Less is more: executive function and symbolic representation in preschool children. *Psychol.Sci.*, *16*, 609-616.
- Carlson, S. M. & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Dev.*, *72*, 1032-1053.
- Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Inf.Child.Dev.*, *11*, 73-92.
- Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *J.Exp.Child Psychol.*, *87*, 299-319.
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr Soc.Res.Child Dev.*, *63*, i-143.
- Casey, B. J., Trainor, R. J., Orendi, J. L., Schubert, A. B., Nystrom, L. E., Giedd, J. N. et al. (1997). A developmental functional MRI study of prefrontal activation during performance of a Go-No-Go task. *J.Cogn Neurosci.*, *9*, 835-847.
- Castelli, F., Frith, C., Happe, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, *125*, 1839-1849.
- Cattell, R. B., Weiß, R. H., & Osterland, J. (1997). *Grundintelligenztest (CFT-1)*. Göttingen: Hogrefe .
- Cavanna, A. E. (2007). The precuneus and consciousness. *CNS.Spectr.*, *12*, 545-552.
- Cavanna, A. E. & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, *129*, 564-583.

- Chasiotis, A. & Kiessling, F. (2004). Bleibt die Spezifität der Beziehung zwischen Theory of mind und inhibitorischer Kontrolle über die Lebensspanne bestehen? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 36, 105-114.
- Chasiotis, A., Kiessling, F., Hofer, J., & Campos, D. (2007). Theory of mind and inhibitory control in three cultures: Conflict inhibition predicts false belief understanding in Germany, Costa Rica and Cameroon. *Int.J.Behav.Devel.*, 30, 249-260.
- Ciesielski, K. T., Lesnik, P. G., Savoy, R. L., Grant, E. P., & Ahlfors, S. P. (2006). Developmental neural networks in children performing a Categorical N-Back Task. *Neuroimage.*, 33, 980-990.
- Constable, R. T., Pugh, K. R., Berroya, E., Mencl, W. E., Westerveld, M., Ni, W. et al. (2004). Sentence complexity and input modality effects in sentence comprehension: an fMRI study. *Neuroimage.*, 22, 11-21.
- Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nat.Neurosci.*, 3, 292-297.
- Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, 58, 306-324.
- Corbetta, M. & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat.Rev.Neurosci.*, 3, 201-215.
- Costa, A., Torriero, S., Oliveri, M., & Caltagirone, C. (2008). Prefrontal and temporo-parietal involvement in taking others' perspective: TMS evidence. *Behav.Neurol.*, 19, 71-74.
- Craik, F. I. M. & Lockhart, R. S. (1972). Levels of Processing: A framework for Memory Research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.

- Curtis, C. E. & D'Esposito, M. (2003). Success and failure suppressing reflexive behavior. *J.Cogn Neurosci.*, 409-418.
- D'Esposito, M. (2000). Functional neuroimaging of cognition. *Semin.Neurol.*, 20, 487-498.
- D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philos.Trans.R.Soc.Lond B Biol.Sci.*, 362, 761-772.
- Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*, 264, 1102-1105.
- Decety, J. & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist.*, 13, 580-593.
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp.Brain Res.*, 91, 176-180.
- Dickstein, S. G., Bannon, K., Castellanos, F. X., & Milham, M. P. (2006). The neural correlates of attention deficit hyperactivity disorder: an ALE meta-analysis. *J.Child Psychol.Psychiatry*, 47, 1051-1062.
- Domes, G., Heinrichs, M., Michel, A., Berger, C., & Herpertz, S. C. (2007). Oxytocin improves "mind-reading" in humans. *Biol.Psychiatry*, 61, 731-733.
- Domes, G., Kumbier, E., Herpertz-Dahlmann, B., & Herpertz, S. C. (2008). [Social cognition in autism. A survey of functional imaging studies]. *Nervenarzt*, 79, 261-274.
- Fine, C., Lumsden, J., & Blair, R. J. (2001). Dissociation between 'theory of mind' and executive functions in a patient with early left amygdala damage. *Brain*, 124, 287-298.

- Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. et al. (1995). Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. *Cognition*, *57*, 109-128.
- Floden, D. & Stuss, D. T. (2006). Inhibitory control is slowed in patients with right superior medial frontal damage. *J.Cogn Neurosci.*, *18*, 1843-1849.
- Flynn, E. (2007). The role of inhibitory control in false belief understanding. *Inf.Child.Dev.*, 53-69.
- Frith, C. D. (2008). Social cognition. *Philos.Trans.R.Soc.Lond B Biol.Sci.*, *363*, 2033-2039.
- Frith, C. D. & Frith, U. (2007). Social cognition in humans. *Curr.Biol.*, *17*, R724-R732.
- Frith, U. & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philos.Trans.R.Soc.Lond B Biol.Sci.*, *358*, 459-473.
- Gallagher, H. L. & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends Cogn Sci.*, *7*, 77-83.
- Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia*, *38*, 11-21.
- Garavan, H., Hester, R., Murphy, K., Fassbender, C., & Kelly, C. (2006). Individual differences in the functional neuroanatomy of inhibitory control. *Brain Res.*, *1105*, 130-142.
- Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165-193.

- German, T. P. & Hehman, J. A. (2006). Representational and executive selection resources in 'theory of mind': evidence from compromised belief-desire reasoning in old age. *Cognition*, 101, 129-152.
- Gilbert, S. J., Bird, G., Brindley, R., Frith, C. D., & Burgess, P. W. (2008). Atypical recruitment of medial prefrontal cortex in autism spectrum disorders: An fMRI study of two executive function tasks. *Neuropsychologia*, 46, 2281-2291.
- Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D. et al. (2006). Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *J.Cogn Neurosci.*, 18, 932-948.
- 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, 1803-1814.
- Goebel, R. & Kriegeskorte, N. (2004). Funktionelle Magnetresonanztomographie. In H.Walter (Ed.), *Funktionelle Bildgebung in Psychiatrie und Psychotherapie* (pp. 22-30). Stuttgart: Schattauer.
- Goel, V., Grafman, J., Sadato, N., & Hallett, M. (1995). Modeling other minds. *Neuroreport*, 6, 1741-1746.
- Gopnik, A. & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Dev.*, 59, 26-37.
- Gopnik, A. & Wellman, H. M. (1994). The theory theory. In L.Hirschfieldd & S. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp. 257-293). New York: Cambridge University Press.

- Grezes, J., Fonlupt, P., Bertenthal, B., Delon-Martin, C., Segebarth, C., & Decety, J. (2001). Does perception of biological motion rely on specific brain regions? *Neuroimage.*, *13*, 775-785.
- Happe, F., Ehlers, S., Fletcher, P., Frith, U., Johansson, M., Gillberg, C. et al. (1996). 'Theory of mind' in the brain. Evidence from a PET scan study of Asperger syndrome. *Neuroreport*, *8*, 197-201.
- Happe, F. G., Winner, E., & Brownell, H. (1998). The getting of wisdom: theory of mind in old age. *Dev.Psychol.*, *34*, 358-362.
- Heinrichs, M. & Domes, G. (2008). Neuropeptides and social behaviour: effects of oxytocin and vasopressin in humans. *Prog.Brain Res.*, *170*, 337-350.
- Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science*, *317*, 1360-1366.
- Hickok, G. & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*, 67-99.
- Hill, E. L. & Bird, C. M. (2006). Executive processes in Asperger syndrome: patterns of performance in a multiple case series. *Neuropsychologia*, *44*, 2822-2835.
- Hughes, C. (1998). Finding your marbles: does preschoolers' strategic behavior predict later understanding of mind? *Dev.Psychol.*, *34*, 1326-1339.
- Hughes, C., Dunn, J., & White, A. (1998). Trick or treat?: uneven understanding of mind and emotion and executive dysfunction in "hard-to-manage" preschoolers. *J.Child Psychol.Psychiatry*, *39*, 981-994.
- Hughes, C. & Russell, J. (1993). Autistic children's difficulty with mental disengagement from an object: its implications for theories of autism. *Dev.Psychol.*, *29*, 498-510.

- Hughes, C. (2002). Executive functions and development: Emerging themes. *Inf.Child.Dev.*, 201-209.
- Hughes, C. & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Dev.Psychol.*, 1447-1459.
- Huttenlocher, P. R. & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *J.Comp Neurol.*, 387, 167-178.
- Iacoboni, M. & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nat.Rev.Neurosci.*, 7, 942-951.
- Iacoboni, M., Koski, L. M., Brass, M., Bekkering, H., Woods, R. P., Dubeau, M. C. et al. (2001). Reafferent copies of imitated actions in the right superior temporal cortex. *Proc.Natl.Acad.Sci.U.S.A*, 98, 13995-13999.
- Johnson, M. H. (2003). Development of human brain functions. *Biol.Psychiatry*, 54, 1312-1316.
- Kain, W. & Perner, J. (2005). What fMRI Can Tell Us About the ToM-EF Connection: False Beliefs, Working Memory, and Inhibition. In W.Schneider, R. Schumann-Hengstler, & B. Sodian (Eds.), *Young Children's Cognitive Development: Interrelationships Among Executive Functioning, Working Memory, Verbal Ability, and Theory of Mind* (pp. 189-218). Mahwah, NJ: Erlbaum.
- Kelly, A. M., Hester, R., Murphy, K., Javitt, D. C., Foxe, J. J., & Garavan, H. (2004). Prefrontal-subcortical dissociations underlying inhibitory control revealed by event-related fMRI. *Eur.J.Neurosci.*, 19, 3105-3112.
- Kennedy, D. P., Redcay, E., & Courchesne, E. (2006). Failing to deactivate: resting functional abnormalities in autism. *Proc.Natl.Acad.Sci.U.S.A*, 103, 8275-8280.

- Kincade, J. M., Abrams, R. A., Astafiev, S. V., Shulman, G. L., & Corbetta, M. (2005). An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *J.Neurosci.*, *25*, 4593-4604.
- Kleinschmidt, A., Nitschke, M. F., & Frahm, J. (1997). Somatotopy in the human motor cortex hand area. A high-resolution functional MRI study. *Eur.J.Neurosci.*, *9*, 2178-2186.
- Kloo, D. & Perner, J. (2003). Training transfer between card sorting and false belief understanding: helping children apply conflicting descriptions. *Child Dev.*, *74*, 1823-1839.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, *297*, 846-848.
- Kolevzon, A., Mathewson, K. A., & Hollander, E. (2006). Selective serotonin reuptake inhibitors in autism: a review of efficacy and tolerability. *J.Clin.Psychiatry*, *67*, 407-414.
- Kourtzi, Z. & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *J.Cogn Neurosci.*, *12*, 48-55.
- Larson, P. S. (2008). Deep brain stimulation for psychiatric disorders. *Neurotherapeutics.*, *5*, 50-58.
- Legerstee, M. (1991). The role of person and object in eliciting early imitation. *J.Exp.Child Psychol.*, *51*, 423-433.
- Leslie, A. M. (1994). Pretending and believing: issues in the theory of ToMM. *Cognition*, *50*, 211-238.

- Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind". *Trends Cogn Sci.*, *8*, 528-533.
- Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognit.Psychol.*, *50*, 45-85.
- Liddle, P. F., Kiehl, K. A., & Smith, A. M. (2001). Event-related fMRI study of response inhibition. *Hum.Brain Mapp.*, *12*, 100-109.
- Lieberman, M. D. (2007). Social cognitive neuroscience: a review of core processes. *Annu.Rev.Psychol.*, *58*, 259-289.
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*, 150-157.
- Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *J.Consult Clin.Psychol.*, *55*, 3-9.
- Maylor, E. A., Moulson, J. M., Muncer, A. M., & Taylor, L. A. (2002). Does performance on theory of mind tasks decline in old age? *Br.J.Psychol.*, *93*, 465-485.
- McEvoy, R. E., Rogers, S. J., & Pennington, B. F. (1993). Executive function and social communication deficits in young autistic children. *J.Child Psychol.Psychiatry*, *34*, 563-578.
- McKinnon, M. C. & Moscovitch, M. (2007). Domain-general contributions to social reasoning: theory of mind and deontic reasoning re-explored. *Cognition*, *102*, 179-218.
- Meltzoff, A. N. & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 74-78.

- Meltzoff, A. N. & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child Dev.*, *54*, 702-709.
- Menon, V., Adleman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related brain activation during a Go/NoGo response inhibition task. *Hum. Brain Mapp.*, *12*, 131-143.
- Mitchell, J. P. (2008). Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cereb. Cortex*, *18*, 262-271.
- Moore, D. G., Goodwin, J. E., George, R., Axelsson, E. L., & Braddick, F. M. (2007). Infants perceive human point-light displays as solid forms. *Cognition*, *104*, 377-396.
- Morawetz, C., Holz, P., Lange, C., Baudewig, J., Weniger, G., Irle, E. et al. (2008). Improved functional mapping of the human amygdala using a standard functional magnetic resonance imaging sequence with simple modifications. *Magn Reson. Imaging*, *26*, 45-53.
- Morris, J. P., Pelphrey, K. A., & McCarthy, G. (2006). Occipitotemporal activation evoked by the perception of human bodies is modulated by the presence or absence of the face. *Neuropsychologia*, *44*, 1919-1927.
- Mostofsky, S. H., Schafer, J. G., Abrams, M. T., Goldberg, M. C., Flower, A. A., Boyce, A. et al. (2003). fMRI evidence that the neural basis of response inhibition is task-dependent. *Brain Res. Cogn Brain Res.*, *17*, 419-430.
- Müller, U., Zelazo, P. D., & Imrisek, S. (2005). Executive function and children's understanding of false belief: How specific is the relation? *Cogn. Devel.*, 173-189.
- Nieminen-von Wendt, T., Metsahonkala, L., Kulomaki, T., Aalto, S., Autti, T., Vanhala, R. et al. (2003). Changes in cerebral blood flow in Asperger syndrome during theory of

- mind tasks presented by the auditory route. *Eur.Child Adolesc.Psychiatry*, 12, 178-189.
- Northoff, G., Heinzel, A., de, G. M., BERPohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain--a meta-analysis of imaging studies on the self. *Neuroimage.*, 31, 440-457.
- O'Connell, R. G., Dockree, P. M., Bellgrove, M. A., Kelly, S. P., Hester, R., Garavan, H. et al. (2007). The role of cingulate cortex in the detection of errors with and without awareness: a high-density electrical mapping study. *Eur.J.Neurosci.*, 25, 2571-2579.
- Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc.Natl.Acad.Sci.U.S.A*, 87, 9868-9872.
- Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The enigmatic temporal pole: a review of findings on social and emotional processing. *Brain*, 130, 1718-1731.
- Onishi, K. H., Baillargeon, R., & Leslie, A. M. (2007). 15-month-old infants detect violations in pretend scenarios. *Acta Psychol.(Amst)*, 124, 106-128.
- Ospina, M. B., Krebs, S. J., Clark, B., Karkhaneh, M., Hartling, L., Tjosvold, L. et al. (2008). Behavioural and developmental interventions for autism spectrum disorder: a clinical systematic review. *PLoS.ONE.*, 3, e3755.
- Ozonoff, S. & Miller, J. N. (1995). Teaching theory of mind: a new approach to social skills training for individuals with autism. *J.Autism Dev.Disord.*, 25, 415-433.
- Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. *J.Child Psychol.Psychiatry*, 32, 1081-1105.

- Ozonoff, S. & Strayer, D. L. (2001). Further evidence of intact working memory in autism. *J. Autism Dev. Disord.*, 31, 257-263.
- Pasternak, T. & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nat. Rev. Neurosci.*, 6, 97-107.
- Pennington, B. F. & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *J. Child Psychol. Psychiatry*, 37, 51-87.
- Perner, J. (1993). *Understanding the Representational Mind*. Cambridge: MIT Press.
- Perner, J. (1998). The meta-intentional nature of executive functions and theory of mind. In P. Carruthers & J. Boucher (Eds.), *Language and Thought* (pp. 270-283). Cambridge University Press.
- Perner, J. & Aichhorn, M. (2008). Theory of mind, language and the temporoparietal junction mystery. *Trends Cogn Sci.*, 12, 123-126.
- Perner, J., Aichhorn, M., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Thinking of mental and other representations: the roles of left and right temporo-parietal junction. *Soc. Neurosci.*, 1, 245-258.
- Perner, J. & Davies, G. (1991). Understanding the mind as an active information processor: do young children have a "copy theory of mind"? *Cognition*, 39, 51-69.
- Perner, J. & Howes, D. (1992). "He thinks he knows": And more developmental evidence against the simulation (role taking) theory. *Mind and Language*, 7, 72-86.
- Perner, J. & Lang, B. (1999). Development of theory of mind and executive control. *Trends Cogn Sci.*, 3, 337-344.
- Perner, J., Lang, B., & Kloo, D. (2002a). Theory of mind and self-control: more than a common problem of inhibition. *Child Dev.*, 73, 752-767.

- Perner, J., Kain, W., & Barchfeld, P. (2002b). Executive control and higher-order theory of mind in children at risk of ADHD. *Inf.Child.Dev.*, 141-158.
- Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am.J.Psychiatry*, 164, 942-948.
- Prasad, H. C., Steiner, J. A., Sutcliffe, J. S., & Blakely, R. D. (2008). Enhanced activity of human serotonin transporter variants associated with autism. *Philos.Trans.R.Soc.Lond B Biol.Sci.*
- Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav.Brain Sci.*, 515-526.
- Reed, M., Pien, D., & Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill Palmer Quarterly*, 30, 131-147.
- Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. *Annu.Rev.Neurosci.*, 27, 169-192.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat.Rev.Neurosci.*, 2, 661-670.
- Rothmayr, C., Baumann, O., Endestad, T., Rutschmann, R. M., Magnussen, S., & Greenlee, M. W. (2007). Dissociation of neural correlates of verbal and non-verbal visual working memory with different delays. *Behav.Brain Funct.*, 3, 56.
- Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). "Theory of mind" impairments and their relationship to executive functioning following frontal lobe excisions. *Brain*, 124, 600-616.

- Rubia, K., Smith, A. B., Brammer, M. J., & Taylor, E. (2003). Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *Neuroimage.*, *20*, 351-358.
- Rubia, K., Smith, A. B., Woolley, J., Nosarti, C., Heyman, I., Taylor, E. et al. (2006). Progressive increase of frontostriatal brain activation from childhood to adulthood during event-related tasks of cognitive control. *Hum.Brain Mapp.*, *27*, 973-993.
- Russell, J., Mauthner, S., Sharpe, S., & Tidswell, T. (1991). The "windows task " as a measure of strategic deception in preschoolers and autistic subjects. *Br.J.Dev.Psychol.*, *9*, 331-349.
- Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind. A comparison of Chinese and U.S. preschoolers. *Psychol.Sci.*, *17*, 74-81.
- Samson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nat.Neurosci.*, *7*, 499-500.
- Saxe, R. (2005). Against simulation: the argument from error. *Trends Cogn Sci.*, *9*, 174-179.
- Saxe, R. (2006). Uniquely human social cognition. *Curr.Opin.Neurobiol.*, *16*, 235-239.
- Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annu.Rev.Psychol.*, *55*, 87-124.
- Saxe, R. & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage.*, *19*, 1835-1842.

- Saxe, R. & Powell, L. J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychol.Sci.*, 17, 692-699.
- Saxe, R. & Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia*, 43, 1391-1399.
- Saxe, R., Schulz, L. E., & Jiang, Y. V. (2006). Reading minds versus following rules: Dissociating theory of mind and executive control in the brain. *Soc.Neurosci.*, 1, 284-298.
- Scheithauer, H., Bondue, R., & Mayer, H. (2008). Förderung sozial-emotionaler Kompetenzen im Vorschulalter: Ergebnisse der Augsburger Längsschnittstudie zur Evaluation des primärpräventiven Programms Papilio® (ALEPP). In T.Malti & S. Perren (Eds.), *Soziale Kompetenz bei Kindern und Jugendlichen. Entwicklungsprozess und Förderungsmöglichkeiten* (pp. 145-164). Stuttgart: Kohlhammer Verlag.
- Schiff, N. D. (2008). Central thalamic contributions to arousal regulation and neurological disorders of consciousness. *Ann.N.Y.Acad.Sci.*, 1129, 105-118.
- Seidman, L. J., Biederman, J., Faraone, S. V., Weber, W., & Ouellette, C. (1997). Toward defining a neuropsychology of attention deficit-hyperactivity disorder: performance of children and adolescents from a large clinically referred sample. *J.Consult Clin.Psychol.*, 65, 150-160.
- Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychol.Sci.*, 16, 114-122.
- Siegal, M. & Varley, R. (2002). Neural systems involved in "theory of mind". *Nat.Rev.Neurosci.*, 3, 463-471.

- Simmonds, D. J., Fotedar, S. G., Suskauer, S. J., Pekar, J. J., Denckla, M. B., & Mostofsky, S. H. (2007). Functional brain correlates of response time variability in children. *Neuropsychologia*, *45*, 2147-2157.
- Simmonds, D. J., Pekar, J. J., & Mostofsky, S. H. (2008). Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia*, *46*, 224-232.
- Sinzig, J., Morsch, D., Bruning, N., Schmidt, M. H., & Lehmkuhl, G. (2008). Inhibition, flexibility, working memory and planning in autism spectrum disorders with and without comorbid ADHD-symptoms. *Child Adolesc.Psychiatry Ment.Health*, *2*, 4.
- Sodian, B. & Frith, U. (1992). Deception and sabotage in autistic, retarded and normal children. *J.Child Psychol.Psychiatry*, *33*, 591-605.
- Sodian, B. & Hülksen, C. (2005). The Developmental Relation of Theory of Mind and Executive Functions: A Study of Advanced Theory of Mind Abilities in Children with ADHD. In W.Schneider, R. Schumann-Hengstler, & B. Sodian (Eds.), *Young children's cognitive development: Interrelationships Among Executive Functioning, Working Memory, Verbal Ability, and Theory of Mind* (pp. 175-188). Mahwah, NJ: Erlbaum.
- Sodian, B., Hülksen, C., & Thoermer, C. (2003). The self and action in theory of mind research. *Conscious.Cogn*, *12*, 777-782.
- Sodian, B. & Thoermer, C. (2006). Theory of Mind. In W.Schneider & B. Sodian (Eds.), *Enzyklopädie der Psychologie, Serie V: Entwicklung, Band 2: Kognitive Entwicklung* (pp. 495-608). Goettingen: Hogrefe.
- Sodian, B. & Thoermer, C. (2008). Precursors to a theory of mind in infancy: perspectives for research on autism. *Q.J.Exp.Psychol.(Colchester.)*, *61*, 27-39.

- Sommer, M., Dohnel, K., Sodian, B., Meinhardt, J., Thoermer, C., & Hajak, G. (2007). Neural correlates of true and false belief reasoning. *Neuroimage*, *35*, 1378-1384.
- Spelke, E., Phillips, A., & Woodward, A. L. (1995). Infants' knowledge about object motion and human action. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: a multidisciplinary debate* (pp. 44-78). Oxford: Clarendon.
- Stevens, M. C., Calhoun, V. D., & Kiehl, K. A. (2005). Hemispheric differences in hemodynamics elicited by auditory oddball stimuli. *Neuroimage*, *26*, 782-792.
- Stevens, M. C., Kiehl, K. A., Pearlson, G. D., & Calhoun, V. D. (2007). Functional neural networks underlying response inhibition in adolescents and adults. *Behav. Brain Res.*, *181*, 12-22.
- Stich, S. & Nichols, S. (1997). Cognitive penetrability, rationality, and restricted simulation. *Mind and Language*, *12*, 297-326.
- Stone, J. L. (1999). Transcranial brain injuries caused by metal rods and pipes over the past 150 years. *J. Hist Neurosci.*, *8*, 227-234.
- Strommen, E. A. (1973). Verbal self-regulation in a children's game: Impulsive errors on "Simon says". *Child Dev.*, *44*, 849-853.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.*, *18*, 622-643.
- Stuss, D. T., Floden, D., Alexander, M. P., Levine, B., & Katz, D. (2001a). Stroop performance in focal lesion patients: dissociation of processes and frontal lobe lesion location. *Neuropsychologia*, *39*, 771-786.
- Stuss, D. T., Gallup, G. G., Jr., & Alexander, M. P. (2001b). The frontal lobes are necessary for 'theory of mind'. *Brain*, *124*, 279-286.

- Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Dev.Psychol.*, *30*, 395-402.
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C. et al. (2001). I know what you are doing. a neurophysiological study. *Neuron*, *31*, 155-165.
- Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O. et al. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage.*, *30*, 1414-1432.
- Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happe, F., Falkai, P. et al. (2001). Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage.*, *14*, 170-181.
- Vogeley, K. & Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends Cogn Sci.*, *7*, 38-42.
- Wager, T. D., Sylvester, C. Y., Lacey, S. C., Nee, D. E., Franklin, M., & Jonides, J. (2005). Common and unique components of response inhibition revealed by fMRI. *Neuroimage.*, *27*, 323-340.
- Watanabe, J., Sugiura, M., Sato, K., Sato, Y., Maeda, Y., Matsue, Y. et al. (2002). The human prefrontal and parietal association cortices are involved in NO-GO performances: an event-related fMRI study. *Neuroimage.*, *17*, 1207-1216.
- Wellman, H. M. & Cross, D. (2001). Theory of mind and conceptual change. *Child Dev.*, *72*, 702-707.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.*, *72*, 655-684.
- Wellman, H. M. & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Dev.*, *75*, 523-541.

Wellman, H. M. & Woolley, J. D. (1990). From simple desires to ordinary beliefs: the early development of everyday psychology. *Cognition*, 35, 245-275.

Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128.

Young, L. & Saxe, R. (2008). The neural basis of belief encoding and integration in moral judgment. *Neuroimage.*, 40, 1912-1920.

Zelazo, P. D., Craik, F. I., & Booth, L. (2004). Executive function across the life span. *Acta Psychol.(Amst)*, 115, 167-183.

Zelazo, P. D., Muller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J. et al. (2003). The development of executive function in early childhood. *Monogr Soc.Res.Child Dev.*, 68, vii-137.

9. Index of Figures

Figure 2.1. Digital reconstruction of brain damage to railroad worker Phineas Gage due to an iron rod propelled through his head. This picture shows that the medial part of the frontal cortex, an area associated with ToM reasoning, had also been affected. Historical accounts whether Gage had subsequently suffered from an inability to impute mental states, however, are vague. Picture taken from Damasio et al. (1994). Reprinted with permission from AAAS.15

Figure 2.2. Two lateral views at MNI coordinates $x = -6$ and $x = 43$ of significantly increased activity in three comparisons conducted in a study by Sommer et al. (2007). Significantly increased activity in the contrast *False-Belief > True-Belief* is depicted in red shading, significantly increased activity in the contrast *True-Belief > False-Belief* is shown in green shading. Areas in yellow color depict common activation for the two contrasts *False-Belief > Baseline* and *True-Belief > Baseline*. Picture reprinted from Sommer et al. (2007) with permission from Elsevier.24

Figure 3.1. Results from the activation-likelihood-estimation (ALE) meta-analysis of 11 IC studies by Simmonds et al. (2008). Brain areas associated with IC are depicted in red color. All of the studies included in the meta-analysis had used a Go / No-go paradigm to measure response inhibition. Numbers above each slice refer to corresponding Talairach coordinates. As can be seen here, IC-related activity is found in a largely right-hemispheric fronto-parietal network. This picture is taken from Simmonds et al. (2008). Reprinted with permission from Elsevier.34

Figure 4.1. Activity in the right TPJ region stemming from a meta-analysis including 18 studies related to attention reorienting (left image) and 24 studies related to ToM reasoning (right image). Image taken from Decety et al. (2007).Reprinted with permission from SAGE publications.48

Figure 4.2. The response selection task used in a study by Saxe et al. (2006). In the compatible condition subjects had to press a button that referred to a line differing in length from the other presented lines. In the incompatible condition subjects had to press a button that was shifted two locations from the location showing the line that differed in length. This figure is taken from Saxe et al. (2006).Reprinted with permission from Psychology Press.51

Figure 4.3. The ToM reasoning task in Saxe et al.'s (2006) reference experiment is depicted here. In both conditions, subjects were required to complete sentences by pressing one of two buttons. In the false-belief condition subjects had to attribute a mental state to a protagonist. The false-photograph condition on the other hand required no such decoupling between a mental state and reality. Reproduced from Saxe et al. (2006) with permission from Psychology Press.51

Figure 4.4. Outline of the response selection task and the ToM task in Saxe et al.'s (2006) main experiment. Both tasks used the exact same visual stimuli. The stimuli presented showed a girl either facing or looking away as a chocolate bar was moved from one box (known to the girl) back

to the same location or to the other box. In the ToM task subjects were instructed to press a button corresponding to the box where the girl would think the chocolate bar was. In the response selection task subjects were instructed to press buttons according to one of two rules. Rule one, for example, required subjects to press a button corresponding to the location of the last box containing the chocolate bar, but to do so only if the girl was facing the two boxes. If the girl was not facing the boxes subjects were instructed to press a button corresponding to the first box containing the chocolate bar. Reproduced from Saxe et al. (2006) with permission from Psychology Press.. 53

Figure 4.5. Activation in the right TPJ region in a study conducted by Mitchell (2008). Shown are two lateral views at MNI coordinates $x=57$ and $x=60$. Depicted in yellow color is activity related to belief-reasoning. Blue areas refer to significantly increased activity associated with attentional reorienting as revealed in an attentional cueing task. The green area refers to activity related to both attentional reorienting and belief-reasoning. Reproduced from Mitchell (2008) with permission from Oxford journals.55

Figure 5.1. Blood-Oxygen-Level-Dependant (BOLD) response as measured in fMRI. After an “initial dip” some 2 seconds after stimulus onset, the BOLD response reaches its peak after about 6 to 8 seconds. This overshoot related to an increase of oxygenated hemoglobin is called “positive BOLD-response”. Before retreating back to its baseline after approximately 12 seconds, the BOLD response briefly dips below baseline (“undershoot”).61

Figure 5.2. Course of the false-belief and the true-belief condition of the belief-reasoning experiment with respective durations. Subjects were instructed to indicate with a button press whether they had expected the person who either left or stayed in the room to perform the behavior shown in the third picture.....64

Figure 5.3. The two conditions of the inhibitory control experiment (inside the grey frame) and respective stimulus durations. The Go condition showed a number of children different from the previous picture and required subjects to press a button with their right index finger. In the No-go condition the number of children in the present picture was the same as in the previous picture. Subjects were instructed to withhold a response in this condition.65

Figure 6.1. Accuracy and corresponding standard deviations in percent of correct trials for the false-belief (FB), the true-belief (TB), the Go and the No-go condition. Paired t-tests revealed no significant difference in accuracy between the FB and the TB condition. However, there was a statistical tendency towards higher accuracies in the Go condition when compared to the No-go condition (t ($df=11$) = 1.997; p = .071).71

Figure 6.2. This figure depicts reaction times (RTs) and standard deviations for the false-belief (FB) and the true-belief (TB) condition in the belief-reasoning experiment as well as for correct Go trials in the IC experiment. There was no significant difference in RTs between the FB and the TB

condition. No RTs were assessed for correct No-go trials since the No-go condition required the withholding of a response.....72

Figure 6.3. Significantly increased brain activity in the true-belief condition when compared to baseline. Activity is shown on a template brain at a voxel-wise threshold of $T = 5$ and a corrected p-value of $p \leq .05$ on cluster-level. A detailed description of activated brain areas can be found in table 6.1.....75

Figure 6.4. Significantly increased brain activity in the false-belief condition when compared to baseline. Activity is shown on a template brain at a voxel-wise threshold of $T = 5$ and a corrected p-value of $p \leq .05$ (cluster-level). For a description of activated brain areas refer to table 6.2.78

Figure 6.5. Brain activity in the contrast *False-Belief > Baseline* (red shading) and *True-Belief > Baseline* (green shading) at a voxel-wise threshold of $T = 5$ and a corrected p-value of $p \leq .05$ (cluster-level). Areas in yellow shading show activity in both comparisons. A detailed description of active brain areas in the contrasts *False-Belief > Baseline* and *True-Belief > Baseline* can be found in tables 1 and 2, respectively.79

Figure 6.6. Significantly increased brain activity in the false-belief condition when compared to the true-belief condition. Significant activity is shown on a template brain at a voxel-wise threshold of $T = 4$ and a corrected p-value of $p \leq .05$ (cluster-level). Please refer to table 6.3 for a detailed description of activated brain areas.....82

Figure 6.7. A sectional view of significant brain activity in the contrast *False-Belief > True-Belief* at MNI coordinates $x = 3$, $y = 60$ and $z = 35$. Brain activity in this contrast is shown at a voxel-wise threshold of $T = 4$ and a corrected p-value of $p \leq .05$ on cluster-level. Numbers correspond to significantly active brain regions (1 = bilateral medial frontal gyrus and superior frontal gyrus (BA 6, 8), 2 = bilateral medial frontal gyrus (BA 9, 10), 3 = bilateral precuneus, 4 = left middle frontal gyrus, 5 = right middle frontal gyrus, 6 = left TPJ, 7 = right TPJ). The color bar depicts corresponding T-values.83

Figure 6.8. Significantly increased brain activity in the No-go condition when compared to the Go condition is depicted here. Activity is shown on a template brain at a voxel-wise threshold of $T = 4$ and a corrected p-value of $p \leq .05$ on cluster-level. For a detailed description of active brain regions please refer to table 6.4.....86

Figure 6.9. This image depicts a sectional view of significant brain activity in the contrast *No-go > Go* at MNI coordinates $x = 3$, $y = 39$ and $z = 38$. Brain activity in this contrast is depicted at a voxel-wise threshold of $T = 4$ and a corrected p-value of $p \leq .05$ on cluster-level. Numbers correspond to significantly active brain regions (1 = right medial and superior frontal gyrus, 2 = right TPJ region

and middle and superior temporal gyrus, 3 = left TPJ region, 4 = right inferior and middle frontal gyrus). The color bar depicts corresponding T-values.87

Figure 6.10. Significant brain activity in the contrasts *False-Belief > True-Belief* (red shading) and *No-go > Go* (green shading) superimposed on a template brain. Areas in yellow shading depict significantly increased activity in both contrasts. A voxel-wise threshold of $T = 4$ and a p-value of $p \leq .05$ on cluster level was set for both comparisons.....89

Figure 6.11. Activity related to response inhibition as revealed in the contrast *No-go > Go* (green shading) and to a decoupling between reality and a false belief as revealed in the contrast *False-Belief > True Belief* (red shading) in the medial prefrontal cortex. Significantly increased activity from the group-analysis is depicted on a sagittal view of the brain at a voxel-wise threshold of $T > 4$ and a cluster-defining threshold of $p \leq .05$90

Figure 7.1. Sagittal view of a template brain depicting activity in the comparison *No-go > Go* (green and yellow shading) and *False-Belief > True-Belief* (red and yellow shading) at a voxel-wise threshold of $T > 4$ and a cluster-defining threshold of $p \leq .05$. As can be seen here, a region in the medial ventral PFC cortex was distinctly activated by the comparison *False-Belief > True-Belief*.
..... 115

Figure 7.2. Activity in the right and the left hemisphere related to IC (green shading) and a decoupling process between reality and a false belief (red shading) as revealed in the current study. Common activity is depicted in yellow shading. Activity is depicted on a template brain at a voxel-wise threshold of $T > 4$ and a cluster-defining threshold of $p \leq .05$ 116

Figure 7.3. Activity related to response inhibition (green shading) and to a decoupling between reality and a false belief (red shading) in the left TPJ region on a coronal (left picture) and a sagittal view of the brain (right picture). Common activity is depicted in yellow shading. Although both cognitive concepts showed significant activations in the left TPJ, there was no substantial overlap between the two in this region. In the right TPJ, however, there was an almost complete overlap between the two respective clusters (only shown in the left picture here). Nevertheless, activity in the right TPJ was more widespread for IC. 117

Figure 7.4. Activity related to response inhibition (blue shading) and a decoupling between reality and mentality (orange shading) in the right parietal and medial prefrontal cortices. Significantly increased activation as revealed in the present study is depicted on a template brain contained in MRicro software (www.mricro.com). This figure shows common activity for both mechanisms in the right TPJ (BA 39, 40) and dorsal medial PFC (BA 6, 8). Distinct activity for a decoupling mechanism was revealed in the ventral medial PFC (BA 10)..... 121

- Figure 7.5.* Schematic outline of the hypothesized neural networks for a decoupling between reality and a false belief (red color) and response inhibition (green color) in the right hemisphere based on the findings from the current study. Yellow ellipses depict common processes and structures possibly shared by both cognitive concepts. 124
- Figure 7.6.* Components and hypothesized functions of the neural networks for inhibitory control and a decoupling between mentality and reality (red shading) in the brain's left hemisphere. Yellow ellipses represent components shared by both processes. 125
- Figure 7.7.* Activity as revealed in several imaging studies investigating self-referential processes. Picture taken from Northoff et al. (2006). Reprinted with permission from Elsevier. 130
- Figure 7.8.* Activity in the ventral medial PFC in the comparison *ToM Stories* > *Physical Stories* in 8 autistic patients and 8 healthy controls in a study conducted by Nieminen-von Wendt et al. (2003). Image is taken from Nieminen-von Wendt et al. (2003). Reproduced with friendly permission from Springer. 132

10. Index of Tables

Table 6.1. Significantly increased brain activity related to the contrast *True-belief > Baseline*. This table includes information on the name and hemisphere of activated brain regions as well as associated Brodmann areas, cluster sizes in voxels, MNI coordinates, and T-values for each cluster's peak voxel. 74

Table 6.2. Brain activity as revealed in the contrast *False-Belief > Baseline*. Given in this chart are the names and hemisphere information for every active brain region as well as corresponding Brodmann areas, cluster sizes, MNI coordinates, and T-values of each cluster's peak voxel. ... 77

Table 6.3. Brain regions showing significantly increased activity in the contrast *False-Belief > True-Belief*. The table contains the names and hemisphere of every activated cluster, associated Brodmann areas, cluster sizes, MNI coordinates, and T-values for each cluster's peak voxel. ... 81

Table 6.4. Significantly increased brain activity in the contrast *No-go > Go*. Depicted here are the names and hemisphere of activated brain regions, associated Brodmann areas, cluster sizes in voxels, MNI coordinates and T-values corresponding to every cluster peak voxel. 85

11. Appendices

Appendix A. Informed consent form for participation in the fMRI experiment.

<h3>Einverständniserklärung für die Teilnahme an einer (funktionellen) Magnetresonanztomographie (fMRT)</h3>		
<p>Liebe Versuchspersonen,</p> <p>beantworten Sie bitte die folgenden Fragen, um eventuelle Einschränkungen oder Gegenanzeigen bezüglich für die Kernspintomographie auszuschließen.</p>		
<p>Name:</p> <p>Geburtsdatum:</p> <p>Körpergewicht:</p> <p>Händigkeit:</p>		
<p>1. Haben Sie Metallteile in Ihrem Körper? (z.B.: Spirale, Gefäßclips, Granatsplitter, Piercing, Zahnspange, Hörgerät, Herzschrittmacher, künstliche Gelenke, Knochenschrauben)</p>	<p>ja <input type="checkbox"/> nein <input type="checkbox"/></p>	
<p>2. Wurden Sie an Herz oder Kopf operiert?</p>	<p>ja <input type="checkbox"/> nein <input type="checkbox"/></p>	
<p>3. Leiden Sie unter Platzangst?</p>	<p>ja <input type="checkbox"/> nein <input type="checkbox"/></p>	
<p>4. Für Patientinnen: Könnten Sie schwanger sein?</p>	<p>ja <input type="checkbox"/> nein <input type="checkbox"/> vielleicht <input type="checkbox"/></p>	
<p>Bevor Sie den Untersuchungsraum betreten legen Sie bitte Uhren, Schmuck, Wertgegenstände, Geldbeutel, Schlüsselbund, Magnetkarten, Kugelschreiber, Kleingeld oder sonstige Gegenstände aus Metall ab!</p>		
<p>Ich bin mit der Teilnahme an der fMRT einverstanden bin. Ich wurde über mögliche Gegenanzeigen informiert.</p>		
<p>..... Datum</p>	<p>..... Unterschrift der aufklärenden Fachkraft</p>	<p>..... Unterschrift der Versuchsperson</p>

Appendix B. Standardized instruction for the Belief-Reasoning and the Inhibitory Control Task. Shown here is the version for the 6 out of 12 subjects starting with the Belief-Reasoning Experiment first. The subjects randomly assigned to start with the Inhibition experiment received an identical instruction however describing the Inhibitory Control task first.

Instruktion A

Liebe Versuchsperson,

bitte lesen Sie sich diese Instruktion aufmerksam durch. Sollten Sie Fragen haben, so wenden Sie sich bitte an den Versuchsleiter.

1

Die folgende Studie besteht aus **zwei** unterschiedlichen Experimenten, die hintereinander im Kernspingerät präsentiert werden.

2

Im **ersten** Experiment werden Ihnen Bildergeschichten präsentiert, die jeweils aus *drei* Bildern bestehen.

3

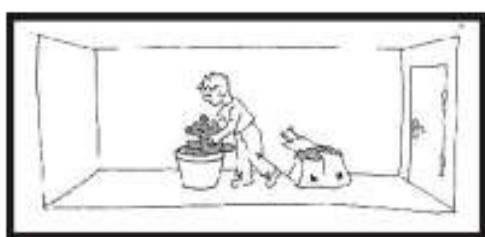
Bild 1



- Im ersten Bild sehen sie immer zwei Kinder. Eines der Kinder legt ein Objekt (hier: Teddybär) in ein Behältnis (hier: Rucksack).

4

Bild 2



oder



- Im zweiten Bild nimmt das andere Kind das Objekt heraus und legt es in ein anderes Behältnis (hier: Eimer)
- Das andere Kind hat
das Zimmer verlassen oder bleibt im Raum und beobachtet das andere Kind

Bild 3



- Im dritten Bild sucht das Kind, das den Raum verlassen hat oder im Raum geblieben ist, das Objekt in einem der beiden Behältnisse.
- Überlegen sie jetzt bitte, ob das Verhalten dieses Kindes **erwartet** oder **unerwartet** ist. Versuchen Sie sich hierzu in das Kind hinein zu versetzen.

6

Bild 3



Haben Sie das Verhalten
des Kindes

erwartet

so drücken Sie bitte so
schnell wie möglich die

linke

Taste.

Haben Sie das Verhalten
des Kindes

nicht erwartet

so drücken Sie bitte so
schnell wie möglich die

rechte

Taste.

7

Beispiel:



Bild 1:
Das Mädchen legt
einen Teddybär in
den Rucksack.



Bild 2:
Das Mädchen verlässt
den Raum.



Bild 3:
Das Mädchen kommt
zurück und sucht im
Rucksack nach dem
Teddybär.

Das Verhalten ist hier also zu erwarten → linke Taste

8

Wichtig:

Bitte drücken Sie die betreffende Taste so
genau und so schnell wie möglich!

9

Nach einer kurzen Pause und einer erneuten kurzen Instruktion (ca. 1 Min.) folgt das **zweite** Experiment:

10




Bild


Pause

Bild


Es werden Ihnen hintereinander Bilder präsentiert, die durch kurze Pausen unterbrochen werden. Diese Bilder stehen, anders als im ersten Experiment, in keinerlei inhaltlichem Zusammenhang.



Bild



Pause




Bild


...

Ihre Aufgabe ist es nun, immer dann so schnell wie möglich den **linken** Knopf zu drücken, wenn sich die Anzahl der Kinder im aktuellen Bild von der Anzahl im vorherigen unterscheidet.


12

Beispiel:






Pause




Tastendruck!


...



Tastendruck!



Pause



Tastendruck!

...

13

Zusammenfassung der beiden Experimente

Experiment 1:



Überlegen Sie sich, ob das Verhalten des Kindes **erwartet** oder **unerwartet** ist. Versuchen Sie sich hierzu in das Kind hinein zu versetzen.

- Haben Sie das Verhalten des Kindes **erwartet**, so drücken Sie bitte so schnell wie möglich die **linke** Taste.
- War das Verhalten des Kindes **unerwartet**, so drücken Sie bitte so schnell wie möglich die **rechte** Taste.

16

Zusammenfassung der beiden Experimente

Experiment 2:



Drücken Sie immer dann so schnell wie möglich die linke Taste, wenn sich die Anzahl der Kinder im aktuellen Bild von der Anzahl im vorherigen Bild unterscheidet.

Unterscheidet sich die Anzahl nicht, so drücken Sie bitte **keine** Taste.

17

Haben Sie jetzt noch Fragen?

12. Curriculum Vitae

Christoph Rothmayr

Date of Birth: November 6, 1979

Place of Birth: Garmisch-Partenkirchen, Germany

Education and Professional Experience:

06/1999 Abitur (High School degree) at the Werdenfels-Gymnasium Garmisch-Partenkirchen.

7/1999 – 04/2000 Military service in Murnau, Sonthofen and Stetten a.k.M.

08/2003- 06/2004 Neuroscience studies at Vanderbilt University in Nashville, Tennessee, USA (Fulbright scholarship).

05/2006 Completion of master's thesis entitled "Functional Magnetic Resonance Imaging of Visual Working Memory" at the University of Regensburg.

09/2006 Master's degree in psychology from the University of Regensburg.

Since 01/2007 Research assistant at the Clinical Neuroscience Center for Emotions and Social Cognition at the University of Regensburg / Department of Psychiatry and Psychotherapy.

Since 04/2007 Training in Cognitive Behavioral Therapy at the "Ausbildungsinstitut für Verhaltenstherapie Regensburg".

07/2009 Ph.D. from the Ludwig-Maximilians-University Munich / Department of Developmental Psychology (Prof. Dr. Beate Sodian). Minor: Psychiatry.