

Stefan Bode: From Stimuli to Motor Responses: Decoding Rules and Decision Mechanisms in the Human Brain. Leipzig: Max Planck Institute for Human Cognitive and Brain Sciences, 2010 (MPI Series in Human Cognitive and Brain Sciences; 123)



Impressum

Max Planck Institute for Human Cognitive and Brain Sciences, 2010



Diese Arbeit ist unter folgender Creative Commons-Lizenz lizenziert:
<http://creativecommons.org/licenses/by-nc/3.0/de>

Druck: Sächsisches Druck- und Verlagshaus Direct World, Dresden

ISBN 978-3-941504-07-3

From stimuli to motor responses: Decoding rules and decision mechanisms in the human brain

**Von der Fakultät für Biowissenschaften, Pharmazie und Psychologie
der Universität Leipzig
genehmigte**

DISSERTATION

**zur Erlangung des akademischen Grades
doctor rerum naturalium,
Dr. rer. nat.,**

vorgelegt

von

**Diplom-Psychologen Stefan Bode
geboren am 26.04.1979 in Celle**

**Dekan: Prof. Dr. Matthias Müller
Gutachter: Prof. Dr. Wolfgang Prinz
Prof. Dr. John-Dylan Haynes**

Tag der Verteidigung: 08.07.2010

Table of Contents

Acknowledgements.....	5
1 Preface.....	7
2 Introduction.....	9
2.1 General introduction.....	9
2.1.1 From stimuli to responses	9
2.1.2 Localizing information about rules and decisions.....	11
2.2 Introduction to methods.....	13
2.2.1 Basic physical principles of MRI.....	13
2.2.2 Functional magnetic resonance imaging (fMRI)	15
2.2.3 Pre-processing and univariate analysis of fMRI data.....	16
2.2.4 The multivariate pattern classification approach.....	18
2.2.5 Methodological principles of pattern classification	20
2.2.6 The application of multivariate pattern classification	23
3 Experimental Section I: Rule-based task preparation.....	27
3.1 Theoretical background.....	27
3.1.1 Task-sets and task switching.....	27
3.1.2 The neural substrate of task-sets	28
3.1.3 The goal of the present study	31
3.2 Methods.....	32
3.2.1 Task-switching experiment	32
3.2.2 Control experiment with fixed stimulus-response mappings...	38
3.3 Results.....	39
3.4 Discussion	45
3.4.1 Prefrontal cortex and rule-guided behaviour.....	45
3.4.2 Parietal cortex and rule-guided behaviour	47
3.4.3 The temporal sequence of information encoding.....	48
3.4.4 Methodological considerations	51
3.4.5 Conclusion.....	52
4 Experimental Section II: Perceptual decisions in task preparation	53
4.1 Theoretical background.....	53
4.1.1 The formal elements of perceptual decision making	53
4.1.2 The neural basis of perceptual decision making	55
4.1.3 Networks for perceptual decision making.....	56
4.1.4 Perceptual decisions about objects.....	58
4.1.5 The goal of the present study	59
4.2 Methods.....	61
4.2.1 Masking experiment.....	61
4.2.2 LOC localizer experiment	67
4.3 Results.....	69
4.4 Discussion	77

4.4.1	Perceptual decisions about highly visible objects	77
4.4.2	Perceptual decisions about poorly visible objects.....	79
4.4.3	A mechanism for internal decisions	81
4.4.4	Alternative mechanisms	83
4.4.5	Conclusion.....	85
5	Experimental Section III: Internal task preparation	87
5.1	Theoretical background	87
5.1.1	The general framework of free decisions	87
5.1.2	The neural basis of free decisions	89
5.1.3	The goal of the present study	90
5.2	Methods	92
5.3	Results	100
5.3.1	Comparison to results of the previous study	104
5.4	Discussion	106
5.4.1	Two modes of perceptual decision making.....	106
5.4.2	From guessing to free decisions	109
5.4.3	Posterior parietal cortex and high self-reference.....	110
5.4.4	Conclusion.....	111
6	General Discussion	113
6.1	Information encoding in parietal cortex	113
6.2	General methodological considerations	116
6.3	Future directions	118
6.4	Concluding summary	121
6.5	Zusammenfassung	123
	References	127
	List of Figures	147
	List of Tables.....	149
	List of Abbreviations.....	151
	Curriculum Vitae	153
	Appendix A (Experimental Section I)	ii
	Appendix B (Experimental Section II)	x
	Appendix C (Experimental Section III)	xxxiv

Acknowledgements

I am convinced that cognitive neuroscience is a ‘team sport’. Conducting studies always involves many people, especially the use of functional magnetic resonance imaging (fMRI) that requires the help of physicists, medical technical assistants and colleagues who are willing to share their experience and knowledge. I could not have done this work without all the people contributing and inspiring me. I sincerely want to thank everyone here and hope I also could give back some of this support by contributing to their work in one way or another.

First of all, I want to thank my supervisor Prof. John-Dylan Haynes, who designed the presented studies with me and gave me valuable input during all stages of my work. None of these projects could have been done without his important contributions, which allowed me to constantly improve my work. It was one of his talks he gave in London a few years ago that inspired me to study multivariate pattern classification and to apply for a position in his lab. During the last years he always trusted in my abilities to learn, gave me the possibility to acquire scanning practice, to meet several important scientists and to attend some excellent conferences. Additionally, I am very grateful for the possibility to write most parts of my thesis in Australia.

I am also thankful for the support, cooperation and input from all the great people and scientists I worked with in the “Attention & Awareness” group at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig. In particular, I want to sincerely thank Anna He, Anita Tusche, Carsten Bogler and Dr. Annette Horstmann for giving me feedback for my thesis; but also Chun Siong Soon, Christian Kalberlah, Chen Yi, Dr. Marcus Grüschow, Daniel Krüger, Saskia Heberling, Susan Szukalski, Sabrina Walter and Kathrin Marx for all their feedback, support and for the great time! With all sincerity, I want to thank my friend Carsten Bogler for his priceless support with getting me started

with fMRI, his patience in explaining why I could have expressed my full page of MATLAB code in two lines, mastering long scanning days together and breaks “on the roof” to improve my motor skills. Most importantly, I cannot express how thankful I am for Anna He’s effort to read the whole thesis and her admirable patience with my persistent ‘Germanisms’. I am more than thankful for all her comments, discussions and huge emotional support during the last years.

I also want to thank everyone who commented on earlier versions of my paper manuscripts or my thesis, discussed my studies with me or gave valuable tips regarding literature and (limits of) interpretations, most importantly Prof. Wolfgang Prinz, Prof. Richard E. Passingham, Prof. Marcel Brass, Prof. John Duncan, Prof. Tobias H. Donner, Prof. Jutta Stahl, Prof. Philip L. Smith, Prof. Stanislas Dehaene, PD Dr. Peter Nickl, Prof. Uwe Mattler, Prof. Michael W. Chee, Dr. Kirsten Jordan, Dr. Franziska Korb, Silvia Kaps and Philipp Bode. I also thank everyone in the Max Planck Institute who helped offering the infrastructure and support. I also want to mention Prof. Hans-Jochen Heinze, Prof. Hermann Hinrichs and Rosemarie Oelschläger from the Department of Neurology, Otto-von-Guericke University Magdeburg; I am grateful for their trust and support.

There are many more people who were important on the long road to the final thesis. I am grateful to Prof. Wolfgang Prinz for the possibility to take part in his research colloquiums and that he agreed to be my second reviewer. I also like to thank Annett Wiedemann, Petra Erz, Simone Wipper, Anke Kummer, Mandy Naumann, Dr. Toralf Mildner and Dr. Jöran Lepsien for their support with scanning. Thanks to Sven Gutekunst for constructing the joystick for the first study and special thanks to Dr. Stefan Zysset for designing the joystick table with me. Additionally, this work could not have been done without all my participants.

Last but not least, I want to thank my family and friends who morally supported me during the last years. Your visits, telephone calls, emails and words kept me going, also through the hard times!

This work is dedicated to Käthe Bilgenroth,
who surely will never read it but always believes in me.

1 Preface

“What we commonly call the mind is a set of operations carried out by the brain. The actions of the brain underlie not only relatively simple motor behaviors such as walking and eating, but all the complex cognitive actions that we believe are quintessentially human, such as thinking, speaking and creating works of art.”

(Eric K. Kandel, in Kandel, Schwartz & Jessell, 2000, Principles of Neural Science, pp. 5-6)

Where does a decision come from? The answer is simple: it is a brain process. Decisions can be very different and occur on all levels of our every-day life: we can make small decisions or we can make life-changing decisions; we can be aware of making decisions or we can even react on something without having consciously made the decision to do so. Usually, we are happy to attribute unconscious decisions, such as finding ourselves opening the door of the fridge without having intended to do so, to automatic brain processes. We do not do this so easily, however, with conscious decisions, in spite of the fact that conscious decisions do not happen “in consciousness” *instead* of the brain (Lamme, 2006). It is not even the brain process that is conscious of itself; but a brain process that somehow *creates* what we perceive as a conscious decision, even though this explanation might not feel intuitively satisfying (Dennett, 1991; Prinz, 2003; 2006a). A conscious decision evolves in the brain (Libet, Gleason, Wright & Pearl, 1983; Soon, Brass, Heinze & Haynes, 2008) and can be evoked by stimulating the brain (Desmurget et al., 2009). One way to look at humans (and other primates) is to say that they are decision-making agents, permanently reacting on stimulation from the environment or from their own somatic and cognitive systems. Investigating different decision processes along the reaction chain from stimuli to responses is one of the most fascinating tasks with which one could possibly be faced. The decision making processes introduced in this work varied in terms of the *degrees of freedom* in different

dimensions that the decision makers had when choosing between responses. Nevertheless, all decisions were very simple; they required nothing more than a transformation from a visual stimulus to a motor response.

The first project addressed decisions in a rule-determined task context. It was comprised of a major functional magnetic resonance imaging (fMRI) study and a smaller fMRI control study (also see Bode & Haynes, 2009). The second project investigated a necessary precondition, namely the discrimination of stimuli under different levels of visibility, in a perceptual decision making study (also see Bode, Bogler, Soon & Haynes, under review). Three small behavioural experiments were used to optimise stimuli and experimental procedures. A major fMRI experiment was conducted as well as an additional fMRI localizer experiment, used for further detailed analyses. Finally, the third project compared guessing in perceptual decision making with intended free decisions (Bode et al., under review). This project was comprised of a small behavioural study and a major fMRI experiment. Despite being interconnected, all three projects make statements on their own right and require context-specific interpretations. They will therefore be presented and discussed in detail in their separate experimental sections. It is intended that each section is highly readable without frequent cross-references to the other sections, which is why each of them also contains an individual detailed methods part. Since all fMRI studies used similar data acquisition parameters and approaches for data analyses, this necessarily leads to some repetition in the description; this, however, hopefully makes each individual section more comprehensible. To further increase the readability of each section, less central parts of the projects (such as behavioural pre-tests, supplementary analyses and supporting results) are reported in individual appendices (A to C). This indeed requires some cross-referencing but hopefully avoids that the reader is too lost in details. Preceding the experiments, a general introduction will explain the basic framework for the presented work; the theoretical context is given in depth in the respective introductions to each experiment. The introduction to methods will then give a short overview about fMRI. It strongly focuses on multivariate pattern classification, which was used for all experiments presented here. This section might interrupt the flow of arguments and could be skipped by the experienced reader. Finally, following the experimental sections, a short general discussion of all studies is given to integrate all results into a common framework – and hopefully to shed light on different mechanisms of decision making, which are used in our daily ventures to find the adequate reaction to all the different stimuli that constantly influence us.

2 Introduction

2.1 General introduction

2.1.1 From stimuli to responses

The present work is dedicated to the question of how reactions to different elements (or stimuli) in the environment are prepared and encoded in the brain. Of course, this might be the broadest question one could possibly ask and most studies in psychology and neuroscience tackle this issue in some way. The number of meaningful elements that are behaviourally relevant is virtually unlimited. It can be a simple flash of light or a complex social interaction, requiring different reactions and involving fundamentally different processes. In its simplest form, however, the process can be depicted like this: a stimulus is presented, which has to be encoded by the neural system; a decision has to be made defining how to react on it and an appropriate (motor) response has to be selected and carried out. One example of such a stimulus is a traffic sign, which has a clearly defined meaning such as “stop here”. The person being confronted with it has to follow a simple rule by activating the corresponding *stimulus-response mapping*. A more challenging situation arises when the same stimulus would lead to a different reaction depending on the *context*. For example, the letter “C” on water taps refers to “cold water” in English speaking countries while in France, it stands for “eau chaude” (meaning hot water). The correct response (turning on the tap or not) requires the activation of abstract rule knowledge, also referred to as *task-sets* (Sakai, 2008). Hence, it is possible to vary the *degrees of freedom* in rule determinism to investigate response preparation. The first experiment used a task-switching approach (e.g. Brass & von Cramon, 2004a; 2004b; Dove et al., 2000) to disentangle the neural representations of fixed and flexible rules in the human brain. Rule information is one important component used by the cognitive control system located in prefrontal and parietal cortex to

flexibly guide behaviour (Miller & Cohen, 2001; Quintana & Fuster, 1999; Stoet & Snyder, 2004). It is an open question, however, which brain regions process rule information during different *processing stages* to the final response.

This first study left the aspect of real *decision processes* untouched. Following the instructed rule was the only required decision. Before reacting according to a rule, however, one has to make sure that the relevant stimulus was perceived correctly. For example, a road sign might easily be misinterpreted in a hail storm. The second study addressed this question and required perceptual decisions about weakly visible stimuli, using a *perceptual decision making* task (Gold & Shadlen, 2007; Heekeren, Marrett & Ungerleider, 2008; Smith & Ratcliff, 2009). Signal detection models of perceptual decision making typically assume that the ability to discriminate two stimuli is governed by the overlap of response distributions they evoke in the sensory system (Green & Swets, 1966). It has been shown that in the sensory system there exist neurons that reliably reflect perceptual *choices* under all visibility conditions (Britten, Newsome, Shadlen, Celebrini & Movshon, 1996). Under this model, the trial-to-trial variability in choice under conditions of low visibility is mainly due to noise fluctuations in the same sensory system. The alternative, however, is that *different neural systems* are involved in perceptual decision making depending on whether people have confidently perceived a stimulus. In line with this, different networks in prefrontal and parietal cortex were shown to be differentially activated for highly or poorly visible stimuli (Dehaene et al., 1998; Heekeren, Marrett, Bandettini & Ungerleider, 2004) and networks can process stimuli differently depending on their visibility (Fang & He, 2005). Also, the activation and information encoding in specialised sensory areas has been shown to be modulated by participants' successful categorisation performance (Grill-Spector, Kushnir, Hendler & Malach, 2000; Williams, Dang & Kanwisher, 2007). The experimental approach taken in the present work was to search for information encoding about the *outcome* of a perceptual decision process separately under high and low stimulus visibility. The dissociation of information about stimuli and choices was considered to be conceptually close to *choice probabilities* used in single-cell studies in monkeys (Britten et al., 1996; Gold & Shadlen, 2000; Kim & Shadlen, 1999; Nienborg & Cumming, 2009; Purushothaman & Bradley, 2005; Shadlen & Newsome, 2001).

Finally, decisions can be made even in the absence of superimposed rules or perceived stimuli. Everyday life, so it seems, permanently requires such stimulus-independent, free decisions. Even though one might never be in a completely stimulus-free environment (Thorndike, 1911), voluntary decisions are defined as internally motivated rather than

externally triggered. The last study therefore addressed the question of how free decisions (Haggard, 2008) compare to guessing in perceptual decision making. These might be fundamentally different since voluntary action selection and stimulus-based action selection have been shown to involve partially different brain networks (Waszak et al., 2005). On the other hand, as will become evident with the second study, perceptual decisions *without* sufficient sensory input (or guessing) might resemble free decisions at least in some important aspects, as both are *internal* decisions. Brain structures in medial prefrontal cortex (PFC) and medial parietal cortex have been shown to be involved in processing intentions and voluntary action selection (Bode et al., in prep.; Cavanna & Trimble, 2006; Cunnington, Windischberger & Moser, 2005; Deiber, Honda, Ibañez, Sadato & Hallett, 1999; Haynes et al., 2007; Jenkins, Jahanshahi, Jueptner, Passingham & Brooks, 2000; Lau, Rogers, Haggard & Passingham, 2004a; Soon et al., 2008) and were candidate regions for a common *internal* decision network. The last study varied decisions from determined to fully internal and by doing so completed the spectrum from the first study that used a fixed task setup. In order to investigate the underlying decision networks, all studies made use of a methodological approach that made it possible to directly decode the decision outcomes and abstract rules from regional cortical activity measured with fMRI (Haynes & Rees, 2006; Norman, Polyn, Detre & Haxby, 2006).

2.1.2 Localizing information about rules and decisions

In the present work, fMRI was used for the investigation of the neural basis of decision processes. This method is a powerful tool for linking cognitive processes to brain activity at a macroscopic level. It has been criticised, however, that neuroimaging does not contribute to the development of models in cognitive psychology (Coltheart, 2006). Functional MRI in particular was characterised as being better suited to formulating data-driven hypothesis rather than unambiguously selecting between them (Logothetis, 2008). Of course, using fMRI is not automatically guaranteed to reveal anything substantial; a *stronger* activation in a brain region for one process compared to a second process by no means rules out the general involvement of the same region in *both* processes. One way to escape this dilemma is offered by recently developed multivariate pattern classification approaches (for reviews see Formisano, De Martino & Valente, 2008; Haynes & Rees, 2006; Mur, Bandettini & Kriegeskorte, 2009; Norman et al., 2006). These methods do not look for simple activation

differences between experimental conditions within single locations but search for *information* encoding in local spatial activation patterns that could be used to predict abstract mental states. This means the *content* of decisions and the information flow in the brain from stimulus processing to response selection become accessible. In the present work, this approach was used to reveal information-encoding brain networks, which would have been overlooked by conventional analyses. It is believed that the decoding of information can indeed be used to test models about the questions outlined above. Demonstrating specific information encoding is superior to revealing unspecific activation, which could be simply related to the general preparation of mental operations or the support of processes in other brain regions. Hence, the conceptual link to the underlying cognitive process becomes closer than in classical fMRI studies. Multivariate methods are more sensitive than conventional analyses because they take into account patterns of brain activity at many locations, they can use weak (but stable) information and they can directly predict the participants' cognitive states (Haynes & Rees, 2006). To allow for a better understanding of the experiments, the next paragraphs will first give a brief overview about the principles of fMRI and data analyses before multivariate pattern classification approaches will be explained in greater depth. For more detailed information see e.g. the textbooks of Buxton (2002), Jezzard, Matthews & Smith (2003) or the review articles of Heeger and Ress (2002), Logothetis and Wandell (2004), Logothetis (2008), Haynes & Rees (2006) and Norman et al. (2006).

2.2 Introduction to methods

2.2.1 Basic physical principles of MRI

MRI is a method that noninvasively acquires images of the brain. This is possible because atomic nuclei possess an intrinsic angular momentum called 'spin'. In very simplified terms, the proton can be thought of as a small magnet, precessing like a spinning top about an axis. More than 70% of the human brain consists of water; and the H⁺ protons of water are spin-1/2 nuclei that dominate the signal measured with MRI (Logothetis, 2008). These normally random spin directions of protons can be aligned parallel to or anti-parallel to an externally applied strong magnetic field in a MRI scanner, denoted B₀ (common field strengths used for functional imaging in humans are 1.5 to 7 Tesla). The majority of protons will preferentially align more parallel to the B₀ field than anti-parallel, because of the lower (and therefore preferred) energy state accompanied with parallel alignment, even though the difference is very small. Importantly, this results in a net magnetization (described as a magnetization vector), which, however, cannot be measured because it is oriented parallel to the magnetic field B₀. Therefore, a radio frequency pulse (RF pulse) is applied perpendicular to B₀ with the amplitude B₁, matching the precession frequency of the protons. The latter is given by the Larmor-frequency (ω), defined by the magnetic field strength (B₀ [T]) and the gyromagnetic ratio (γ [MHz / T]), which is different for each kind of nucleus:

$$\omega = B_0 * \gamma$$

Due to the protons' resonance, the RF pulse causes the protons to absorb energy and has two important effects: first, it tilts the magnetization vector from the horizontal plane (usually described as dimension z) to the transversal plane (xy). Second, it aligns the precession of the spins, which means that the protons' rotations are in phase. This transversely rotating magnetization vector can then be recorded as a signal by a receiver coil placed around the head. The technique makes further use of the fact that after switching off the RF pulse, the transverse magnetization decays quickly, a process described as *relaxation*. The dipoles start to align with the magnetic field B₀ again and release their energy to the lattice. This process is called longitudinal relaxation or spin-lattice relaxation. It is described by the time-constant T₁ (the time it takes for 63% recovery of magnetization along the z -axis). The time-constant T₂ describes the time it takes for 37% of the transversal magnetization (xy -

plane) to decay, which is called spin-spin relaxation. The T_2 relaxation occurs because spins start precessing out of phase again, due to random, fluctuating fields caused by other spins in the surround. A faster transversal relaxation as predicted by T_2 , however, can be observed because of the additional effects of constant, static field inhomogeneities, described by the time-constant T_2^* . In order to acquire images, MRI makes use of these relaxation processes, which are differentially fast depending on the tissue. Different tissues therefore are visible as different contrasts in the images.

In order to acquire an image of the whole brain (also referred to as *volume* or *scan*), signals from each position in the brain (cuboid *voxels*) have to be received; the size of each voxel can vary between studies depending on research interests and technical requirements. Several steps are necessary for the spatial reconstruction of the signal. First, a field gradient is applied to B_0 (axial z -axis: inferior to superior), which makes it possible to excite only a small slice of tissue with a RF pulse of a given frequency and therefore to *select the slice position*. The decoding of the y -position is achieved by another short gradient applied along the coronal y -axis (anterior to posterior). When switched off, the positions along the y -direction have locked-in phase differences, because they precessed with a different speed for a very short time and can therefore be identified (*phase encoding*). Finally, another negative gradient is used along the sagittal x -axis (left to right) after the RF pulse is applied, followed by a positive gradient such that a *gradient echo* (GE) occurs after half of the signal read-out time (used to re-phase the spins). The frequency or precession therefore also varies along the x -axis (*frequency encoding*). With each measurement, the mean signal contains the sum of frequencies; the individual signals, which correspond to the spatial positions along the axes, can then be decoded using Fourier Transformation. The time between two RF pulses is called *repetition time* (TR); the time between excitation and measurement is called *echo time* (TE). These parameters can be chosen such that they determine the influence of T_1 and T_2 on the image contrast. Finally, different imaging sequences can be used, differing in their use of RF pulses, gradients and therefore in their T_1 -, T_2 - and T_2^* -weighting. All the presented studies used gradient-echo EPI sequences, which are very fast and efficient because they allow the acquisition of a full image using only one RF excitation pulse (for details refer to e.g. Buxton, 2002; Logothetis, 2008).

2.2.2 Functional magnetic resonance imaging (fMRI)

In order to investigate rules and decision processes as in the present studies, brain *activity* had to be measured and related to the *cognitive tasks*. This can be done with *functional MRI* (fMRI). It makes use of the blood-oxygen-level-dependent (in short: BOLD) contrast, which was discovered by Ogawa and colleagues. They found that MRI images displayed dropouts at the location of big blood vessels (Ogawa, Lee, Nayak & Glynn, 1990a; Ogawa, Lee, Kay & Tank, 1990b) and related this effect to *regional neural activity*; this assumption was experimentally confirmed shortly after (Turner, Le Bihan, Moonen, Despres & Frank, 1991). In short, oxygenated blood, oxyhaemoglobin (Hb), is diamagnetic, which enhances the signal. Deoxyhaemoglobin (dHb) is paramagnetic and introduces field distortions (susceptibility artefacts), which lead to less signal. Neural activity is accompanied by a *local increase in blood oxygenation*, which is needed for glucose metabolism. The increase in oxygen extraction from the blood is achieved by vascular changes, namely an increase in regional blood flow, in blood volume and in blood velocity, all of which are related in a complex way (see Buxton, 2002 for details). Importantly, even though more oxygen is extracted from the blood during neural activity, this is accompanied by a local *oversupply* in blood oxygen and therefore a *better BOLD signal* (for a discussion of the oversupply see e.g. Buxton, 2002; Logothetis & Pfeuffer, 2004). This increased BOLD signal occurs with a latency of 3-6 s after the onset of neural activity, depending on the task and the brain region involved (for details see Heeger & Ress, 2002). The location of the signal can be revealed with fMRI with high spatial resolution, whereas the temporal resolution is rather poor (e.g. compared to electroencephalography, EEG) because of the hemodynamic delay. Since 1992 the BOLD signal has been used in experiments in humans (Bandettini, Wong, Hinks, Tikofski & Hyde, 1992; Kwong et al., 1992; Ogawa et al., 1992).

It is crucial to understand what exactly the measured signal reflects in order to relate it to the cognitive mechanisms of interest. Is it valid to conclude that the BOLD signal simply reflects a local increase in neural activity in terms of a higher *firing rate* (action potentials, the output signals of neurons)? At least, this is usually the implicit assumption. The BOLD signal, however, does not perfectly correlate with the neurons' firing rate as measured with single-cell recordings in monkey cortex. Perisynaptic activity from neurons in the surround up to 3 mm from the electrode tip (*local field potentials*, LFP) was found to be the better predictor (Logothetis 2002; 2003; 2008; Logothetis & Pfeuffer, 2004; Logothetis and

Wandell, 2004). Moreover, LFPs predicted the BOLD signal even in the absence of spiking activity (Logothetis, Pauls, Augath, Trinath & Oeltermann, 2001), confirming the importance of *pre-synaptic activity* for the BOLD signal rather than simple output signals to other regions (Logothetis, 2008; Logothetis & Wandell, 2004). Unfortunately, the LFPs also represent an ambiguous mixture of activity. Logothetis (2008) pointed out that neural activity in general cannot be explained by a simple feed-forward integration model in which all information is just transmitted to the next “higher” brain area; connections between regions are feed-forward *and* feedback. Moreover, it is not the single neuron that processes the information but micro-units composed of many interconnected neurons, resulting in local mass activation. The local connectivity of such micro-units reveals strong excitatory and inhibitory recurrence, coining the term excitation-inhibition networks (EIN) (Logothetis, 2008). The BOLD signal could therefore not only arise from an increase in firing of task- or stimulus sensitive neurons but also from an increase in balanced excitatory and inhibitory conductance, which does not necessarily cause a net increase of stimulus-related cortical output. In conclusion, it seems that the BOLD signal always contains a mixture of sources and it can hardly be used to draw conclusions about what exactly happens within a single fMRI voxel. The magnitude of the fMRI signal therefore cannot be quantified to accurately reflect differences *between* brain regions (Logothetis & Wandell, 2004). Even the comparison between tasks *within* the same region might be problematic, which makes it theoretically impossible to deduce the exact role of an area for a task at hand. Nevertheless, today BOLD fMRI is one of the best tools for the investigation of the neural basis of cognitive functions in humans (Logothetis, 2008), but its restrictions have to be kept in mind: it will not be possible to unravel the exact mechanisms of the cognitive processing steps using fMRI. The present studies, however, still go beyond conventional fMRI studies by directly demonstrating the *encoding of information* within brain regions.

2.2.3 Pre-processing and univariate analysis of fMRI data

The classical approach of fMRI data analysis is briefly described in this section. It explains how to get from the acquired images to conclusions about cognitive tasks; but it is also crucial to understand the differences to the multivariate approach that was used in the present work. Detailed overviews about various other ways of data analyses can be found in

several textbooks (e.g. Faro & Mohamed, 2006; Friston, Ashburner, Kiebel, Nichols & Penny, 2006; Jezzard et al., 2003).

The signal in fMRI time-series contains a mixture of changes in neural activity as well as correlated and uncorrelated noise (Friston et al., 1995b). An important limitation for task-related signals is the *signal-to-noise ratio* (SNR) that defines how much signal can be attributed to the experimental condition. Averaging across *multiple* functional images from the same experimental condition is required to eliminate the (unsystematic) noise components. Therefore, a series of t images is recorded (which also can be expressed as v voxels with a *time-series* of t time-points). During data pre-processing, the images of each participant first have to be *realigned* to one image of the series. This procedure corrects for small movement that might have occurred during the scanning session (Brett, Johnsrude & Owen, 2002). For this, the brain is treated as a rigid body and six motion parameters are used for realignment (translational movement and rotation in three spatial dimensions). It is also possible to *co-register* the functional images to a structural, high-resolution image recorded during the same session. Additionally, it has to be considered that the functional images are multi-slice images and each slice samples a slightly different point in time. *Slice acquisition time correction* can be applied to compensate for this effect. In order to compare individual brain images between participants, the images must then be transformed into a common space: *spatial normalisation* uses computational warping to match the individual images to a standard brain (see e.g. Aguire, 2006). Template images are provided e.g. by the Montreal Neurological Institute (MNI) or the stereotactic brain atlas by Talairach and Tournoux (1988). Commonly, the functional images are also *spatially smoothed* using Gaussian kernels. This step is useful since it is physiologically implausible that sharp drops of activation occur at the boundaries between voxels and if these are observed in the data, they rather display the influence of measurement noise. It therefore helps to increase the normalities of the data, to overcome residual differences between participants and to reduce sensitivity for false positive results in single voxels' time-series (Aguire, 2006).

The statistical analysis aims at identifying voxels' time-series in which the signal of interest is greater than the noise level (Smith, 1994). The experimental conditions are contrasted with a baseline condition or with each other, depending on the experimental protocol and research question. The most common way to achieve this is to use a *General Linear Model* (GLM) (Friston et al., 1995a,b). The events (or different experimental blocks) are used as regressors to explain variance in the data. For each voxel's time-series (*univari-*

ate analysis), the fit to the model is quantified by a *parameter estimate* (β -value). In a simple way, for one stimulation condition this can be expressed as:

$$y(t) = \beta * x(t) + c + e(t)$$

In this equation, $y(t)$ corresponds to the observed data (with a vector of signal intensity values for each time point), $x(t)$ corresponds to the model (with a vector of ones and zeros representing the presence or absence of stimulation conditions), c is a constant (e.g. baseline intensity) and $e(t)$ corresponds to the error. Therefore, β is the value with which the model has to be multiplied to fit the data. In order to get a better fit of the model, the time-series can be convolved with the *haemodynamic response function* (HRF) to mimic the BOLD response in a more realistic way. The parameter estimates obtained from the GLM are fed into group-level statistical analyses, which then result in statistical maps, displaying the uncertainty of their estimation (t -values). This allows the identification of voxels that are believed to be significantly activated for each experimental condition. A statistical threshold (expressed as p -value) has to be applied to these maps, taking into account a correction for multiple comparisons. Often, additional voxel-cluster thresholds are used to further limit the possibility of false positives. *Differences* between the relevance (or activation) of voxels for experimental conditions can be inferred by subtracting parameter estimates from each other. Finally, the activated areas have to be labelled, which can be done on different levels (Brett et al., 2002).

2.2.4 The multivariate pattern classification approach

In recent years, new approaches were taken in analysing fMRI data. As outlined above, conventional univariate analysis compares each voxel in the brain separately with respect to *activation differences* between experimental conditions. This approach has proven to be very useful for identifying areas involved in cognitive processes; however, at the cost of information reduction. Differences between participants are treated as noise and fine-grained information in activation patterns are blurred or cancelled out. Only the “tip of the iceberg” (Formisano et al., 2008) of shared activity becomes visible. In contrast, pattern classification methods make use of unique fine-grained spatial activation patterns of individual participants’ data. The idea behind this approach is to extract *information* related to different cognitive processes by testing the *correlation* of activation patterns between

experimental conditions or by using them to *predict* experimental conditions directly (Formisano et al., 2008; Haxby et al., 2001; Haynes & Rees, 2006; Kriegeskorte, Goebel & Bandettini, 2006; Mur et al., 2009; Norman et al., 2006). This approach has been labelled “information based fMRI” (Kriegeskorte et al., 2006; Kriegeskorte & Bandettini, 2007a,b; Mur et al., 2009), “mind reading” or “brain reading” (Cox & Savoy, 2003), “multi-voxel pattern analysis” (MVPA) (Norman et al., 2006) or “multivariate decoding” (Haynes, 2009; Haynes & Rees, 2006).

The great advantage of this class of methods is that it accounts for the possibility that regions carry information about cognitive processes even though this might not be reflected in differences in the average signal (Haynes & Rees, 2006). Consider for example the neural responses to mental processes such as the processing of very similar visual stimuli, or abstract rules as in the present studies. In principle, one would expect that different neurons (or micro-units) might be involved in each computation but one would also expect these neurons to be located in the same macro-anatomical regions. Different neurons within a single voxel might code for one mental process (e.g. detecting the first stimulus), some might code for the other process (detecting the second stimulus), while other neurons might not code for any task related processes at all. A region of a relatively small size such as 1 mm² covering cortical surface will still contain 90.000 to 100.000 neurons (Logothetis, 2008). It is therefore very likely that conventional analysis will fail to detect small activation *differences* within this voxel (and neighbouring voxels) between conditions. If these voxels display small but *stable biases*, however, multivariate methods can link the resulting activation patterns to experimental conditions. This method could therefore be regarded as a test for *pattern stability*. It has been claimed that this approach is sensitive to differences at a finer scale than the voxel size (Haynes & Rees, 2005a; Haynes & Rees, 2006; Kamitani & Tong, 2005). This view has recently been challenged by simulations, which showed that supra-voxel information sources rather than sub-voxel information would account for decoding (Op de Beeck, 2010). Others, however, demonstrated that these simulations were partly based on inaccurate assumptions and could not dismiss potential sub-voxel information sources (Kamitani & Sawahata, 2010). Also at odds with Op de Beeck (2010), strong orientation-specific signals at the scale of cortical columns (below and up to 1 mm) in primary visual cortex (V1) of cats have been demonstrated using high-field and high-resolution fMRI; these orientation biases (which were also found in humans) were still detectable when the spatial scale was increased to several millimetres (Swisher et al., 2010). Likewise, the build-up of free decisions could be decoded from prefrontal cortex using a

standard resolution of $3 \times 3 \times 3 \text{ mm}^3$ voxels (Soon et al., 2008) as well as using a much smaller resolution of $1 \times 1 \times 1 \text{ mm}^3$ voxels (Bode et al., in prep.). Hence, these small biases are likely to be the basis for pattern classification approaches; however, most likely in combination with some large-scale global biases and effects of blood vessels (Kamitani & Sawahata, 2010; Swisher et al., 2010).

Multivariate pattern classification can potentially demonstrate involvement of brain regions beyond *unspecific* preparatory activation or support for processes taking place elsewhere in the brain. For example, several studies found that parietal and prefrontal cortex were both involved in task switching (e.g. Brass & von Cramon, 2004a; Crone, Wendelken, Donohue & Bunge, 2006; Rowe, Hughes, Eckstein & Owen, 2008). In order to decide if one region performs more abstract mental operations than the other (e.g. processing task rules), the direct test would involve looking at *information encoding* in the respective areas (see Experimental Section I). It has been stressed that only revealing the *representational content* can take the understanding of mental processes to the next level (Haynes, 2009; Mur et al., 2009). The basic methodological principles of multivariate pattern classification techniques will be described in the following paragraph. A more detailed description of the exact method used in the present studies will be given in each methods section separately.

2.2.5 Methodological principles of pattern classification

A number of different approaches for the analysis of fMRI activation patterns have been introduced in recent years, the most important ones will be described here. The fMRI data usually undergoes individual movement correction but no further normalisation or smoothing in order to preserve as much of the original information as possible (Formisano et al., 2008; Haynes & Rees, 2006; Mur et al., 2009; see Op de Beeck, 2010, and Kamitani & Sawahata, 2010, for a discussion about the effects of spatial smoothing). The next step always involves defining a set of voxels that constitutes the patterns (*feature selection*). The activation values from the selected voxels are combined to a *feature vector* (for an overview see Pereira, Mitchell & Botvinick, 2009). In principle, it is possible to use activation patterns from all voxels in the brain for classification (e.g. Polyn, Natu, Cohen & Norman, 2005). This approach, however, results in a very high-dimensional feature vector and therefore in a great proportion of noise, considering that only the minority of voxels will

carry task-related information; it might also lead to over-fitting. Several approaches have been suggested to reduce the dimensionality of high-dimensional feature vectors, including principal components analysis (PCA) (Mourão-Miranda, Bokde, Born, Hampel & Stetter, 2005), taking the most discriminative voxels only or taking the most activated voxels only (Mitchell et al., 2004). The latter approach risks missing voxels that display weak but stable biases, which could have enhanced the classification performance if included. One could also define a region of interest (ROI) according to anatomical or functional criteria (Mur et al., 2009). Others used a combination of anatomically defined regions and voxel sampling by activation strength (Haynes & Rees, 2005a; Kamitani & Tong, 2005). It has to be ensured, however, that voxel selection and pattern classification are performed on independent data-sets to avoid circularity (Kriegeskorte, Simmons, Bellgowan & Baker, 2009). Unfortunately, these approaches do not allow a search for local information encoding without making presumptions. An elegant way to overcome this was introduced by Kriegeskorte and colleagues (2006). These authors constructed a *searchlight* cluster with a fixed radius around each voxel in the brain and analysed the activation patterns contained within them. By doing so, the classification accuracy achieved from each of these clusters could be examined. This approach results in 3-dimensional brain *maps* of decoding accuracies, which are then normalised to standard space, spatially smoothed and can be subjected to group-level analyses (Haynes et al., 2007; Mur et al., 2009; Soon et al., 2008). This method was used for the present studies and will be described in greater detail in the separate method sections.

For the analysis, it is generally possible to take the data from each individual trial and sort them according to the experimental conditions (Haynes & Rees, 2005a; Haxby et al., 2001; Kamitani & Tong, 2005) or to apply a GLM for each voxel first, (using HRF or Finite Impulse Response, FIR, based modelling) and perform pattern classification on the resulting parameter estimates (beta-images) (Haynes et al., 2007; Mourão-Miranda, Reynaud, McGlone, Calvert & Brammer, 2006; Soon et al., 2008). Additionally, each time point in a trial can be included as a different feature (Mitchell et al., 2004; Mourão-Miranda, Friston & Brammer, 2007). Subsequently, the complete data-set is split into a “training data set” and a “test data set”, which have to be acquired independently (Haynes & Rees, 2006; Norman et al., 2006). In its simplest case, the data-set is split in half, e.g. odd and even runs. A correlation analysis is computed between the pattern vectors from both sets and each pattern is assigned to the condition with which it displays the highest correlation (Haxby et al., 2001). Other approaches more strongly stress the *prediction* aspect. The pattern vectors

from the training data-set are subjected to a classifier or decoder, which learns to distinguish between activation patterns relating to different experimental conditions. In order to do so, the patterns are treated as points in multidimensional space with as many dimensions as voxels contained in the pattern. The classifier has to estimate a decision boundary, or more precisely a decision hyperplane, in order to optimally separate pattern exemplars of different conditions from the training data set. In principle, the learned classifier is a model for the relationship between features and class labels (Pereira et al., 2009). In a second step, the test data set is used to test if the decision hyperplane can successfully classify independent data (Haynes & Rees, 2006; Mur et al., 2009; Norman et al., 2006). Cross-validation has to be performed to obtain a better estimation of the classification accuracy and to control for over-fitting. This can be done by dividing the data into independently acquired sub-sets from each functional run such that each run serves as the independent “test run” once. The final *decoding accuracy* is calculated as the average percentage of correct classifications from all cross-validation steps (Haynes & Rees, 2006). For a detailed illustration of this procedure see Figure 3-2 of the first study.

There are several mathematical algorithms used for decoding, each designed to find an optimal decision hyperplane for the classification problem (for detailed overviews see Formisano et al., 2008; Mur et al., 2009; Pereira et al., 2009). The most common classifiers are linear classifiers because these are simpler and non-linear approaches have not been shown to be superior (e.g. Cox & Savoy, 2003) and are additionally difficult to interpret (Pereira et al., 2009). Simple correlation approaches rely on minimum distance classification. They use the distance (e.g. Euclidian distance) between the test data example and the training data sets for each experimental condition (Pereira et al., 2009). This can be depicted by imagining the single vectors as dots in multidimensional space (often illustrated in a two-dimensional coordinate system; see Mur et al., 2009). The centroids from both sets (vector ‘clouds’ for each experimental condition) are simply connected using a straight line. A linear decision hyperplane (boundary) is placed orthogonal to the centroid connection line and divides it in the middle.

A different approach is linear discriminant analysis (LDA) (e.g. Haynes & Rees, 2005a; Kamitani & Tong, 2005) that projects each vector on a single linear discriminant dimension, which replaces the connection line between centroids. To obtain the decision hyperplane, the covariance of the vectors is taken into account such that it optimises the ratio of between-condition variance (which should be high) and the within-condition

variance (which should be low). This method, however, requires multivariate-normal distributions of each condition (Mur et al., 2009).

In the past years, support vector machine (SVM) classifiers have proven to be powerful in classification of fMRI activation patterns (Haynes & Rees, 2006; Haynes et al., 2007; Mourão-Miranda et al., 2005; 2006; Soon et al., 2008). This method works by a similar principle to the LDA, but does not assume multivariate normality. It differs, however, in how the decision hyperplane is set. Vector weights are computed, indicating the importance of each vector for the prediction. The *support-vectors*, the closest points in space between conditions, are taken to define a *margin* around the decision hyperplane such that the distance between the separating hyperplane to the nearest training samples is maximized (Formisano et al., 2008). Optimally, the decision hyperplane is then set in the middle of the margin, which, in turn, is set in a way that it extends to the closest vectors on both sides without including them. If no solution can be found that completely separates both conditions, it is possible to flexibly define the number of vectors to be ignored that would be misclassified in the training data-set. This method is not much affected by changes far away from the decision boundary but strongly affected by changes of the support vectors (Mur et al., 2009). Classification using SVMs can also be extended to more than two conditions, e.g. by calculating all pair-wise classifications and indentifying the most frequently chosen one.

2.2.6 The application of multivariate pattern classification

Pattern classification on fMRI data was first applied in the visual domain. Specifically, the question of how the brain encodes objects in the occipital and ventral temporal cortex led to great debate (e.g. Aguirre, Zarahn & D'Esposito, 1998; Downing, Jiang, Shuman & Kanwisher, 2001; Edelman, Grill-Spector, Kushnir & Malach, 1998; Epstein & Kanwisher, 1998; Epstein, Harris, Stanley & Kanwisher, 1999; Kanwisher, McDermott & Chun, 1997; Spiridon & Kanwisher, 2002). It was the use of pattern classification, however, that revealed the distributed nature of object category encoding in those regions (Haxby et al., 2001). These authors concluded that an unlimited number of categories and single objects could be encoded in the activation patterns with many voxels contributing to the coding. Cox and Savoy (2003) used SVM classifiers and demonstrated that decoding accuracies

were even higher if patterns were not restricted to voxels that showed the strongest BOLD signal during object presentation. Furthermore, training and testing on *different exemplars* of objects from the *same category* led to comparably good results (Cox & Savoy, 2003). In the following years, pattern classification has been increasingly used in the perceptual domain of object-, face-, scene- and motion recognition (e.g. Eger, Ashburner, Haynes, Dolan & Rees, 2007; Downing, Wiggett & Peelen, 2007; Haushofer, Livingstone & Kanwisher, 2008b; Kay, Naselaris, Prenger & Gallant, 2008; Kriegeskorte, Formisano, Sorger & Goebel, 2007; Kriegeskorte et al., 2008; Op de Beeck, Torfs & Wagemans, 2008b; O'Toole, Jiang, Abdi & Haxby, 2005; Peelen, Wiggett & Downing, 2006; Peelen & Downing, 2007; Schwarzlose, Swisher, Dang & Kanwisher, 2008; Shinkareva et al., 2008; Spiridon & Kanwisher, 2002; Sterzer, Haynes & Rees, 2008; Walther, Caddigan, Fei-Fei & Beck, 2009; Weber, Thomson-Schill, Osherson, Haxby & Parsons, 2009; Williams et al., 2007; 2008), as well as saliency (Bogler, Bode & Haynes, in prep.), colour encoding (Brouwer & Heeger, 2009) and odour perception (Howard, Plailly, Grueschow, Haynes & Gottfried, 2009).

Another fascinating application was in the investigation of subjective experiences, such as the awareness of visual contents. Kamitani and Tong (2005) used multivariate decoding to predict the orientation of simple stimuli composed of lines from activation patterns in early visual cortex during passive viewing. Furthermore, these authors could predict participants' covert attention to one of two overlaid line orientations. Using a similar approach, Haynes and Rees (2005a) could even predict the orientation of stimuli that were rendered invisible by masking from activation patterns in primary visual cortex (V1). This was a powerful demonstration that information can be encoded in – and decoded from – brain activity without the participants having conscious access to that information. The same authors have also demonstrated the converse by predicting changes in the content of conscious experiences while visual stimulation was kept constant using a binocular rivalry task (Haynes & Rees, 2005b).

Recently, the application of multivariate pattern classification was extended to higher cognitive processes, such as memory (Harrison & Tong, 2009; Polyn et al., 2005; Serences, Ester, Vogel & Awh, 2009), navigation (Hassabis et al., 2009), deception (Davat-zikos et al., 2005), decision making (Hampton & O'Doherty, 2007; Li, Ostwald, Giese & Kourtzi, 2007; Pessoa & Padmala, 2007; Serences & Boynton, 2007; Tusche, Bode & Haynes, 2010), number representation (Eger et al., 2009) and the storage and formation of intentions (Bode et al., in prep.; Haynes et al., 2007; Soon et al., 2008; Soon, He, Bode &

Haynes, in prep.) While the functional columnar organisation of early visual cortex is comparably well investigated, the functional organisation of higher cortical areas (e.g. involved in cognitive control) is still poorly understood. These recent studies, however, gave rise to the notion that higher cognitive processes, which were of interest for the present work, can indeed be decoded from local patterns of brain activity. Here, it was concluded that this approach would allow the investigation of rule-based behaviour as well as perceptual decision making, given its sensitivity for information about objects, the content of conscious experiences and decisions. By using a searchlight approach (Haynes et al., 2007; Kriegeskorte et al., 2006) the search for information encoding did not have to be restricted to pre-selected brain regions. The following experimental sections will link these detailed methodological considerations back to the concrete research questions.

3 Experimental Section I: Rule-based task preparation

“Life is the only game in which the object of the game is to learn the rules”

(Ashleigh Brilliant; <http://www.ashleighbrilliant.com>)

3.1 Theoretical background

3.1.1 Task-sets and task switching

One simple form of responding to a stimulus, which will be the starting point for the present work, is following a rule. In this case, there are no degrees of freedom for the acting agent because the response is clearly predefined by the context and the stimulus. Hence, if the rule is not willingly violated, there is no real decision process involved. For example, in road traffic (context) we respond to the red traffic light (stimulus) by stopping (response) and to the green light (alternative stimulus) by walking (alternative response). In our every day life, rule-based behaviour is very common; it allows for highly automated processes based on useful associations between context cues and the adequate responses (Bunge, 2004). In most situations, however, our behaviour is not as predefined as in the case of traffic lights. Crossing the street in Germany, for example, requires first looking to the left, whereas in Australia it requires looking right. Hence, responses to stimuli are often context-dependent, requiring a switch between multiple motor behaviours associated with the same stimulus. The term *task-set* is used to refer to this prospectively configured combination of perceptual, mnemonic and motor processes in an abstract form (Sakai, 2008). Flexible rule usage, swift interpretation of environmental cues and the ability to switch between adequate responses are key features of human (and primate) behaviour; hence much research has

focused on *task-switching*. In task switching, the agent needs to flexibly adjust its behaviour according to the current rule. This ability requires flexible reconfiguration of the flow of information through the brain. Perceptual information that enters the primary sensory areas needs to be routed into different motor responses depending on current tasks (or other context information), which also have to be encoded in the brain.

Early behavioural studies showed that participants needed additional time whenever they switched from one simple task to another (Jersild, 1927). These *switch costs* were reduced but not eliminated when the preparation of the alternative task was facilitated by longer stimulus-response delays using fixed task sequences (Monsell, 1996; Rogers & Monsell, 1995; Spector & Biederman, 1976) or variable cue-target intervals (Meiran, 1996). These residual costs were often attributed to the process of *task-set reconfiguration* but the exact sources are still under debate (see e.g. Rogers & Monsell, 1995; Rubinstein, Meyer & Evans, 2001). One explanation states that inhibition effects from the preceding trial resulting from strong associations between the same stimulus and the *other* motor response are affecting the current trial (Allport, Styles & Hsieh, 1994; Allport & Wylie, 1999). Importantly, many of these interpretations point to the conclusion that task-set reconfiguration can only be finally completed *after* the stimulus is presented and not beforehand (De Jong, 2000; Monsell, 2003). Hence, it can be assumed that *different processes* operating on the abstract level of task-sets contribute to rule-guided task preparation at different points *in time*.

3.1.2 The neural substrate of task-sets

In recent years, task switching has become a frequently used method for research on the neural mechanisms underlying cognitive control and rule-based behaviour (Bunge, 2004; Bunge & Wallis, 2007; Sakai, 2008). It has been suggested that the lateral prefrontal cortex (PFC) is involved in guiding flexible behaviour (Badre & D'Esposito, 2007; 2009; Badre, Hoffman, Cooney & D'Esposito, 2009; Constantinidis, Williams, Goldman-Rakic, 2002; Desimone & Duncan, 1995; Miller & Cohen, 2001). Signals from PFC could bias competing activity in other regions of the brain, thereby guiding behaviour from top-down (Miller & Cohen, 2001). In support of this, the human PFC is approximately twice as large as would be expected for nonhuman primates with a neocortex of equal size and it is also more

complexly folded (Passingham, 1973; Striedter, 2004; Zilles, Armstrong, Moser, Schleicher & Stephan, 1989), pointing to an involvement in high-level cognitive functions. Additionally, PFC is highly interconnected with all sensory association cortices, motor structures, basal ganglia, thalamus and several regions in temporal and parietal cortex, which may allow it to play a key role in the integration of information (Miller & Cohen, 2001).

Monkeys are also generally capable of performing tasks that require high levels of cognitive control and have therefore been intensively studied using single cell recordings in several sub-regions of PFC. In such studies, monkeys usually perform a simple task (e.g. tilting a joystick in one of two directions or performing a delayed saccade to a preceding stimulus) according to one or several given task rules while the activity of single neurons is recorded. The *tuning* of such neurons is then assessed according to their firing rates during particular phases of the trial. For rule-guided behaviour, PFC neurons were identified that showed tuning to the visual properties of behaviourally relevant stimuli (Funahashi, Chafee & Goldman-Rakic, 1993; Hoshi, Shima & Tanji, 2000; Hoshi & Tanji, 2004), even when these had to be kept in memory during a delay period (Constantinidis, Franowicz & Goldman-Rakic 2001a; 2001b), as well as neurons tuned to visual properties of stimuli that were not even relevant for the selection of behaviour (Quintana & Fuster, 1999). Other PFC neurons were tuned to motor responses while a different set of neurons encoded the predictability of the correct response (Quintana & Fuster, 1999). In other studies, PFC neurons were found that changed their tuning properties according to the abstract task rules (Amemori & Sawaguchi, 2006; Asaad, Rainer & Miller, 2000), encoded reward expectations and upcoming responses (Hoshi, Shima & Tanji, 1998; Hoshi et al., 2000; Wallis & Miller, 2003), encoded the monkeys' certainty about the correct sequence of action (Averbeck, Sohn & Lee, 2006) or directly encoded the abstract rule (Johnston, Levin, Koval & Everling, 2007; Kusunoki, Sigala, Gaffan & Duncan, 2009; Sigala, Kusunoki, Nimmo-Smith, Gaffan & Duncan, 2008; Wallis, Anderson & Miller, 2001). Remarkably, PFC neurons in a wide range of sub-regions seem to be capable of encoding all possible aspects of rule-guided task preparation, making it difficult to pinpoint – and generally questions – any functional regional specialisation (Duncan, 2001; Sigala et al., 2008).

There is also evidence that posterior parietal cortex (PPC) plays an important role in cognitive control (Quintana & Fuster, 1999). Parietal cortex has been less intensively investigated than prefrontal areas in this context but recent studies demonstrated the existence of neurons sensitive to abstract rules in several parietal sub-regions in the monkey brain (Freedman & Assad, 2006; Gail & Andersen, 2006; Gottlieb, 2007; Oristaglio,

Scheider, Balan & Gottlieb, 2006; Stoet & Snyder, 2004, 2007; Toth & Assad, 2002). This also led to new hypotheses that parietal cortex plays a much more active role in task reconfiguration than previously thought (Singh-Curry & Husain, 2009).

In humans, several brain areas have been shown to be activated during preparation of externally cued tasks, including lateral PFC, intraparietal sulcus (IPS), the pre-SMA and the anterior cingulate cortex (ACC) (Brass & von Cramon, 2002, 2004a, 2004b; Brass et al., 2003; Bunge, Hazeltine, Scanlon, Rosen & Gabrieli, 2002; Braver, Reynolds & Donaldson, 2003; Crone et al., 2006; Dosenbach et al., 2006; Dove et al., 2000; Forstmann, Brass, Koch & von Cramon, 2005; Kimberg, Aguirre & D'Esposito, 2000; Rowe et al., 2008; Rushworth, Paus & Sipila, 2001, Rushworth, Hadland, Paus & Sipila, 2002a; Rushworth, Passingham & Nobre, 2002b; Sohn, Ursu, Anderson, Stenger & Carter, 2000; Toni, Schluter, Josephs, Friston & Passingham, 1999; Yeung, Nystrom, Aronson & Cohen, 2006). In line with findings in monkeys, a fronto-parietal network emerges as being important for rule-guided behaviour. However, it is still heavily debated as to which brain areas correspond to which processing stages from stimulus representation to response execution. For example, using fMRI ventrolateral prefrontal cortex (VLPFC) and dorsolateral prefrontal (DLPFC) were found to be differentially activated depending on the task domain (Sakai & Passingham, 2003; Sakai & Passingham, 2006; Yeung et al., 2006). Others suggested that DLPFC is generally more involved in rule-based response selection while anterior VLPFC plays a key role in the retrieval of rule knowledge; posterior VLPFC, pre-SMA and parietal cortex may be involved in rule maintenance and the transformation of rules into action codes (Bunge, 2004). Brass and von Cramon (2002; 2004a; 2004b) dissociated the visual properties of a rule cue and its abstract meaning and could show that a sub-region of VLPFC, the inferior frontal junction (IFJ), was important for the selection of task relevant information. These authors assumed that the IFJ might be a key region in cognitive control across different tasks (Brass, Derrfuss, Forstmann & von Cramon, 2005a; Derrfuss, Brass & von Cramon, 2004; Derrfuss, Brass, Neumann & von Cramon, 2005). More abstract operations, such as task management or abstract mental operations, were suggested to be carried out by anterior frontopolar cortex (FPC) (Badre & D'Esposito, 2009; Bunge, 2004; Koechelin & Hyafil, 2007), while ACC and the anterior insula / operculum were discussed to constitute a core task-set system that regulates activity in other brain regions (Dosenbach et al., 2006). The specific role of the prefrontal sub-regions, however, is far from clear; competition and interaction of representations of rules, stimuli and responses may also take place within the same brain regions (Sakai, 2008).

Parietal cortex is most commonly thought to play a subordinate role in cognitive control, mainly linked to storage of stimulus-response (S-R) associations, attention and response preparation (Brass & von Cramon, 2002, 2004a, 2004b; Brass, Ullsperger, Knoesche, von Cramon & Phillips, 2005b; Bunge, 2004; Crone et al., 2006; Rushworth et al., 2001). There is also much debate regarding the roles of different prefrontal and parietal areas *in time*, i.e., from the retrieval of a general rule, to its associations with response representations and to the final motor preparation (Brass and von Cramon, 2004a, 2004b; Brass et al., 2005a; Braver et al., 2003; Bunge, 2004; Bunge et al., 2005). In most studies, it has also remained unclear whether brain activity related to task switching reflected the specific *encoding* of the task, or whether it instead reflected unspecific, global processes that did not differentiate between tasks.

3.1.3 The goal of the present study

The present fMRI study investigated how information about task-sets builds up across time when participants are cued to use different rules to guide their behaviour. Being able to pinpoint the *temporal sequence of information encoding* would allow conclusions to be drawn about two important aspects of task preparation: first, identifying the *content* of information stored in a specific brain region could shed light on the specific processes in which a particular region is engaged. Second, knowing the *time* at which a specific region holds that information would help clarify its role in preparing rule-guided behaviour. Specifically, multivariate pattern classification (Haynes & Rees, 2006; Norman et al., 2006) was used to disentangle the encoding of task-sets, target stimuli and motor responses (Passingham, Toni & Rushworth, 2000). Multivariate decoding has been successfully applied to analyse high-order cognitive functions beyond visual representations from patterns of brain activity including prefrontal cortex (Hampton & O'Doherty, 2007; Haynes et al., 2007; Li et al., 2007). Additionally, it can be used to track the built-up of decisions across multiple brain regions over time (Soon et al., 2008). In the present study, a task was used that required participants to alternate between two simple task-sets based on a visual cue presented prior to each trial. The task was temporally staggered and required two different mappings of visual stimuli to motor responses. Using this procedure, it was possible to identify different stages of information processing during rule-based response preparation.

3.2 Methods

3.2.1 Task-switching experiment

Participants

Seven male and seven female participants took part in the study and gave written informed consent to the test procedure. The experiment was approved by the local ethics committee and was conducted according to the Declaration of Helsinki. All participants were right-handed and had normal or corrected to normal visual acuity. Data from two participants were excluded due to poor task performance in one case and technical problems in recording responses in the other case. The final sample consisted of five male and seven female participants (mean age 26.4, range 24-30 years).

Stimuli

All stimuli were created using MATLAB 7.0 (The MathWorks, Inc.) in combination with the Cogent toolbox (<http://www.vislab.ucl.ac.uk/Cogent>). During the experiment, stimuli were presented via a projector (resolution 1024x768 pixel, 60Hz) that projected from the head-end of the scanner onto a screen. Participants viewed the projection through a mirror fixed on the head coil. The visual angle was $\alpha = 0.72^\circ$ for cues and $\alpha = 7.15^\circ$ for target stimuli. The target stimuli consisted of dynamic colour patterns that yielded strong responses in colour and motion responsive regions without semantic associations (Figure 3-1). Each pattern was based on a superposition of phase-randomised sinusoids that was used to index colours in a circular colour map. These moving coloured patches were animated by smoothly varying the phase, which created the impression of pulsation: for each stimulus, ten single frames of these patterns were shown in rapid succession, each a slightly warped rendering of the preceding one. This series of 10 frames (700 ms) was repeated six times for the duration of 4200 ms, thus creating the impression of smooth, wave-like movement. For manual responses, a custom-made MR-compatible joystick was used which allowed left and right movement only. The joint of the joystick was adjusted until participants reported a comparable amount of effort to move it in each direction. The joystick was fixed on an MR-

compatible plexiglass construction, the surface functioning as a table. This construction allowed for the adjustment of the joystick table in every desired angle to ensure that each participant's arm rested in a comfortable position and to avoid any additional movement artefacts (Appendix A, Figure A-1).

Experimental procedure

The experiment used a variant of a classical task-switching paradigm (Dove et al., 2000). One of two possible cues was given at the beginning of each trial. This cue represented the rule that specified which stimulus-response mapping had to be established between the two possible target stimuli and the two responses (therefore it defined the task-set). This yielded a total of four conditions (see Figure 3-1). Every run started with a 5600 ms presentation of a black screen with a white cross in the centre, upon which the participant had to fixate. At the beginning of each trial, one of the two task cues was presented for 1400 ms in the form of a letter ("A" or "B") indicating the active rule. Cue "A" instructed the participant to move the joystick to the left for target 1 and to the right for target 2. Conversely, cue "B" reversed the mapping, indicating that participants should move a joystick to the right for target 1 and to the left for target 2 (see Figure 3-1). The cue was followed by a 2800 ms delay during which the fixation cross was shown. After this interval, one of the two target stimuli was presented for 4200 ms in the centre of the screen with the fixation cross superimposed. Participants were instructed to observe the target stimulus for the entire period of its presentation, but not to respond yet. Subsequently, once the target stimulus disappeared, participants were allowed to respond according to the task rule. The response period lasted for 2800 ms. This procedure allowed the de-correlation (and hence dissociation) of task-sets, stimuli and responses because either stimulus could lead to either response, depending on the instructed rule. To familiarise the participants with the joystick and the test procedure, a 20-minute training session outside of the scanner was conducted one or two days prior to the fMRI experiment. In the MRI-scanner, the participants were instructed to fixate throughout all stages of the experiment. The experiment consisted of four runs, each repeating the four different conditions 16 times, resulting in 64 trials in total. The order of the four conditions was repeatedly pseudo-randomised in blocks. After every 16th trial (4 blocks) there was a pause of 14 s.

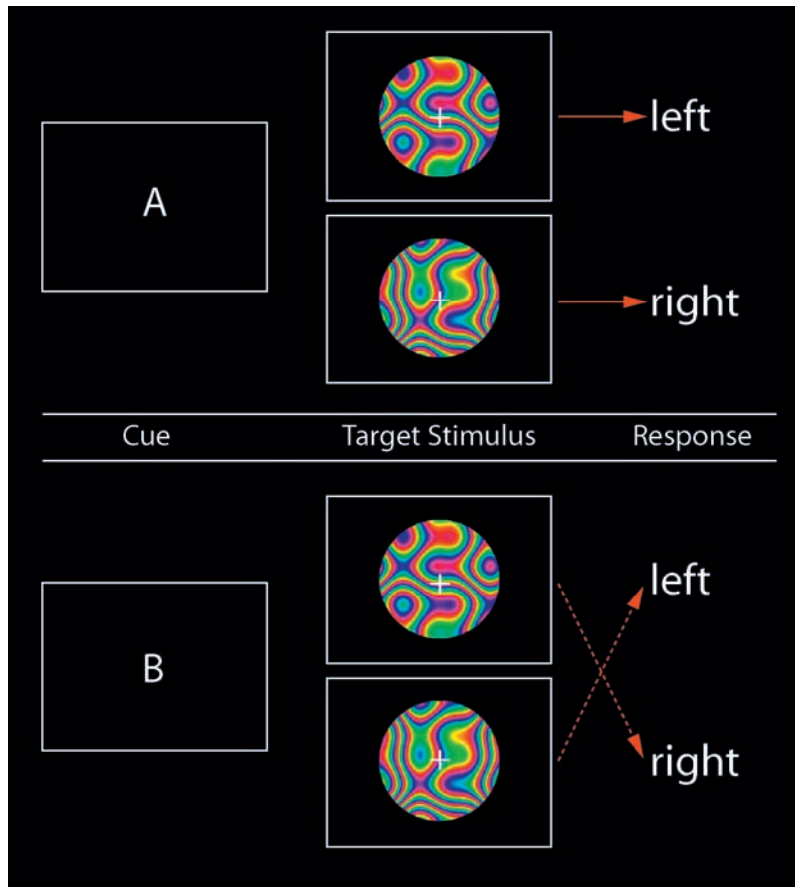


Figure 3-1: Experimental paradigm. Every task-set contained two stimulus-response mappings. Cue “A” indicated that a joystick should be moved to the left after presentation of target stimulus 1 and to the right after presentation of target stimulus 2 (picture above). Cue “B” indicated the reverse mapping (picture below). In each trial, one of the two cues (“A” or “B”) was presented for 1400 ms to indicate the active rule, followed by a delay of 2800 ms. Subsequently, one of the two target stimuli was presented for 4200 ms. Participants responded immediately following the stimulus presentation (response delay 2800ms).

Functional imaging

A Siemens TRIO 3T scanner (Erlangen, Germany) with standard head coil was used to acquire gradient-echo EPI functional MRI volumes of the whole brain (42 axial slices, repetition time TR = 2800 ms, echo time TE = 30 ms, resolution 3x3x2 mm³ with 1 mm gap). Within each of the four runs, 278 images were acquired for each participant. The first two recorded images were discarded to allow for magnetic saturation effects.

Data analysis

The data were subjected to three pattern classification analyses. Each of them was designed to identify brain regions carrying information about one of the three routes of information processing, namely the target stimuli, motor responses or task-sets. This approach goes beyond conventional univariate analysis, since it was designed to discriminate between the individual representations *within* a class of experimental conditions, which may be too similar for the mean signals to show any differences (see Appendix A, Univariate analysis).

A Finite Impulse Response (FIR) model was used as implemented in SPM2 (<http://www.fil.ion.ucl.ac.uk/spm/>). A high-pass filter with a cut-off of 128 s removed low frequency drifts in the time series at each voxel. The model was conducted based on motion and slice timing corrected data that was neither normalised nor smoothed. By not smoothing the data, the sensitivity for information encoded in fine-grained spatial voxel patterns was maximized (Haynes and Rees, 2005a; Kamitani and Tong, 2005). For the FIR model each trial was subdivided into eight distinct time bins of 2800 ms (= 1 TR) covering a total of 22400 ms. In each of these time bins, the cortical response was estimated separately by one parameter. This allowed the build-up of informative pattern signals to be analysed across time. A “searchlight” approach was implemented to analyse activity patterns at each brain location (Haynes et al., 2007; Kriegeskorte et al., 2006). This approach allowed the search for informative voxels in an unbiased fashion across the whole brain. Modifying the approach by using a FIR model, it was possible to estimate the information not only for each location but also for each time point separately (Soon et al., 2008). The general procedure is illustrated in the following for one time bin but was computed separately for each of the eight time bins.

First, the task-sets were decoded from brain activity. A spherical cluster of N surrounding voxels ($c_{1..N}$) within a radius of four voxels was created around a voxel v_i . The FIR-parameter estimates for these voxels were extracted separately for both task-set conditions at a given delay for each run. They were then transformed into pattern vectors for each condition for each run of each participant. These vectors represented the average spatial response patterns to the given task-set condition from the chosen cluster of voxels (Figure 3-2a). In the next step, multivariate pattern classification was used to assess whether information about the experimental condition was encoded in the spatial response patterns. For this purpose, the pattern vectors for three of the four runs were assigned to a “training data

set” that was used by a linear support vector machine (SVM) pattern classifier (Müller, Mika, Rätsch, Tsuda & Schölkopf, 2001) with a fixed regularisation parameter $C = 1$. First, the classifier was trained on this data to identify patterns corresponding to each of the two task-set conditions (LIBSVM implementation, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>). It was then used to classify independent data from the last run (“test data set”). Cross-validation (4-fold) was achieved by repeating this procedure independently with each run acting as the test data set once, while the other runs were used as training data sets (Figure 3-2b). The decoding accuracy was assessed by averaging the results of all four classification iterations and was assigned to the central voxel v_i of the cluster. It therefore reflected the accuracy of classification based on the given spatial activation patterns of this particular local cluster. Classification accuracy significantly above chance implied that the local cluster of voxels encoded information about the task-sets, whereas chance level performance implied no information. The same analysis was then repeated with the next spherical cluster, created around the next spatial position at voxel v_j . Again, an average decoding accuracy for this cluster was extracted and assigned to the central voxel v_j . By repeating this procedure for every voxel in the brain, a 3-dimensional map of decoding accuracies for each position could be created (Figure 3-2c). Furthermore, because the complete classification process was performed for each time bin separately, independent accuracy maps were available for each time bin.

The resultant accuracy maps were normalised to a standard stereotaxic space (Montreal Neurological Institute EPI template), re-sampled to an isotropic spatial resolution of $3 \times 3 \times 3 \text{ mm}^3$ and smoothed with a Gaussian kernel of 6 mm full width at half maximum (FWHM) using SPM2. This allowed a random-effects group-level analysis, computed on a voxel-by-voxel basis, to test the decoding accuracy for each position in the brain statistically across all participants (Haynes et al., 2007) and to track the evolution of decoding accuracies across time in separate brain regions (Soon et al., 2008).

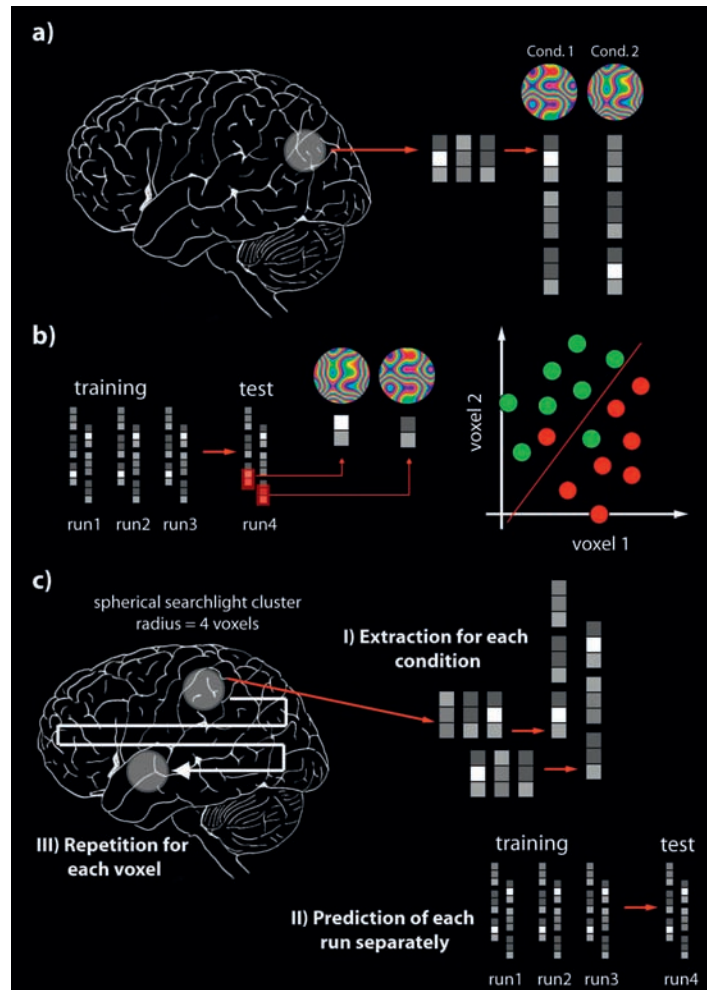


Figure 3-2: Multivariate pattern classification: a) Extraction of voxels for classification: Parameter estimates for all voxels in spherical cluster (radius = 4 voxels) around voxel v_i were calculated separately for each of the two conditions (here illustrated for target stimuli, functioning likewise for task-sets and responses). Subsequently, they were transformed into two pattern vectors. b) Classification: Pattern vectors of all runs but one were assigned to a “training data set”, the remaining run was assigned to a “test data set”. A linear support vector pattern classifier was trained on the “training data set” and subsequently applied to classify the “test data set” as belonging to one of the two conditions (here illustrated for two voxels where the decision boundary is located in 2-dimensional space. With N voxels in a spherical cluster, the decision boundary is located in N -dimensional space). For cross-validation, the classification procedure was repeated for all runs providing the “test data set” separately. The decoding accuracy was the average accuracy [%] across all cross-validation steps. c) Searchlight approach: After data extraction (I), pattern classification and cross-validation (II), a new local spherical cluster (radius = 4 voxels) was constructed around the adjacent voxel v_j . The classification process was repeated for all voxels serving as the central voxel of a “searchlight”-sphere once (III). Using this procedure, a 3-dimensional brain map of decoding accuracy values was created.

3.2.2 Control experiment with fixed stimulus-response mappings

A small control study was conducted in order to investigate whether regions engaged in cognitive control (beyond visual and motor cortex) were also involved when task-sets were *fixed* and no flexible switching between rules was required. Notably, in this case stimuli and responses were always correlated and their encoding cannot be dissociated.

Participants and experimental procedure

Five healthy, right-handed participants with normal visual acuity took part and gave written informed consent to the test procedure. For one participant, functional data was not completely recorded because of technical problems. Data from the remaining four participants (2 female; mean age 28.5; range 25-34 years) is reported here.

The experiment consisted of five runs in which four different visual target stimuli were presented, similar to those used in the task-switching experiment (see Figure 3-6a). Every run started with a fixation period of 5600 ms, followed by the presentation of one of the four stimuli in the centre of the screen for 11200 ms with the fixation cross superimposed. Participants were instructed to fixate and not to execute a response until the target stimulus disappeared. Subsequently, participants were instructed to respond using one of four buttons on a response device (operated by the index- and middle fingers of both hands; response period 5600 ms). The mapping between target stimuli and response buttons was fixed during the experiment. The presentation order of the four stimuli was pseudo-randomised such that no stimulus was repeated on the consecutive trial. In sum, all stimuli were shown ten times in each run, resulting in 40 stimuli per run and 200 stimuli in total.

The scanning procedure was identical to the task-switching experiment, except that 242 images were acquired for each run. For analyses, the fMRI data were again motion corrected and re-sampled ($3 \times 3 \times 3 \text{ mm}^3$). A GLM was used as implemented in SPM2 that consisted of eight boxcar regressors, each convolved with a canonical HRF. The first four regressors modelled the target stimuli, and the last four regressors modelled the motor responses (button presses). Again, a searchlight pattern classification analysis was implemented as described above to decode the four stimuli and their correlated motor responses.

3.3 Results

In the task-switching experiment, participants performed well with a low error rate of 2.6% (standard deviation SD 5.0; range 0 to 5.9% of 64 trials per run). Only trials with correct responses were included in the following analyses.

The present study investigated whether distinct stages of information processing could be decoded from brain responses. Specifically, multivariate pattern classification was applied to each area in turn to assess whether it was possible to decode which of the two task-sets, target stimuli and motor responses were engaged on a particular trial (chance level was always 50%). First, the encoding of the task-sets was investigated. Each trial began with the presentation of a task cue indicating which rule to use in determining the response. A sequence of encoding stages for the task-sets could be found (Figure 3-3, top row). In an early phase, only visual cortex encoded task-cues above chance, peaking with up to 63% decoding accuracy. Subsequently, it was possible to decode which specific task-set was engaged from high-level control regions in parietal and prefrontal cortex. Specifically, decoding was first possible from left IPS (peaking at 60% accuracy) even before target presentation, followed by left posterior VLPFC (peaking at 56%). The third region encoding the task-sets was located in the anterior part of left VLPFC and showed significant accuracy only very late (peaking at 58%), around the time of response execution (Figure 3-4). The position of peak decoding accuracy in posterior VLPFC was anterior and ventral to the IFJ found by Brass and colleagues (Brass and von Cramon, 2002; 2004a; Derrfuss et al., 2005). The two regions showed no overlap.

The next analysis was conducted to reveal brain regions encoding the identity of the target stimuli over time (Figure 3-3, second row). Regions showing above chance decoding accuracy throughout target stimulus presentation were located in bilateral visual cortex, with a peak accuracy of 72%. Interestingly, a region in prefrontal cortex, the left anterior VLPFC, also showed significant decoding accuracy, again in a later processing stage, peaking at 56% (for details see Appendix A, Figure A-2).

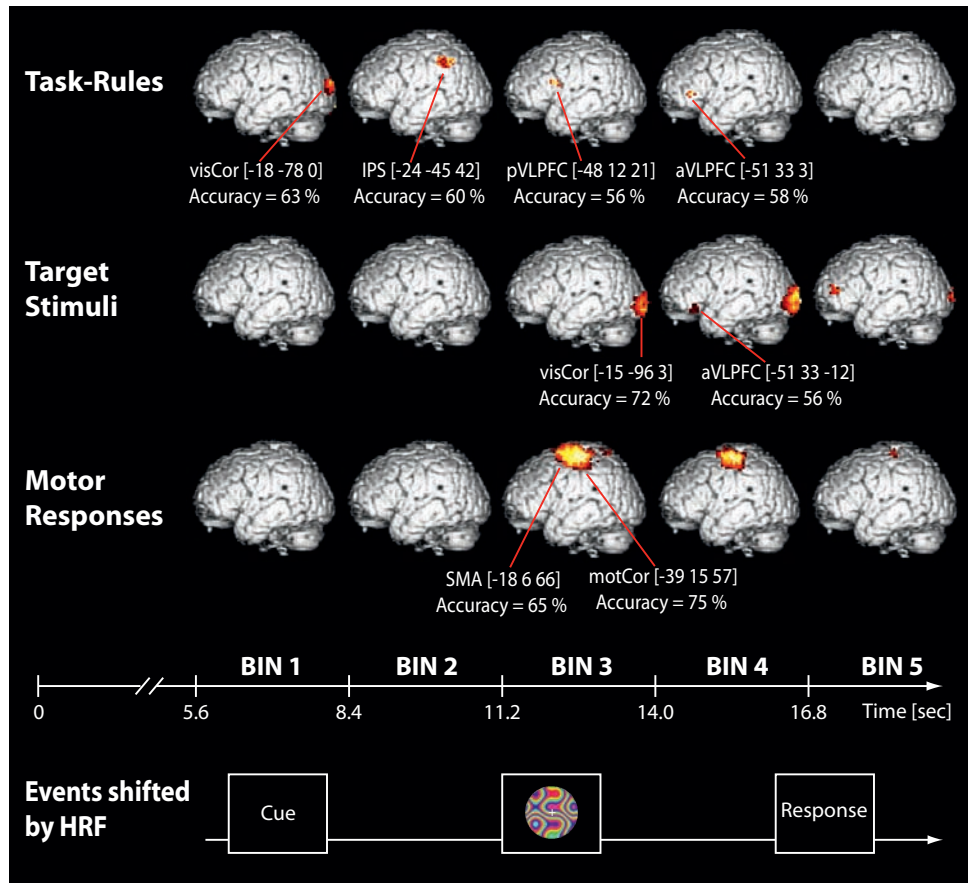


Figure 3-3: Decoding the task-sets, target stimuli and motor responses over time. The parameter estimates from the FIR model were used for 3 independent multivariate searchlight pattern classification analyses (spherical voxel clusters with radius = 4) to decode information about task-rules (upper row), target stimuli (second row) and motor responses (third row) over time. The FIR model consisted of eight time bins, each of which had a length of 2800 ms to model the entire duration of each trial. The cue was presented at the beginning of each trial (onset 0 sec), followed by the target stimulus (onset 4.2 sec) and the motor response (onset 8.4 sec). To account for the temporal delay of the BOLD signal the decoding time bins were shifted by 2 volumes (i.e. time bin 1 is the earliest that could reflect cue related activity). Time bin 3 coincided with the target stimulus presentation, and time bin 4/5 with the motor response (bottom row). A perfect match between events and time bins cannot be postulated since the time bins cover a period of 2.8 seconds each, and the delay of the BOLD response can only be coarsely estimated. The task cues were found to be encoded bilaterally in visual cortex (visCor) (bin 1; $p < .00001$ uncorrected; here displayed for left hemisphere). Information about task-sets then shifted to the left IPS (bin 2; $p < .0001$ uncorrected), the pVLPFC (bin 3; $p < .001$ uncorrected) and finally the aVLPFC (bin 4; $p < .001$ uncorrected). Target stimuli were encoded bilaterally in visual cortex (bin 3-5; $p < .05$ FWE corrected; here only displayed for left hemisphere) and left aVLPFC (bin 4-5; $p < .001$ uncorrected). Responses could be decoded bilaterally (here only displayed for left hemisphere) from SMA, primary motor cortex (motCor) and left medial occipitotemporal sulcus (not displayed) (bins 3-5; $p < .05$ FWE corrected). For

better visualisation, decoding accuracies are displayed with a threshold of $p < .001$ uncorrected for task-rules and target stimuli as well as $p < .00001$ uncorrected for motor responses. Coordinates displayed are MNI coordinates.

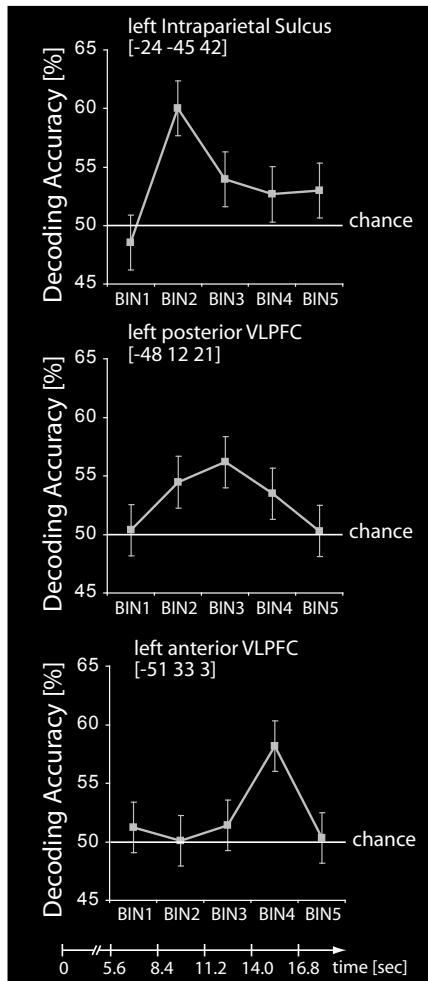


Figure 3-4: Information time courses for task-set decoding. Displayed are mean decoding accuracy and standard errors for the left IPS (upper part), the left posterior VLPFC (middle part) and the left anterior VLPFC (lower part) over time. The cue was presented at the beginning of each trial (onset 0 sec), followed by the target stimulus (onset 4.2 sec) and the motor response (onset 8.4 sec). To account for the temporal delay of the BOLD signal the decoding time bins were shifted by 2 volumes (i.e. time bin 1 is the earliest that could reflect cue related activity). The IPS showed an early peak in decoding accuracy before the presentation of the target stimulus (time bin 2; $p < .0001$ uncorrected). Decoding accuracy in posterior VLPFC slowly increased in the same time bin without reaching significance and peaked during early stimulus target presentation (time bin 3; $p < 0.001$ uncorrected) and gradually decreased to baseline in the next bins. Task-sets could be decoded from left anterior VLPFC during the fourth time bin only ($p < 0.001$ uncorrected). The time course from visual cortex related to the cue presentation is not shown in this figure. Coordinates displayed are MNI coordinates.

Motor responses could be decoded from spatial activation clusters in motor-related brain regions including bilateral primary motor cortex and bilateral premotor cortex with a peak accuracy of 75% (Figure 3-3, third row; for details see Appendix A, Figure A-3). All regions showed a similar characteristic information time course: the peak in accuracy corresponded to the period of early stimulus decoding, and then slowly decreased back to chance level. This result is not surprising since participants could prepare the response from

the moment the target stimulus was presented, but were instructed not to execute until the target stimulus presentation had ceased.

Comparing the amplitude of the BOLD signal across time bins to the time course of information encoding (decoding accuracies) revealed that these did not necessarily correspond (Figure 3-5). Especially for motor responses it could be observed that while *information* about the action to be performed was already present in motor areas during early stimulus presentation, the BOLD amplitude showed a peak not before response execution. Prefrontal and parietal regions showed an even stronger de-correlation between activation and information time courses (also see Appendix A for detailed univariate analysis). The process-relevant information at a given time point was therefore encoded in fine-grained spatial activation patterns rather than in the mean signal in single voxels.

Given previous findings by Asaad, Rainer and Miller (1998) that lateral PFC can encode *combinations* of stimuli and rules, it was also investigated whether the four individual rule-stimulus combinations could be decoded from prefrontal and parietal cortex. Spatial activation patterns in prefrontal and parietal areas that were predictive for task-sets and stimuli were found to exhibit a tendency to encode specific combinations of rules and target stimuli. This trend, however, did not reach significance (for details see Appendix A, Table A-4 and Figure A-4).

In order to assess whether the information in prefrontal cortex about *target stimuli* was also present in a simpler visually guided experiment, the data of the control experiment was analysed. In this study, the stimulus-response mappings were kept constant, meaning that no switch between task-rules was required. The error rate was very low with 0.63 % for all trials in the whole experiment (mean 1.2 errors out of 200 trials; SD = 1.30). Target stimuli and correlated motor responses could only be decoded from bilateral visual cortex (65% decoding accuracy) and motor cortex (54% decoding accuracy) where chance level was 25% (Figure 3-6). Contrary to the task-switching experiment, there were no regions in prefrontal cortex from which decoding of visual stimuli was possible, suggesting that prefrontal information about target stimuli was specific to tasks that required alternation between different stimulus-response mappings. Furthermore, these results suggest that in the task-switching experiment, the encoding of visual properties of the *cue* itself cannot account for the information present in the PFC and IPS.

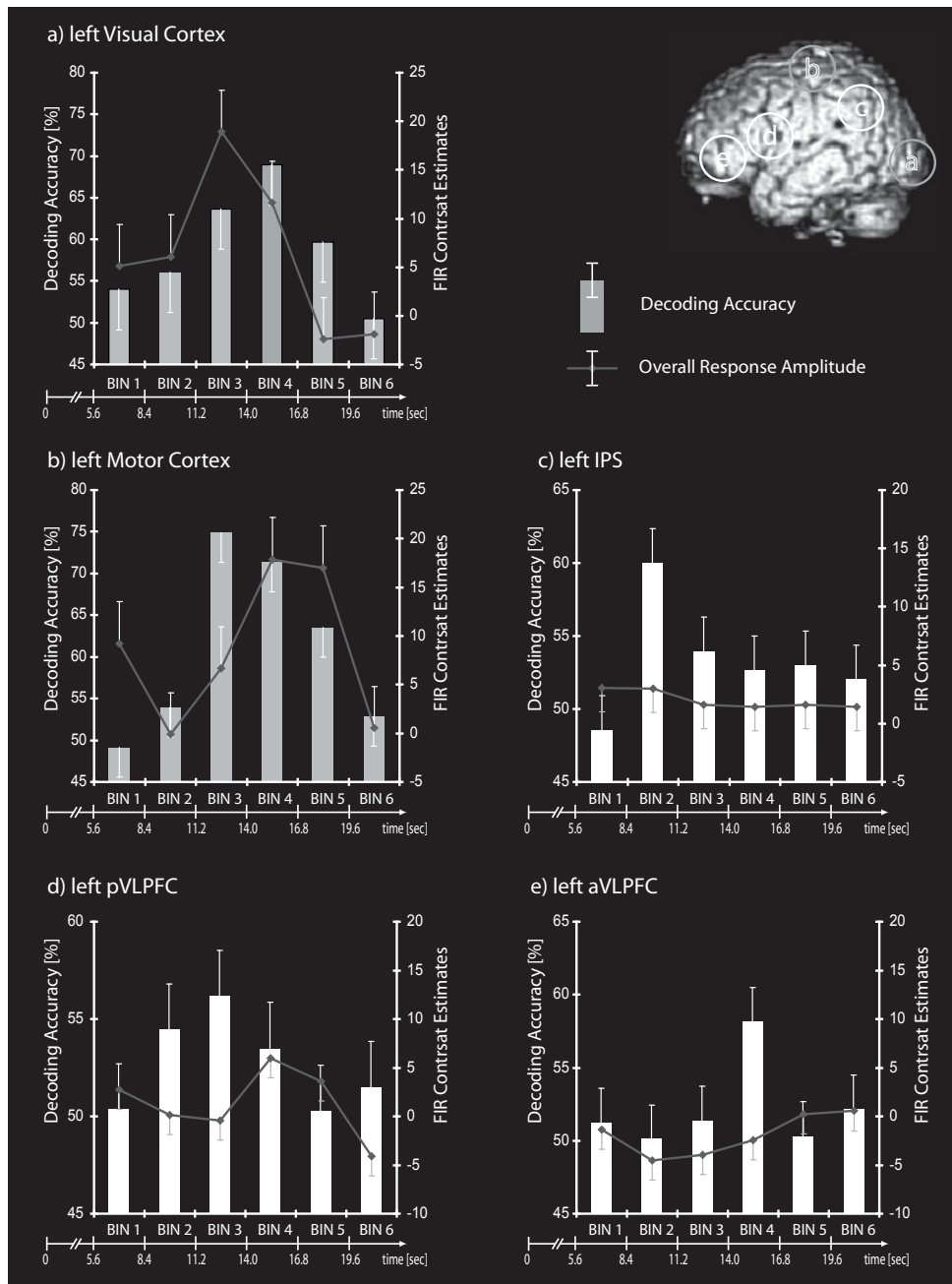


Figure 3-5: Averaged decoding accuracies and FIR contrast estimates. The graph shows the overall BOLD response amplitude (solid grey lines; +/- SE) for those central searchlight voxels that demonstrated the highest accuracy for decoding stimuli and motor responses (dark bars; +/- SE (a) visual cortex [-15 -96 3]), (b) motor cortex [-39 -15 57]), and task rules (bright bars; +/- SE (c) IPS [-24 -45 42], (d) pVLPFC [-48 12 21], (e) aVLPFC [-51 33 3]) averaged across all participants. The cue was presented at the beginning of each trial (onset 0 sec), followed by the target stimulus (onset 4.2 sec) and the motor response (onset 8.4 sec). To account for the

temporal delay of the BOLD signal the decoding time bins were shifted by 2 volumes (i.e. time bin 1 is the earliest that could reflect cue related activity). The visual cortex showed strong activations with the presentation of the target stimulus (bin 3) while the motor cortex was strongly activated during the response period (bin 4-5). Interestingly, the information time course in motor areas already peaked in time bin 3, corresponding to the time of motor preparation and not execution. This finding supports a partial dissociation of the time course of information and overall response amplitude. Regions encoding the task rule did not show a strong increase in BOLD signal across the trial. This emphasizes the notion that information about rules was encoded in fine-grained spatial activation patterns rather than in the mean BOLD signal in single voxels. This comparison of time courses (information vs. BOLD signal) demonstrates the additional benefit of using multivariate analyses to pinpoint different roles of the same brain regions in cognitive processes over time.

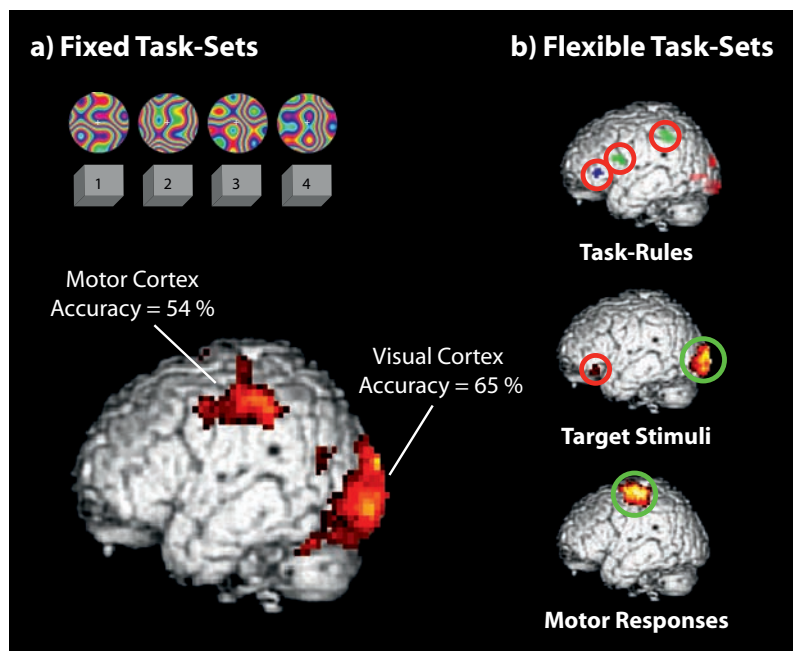


Figure 3-6: Decoding fixed task-sets in the control experiment. a) Paradigm of fixed task-sets experiment: One of four different stimuli was presented, each associated with a fixed response button (operated by middle and index finger of each hand). Decoding the four different target stimuli and their correlated motor preparation revealed accuracies above chance (25%) in bilateral visual cortex (peak at MNI [21 -90 0] with 65% decoding accuracy; SE 4.16; $p < .000001$ uncorrected) and bilateral motor cortex (peak at MNI [33 -12 48] with 54% decoding accuracy; SE 1.34; $p < .05$ FWE corrected). For better visualisation all decoding accuracies are displayed with $p < .001$ uncorrected. b) Flexible task-sets (taken from task-switching experiment): Other regions such as aVLPFC, pVLPFC and IPS, which encoded stimuli or task-rules in the main experiment (red circles) using flexible task-sets, were not found to predict the fixed task-sets in the control experiment. Only visual and motor areas were the same in the task-switching experiment and the control experiment (green circles).

3.4 Discussion

Owing to multivariate pattern classification, this study provides new insights into the information flow through the brain during the preparation of rule-guided behaviour. It demonstrated a shift of task-set information encoded in the brain over time. First, with the presentation of one of two visually distinct task cues, their identity could be decoded from visual cortex. After cue presentation, the information about task-sets shifted to different brain regions. The left IPS showed an early, transient peak in encoding of task-set information, while a slower build-up of information encoding could be observed in left posterior VLPFC, approximately coinciding with the presentation of the target stimulus. Finally, approximately during the execution of the motor response, task-sets could be decoded from a distinct area in left anterior VLPFC. Interestingly, it was also possible to decode the identity of visual target stimuli from anterior VLPFC. However, this was only possible when participants needed to flexibly switch between different task rules, and not when the task rule was fixed as in the control experiment.

3.4.1 Prefrontal cortex and rule-guided behaviour

Several studies on task-switching demonstrated the involvement of a parietal and lateral prefrontal network in cognitive control, including the areas identified in the present study (Asari, Konishi, Jimura & Miyashita, 2005; Chiu & Yantis, 2009; Brass & von Cramon, 2002, 2004a, 2004b; Brass et al., 2003; Braver, et al., 2003; Bunge et al., 2002, Bunge, Kahn, Wallis, Miller & Wagner, 2003; Dove et al., 2000; Rowe et al., 2008; Sakai & Passingham, 2003, 2006; Sohn et al., 2000). Activation related to cognitive control in PFC has been reported bilaterally but more often in the left hemisphere, matching the present results (e.g. Badre & Wagner, 2007; Brass & von Cramon, 2002, 2004b; Dove et al., 2000). Lateral prefrontal cortex is thought to be a generic region for executive control, guiding behaviour from top-down (Miller & Cohen, 2001). The specific roles of prefrontal sub-regions, however, are highly debated. The VLPFC in general has been linked to context processing (Badre & D'Esposito, 2009). In particular, it has been suggested that *anterior* VLPFC is responsible for the controlled *retrieval* of abstract rules (Badre and Wagner,

2005, 2007; Bunge et al., 2003; Bunge, 2004) or general maintenance of cognitive control (Braver et al., 2003). Posterior VLPFC, on the other hand, is thought to process contextual information of a rule (Brass & von Cramon, 2004b) or to represent and maintain the task-sets (Braver et al., 2003; Bunge, 2004; Sohn et al., 2000). In line with these interpretations, in the present study posterior VLPFC might have maintained the active task-set for the final response selection. This could be demonstrated by directly showing that abstract rule information was *encoded* in posterior VLPFC during the time period in which the decision for a response had to be made.

Other prefrontal regions that have been reported to be involved in cognitive control but were not found in the present study are DLPFC and FPC (Badre & D'Esposito, 2009; Bunge, 2004; Koechlin, Corrado, Pietrini & Grafman, 2000; Koechlin & Hyafil, 2007; Sakai & Passingham, 2003, 2006). It has been suggested that DLPFC is also involved in rule maintenance, while FPC is involved in moderating task-specific activations in other prefrontal regions (Sakai and Passingham, 2003, 2006). The latter studies, however, required participants to switch between different task *domains*, such as spatial versus verbal tasks, or semantic versus visual processing, and might therefore not be directly comparable to the present study. Others also found pre-SMA and SMA to be involved in task switching and cognitive control (Brass & von Cramon, 2002; Braver et al., 2003; Dosenbach et al., 2006; Rushworth et al., 2002). These regions have been discussed as core regions for motor preparation and sensory-motor associations (Picard & Strick, 2001), for linking cognition and action (Nachev, Kennard & Hussain, 2008) as well as for the initiation of voluntary action (Cunnington, Windischberger, Deecke & Moser, 2002; Cunnington et al., 2005; Haggard, 2008). However, abstract task-set information was not found to be encoded in the pre-SMA or the SMA in the present study, instead it encoded the motor responses. In agreement with this, it has recently been shown that the pre-SMA only encoded the *timing* of freely paced decisions rather than the *content* of such (Soon et al., 2008). It has also been argued that pre-motor cortex performs the execution of rules on a *non-abstract level* (Badre & D'Esposito, 2009), which is in line with the present findings.

3.4.2 Parietal cortex and rule-guided behaviour

Compared to PFC, the role of IPS in cognitive control is less well established. Classical models on posterior parietal cortex (PPC) functions do not focus on cognitive control but on the dorsal visual pathway, processing spatial information (Ungerleider & Mishkin, 1982), providing “vision for action” (Milner & Goodale, 1993) or being related to different kinds of attention (Rizzolatti & Matelli, 2003). For cognitive control, the PPC is discussed as playing a subordinate role and either being related to response preparation (Bunge et al., 2002; Bunge, 2004) or to attention (Forstmann, Brass, Koch & von Cramon, 2006; Rushworth et al., 2001). It has been suggested that the IPS provides stimulus-response associations rather than abstract rules (Brass & von Cramon, 2002, 2004a, 2004b). Several recent findings, however, question this view as the present study does.

It has been shown that a region in left parietal cortex, overlapping the IPS found in the present study, is directly involved in switching between task rules (Crone et al., 2006). The authors suggested that parietal cortex might be involved when there is a need to control *sets* of stimulus-response mappings (Crone et al., 2006; see also Braver et al., 2003). Furthermore, it has been difficult to identify a clear dissociation between prefrontal and parietal regions in establishing task-sets using classical univariate fMRI analyses. The IPS showed an activation profile very similar to that of prefrontal regions during task preparation (Brass and von Cramon, 2002, 2004a, 2004b; Crone et al., 2006; Rowe et al., 2008). Rowe and colleagues (2008) found parietal as well as lateral prefrontal areas to be involved in switching between rules and switching between responses made under single rules without any hierarchical order between the two areas. Instead, these authors suggested that both prefrontal and parietal cortex guide behaviour together. Bilateral posterior parietal cortex, predominantly the left IPS, has also been found to be involved in cognitive set-shifting using the Wisconsin Card Sorting Test (WCST), even when spatial task components were eliminated (Asari et al., 2005). Recently, it has been suggested that a domain-independent mechanism for reconfiguration of task-sets and cognitive control is located in PPC (Chiu & Yantis, 2009). Current models of PPC functions have been revised based on recent meta-analyses (Singh-Curry & Husain, 2009). These authors conclude that classical models focussing on visual and attentional functions of PPC (predominantly the inferior parietal lobe, eventually extending to the IPS in humans; see Orban et al., 2006) fail to account for the variety of tasks in which parietal regions have been demonstrated to be

engaged. According to their model, the PPC functions as the centre of a control network in which different information competes, and sensory information, motor information as well as task-goals and reward-related information are integrated. Its primary purpose is to keep the cognitive system in a goal-focussed and task-engaged state (Singh-Curry & Husain, 2009). In order to fulfil these functions, which were classically attributed to the PFC alone, PPC must encode abstract rule information, as demonstrated by the present study.

3.4.3 The temporal sequence of information encoding

Extending previous findings, the present study was able to reveal the time at which information about task-sets emerged in different regions of parietal and prefrontal cortex. A temporal sequence of task-set information encoding could be demonstrated, starting with an early, transient encoding in IPS *before* target presentation and then followed by a more sustained encoding of task-sets in posterior VLPFC. The finding that abstract rules could be read out from the IPS first suggests that the IPS might be able to provide abstract input to prefrontal cortical regions rather than only being engaged in response processing (Brass et al. 2005b; Bunge 2004). Evidence from human EEG and fMRI studies has so far yielded conflicting results about the relative timing of parietal and prefrontal cortex in establishing task-sets (e.g. Brass et al., 2005b; Buschman & Miller, 2007; Forstmann, Ridderinkhof, Kaiser & Bledowski, 2007). In contrast to the present results, a previous EEG study found that activity in PFC might arise earlier than in IPS, suggesting that the IPS would be subordinate to PFC in cognitive control (Brass et al., 2005b). On the other hand, there is evidence from a study using event-related potentials in a free selection task that parietal activation arises earlier than frontal activation (Forstmann et al., 2007). This questions the notion of a strict hierarchy from frontal to parietal areas. The inconsistent findings regarding the temporal order of signals in prefrontal and parietal cortex might simply reflect differences in the tasks studied. Moreover, none of these previous studies investigated the *encoding* of task-related information, hence, could not distinguish between unspecific activation (for example, related to preparatory arousal) and specific encoding of task-sets in these regions.

The present finding of early task-set information in parietal cortex is supported by single cell electrophysiology studies in monkeys that demonstrated rule sensitive neurons in parietal cortex (Freedman & Assad, 2006; Gail & Andersen, 2006; Gottlieb, 2007;

Oristaglio et al., 2006; Stoet & Snyder, 2004, 2007; Toth & Assad, 2002). Neurons in medial and lateral IPS could be found that only responded to abstract rules in task switching, showing their highest selectivity in a delay after cue presentation and *before* target stimulus presentation (Stoet & Snyder, 2004, 2007), comparable to the present study. The authors argued that these neurons directly encode the abstract upcoming task (“cognitive set”). Likewise, Gail and Andersen (2006) implemented a task switching paradigm to perform multivariate decoding on a population of neurons in the monkey IPS. They found that the abstract rule could already be predicted *before* the current target stimulus was presented, suggesting a direct involvement of IPS in cue-to-goal transformations at an abstract level. Notably, the temporal flow of information in the fronto-parietal *network* has only rarely been addressed in electrophysiological studies. Recordings were typically made only from one brain region, either lateral PFC (Wallis et al., 2001) *or* the IPS (Gail & Andersen, 2006; Stoet and Snyder, 2004, 2007). fMRI studies in humans might also fail in this respect if conventional voxel-wise analyses of activation are used. In contrast, substantial contributions of different areas in cognitive control over time could instead be encoded in spatial patterns of activation, which do not necessarily coincide with supra-threshold activation (Haynes & Rees, 2006; Norman et al., 2006). This interpretation is supported by the finding that only multivariate pattern classification, rather than univariate analyses, revealed the time courses of areas involved in cognitive control in the present study. Importantly, early encoding of task-set information in the IPS does not rule out the possibility that parietal cortex is later engaged in processes related to attention and response preparation as suggested by others (Brass & von Cramon, 2002; Brass et al., 2005b; Bunge et al., 2002; Bunge, 2004; Crone et al., 2006; Rushworth et al., 2001).

The activity levels and temporal dynamics in the fronto-parietal network might also depend on how well-learned the rules are. It has been shown that monkey PFC neurons are more strongly activated for new rules compared to familiar ones (Asaad et al., 1998). Using fMRI in humans, the right (but not left) VLPFC was found to be more activated during the processing of newly learned rules compared to well-known ones (Donohue, Wendelken, Crone & Bunge, 2005). Similarly, Wallis and Miller (2003) found that the contribution of PFC was weaker and occurred later compared to pre-motor cortex using a paradigm in which rules were well practised. These findings suggest that PFC is less involved with well-learned rules. In the present study, the rules were very well learned and participants hardly made any mistakes. This might explain why, despite being highly reliable, the overall decoding accuracy in VLPFC was lower than in other informative regions. The results

might also be a reflection of differences in the anatomical-functional micro-architecture of cortical columns. The structure to function relationship in lateral PFC might involve dynamic, adaptive re-coding as suggested previously (Duncan, 2001; Kusunoki et al., 2009). Thus, the exact locus of information encoding in PFC might vary more strongly across participants compared to other regions.

One functional interpretation of the early task-set encoding in parietal cortex might be that PPC acts as a control hub directly engaged in reconfiguring task goals (Singh-Curry & Husain, 2009). In order to keep attention directed to the task goal or to actively reconfigure mappings, a brain region necessarily has to encode highly abstract rule information at a very early stage of task preparation. This is exactly what was found in the present study. Rubinstein and colleagues (2001) suggested a formal model of task switching in which two separable components of cognitive control can be identified, namely *goal shifting* and *rule selection*. Goal shifting is thought to provide the information about the current task for other components of the cognitive system. This can take place in a potential delay between the indication of the upcoming task and the target stimulus. Subsequently, rule selection is externally triggered by the presentation of the target stimulus and is thought to load the current rule into working memory. This stepwise processing of rule information is thought to prevent interference that might occur if more rules were active in working memory (Rubinstein et al., 2001). With respect to this framework and considering the PPC model of Singh-Curry and Husain (2009), it could be speculated that the IPS is involved in goal shifting. Subsequently, task-set information from the IPS could serve as input for further processing in VLPFC. This region might then perform the rule selection and utilise the current rule to select a response once the target stimulus is successfully identified (Brass & von Cramon, 2004a, 2004b; Bunge, 2004). This interpretation is also supported by the finding that all motor regions encoded the upcoming motor response simultaneous to the peak of task-set encoding in posterior VLPFC and not in IPS.

Finally, anterior VLPFC encoded the task-set as well as the identity of the target stimulus around the time of the motor response. The motor response could not be decoded from VLPFC, matching recent findings from single-cell recordings in monkeys for fixed and flexible task rules (Kusunoki et al., 2009). The coincidence of visual information and rule information in VLPFC suggests that it plays an important role in making the final decision to act. Anterior VLPFC has been suggested to be involved in top-down controlled retrieval of plans or relevant semantic knowledge from memory (Badre & Wagner, 2007). Hence, this region might initiate a final check of the validity of the planned motor response

with respect to the target stimulus and the rule. Recently, anterior *medial* PFC has been demonstrated to encode decisions during the preparation of *voluntary action* (Haynes et al., 2007; Soon et al. 2008). One might speculate that in a similar manner, during the preparation of *exogenously cued action*, the anterior *lateral* PFC is involved in retrieving the stimulus and the rule from memory when it comes to the final decision to act. Another explanation for the present result would be an involvement of anterior VLPFC in a post-execution assessment of the appropriateness of the motor response with respect to the rule.

3.4.4 Methodological considerations

An important limitation of fMRI compared to electrophysiology is that its poor temporal resolution requires specific experimental paradigms to reveal onset differences between cognitive processes. One option to optimise the deconvolution of task components would have been the usage of a conventional GLM fitting the HRF instead of using FIR modelling. This, however, would also have excluded the possibility of time-resolved decoding. Interestingly, it has been shown in monkeys that while populations of prefrontal neurons maintained their selectivity for task rules during different phases of a task-switching experiment, the underlying patterns were only reliable *within* single phases and were orthogonal across task phases (Sigala et al., 2008). In other words, the use of pattern classification should greatly benefit from time-resolved modelling of separate task-phases as achieved by FIR, because patterns (for each task component) cannot be expected to stay the same across different stages of task preparation.

To reveal onset differences between cognitive processes in the present study, the pacing of events was deliberately slowed by introducing long delays (4.2 s) between the onsets of cues and targets as well as between targets and responses. Although fMRI is not perfectly suited to resolving sub-second processes, it has previously been shown for longer time scales that the latency of the haemodynamic response appropriately reflects the information flow in different regions in cortical networks (Formisano & Goebel, 2003). Thus, latency differences in the order of a second or more are considered to reflect the true neural order (Formisano & Goebel, 2003; Huettel & McCarthy, 2001). Given that the present study used onset differences as large as 4.2 s, it can be assumed that any small differences in haemodynamic latency in different brain regions cannot disqualify the observed temporal

order of their involvement. Since fine-grained spatial activation patterns form the basis for decoding, the same conclusion holds for the temporal order of information encoded in these regions. Nevertheless, because the haemodynamic delay can only be estimated, it cannot be excluded that some activation patterns were already established in one time bin but were at that stage still too weak to be detected by the analyses. This raises the question of whether the later onset of information in prefrontal cortex might be due to poor sensitivity of the method to early signals. The interpretation of negative results must therefore be made with care. Although it is possible that pattern classification lacked the sensitivity to reveal any early signals in PFC, the method clearly does not lack sensitivity in general because it does reveal information in PFC at later stages of processing. This again speaks for the validity of the present findings.

3.4.5 Conclusion

The present results substantially extend previous findings by resolving the build-up of *specific* task-related information over time. This goes beyond the mostly indirect evidence of other human fMRI studies and shows that activation in these areas does not simply reflect unspecific processes related to task-learning and task-switching. Hence, it could be demonstrated that cued rules can be decoded from human parietal and prefrontal cortex. The data suggest a tentative model of information flow in establishing task-sets. IPS first provided updated representations of the active abstract rule. This information was then shifted to posterior VLPFC, which maintained the representation and biased the response selection as soon as the target stimulus was presented. From there on, the upcoming response was encoded in motor areas. Around the time of the final motor response, patterns in anterior VLPFC in turn encoded the task-set and the target stimulus, suggesting a role in making the final decision to act. Importantly, using multivariate decoding the substantial involvement of parietal cortex at an early stage of task preparation could be revealed, which might have been overlooked with conventional analyses. In summary, the present study demonstrated that a network of lateral prefrontal and posterior parietal cortex encodes abstract rule information when task-sets have to be flexibly chosen according to a priori established rules.

4 Experimental Section II: Perceptual decisions in task preparation

“Devine, si tu peux, et choisis, si tu l’oses.”

(Guess if you can, and choose if you dare.)

(Pierre Corneille, Léontine, Héraclius, 1646, act IV, scene IV)

4.1 Theoretical background

4.1.1 The formal elements of perceptual decision making

In every-day life, we have to flexibly adapt our behaviour to the objects we perceive in the environment around us. An appropriate reaction on stimuli firstly requires successful perception, thus perceptual decision making. For example, we would not sit on an object that we identified as a piano but rather on a chair. It is unclear, however, if there is *one* perceptual decision making system that operates under all conditions of perceptual ambiguity. Models of perceptual decision making typically assume a sequence of neural processing stages that begins with a sensory representation of the stimulus and then proceeds to compute the perceptual choice at some higher level (Britten, Shadlen, Newsome & Movshon, 1992; Britten et al., 1996; Gold & Shadlen, 2000; 2001; 2007; Heekeren et al., 2008; Leon & Shadlen, 1998; Mountcastle, Talbot, Sakata & Hyvärinen, 1969; Parker & Newsome, 1998; Shadlen, Britten, Newsome & Movshon, 1996; Smith & Ratcliff, 2009). This idea has been formalised using probabilistic models of statistical inference and likelihood functions (Gold & Shadlen, 2007; Jazayeri & Movshon, 2006; Smith & Ratcliff, 2004; 2009). One common model assumes that perceptual hypotheses (“I see a piano” or “I see a chair”) are

tested using the noisy representations in visual areas (evidence), considering a certain probability or likelihood given by the context, to calculate the decision. A decision variable (DV) represents the aggregation of evidence, prior hypotheses and values. This DV is shifted during the decision process towards a criterion (for an overview about how to express this model mathematically see Gold & Shadlen, 2001). The decision criterion is the rule used to determine how much evidence has to be collected before a decision can be made one way or another (Gold & Shadlen, 2007). In some cases, additional evidence processed in a short time-window after the criterion is reached might even be used to reverse the decision (Resulaj, Kiani, Wolpert & Shadlen, 2009). Models typically assume that incoming evidence will shift a single DV from a neutral mid-position to the one side or the other over time until it reaches the criterion for one decision (*symmetric random walk model* or Wiener diffusion model). Alternatively, separate processes for individual DVs for each decision could compete for reaching their respective criterion (*race model* or dual diffusion model; for an overview about models see Gold & Shadlen, 2003; 2007; Smith & Ratcliff 2004; 2009; for urgency-gating models see Cisek, Puskas & El-Murr, 2009; Ditterich, 2006).

The implicit assumption is that perceptual decision making under different levels of visibility always operates based on the *same* sensory mechanism. One of the simplest models is the classical signal detection theory (SDT) model of perceptual decision making (Green & Swets, 1966; Parker & Newsome, 1998; Swets, 1961). It assumes that even under the same presentation conditions, the internal representation will never be exactly the same, due to external (e.g. exact number of photons reaching the retina) and internal noise factors (e.g. non-zero level neural responses even without stimulation). The discriminability of two stimuli therefore depends on the overlap between the sensory distributions evoked by these different stimuli in the visual system. For clearly distinguishable stimuli, these distributions will be relatively distinct and not strongly overlapping, which facilitates accurate perceptual choices by placing the decision criterion in an optimal position along the sensory continuum. When stimuli are not discriminable, they are believed to have overlapping sensory distributions and the trial-by-trial noise fluctuations in the sensory signal rather than marked differences in sensory signal determine whether it falls to one or the other side of the criterion (Shadlen et al., 1996; Swets, 1961). The decision process will then produce errors and close-to-chance-level performance. If this model were true, it would mean that the sensory signals predictive for high visibility choices would also be predictive for low visibility choices. Before an alternative model will be considered, the next paragraph gives

an overview about neuroscientific research on perceptual decision making and its reference to this framework.

4.1.2 The neural basis of perceptual decision making

Studies using single-cell recording in monkeys used the concept of *choice probability* (CP) to determine the neurons' relationship to the monkeys' choices independent from the presented stimuli (e.g. Britten et al., 1996; Newsome, Britten & Movshon, 1989; Nienborg & Cumming, 2009; Shadlen et al., 1996; Uka, Tanabe, Watanabe & Fujita, 2005). The neuron's CP indicates the probability with which an independent observer would be able to correctly predict the monkey's perceptual choice from this neuron's firing rate on a given trial, knowing only the neuron's firing characteristics for the possible decisions. In the classical task, monkeys were presented with random dot motion for one or two seconds. The monkeys watched dots on a screen; the dots were replaced with every monitor refresh. The coherence in replacement of dots was used to systematically vary the magnitude of perceived movement to the left or to the right, usually including a neutral (non-coherence) condition (Britten et al., 1996). In early studies, recordings were made in extrastriate medial temporal area (MT) in the visual system. Weak but significant positive CPs were found for MT neurons even if stimuli did not contain any real motion, suggesting a link between the firing of neurons and the monkeys' perceptual choices independent from the real stimulus motion. Others extended these findings by identifying neurons for very fine perceptual decisions (Jazayeri & Movshon, 2006; 2007; Purushothaman & Bradley, 2005). Accordingly, Salzman, Murasugi, Britten & Newsome (1992) showed that micro-stimulation of neurons in monkey area MT favoured perceptual choices about random dot motion for the preferred direction of the stimulated neuron. Given the independence of neurons' firing from real stimulus motion in these studies, the findings can be interpreted as evidence that the *same* neural system computes perceptual choices for stimuli of all visibility levels.

However, the decision making system for motion is unlikely to be solely located in area MT. The choice probabilities found by Britten and colleagues (1996) were rather weak, suggesting that other regions contribute to a distributed decision making system that is more strongly involved in the decision process itself. Accordingly, others demonstrated that sensory areas are likely to be modulated by top-down processes during early processing

stages (Nienborg & Cumming, 2009). In line with these findings, it was assumed that further decision-related computations are carried out by frontal and parietal areas, as well as by motor structures, which might plan the upcoming actions (Gold & Shadlen, 2000; Kim & Shadlen, 1999; Shadlen & Newsome, 2001). Most importantly, the monkey lateral intraparietal area (LIP) was shown to be involved in the integration of collected evidence over time and in target selection for saccades (Huk & Shadlen, 2005; Roitman & Shadlen, 2002; Shadlen & Newsome, 2001). For other modalities, similar distributed neural systems have been proposed, such as primary and secondary somatosensory cortex and premotor cortex for vibro-tactile frequency (flutter) discrimination (Deco, Pérez-Sanagustín, de Lafuente & Romo, 2007; de Lafuente & Romo, 2005; Hernández, Zainos & Romo, 2000; 2002; Romo et al., 2002; Romo, Hernández & Zainos, 2004; Salinas, Hernández, Zainos & Romo, 2000; for a review see Romo & Salinas, 2003) and MT and the inferior temporal cortex (IT) for fine grained depth perception (e.g. Uka & DeAngelis, 2003; 2004; Uka et al., 2005). A clear-cut classification of sensory areas, motor areas and abstract decision areas *within* these perceptual decision making networks, however, seems to be an oversimplification; integration towards a decision might require direct ongoing modulation of sensory input (Kim & Shadlen, 1999; Romo & Salinas, 2003; Shadlen & Newsome, 2001).

4.1.3 Networks for perceptual decision making

An often neglected possibility is that *different neural processes* might determine the response when participants are confident about their perception of a stimulus, as compared to conditions when they are not. For instance, for visual stimuli presented under high versus low visibility conditions, different networks in prefrontal and parietal cortex became more active (Dehaene et al., 1998; Heekeren et al., 2004). It has also been shown that different processing streams can be differentially influenced by stimulus visibility (Fang & He, 2005). Under high visibility conditions there is also higher subjective confidence in the accuracy of a decision (Kiani & Shadlen, 2009; Kunitomo, Miller & Pashler, 2001) and less conflict between different response alternatives (Botvinick, Braver, Barch, Carter & Cohen, 2001), whereas low visibility and uncertainty might involve different brain areas, require additional processes and introduce decision latencies (Philiastides & Sajda, 2006; Philiastides, Ratcliff & Sajda, 2006). There is also evidence for more recurrent processing (Philiastides & Sajda, 2006; Supèr, Spekreijse & Lamme et al., 2001; VanRullen & Koch, 2003)

and higher functional integration within visual cortex for high compared to low visibility (Haynes, Driver & Rees, 2005). The picture is further complicated by the fact that not all sensory information in the brain is available to guide a response (Haynes & Rees, 2005a) and that behaviour can be influenced by stimuli that fail to reach awareness (Vorberg, Mattler, Heinecke, Schmidt & Schwarzbach, 2003). Taken together, this raises the question of whether the brain might compute decisions in a fundamentally different way under conditions of high and low visibility. Although several previous electrophysiological studies have investigated the influence of visibility on decision-making signals, these typically focussed on only few selected brain regions and are thus not suitable for identifying any differential roles between large-scale networks under high versus low visibility (e.g. Britten et al., 1996; Gold & Shadlen, 2000; Kim & Shadlen, 1999; Nienborg & Cumming, 2009; Purushothaman & Bradley, 2005; Shadlen & Newsome, 2001).

Studies using fMRI in humans could demonstrate the involvement of larger networks in perceptual decision making (Heekeren et al., 2004; Ho, Brown & Serences, 2009; Philiastides & Sajda, 2007; Shulman, Ollinger, Linenweber, Petersen & Corbetta, 2001; Tosoni, Galati, Romani & Corbetta, 2008; for a review see Heekeren et al., 2008). Some studies used multivariate pattern classification on fMRI data to decode participants' perceptual choices (Li et al., 2007; Pessoa and Padmala, 2007; Serences & Boynton, 2007). In one study, participants were presented with random dot motion, which was dissociated from real motion or attended motion. When the motion was unambiguous (clearly visible), nearly all visual areas from V1 to the human MT complex (hMT+) encoded the perceived motion (Serences & Boynton, 2007). Perceived motion for ambiguous stimuli that did not contain real motion could only be decoded from hMT+ (to V3a), confirming findings of electrophysiology studies (Britten et al., 1996). It is important to note, however, that motion stimuli require the integration of information across longer time periods (Leon & Shadlen, 1998; Mazurek, Roitman, Ditterich & Shadlen, 2003) and might easily lead to illusionary perception. This renders them problematic for the investigation of decision making under different levels of stimulus perceivability. Other fMRI studies in humans also have focused on selected brain regions from the visual system and prefrontal cortex when comparing high and low visibility conditions (Serences & Boynton, 2007; Williams et al., 2007). Thus, it remains unclear whether visibility might change the influence of specific brain regions on participants' decisions.

4.1.4 Perceptual decisions about objects

In order to reduce the complexity of perceptual decisions processes, the component of long temporal integration can be eliminated by using static images, e.g. of different object categories (Philiastides et al., 2006; Philiastides & Sajda, 2006; 2007). Object classification still involves several processing stages (e.g. Humphreys, Price & Riddoch, 1999; Riddoch & Humphreys, 2003) but the integration of sensory information in static images towards a perceptual decision will take place within ~ 500 ms, depending on the difficulty (Philiastides & Sajda, 2006). Using static object images also has the advantage that the respective sensory regions for objects, the ventral and lateral-occipital complex (LOC) in occipital and inferior temporal cortex, are well investigated (e.g. Haushofer et al., 2008b; Haxby et al., 2001; Ishai, Ungerleider, Martin, Schouten & Haxby, 1999; Malach et al., 1995; Op de Beeck et al., 2008b; for reviews see Grill-Spector, 2003; Grill-Spector & Malach, 2004). There are other ventral high-level visual areas adjacent to the LOC that were discussed to be specialised for subclasses of object stimuli, e.g. the fusiform face area (FFA), the parahippocampal place area (PPA) and an area for body parts (Aguirre et al., 1998; Epstein & Kanwisher, 1998; Epstein et al., 1999; Downing et al., 2001; Kanwisher et al., 1997; Spiridon & Kanwisher, 2002; Reddy & Kanwisher, 2006; Spiridon, Fischl & Kanwisher, 2006). There is much debate, however, whether the whole object sensitive occipito-temporal cortex is organised in distinct modules for different object categories, clustered by *shared processes* (Gauthier, Tarr, Anderson, Skudlarski & Gore, 1999; Gauthier, Skudlarski, Gore & Anderson, 2000; Tarr & Gauthier, 2000), *topographical eccentricity* (Malach, Levy & Hasson, 2002) or *object form*, encoded in a distributed and overlapping fashion (Cox & Savoy, 2003; Eger et al., 2008; Haxby, Hoffman & Gobbini, 2000; Haxby et al., 2001; O'Toole et al., 2005). It was also suggested that the LOC represents objects mainly by the arrangement of their parts (Haushofer, Baker, Livingstone & Kanwisher, 2008a; Hayworth & Biederman, 2006; Op de Beeck et al., 2008b). A recent review comes to the conclusion that a non-additive combination of distinct weak maps of all these different properties could result in very pronounced selectivity profiles in some sub-regions of the ventral object-vision pathway. The combinations of these maps might then appear as specialised modules in functional imaging studies (Op de Beeck, Haushofer & Kanwisher, 2008a).

It has further been proposed that in *object recognition*, inferior temporal cortex might act as a 'connection centre', or multiplexer, for context associations that activate

candidate object representations from which high-level areas could choose in a second processing step (Bar, 2004; Bar et al., 2006). Some fMRI studies also linked the LOC cortex directly to perceptual decisions about objects (Grill-Spector et al., 2000; Li et al., 2007; Philiastides & Sajda, 2007; Walther et al., 2009; Weber et al., 2009; Williams et al., 2007). In one study, the correlation between the classifier's and participants' classification errors of natural scenes, which contained typical objects, was high in LOC, pointing to a direct behavioural relevance of information encoded in LOC (Walther et al., 2009). In another study, Grill-Spector and colleagues (2000) presented their participants with masked images from different categories, using different image durations to achieve different levels of visibility. Activation in the LOC showed the highest correlation with recognition performance and was therefore assumed to be necessary for successful object perception (Grill-Spector et al., 2000). Williams and colleagues (2007) were able to predict the category of successfully identified masked object images from the LOC but not when identification failed. These findings point towards a *differential* involvement of the LOC in perceptual decisions for different visibility conditions. No study, however, directly addressed the *encoding of the decision outcome* in comparison to the presented objects. Therefore, these studies again did not reveal whether a sensory region such as the LOC is always involved in perceptual decisions or if different neural systems might be engaged for high and low visibility.

4.1.5 The goal of the present study

The present study sought to directly investigate whether visibility has an effect on which brain regions determine a participant's perceptual choice. In the present study, participants were asked to make perceptually based decisions about the category of weakly or strongly masked visually presented objects from three different categories. Static images were used to avoid extended temporal integration. To investigate the relationship between stimulus visibility and choices, the true object categories were dissociated from participants' choices in the low visibility condition. Additionally, the motor responses were dissociated from the category decisions. This avoided confounds of choice and response direction that were inherent in most classical perceptual decision making studies in monkeys (e.g. Britten et al., 1996; Huk & Shadlen, 2005; Shadlen & Newsome, 2001). Multivariate pattern classification (Haynes & Rees, 2006; Norman et al., 2006) was used to predict the *outcome* of

perceptual decisions from brain activity under high and low visibility. This approach allowed neural encoding and choice performance to be directly linked and is conceptually close to choice probabilities in single-cell studies of perceptual decision making (e.g. Britten et al., 1996). First, a searchlight approach (Haynes et al., 2007; Kriegeskorte et al., 2006) was applied to identify those brain regions that were maximally informative of choices under different visibility conditions. In a second step, the results were further elaborated using independent localizer scans to investigate the encoding in object-sensitive voxels in LOC and their contribution to perceptual decision making in greater detail.

4.2 Methods

4.2.1 Masking experiment

Participants

Nineteen participants (10 female, mean age 26.1; range 23-31) took part in the study. All were right-handed, had normal or corrected to normal visual acuity and gave written informed consent to the fMRI procedure. Data from two participants were excluded from all analyses due to excessive head movement. Data from another three participants were excluded because their categorisation performance did not differ between the two visibility conditions, suggesting a lack of attention to the task or the use of illegitimate strategies (e.g. one participant reported having used rapidly blinking in order to render the mask ineffective). The data from the remaining fourteen participants (8 female, mean age 25.8; range 23-31) were used for the analyses.

Stimuli

The stimuli were pictures of objects from two categories – pianos and chairs – as well as phase-randomised noise images serving as a non-object category. The categories were selected as they demonstrated comparable and optimal visibility characteristics under different masking conditions in behavioural pre-tests (Appendix B, Behavioural pre-tests I). Object stimuli were created from freely available pictures from the Internet or custom-made photographs using Adobe® Photoshop Version 7.0. Following Grill-Spector and colleagues (2000), the images showed objects in different natural backgrounds. All images were transformed into grey-scaled versions with a size of 400 x 400 pixels. The overall contrast (object to background) was kept approximately constant for all stimuli by matching the spectra and eliminating extreme contrast values. Thirty stimuli were created for each object category from which 24 were selected for the experiment (see Appendix B, Behavioural pre-tests III and Figure 4-1a for examples of the stimuli). The noise category also contained 24 stimuli consisting of scrambled phase textures with the same power spectra as the

pictures from both object categories (see Appendix B, Figures B-3 to B-5 for all stimuli). To create the noise images, a two-dimensional Fourier transformation was performed on all stimuli from each object category (pianos, chairs). As in previous studies their phase maps were scrambled by adding a random value of $\pm 1.75 \cdot \Pi$ to each phase angle (Malach et al., 1995). The resulting phase maps were then transformed back to images and were contrast-normalised. Two different scrambled masks were used as pre- and post-mask. Masks were constructed by dividing every target image into 10 x 10 squares, each 40 x 40 pixels. Subsequently, 100 of these squares were randomly chosen from the pool of all scrambled target images, such that they did not contain any clearly identifiable parts of objects. For both masks, 100 squares were reorganised into new 400 x 400 pixel scrambled images with 10 squares in every row and every column (Figure 4-1b).

Experimental procedure

During the fMRI experiment, stimuli were presented in the middle of the screen. Participants were required to fixate a white cross that was superimposed at the centre. Each trial consisted of a brief “standing wave” (Macknik & Livingstone, 1998) of the same mask-target-mask sequence repeated four times to increase signal-to-noise level, extending the total duration of visual stimulation to 2668 ms (Figure 4-1c; see also Appendix B, Behavioural pre-tests II). In both high and low visibility conditions the target images were preceded by a 167 ms pre-mask and followed by a post-mask of variable duration. The target image was shown for 16.7 ms (1 frame) under low visibility followed by a 483.3 ms post-mask. Under high visibility it was shown for 66.7 ms (4 frames) followed by a 433.3 ms post-mask. Thus, the visibility conditions were defined by the ratio of target and post-mask duration (Grill-Spector et al., 2000). Each trial was followed by a response mapping screen that displayed three letters as response options (“K” Klavier = piano, “S” Stuhl = chair, “M” Muster = noise). Participants had to indicate the category of the object they *believed they had seen* by pressing one of three buttons using the index-, middle- and ring finger of the right hand. The position of letters assigned to the buttons was pseudo-randomised in order to avoid confounds of categorical choices with motor responses. Each button served as the correct response button equally often, separately for each visibility condition. Note that using this technique, motor responses could not be prepared until the response mapping screen appeared after stimulus presentation. Participants were encouraged to take their best

guess if they could not identify or were unsure about an object's category, rather than not responding at all. The response mapping screen was presented for a pseudo-randomised duration, completing the total trial duration of 5000 ms, 7000 ms or 9000 ms. In each run all 24 images from each of the three categories (pianos, chairs and noise) were shown in random order, under either of the two visibility conditions. Thus, participants were presented with 72 trials per run which was repeated for 6 runs in randomised order resulting in 432 trials for the experiment. Altogether, each image was presented three times under each visibility condition in the entire experiment. Presentation during the experiment was controlled by the Cogent toolbox (<http://www.vislab.ucl.ac.uk/Cogent>) for MATLAB 7.0 (The MathWorks, Inc.). The stimuli were presented via a projector (resolution 1024x768 pixel, 60Hz) that projected from the head-end of the scanner onto a screen. Participants viewed the projection via a mirror fixed onto the head coil. Using stimuli 10 x 10 cm in size and a distance of approximately 70 cm between the participants' head and the screen, the estimated visual angle was $\alpha \sim 7.15^\circ$ for all stimuli and masks.

Functional imaging

Functional MRI volumes of the whole brain were acquired using a Siemens TRIO 3T scanner (Erlangen, Germany) with a standard head coil (42 axial slices, TR = 2800 ms, echo time TE = 30 ms, resolution 3 x 3 x 2 mm³ with 1 mm gap). Within each of the six runs, 182 volumes were acquired for each participant using gradient-echo EPI. The first two volumes of every session were discarded by default to allow for magnetic saturation effects.

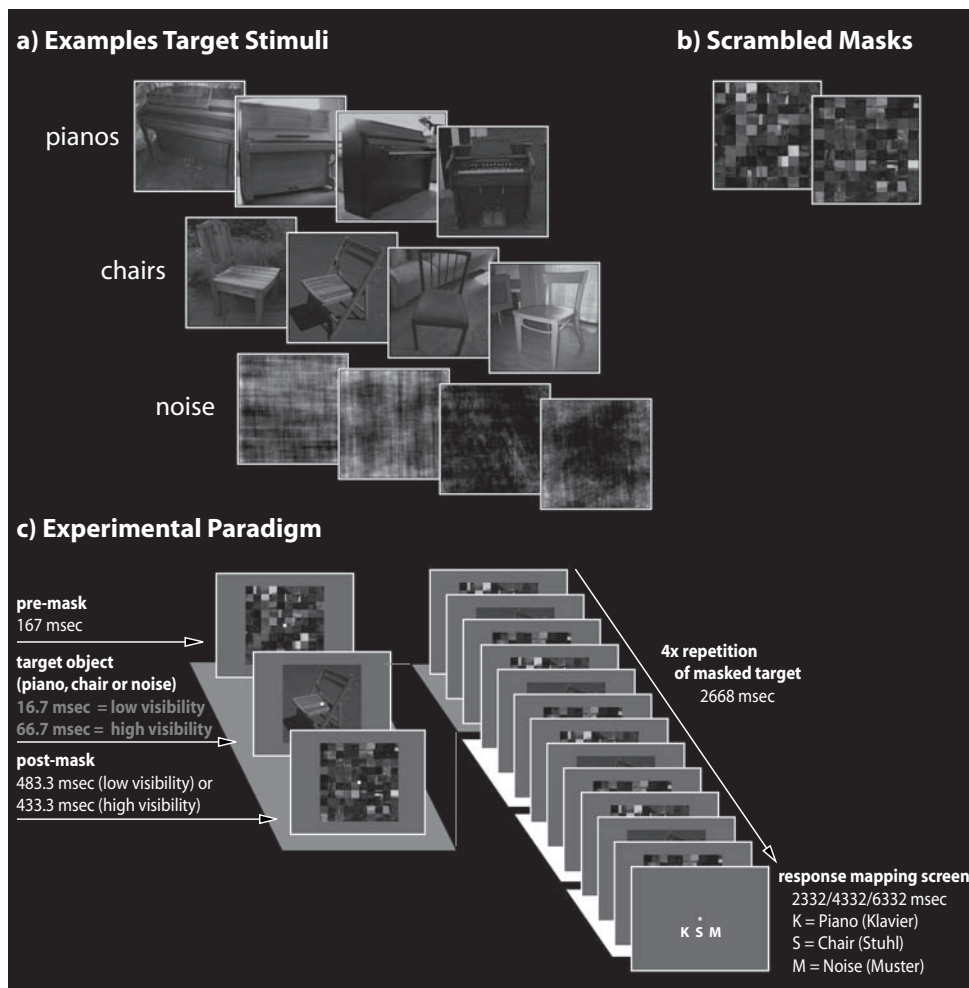


Figure 4-1: Experimental paradigm and examples of stimuli and masks. a) 24 different stimuli from three categories (pianos, chairs, noise) were used in the experiment (four example images of each category are displayed here). b) Masks were created by randomly scrambling the target images and constructing two new mask images of 10 by 10 tiles. c) Target images were preceded by a 167 ms pre-mask and followed by a post-mask of variable duration. The target image was shown for 16.7 ms (1 frame) under low visibility followed by a 483.3 ms post-mask. Under high visibility it was shown for 66.7 ms (4 frames) followed by a 433.3 ms post-mask (left side). The flashed sequence of masked target images was presented four times in direct succession (right side), resulting in 2668 ms total presentation duration. Each trial was followed by a response mapping screen with a pseudo-randomised order of response buttons associated with the categories in order to avoid confounds of categorical choices with motor responses. The response mapping screen was presented for a pseudo-randomised duration, completing the total trial duration of 5000 ms, 7000 ms or 9000 ms.

Analysis of behavioural data

Behavioural data was analysed by calculating hit rates as well as a visibility index for each category and each visibility condition, similar to the d-prime (d') for the two-condition case in SDT (Green & Swets, 1966). For this, the percentage of hits and false alarms for each condition of each category was transformed into z-values.

$$d' = z_{(\text{hits})} - z_{(\text{false alarms})}$$

High values corresponded to high visibility and low values indicated low visibility. The maximum value that could be achieved by 100% hits and no false alarms was $d' = 7.44$. On the other hand, a minimum in visibility cannot be easily specified because possible systematic confusion between categories would lead to a high number of false alarms, which, in turn, would produce negative values for d' . True invisibility, on the other hand, would produce a hit and false alarms rate at chance level, which would be indicated by a $d' = 0$. Note that this only holds true if there are no missing responses, so in general d' should be interpreted as an estimation rather than a direct measurement of visibility.

Multivariate pattern classification

Multivariate pattern classification (Haynes & Rees, 2006; Norman et al., 2006) was performed on the fMRI data in order to identify brain regions encoding information about *participants' choices* and *object categories*. For both procedures, data were subjected to two individual pattern classification analyses, one for high and one for low visibility. For the analyses of participants' category choices, one participant had to be excluded as she never selected the "noise" category in the low-visibility condition during one run.

As a precursor to the decoding analyses, a general linear model (GLM) was estimated on an individual subject level based on motion corrected, non-normalised and unsmoothed data using SPM2 (<http://www.fil.ion.ucl.ac.uk/spm/>). This GLM yielded parameter estimates for *each voxel* in *each participant* for all *three category decisions* (pianos, chairs, noise), *for each visibility condition* and *for each run*. These parameter estimates were used for the following decoding analyses. Both high and low-visibility decoding analyses were carried out quasi identically, again using a variant of the "search-

light approach” (Haynes et al., 2007; Kriegeskorte et al., 2006). This method searched for information-encoding regions in the brain in an unbiased fashion, as it tested spatial clusters around every voxel separately for their information content. For a given position in the brain, a searchlight cluster was defined as a sphere of N voxels ($c_{1...N}$), with radius of 4 voxels, constructed around the central voxel v_i . For these voxels the GLM parameter estimates were extracted for each stimulus category and run, separately for each participant and for each visibility condition. A multivariate pattern classification algorithm was then applied to decode the stimulus category from these pattern vectors. Specifically, a linear support vector pattern classifier (Müller et al., 2001) with a fixed regularisation parameter $C = 1$ was trained on patterns vectors from all runs but one (“training data set”) using the LIBSVM toolbox. Based on the training data the classifier estimated an optimal linear decision hyperplane to separate the pattern vectors associated with choices for the different object categories. Subsequently, this hyperplane was used to predict the categories from the pattern vectors of an independent “test data set” taken from the remaining run. A six-fold cross-validation was performed by repeating the classification process independently with the pattern vectors of each run as the “test data set” and the remaining five runs as the “training data set”. The use of independent training and test data sets avoided problems of circular inference (Kriegeskorte et al., 2009). The mean probability of correct classification was assigned to the central voxel v_i of the searchlight cluster and indexed the local decoding accuracy. This process was performed for every voxel in the brain by moving the searchlight cluster through all spatial positions. Using this technique, two 3-dimensional maps of local decoding accuracy values were obtained (one for “high visibility” and one “low visibility”) for each participant. These maps were then normalised to standard stereotaxic space (MNI EPI template) and smoothed with a Gaussian kernel of 6 mm FWHM to account for anatomical variability. Random-effects group-level analyses were then computed on a voxel-by-voxel basis to identify searchlight positions where local patterns had significant decoding accuracy across all participants (Haynes et al., 2007).

A very similar pattern classification analysis was conducted in order to predict the category of the *presented* objects. The classification procedure was the same as described above, only that parameter estimates were obtained from a different GLM that estimated regressors for the different *presented categories*, again individually for each participant and each run.

The second multivariate pattern classification control analysis was conducted to decode the motor responses (button presses) instead of the assigned categories, again separately for both visibility conditions. The decoding analysis therefore aimed to confirm that the dissociation between response buttons and categories was successful. Again, the classification procedure was the same as described above, based on GLM parameter estimates for the *motor responses*, individually for each participant and each run.

Finally, the procedure was repeated for decoding *errors* versus *correct responses*. This could only be achieved in the low visibility condition because the number of errors in the high visibility condition was too small because stimuli and responses were highly correlated. Taken together, these analyses ensured that all combinations of a typical confusion matrix in perceptual decision making (see Appendix B, Figure B-6) were systematically investigated.

4.2.2 LOC localizer experiment

An additional experiment was performed, nominated a priori on the lateral occipital complex (LOC) region. It should be assessed whether the results might change in case the analyses were restricted to specialised object processing voxels, using a conventional individual-specific LOC localiser (Malach et al., 1995). Independent functional localizer scans were used to predefine the LOC region in nine participants from the original sample (4 female, mean age 26.3, range 23-31), conducted some weeks later. Regions of interest (ROIs) were constructed for the original data from the main experiment and parameter estimates of these voxels were re-sampled as a different approach to feature selection for multivariate pattern classification.

Stimuli, experimental procedure and data analysis

Stimuli were presented in alternating blocks showing either object stimuli or phase-randomised noise images (Figure 4-2). Blocks lasted 10000 ms and consisted of 10 object or noise images shown in random order, each for 900 ms with a 100 ms gap. Each block was followed by a fixation-only period of 10000 ms, prolonged by a pseudo-randomised delay

of 3000 ms, 5000 ms or 7000 ms. In one third of all object and noise blocks, one image was left out and instead the preceding image was directly repeated. The participants' task was to detect these image repetitions (1-back task). This procedure ensured that participants' attention to the target images was sustained during the whole scanning session. Scanning parameters were the same as for the masking experiment, but using 163 volumes acquired across 3 functional runs. Each run consisted of 18 blocks: 9 with object images and 9 with noise images.

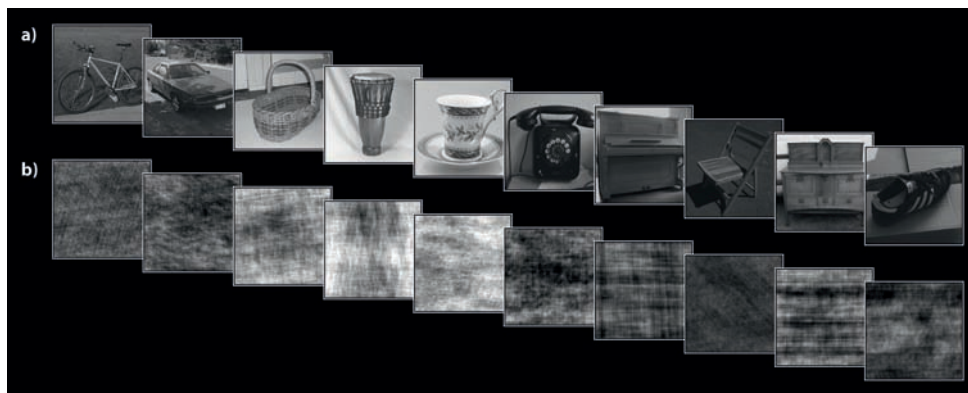


Figure 4-2: Stimuli used in the localizer experiment. a) Object stimuli b) Phase-randomised noise stimuli. In separate object blocks and noise blocks, all images were each shown for 900 ms in a random order, separated by 100 ms blanks so that each image emerged discretely. In one third of all blocks, one image was left out and instead a different image was repeated in direct succession, which was to be detected. This procedure ensured that participants' attention to the target images was constantly sustained.

Data from this independent object localizer were motion-corrected and then co-registered to the non-normalised images of the individual participants from the masking experiment using SPM2. Two regions were identified that were activated stronger by object images than noise images ($p < .05$ FWE corrected) in each participant. These were the left and right LOC (left LOC: average 83 voxels, range 45-159; right LOC: average 79 voxels, range 40-124). The decoding analyses (*participants' categorical choices* and *presented object categories*) performed on the data from the masking experiment were then repeated as described above but restricted to these conventional LOC voxels. Detailed analyses were performed to decode each possible pair-wise classification combination (pianos-chairs, pianos-noise, chairs-noise).

4.3 Results

Behavioural data was analysed by means of d' values. Following an ANOVA for repeated measurements that included all combinations of categories and visibility conditions as factors [$F(5,65) = 30.08$; $\text{Eta}^2 = 0.59$; $p < .001$], the differences in d' values between “high visibility” and “low invisibility” reached statistical significance for all object categories in post-hoc Scheffé-tests correcting for multiple comparisons (pianos: high visibility $d' = 2.90$, $\text{SE} = 0.30$; low visibility $d' = 0.69$, $\text{SE} = 0.10$; $p < .001$; chairs: high visibility $d' = 3.13$; $\text{SE} = 0.30$; low visibility: $d' = 1.02$, $\text{SE} = 0.11$; $p < .001$; noise: high visibility $d' = 2.16$, $\text{SE} = 0.23$; low visibility $d' = 1.11$, $\text{SE} = 0.19$; $p < .01$) (Figure 4-3a). The same analysis conducted using hit rates instead of d' values showed highly comparable results (Appendix B, Figure B-7). It can therefore be assumed that distinct visibility characteristics for each condition were successfully achieved, as was the comparability of the object categories within each visibility condition. None of the participants was better at categorising poorly visible objects compared to highly visible objects from the same category (Figure 4-3b). A control experiment with an independent sample demonstrated that high visibility was strongly associated with high decision confidence while low visibility was associated with low confidence in the decision (Appendix B, Figure B-13).

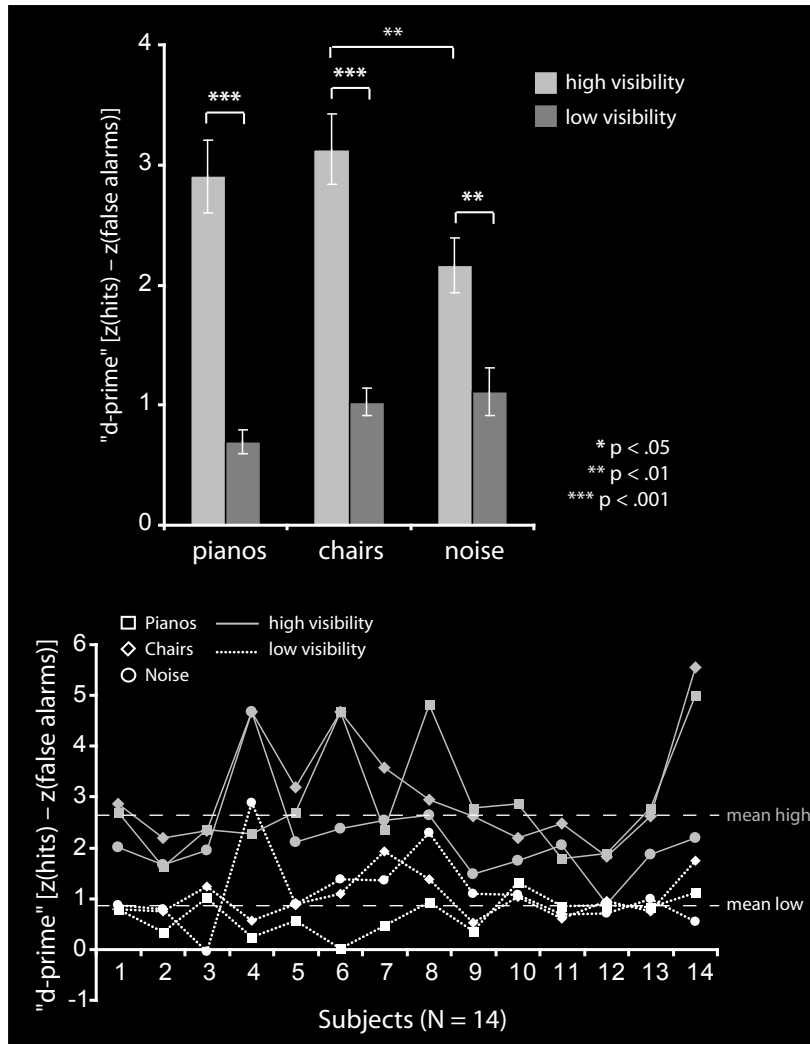


Figure 4-3: Visibility of stimuli. a) Mean visibility indices for both visibility conditions and all three categories. Hits and false alarms [%] were transformed into z-values for each category in each visibility condition. The difference between $z(\text{hits}) - z(\text{false alarms})$ then served as the index of visibility d' . Displayed are mean d' values and standard errors for the final sample (N=14). Greater d' values for the “high visibility” compared to the “low visibility” condition were achieved for all categories. b) Individual visibility indices for all participants for both visibility conditions and all three categories. The graph displays individual d' values for each participant. Solid lines: categories in “high visibility” condition (mean 2.73; SE 0.27); dotted lines: categories in “low visibility” condition (mean 0.94; SE 0.14). Nearly all participants showed a clear separation between visibility conditions (exceptions were participant 4 and 12 in one case). The same category was never perceived better under “low visibility” compared to “high visibility” by any of the participants.

In the next step, searchlight decoding was used to identify brain regions that encoded *participants' classification choices* (chance level 33%). For the “high visibility” condition, the chosen category could be decoded from bilateral LOC (accuracies: left 45%; right 46%; Figure 4-4a, green regions). For the low “visibility condition” participants’ choices could not be decoded from LOC. Instead, their choices could be decoded from the precuneus, which showed a significant decoding accuracy of 45%. This region extended to lateral inferior parietal cortex (Figure 4-4a, red regions). The activity patterns underlying the category choices were highly unique for each participant (Figure 4-4b; additional illustrations for all participants can be found in Appendix B, Figures B-11 and B-12).

To further elaborate upon the relationship between the pattern classifier’s classification performance and participants’ categorisation performance, correlation analyses of confusion matrices (see Walther et al., 2009, for a similar approach) were performed (see Appendix B; Figure B-10a for details). This analysis confirmed that regions that displayed a high similarity between the *pattern classifier’s performance* and the *participants’ performance* (therefore reflecting similarities in classification errors) were located in bilateral visual cortex and LOC when objects were highly visible. For poorly visible objects the correlation was highest in a cluster in the precuneus, located close to the cluster found for searchlight decoding (Appendix B; Figure B-10b).

Pattern classification analyses were also conducted to decode the *true categories of presented objects*. These could only be decoded from bilateral visual cortex (accuracies: left 47%; right 44%), left LOC (50% accuracy) and right LOC (48% accuracy) when objects were highly visible (Figure 4-5a). In the low visibility condition the categories could not be decoded from any region. The LOC regions for high visibility strongly overlapped with those found for decoding of category choices. This was expected since participants’ choices for highly visible objects corresponded well with the category of presented objects. Early visual cortex, including V1, also encoded the true category of highly visible objects in the present study. The wide spread of decoding accuracy in early visual regions might mostly be due to stronger differences in low-level features between noise images and real object images compared to differences between the real object images (pianos and chairs). In line with this assumption, control analyses using pair-wise decoding of only two categories at a time confirmed that only the LOC allowed real object images to be decoded. Patterns in wide spread regions of the ventral visual stream, however, allowed decoding of noise versus real object categories (Figure 4-5b; for details see Appendix B, Table B-3).

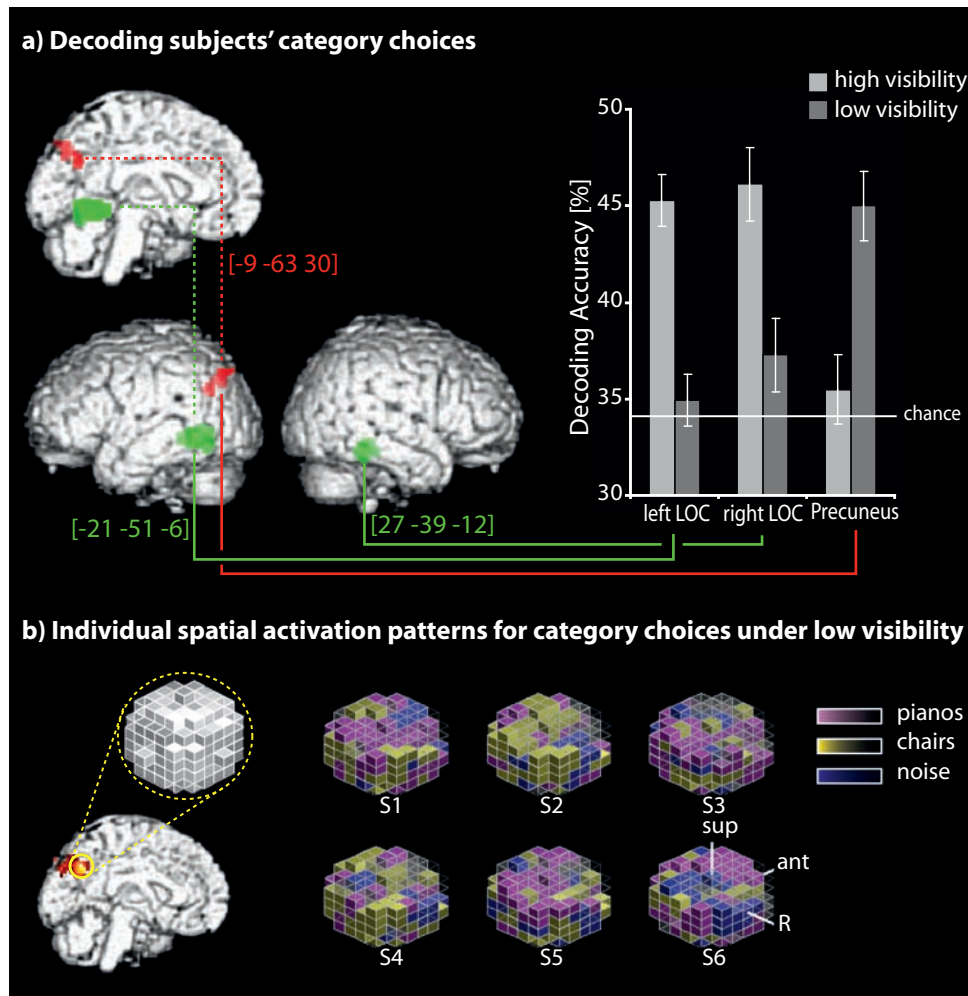


Figure 4-4: Whole-brain searchlight decoding results for *categorical choices*. a) Participants' perceptual choices for all categories (pianos, chairs, noise) were used for pattern classification. Displayed in green are regions where a searchlight allowed decoding above chance level for high visibility conditions (chance level 33% for three categories). This was possible from left (45% accuracy) and right LOC (46% accuracy) with a threshold of $p < .05$ (FWE corrected). Displayed in red are regions where choices could be decoded for the low visibility condition. The choices could be decoded from precuneus (accuracy 45%; $p < .000001$). For better visualisation, informative regions are displayed with a threshold of $p < .0001$. Note the double dissociation between regions encoding of choices under high and low visibility. b) Examples of individual spatial activation patterns for participants' choices in the low visibility condition for six participants. Displayed are the searchlight clusters (radius = 4 voxels) with the best average decoding accuracy from the main analysis (MNI: -9; -63; 30). The colour code indicates which category the voxels preferentially respond to (magenta for pianos, yellow for chairs, blue for noise; sup = superior, ant = anterior, R = right). Grey voxels did not show a category-preference or were not located in grey matter. Colours are scaled for better visualisation. Informative patterns were unique for each participant. Bars indicate standard errors. Coordinates are MNI coordinates.

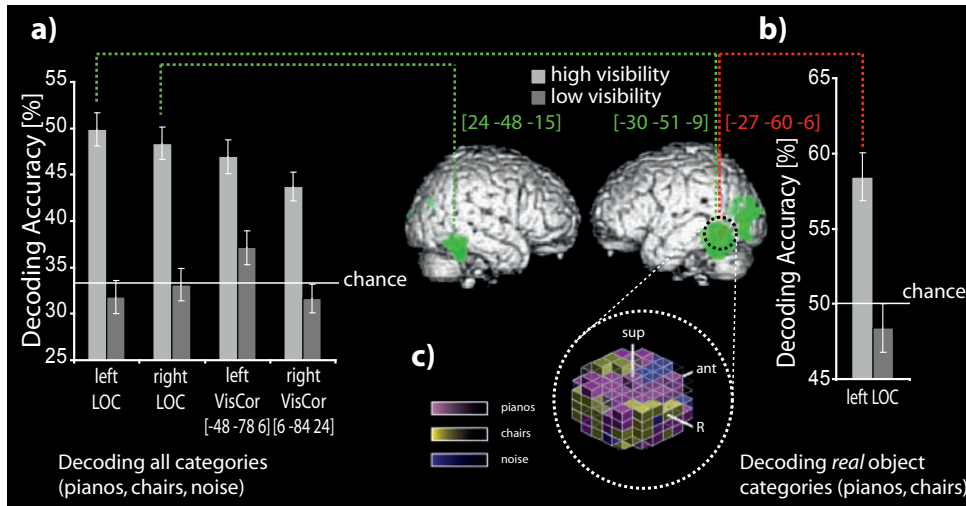


Figure 4-5: Whole-brain searchlight decoding for object categories. a) Decoding of sensory stimulus categories (pianos, chairs, noise). Displayed in green are clusters from which decoding above chance level (33%) was achieved; peak accuracies and standard errors are given for each cluster. In the high visibility condition the category of sensory stimuli could be decoded from bilateral LOC ($p < .05$ FWE corrected; displayed with $p < .00001$ uncorrected for better visualisation) extending into visual cortex. No region could be found to encode the category of sensory stimuli in the low visibility condition. b) When the decoding analysis was restricted to real object categories (pianos, chairs; plotted here in red; see Appendix B) the informative brain regions were even more closely confined to LOC (chance level 50% for two categories) and again only for the high visibility condition ($p < .00001$ uncorrected). c) Example of a spherical voxel cluster ("searchlight") of one participant for the left LOC. Voxels responding preferentially to one category are colour-coded (magenta for pianos, yellow for chairs, blue for noise; sup = superior, ant = anterior, R = right). Grey transparent voxels did not show category-preference or were not located in grey matter. Colours are scaled for better visualisation. As expected, informative patterns were different for each participant. All coordinates are MNI coordinates.

To ensure that decoding of participants' choices was not based on motor decisions, additionally pattern classification analyses were performed to decode the motor response (button press) instead of the associated category. As expected from the randomised response mappings, no choice-related region showed significant decoding accuracy. The only region which encoded the motor responses (peak accuracy 45%) was left motor cortex (since responses were given with the right hand only) under both visibility conditions (Figure 4-6).

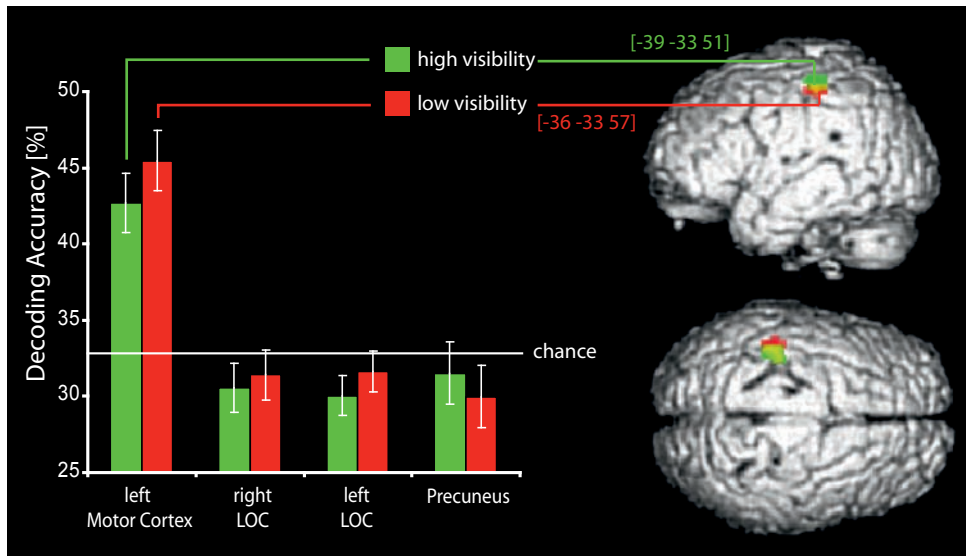


Figure 4-6: Whole-brain searchlight decoding results for motor responses. Response buttons were randomised with respect to the assigned categories to avoid confounds of motor responses with categorical choices. The successful dissociation between encoding of choices and motor responses is highlighted by the results of a searchlight decoding of participants' button presses (index/middle/ring finger of the right hand; chance level 33%). Only the left motor cortex allowed decoding of motor responses for both, the high visibility condition (accuracy: 43%; $p < .0001$ uncorrected, 10 voxels threshold) and the low visibility condition (accuracy: 45% $p < .000001$ uncorrected, 10 voxels threshold) with comparable accuracy (all displayed with $p < .0001$ uncorrected for better visualisation). The regions which were found to encode participants' abstract category *choices* (left and right LOC for high visibility and precuneus for low visibility) did not allow decoding of motor responses above chance. Bars are standard errors. Coordinates are MNI coordinates.

The next step was to compare the average BOLD signal in those voxels that were the centre of the most successful decoding searchlight clusters from the LOC and the precuneus. Both left and right LOC showed stronger activation (as assessed by BOLD contrast estimates) for real objects compared to noise images when they were highly visible. On contrast, the activation was lower for poorly visible objects from all three categories, confirming earlier findings (Grill-Spector et al., 2000). Despite containing choice-related information, the precuneus did not show significant overall differences in average BOLD signal between objects or conditions. (Figure 4-7).

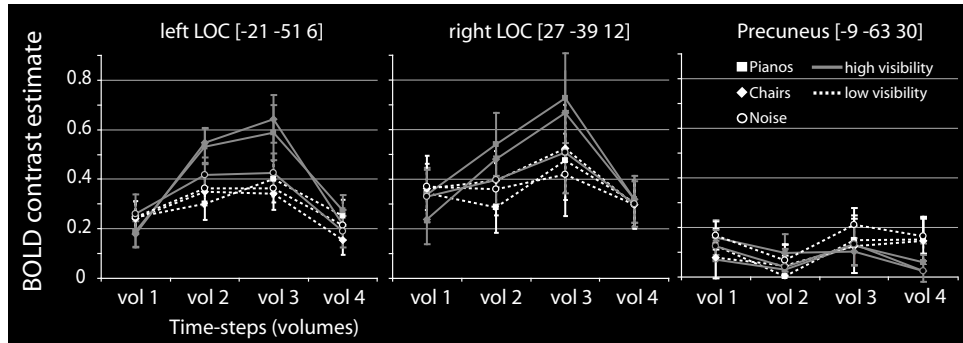


Figure 4-7: BOLD contrast estimation profiles for all conditions. Event-related differences in BOLD signal between all six conditions (pianos, chairs, noise in the high visibility (solid lines) and the low visibility (dashed lines) condition) averaged across all 14 participants with standard errors (SE). Time-steps are given in volumes (1 volume = 1 TR = 2800 ms), the first volume is time-locked to the onset of stimulus presentation. Time-series (here based on data smoothed with a Gaussian kernel of 8 mm FWHM) were extracted for the voxels with the highest decoding accuracies for participants' category choices in the high visibility condition (left and right LOC) and the low visibility condition (precuneus). Both LOC regions showed a stronger BOLD response for highly visible objects (pianos, chairs) compared to highly visible noise as well as compared to all categories in the low visibility condition. The precuneus was not strongly activated for images from any visibility condition.

Comparing the BOLD response (as estimated by a GLM) between visibility conditions throughout the whole brain, only the rostral cingulate zone (RCZ) in medial prefrontal cortex showed stronger activation for objects from all categories presented under low visibility compared to high visibility. Stronger activation for objects presented under high visibility compared to low visibility could only be found in LOC and anterior temporal cortex (see Appendix B, Univariate Analysis, Figure B-8). No brain region showed a difference between correct trials and error trials under low visibility, even with a very liberal threshold ($p > .01$ uncorrected), or when the analyses were restricted to ROIs in LOC and precuneus (Appendix B, Figure B-9). Correct versus error trials could also not be decoded from any brain region, even when a lower threshold than for the all other analyses was used ($p > .0001$, uncorrected). These results confirm that it is very unlikely that substantial residual visibility for strongly masked images was present; this should have caused some brain region to reflect incorrect categorisation.

Finally, the decoding analyses were repeated for participants' category choices and object categories but restricted to voxels from conventional LOC ROIs as obtained by the independent localizer scans. The results confirmed that the objects' true categories as well as category choices could only be decoded from LOC when objects were highly visible,

closely matching the results from the searchlight decoding analyses (see Table 4-1; for details see Appendix B, Decoding results from conventional LOC ROIs). The lower decoding accuracies for the classification of pianos and chairs might be due to greater similarities between low-level features of pianos and chairs compared to both object categories. This is in line with the lower statistical values obtained from searchlight decoding of pianos and chairs when leaving out the noise trials (see Appendix B, Two-class pattern classification for presented objects). Given the smaller sample size of the localizer experiment, this analysis also had lower statistical power. Additionally, it has been demonstrated before that decoding accuracy drops when only object-selective voxels (defined by univariate analyses) are selected for pattern classification, possibly due to redundancy effects within the selected voxel set and the elimination of information from weakly activated voxels, which otherwise importantly contribute to the patterns (Cox & Savoy, 2003).

Table 4-1: Decoding from LOC ROIs

Decoding	Anatomical area	L/R	Accuracy		T	p <
			M	SE		
True Category						
High Visibility	LOC	L	69	4.2	7.76	.001
		R	60	4.1	3.42	.01
Low Visibility	LOC	L	50	4.0	0.01	.99
		R	54	3.7	1.55	.16
Chosen Category						
High Visibility	LOC	L	65	4.7	5.94	.001
		R	61	6.0	2.10	.07
Low Visibility	LOC	L	52	5.2	0.61	.65
		R	56	3.6	1.68	.13

Note: LOC = Lateral-occipital complex (based on individual localizer scans); the results are averaged across all pair-wise decoding analyses, separately for high and low visibility (chance level 50%); L = left hemisphere, R = right hemisphere; M = mean; SE = standard error.

In summary, it was only for highly visible objects that their identity as well as participants' (correct) category choices could be decoded from the LOC region in inferior temporal cortex. In the case of poor visibility, achieved by strong masking, the object category could not be decoded from brain activity - but the participants' categorical choices could nonetheless be predicted from the precuneus.

4.4 Discussion

This study shows that stimuli and choices during perceptual decision making can be reliably decoded from patterns of brain signals. Importantly, it could be demonstrated that different mechanisms were responsible for perceptual choices about highly and poorly visible objects. The maximally predictive brain regions depended strongly on stimulus visibility; a clear double dissociation could be revealed between the visibility conditions. The LOC encoded the true category of highly visible objects as well as decision outcomes of such. It did not, however, encode the object categories or decision outcomes when objects were rendered nearly invisible. The drop in information presumably explains the drop in performance for conditions of low visibility. No regions could be found to encode the true category of presented objects when these were strongly masked. The precuneus, however, encoded *choices about poorly visible objects* but not highly visible ones. The BOLD signal strength in this region did not reflect preferences for any object category. Decisions were only encoded in finer grained stable activation patterns. These results demonstrate that under certain circumstances, it is possible that perceptual decisions are *not* based on the same sensory signal dimension.

4.4.1 Perceptual decisions about highly visible objects

The present results demonstrate that categorical choices were encoded in object sensitive cortex in the LOC for highly visible objects. This is in line with studies showing that the LOC can be regarded as a sensory processing area for objects and categories (e.g. Cox & Savoy, 2003; Grill-Spector, 2003; Haushofer et al., 2008a; Haxby et al., 2001; Op de Beeck et al., 2008a,b; Reddy & Kanwisher, 2006; Schwarzlose et al., 2008). This region extended to the medial temporal lobe (MTL), which is highly interconnected with the lateral LOC. It has been demonstrated using single-unit recordings in humans that the MTL contains highly selective neurons for a variety of categories, pointing to a sparse, invariant and explicit representation in the MTL related to memory and conscious perception (Mormann et al., 2008; Quiroga, Reddy, Kreiman, Koch & Fried, 2005; Quiroga, Reddy, Koch & Fried, 2007). Given the greater extent in the LOC the whole area will be simply referred to as LOC in the following.

When a stimulus can be easily be seen, it is difficult to dissociate sensory from choice-related signals because participants will choose correctly on most trials and the two measures are therefore highly correlated. Accordingly, regions encoding presented objects and chosen objects strongly overlapped in the high visibility condition. Highly visible objects could additionally be decoded from early visual cortex. Previously, it has been demonstrated that activation patterns from V1 to V3 are sufficient to encode literally every possible object image (Kay et al., 2008). Others also found that category information could be decoded from earlier areas of the visual system (Eger et al., 2008; Walther et al., 2009; Williams et al., 2008). In the present study, this can be easily explained by differences in low-level visual features, especially between the noise category and the real object categories (pianos and chairs). It seems most likely that areas from early visual cortex up the ventral visual pathway were recruited to process information about the objects, resulting in representations about object categories in the LOC.

The LOC might therefore also be a high-level integration area for *perceptual decisions about object categories*. Others suggested the existence of a “domain general accumulator” for sensory evidence located in the insula (Ho et al., 2009). Similarly, a “general mechanism” for perceptual decision making has been suggested, reflected by a stronger difference signal for objects under high visibility compared to low visibility in DLPFC (Heekeren et al., 2004). These studies, however, did not explicitly look at choice prediction in various brain regions and used different presentation and visibility conditions compared to the present study. There is little evidence for the insula or lateral prefrontal cortex acting as a general decision maker in the present study (only the correlation analysis of confusion matrices showed significant results in DLPFC; see Appendix B) (but see Rorie & Newsome, 2005; Tosoni et al., 2008 for a general discussion of this model).

The present findings extended earlier studies substantially by demonstrating a case in which the encoding of choice outcomes in LOC was *only* possible under *high visibility*. By varying the visibility of the visual input, others found that the LOC only encoded object information on trials in which participants successfully classified the objects while earlier visual regions also contained object information if objects were not correctly identified (Williams et al., 2007); but these authors did not directly investigate the *encoding of decision outcomes*. In a different fMRI study, activation in LOC was found to strongly correlate with the visibility of objects and therefore with participants’ ability to correctly identify the objects’ category (Grill-Spector et al., 2000). Using very short target durations of 20 ms, no object-related activity was reported for the LOC after subtracting activity related to the

mask. The direct encoding of choice-related signals was again not addressed. Another study that used pattern classification to compute choice probabilities based on EEG signals also did not specifically address any potential changes in choice-predictive topographies depending on visibility (Philiastides & Sajda, 2006). The present study, however, could demonstrate that for very short target durations, the LOC did not encode object categories or decision outcomes. This clearly speaks against the existence of one self-contained neural system for perceptual object decisions of all levels of ambiguity.

4.4.2 Perceptual decisions about poorly visible objects

Several previous near-threshold experiments have studied which brain signals are predictive of perceptual choices, either by comparing hits/false alarms with misses/correct rejections or by explicitly computing choice probabilities (Bradley, Chang & Andersen, 1998; Britten et al., 1996; Gold & Shadlen, 2000; Huk & Shadlen, 2005; Nienborg & Cumming, 2009; Kim & Shadlen, 1999; Ress & Heeger, 2003; Serences & Boynton, 2007; Shadlen et al., 1996; Shadlen & Newsome, 2001; Thompson & Schall, 1999; Williams et al., 2007). Importantly, however, the degree to which different networks might be involved in perceptual choices under high and low visibility has remained unclear. The most important finding in the present study is that the LOC did not encode the perceptual choices for poorly visible objects; they were instead encoded in the precuneus. This clear double-dissociation in decoding demonstrated that *different* neural systems were involved. This contradicts the assumptions of SDT (Green & Swets, 1966), which predicts that decisions about perceptual entities should be based on the same dimension of signal, even in the absence of useful sensory information. In temporal accumulation models, choices under low visibility are dominated by a random walk that is based on trial-by-trial fluctuations in the underlying sensory signals (Mazurek et al., 2003). The variability in decisions is then thought to be dominated by “noise” (Faisal, Selen & Wolpert, 2008). This noise could be related to changes and random processes in single neurons (e.g. electrochemical and biochemical fluctuations) and in the respective sensory network (Faisal, et al., 2008), driving the random walk towards one decision threshold or another. Accordingly, using choice probabilities, it has repeatedly been demonstrated that decisions about random dot motion could be predicted from the same neural populations in area MT, LIP and DLPFC, independent of the real stimulus motion and even in the *absence* of any real motion in the stimulus (Britten et

al., 1996; Huk & Shadlen, 2005; Kim & Shadlen, 1999; Shadlen & Newsome, 2001). These findings suggested one self-contained neural system for all levels of perceptual ambiguity (Gold & Shadlen, 2007), which is not supported by the present results.

The applicability of one single system for perceptual decisions is, however, subject to question. First, in the abovementioned studies, recordings in monkeys were usually made in one cortical region only (Gold & Shadlen, 2000; Huk & Shadlen, 2005; Kim & Shadlen, 1999; Shadlen & Newsome, 2001). The involvement of a different decision system for poor visibility therefore might have been easily overlooked. Second, recent models acknowledge that the information flow between sensory regions, parietal and prefrontal integration areas and motor regions appears to be continuous and reciprocal such that the decision is not the end-product of the computation of a central decision maker (Gold & Shadlen, 2007; Nienborg & Cumming, 2009; Tosoni et al., 2008). Hence, perceptual decision making appears to be more dynamic than initially thought, which might also permit interactions with a different network for poor visibility. Additionally, unlike the present study, signal detection models are designed for stimuli that require temporal accumulation of evidence over time (Cisek et al., 2009; Mazurek et al., 2003). Moving dot patterns in most studies were always visible for several seconds, which gave the participants the possibility to gather a *relevant percept* that could be interpreted, even illusory. The static object images used in the present study required little spatio-temporal integration (for a similar argument see Philiastides & Sajda, 2006) as object classification is usually a very fast and highly automated process (Bowers & Jones, 2008; Grill-Spector & Kanwisher, 2005; Mack, Gauthier, Sadr & Palmeri, 2008). It was demonstrated with the present study that under these conditions, decision making under high and low visibility can involve a switch between *different* cortical networks.

Several possible confounds have to be considered before a possible mechanism for perceptual decision making with insufficient visual input is discussed. First, the strongly masked objects might have been visible enough to be processed at a conceptual level, given that hit rates and d' values slightly exceeded chance level in some cases. It has been argued that only a d' value of zero is sufficient for true invisibility (Schmidt & Vorberg, 2006). Additionally, regions in parietal cortex, as part of the dorsal visual stream, can also process shape (Sereno & Maunsell, 1998) and object information (Konen & Kastner, 2008), thereby might reflect some residual visibility in the positive decoding results. Contradicting this argument, however, the masking procedure used in the present study was comparable to other studies in which objects were successfully rendered invisible (Grill-Spector et al.,

2000; Quiroga, Mukamel, Isham, Malach & Fried, 2008). It might have been even stronger due to the use of shorter target durations and stronger sandwich-masking (Appendix B, Figure B-2). Moreover, if the precuneus did indeed process the barely visible objects on a visual basis (and furthermore at a category level) it should have shown some information encoding for *poorly visible objects* or *categorisation errors*, which was not observed. All participants reported they often felt they were purely guessing as they did not see the objects. The subjective experience of *low decision confidence* might therefore activate a different network (see Appendix B, Figure B-13).

Another concern might be that participants gave up on the task and responded in a random manner. Hence, the double dissociation between encoding in LOC and precuneus would only be due to different kinds of decisions, namely object related decisions under high visibility and motor decisions under low visibility. This explanation, however, can be ruled out because motor responses could not be prepared beforehand and decisions and motor responses were dissociated due to the use of pseudo-randomised response mapping screens. Additional control analyses clearly demonstrated that motor responses could not be decoded from choice-related brain areas under any visibility condition. It can therefore be concluded that participants always made *category* decisions instead of random button presses.

Furthermore, it could also be argued that sensory regions must be the origin of the perceptual decision because it is possible to evoke perception in the absence of sensory information via micro-stimulation in sensory cortex (e.g. Britten & van Wezel, 1998; Murphey & Maunsell, 2007; Murphey, Maunsell, Beauchamp & Yoshor, 2009; Salzman, Britten & Newsome, 1990; Salzman et al., 1992). In successful direct cortical stimulation, however, the information is actually present because it is artificially *induced* in the sensory area and is again *sufficient* to make a perceptual decision. Hence, this argument does not contradict the present findings since decisions were made in the *absence* of such information in decision-related areas in the present study.

4.4.3 A mechanism for internal decisions

It is proposed here that in the low visibility condition, participants might have purely guessed when stimuli were presented close to the perceptual threshold. Without access to

substantial sensory evidence for either choice alternative, they faced a symmetry situation. Thus, trial-by-trial choices could be based on a dedicated *symmetry-breaking network* rather than on noise fluctuations in sensory signals. This network would come into play as soon as the participants believe the information to be insufficient. Previous studies have shown differences in large-scale cortical activity patterns under different visibility conditions (Dehaene et al., 1998; Fang & He, 2005; Heekeren et al., 2004; Philiastides et al., 2006; Philiastides & Sajda, 2006). These different networks are often believed to reflect uncertainty or effort but instead they might reflect neural mechanisms that are actively involved in symmetry-breaking. The present study is therefore compatible with the data obtained in these studies. The precuneus could be part of the neural substrate of the symmetry-breaking network and produce random *internal* decisions. Accordingly, recent fMRI studies using multivariate pattern classification found that the precuneus encoded decision outcomes for intended *free decisions*, even in the absence of a visual task (Soon et al., 2008; in prep.). Participants had to freely decide whether they wanted to press the left or right button at a time of their choosing (Soon et al., 2008) or decide upon a simple mathematical operation to be performed (Soon et al., in prep.). It was possible to predict the decision outcomes from activation patterns in the precuneus / PCC as well as from anterior medial prefrontal cortex (MPFC), even seconds before participants were aware of making a decision. It is worth noting that the precuneus cluster found by Soon and colleagues (2008) was located anterior to the present results but demonstrated substantial overlap when the same statistical thresholds were applied (see Appendix B, Figure B-16). Others found regions in inferior parietal lobe / precuneus involved in internally selected actions compared to externally cued ones (Jenkins et al., 2000) or focussing on one's own intentions (den Ouden, Frith, Frith & Blakemore, 2005). It also has been demonstrated that posterior parietal areas are involved in the formation of free decisions in monkeys (Cui & Andersen, 2007; Pesaran, Nelson & Andersen, 2008) and intentions in humans (Desmurget et al., 2009). The precise localization of effects within this region can be rather widespread depending on the exact nature of the task (e.g. den Ouden et al., 2005). These studies support the hypothesis that low visibility required participants to switch to a mechanism that breaks the symmetry, allowing a decision to be made when sensory input is unreliable, absent or irrelevant. The precuneus could act as a generator for internal random decision outcomes. It is well suited for this purpose, because it has one of the most complex columnar cortical organisations and has been discussed to be involved in the implementation of a wide range of higher-order cognitive functions (Cavanna & Trimble, 2006). The precuneus has reciprocal connections to several

regions in medial and lateral parietal cortex as well as prefrontal cortex including DLPFC, SMA and pre-SMA (Cavanna & Trimble, 2006), the latter of which were also discussed to be involved in the generation of self-initiated action (Haggard, 2008; Soon et al., 2008).

The present findings require a modified conceptual framework for perceptual decision making. The classical model assumes that sensory evidence for the given alternatives is collected, leading to the decision. The decision, in turn, triggers the (motor) response (e.g. Gold & Shadlen, 2007). Others, however, assumed that central (rather than peripheral) noise fluctuations between large-scale attractor networks are essential during near-threshold decision making (Deco & Romo, 2008). When no sufficient evidence can be collected, or sensory information is equal for each alternative, the response selection process might still be routed into one decision or the other, such that a random decision must be made (Romo & Salinas, 2003). Note that this model does not require the temporal accumulation of signals towards a threshold in sensory regions. Decisions with negligible input might rather be based on the current state of a different system, which, in turn, is determined by random noise factors. It could further be speculated that the switch from perceptual decisions to random decisions could be achieved by neurons signalling the absence of a detectable stimulus (e.g. Deco et al., 2007). It has been suggested that the positive detection of sub-threshold input might allow systems that rely on strong noise fluctuations to determine a response (Faisal et al., 2008). Such a scenario would be plausible when no extended temporal integration of information takes place and no substantial build-up of decision biases can be expected. Interestingly, this model is similar to a model proposed for voluntary decision making, which assumes that if sensory stimuli fail to generate sufficient information for a response, a free decision is required (Haggard, 2008).

4.4.4 Alternative mechanisms

There are, however, some alternative accounts for the precuneus involvement in the present decision task. The medial part of inferior parietal cortex has been reported to work in concert with its lateral counterparts in spatial and abstract mental imagery (Fletcher et al., 1995; Knauff, Fangmeier, Ruff & Johnson-Laird, 2003). Other work also identified the precuneus as being related to episodic memory retrieval (Addis, McIntosh, Moscovitch, Crawley & McAndrews, 2004; Fletcher, Shallice, Frith, Frackowiak & Dolan, 1996; Tulv-

ing et al., 1994). The involvement of the precuneus could alternatively be explained by the generation of mental images of objects as a substitute for visual information. It could be speculated that these images were then compared with templates from recent memory traces. This explanation, however, has several shortcomings. First, it contradicts self-reports of the participants who agreed in great consensus on simply having guessed if they did not see the objects. This makes it unlikely that any strong mental images were created. Second, the precuneus was more often found to be involved in *self-related episodic memory* (Lundstrom et al., 2003; Lundstrom, Ingvar & Petersson, 2005) or *self-related prospective memory* (Burgess, Quayle & Frith, 2001) rather than simple item recall. Third, visual imagery is believed to involve networks that are required for perception (Kosslyn, Ganis & Thompson, 2001; O'Craven & Kanwisher, 2000). If visual imagery were relevant, one would expect an involvement of LOC rather than a near-perfect double dissociation between visible and invisible conditions. Therefore, it seems more likely that the decisions themselves were internally generated, thus made with a high self-reference.

If symmetry-breaking leads to internal decisions, partially resembling free decisions, one would also expect prefrontal areas to be involved in this process; specifically, the medial and anterior PFC were often found to be involved in free decision making and internal action selection (Cunnington et al., 2002; 2005; Haggard, 2008; Haynes et al., 2007; Lau et al., 2004a; Lau, Rogers & Passingham, 2007; Soon et al., 2008; in prep.). However, while activation in medial prefrontal cortex at the border of RCZ and SMA / pre-SMA was stronger for the low visibility condition than for the high visibility condition in the present study, this area did not encode the decision outcomes. This is in line with results from Soon and colleagues (2008; in prep.) where the pre-SMA encoded the *timing*, but not the outcome of the decisions. In the present study, timing aspects could not be teased apart because of the temporal proximity of visual presentation and decisions, excluding the use of time-resolving FIR models. It is likely that the SMA / RCZ activation in the present study was simply related to decision uncertainty under low visibility (Nachev et al., 2008). In summary, it can be assumed that internal guessing decisions and intended free decision share some neural substrate but they might not be completely identical. This question is investigated in detail with the next experiment (Experimental Section III).

4.4.5 Conclusion

It was demonstrated that the brain might switch between two different decision networks depending on whether visual stimuli are easy or difficult to see. For clearly visible objects, decision-relevant information might be contained in specialised sensory regions in the LOC. In the absence of sufficient perceptual input, the decisions rely on neural populations in the precuneus, distinct from those for perceptual decisions made with certainty. This interpretation is in line with findings showing that the precuneus was also a key region for free decisions. The precuneus therefore might act as the brain's symmetry breaker and generator of internal random decisions when participants believe themselves to be guessing.

5 Experimental Section III: Internal task preparation

“So I only mean I am free in the sense that it’s not you who is determining my actions; it’s not blind force or destiny; it’s my upbringing, and my genes, and my predilections, and my desires. All of this, plus some random component depending on fluctuation and noise in my brain, comes together in making a decision one way or the other way.”

*(Christof Koch in S. Blackmore (2006),
Conversations on Consciousness, p. 131)*

5.1 Theoretical background

5.1.1 The general framework of free decisions

Imagine a card game in which one wins by picking the higher card of two alternatives. Most of us would agree that, if we had to choose between two *covered* cards without seeing their faces, it would be a completely free decision. Would it be different, however, if the faces were shown so briefly that we could not identify the cards and had to guess? The previous study (Experimental Section II) demonstrated a case in which visual information about objects was so well masked such that it became insufficient for perceptual decision making. Contrary to assumptions derived from SDT (Green & Swets, 1966), it could be shown that those decision outcomes were not encoded in the sensory system (LOC) anymore but in a different region, namely the precuneus. The precuneus was assumed to be part of a symmetry-breaking network, which comes into play in the absence of sufficient visual input. Given that other studies linked the precuneus to free decisions (e.g. Soon et al., 2008) it was

assumed that the participants *guessed* instead of using residual visual input. Guessing might therefore be an *internal* decision, similar to free decisions. However, it has also been demonstrated before that free decisions were additionally encoded in medial and anterior PFC (Haynes et al., 2007; Soon et al., 2008), which was not observed in the previous study. The goal of the present study was to investigate these differences and similarities and to directly compare guesses with free decisions.

When talking about free decisions, it is necessary to give a clear definition of “free” decisions in a neuroscientific context, as this is subject of a great, sometimes highly passionate debate in natural science and philosophy (e.g. van der Heiden & Schneider, 2007; Blackmore, 2006; Chiang, 2005; Dennett, 2003; Haynes, submitted; Heisenberg, 2009; Libet, 2004; Prinz, 2006a; Searle, 2004). The everyday meaning of “free” in this context is often related to a decision that “I” (meaning the experienced “self”) make as opposed to (automatic) reactions of the brain and body. This dualistic dichotomy is scientifically unlikely and is more related to properties of our language as well as to how we *experience* our selves as socially acting agents (Dennett, 1991; Prinz, 2003; 2006a). In the context of neuroscientific research, free decisions were defined as decisions that are maximally non-stimulus driven, the opposite to simple reflexes (Haggard, 2008). It has been pointed out that in experimental situations that lack any stimulation, participants also have no *motivation* to choose one alternative over the other, which is a rather unnatural case (Haggard, 2009). On the other hand, only a situation in which no choice alternative is more valuable than the other satisfies the requirements for the definition of a free decision (Haggard, 2008). The assumption of a determined world seems to contradict “truly” free decisions (Haynes, submitted), but even determinism does not rule out the possibility of free decisions (depending on the definition of a free decision) (Dennett, 2003). However, as already acknowledged by Thorndike (1911; cited by Haggard, 2008), voluntary behavior might be organised in loops rather than in single chains of events, as they do not need an initial cause; a free decision might be best described as “intelligent interaction with the animal’s current and historical context”. Nevertheless, certain brain regions are involved in computing these interactions, thereby acting as generators of free decisions.

5.1.2 The neural basis of free decisions

The starting point of neuroscientific research on free decisions was the discovery of the *Bereitschaftspotential* (BP), or readiness-potential (Kornhuber & Deecke, 1965). It describes a slow negative potential shift in EEG activity preceding voluntary movement. It has been shown that the BP also precedes the conscious awareness of initiating the movement by several hundreds of milliseconds (Libet, et al., 1983; Libet, 1985). The origin of the BP seems to be located in the SMA / pre-SMA and anterior cingulate motor areas (CMA) (Ball et al., 1999; Deecke, Lang, Heller, Hufnagel & Kornhuber, 1987). Several studies further demonstrated that the pre-SMA is a key region for focusing on self-initiated actions (Cunnington et al., 2005; Haggard & Eimer, 1999; Lau et al., 2004a; Lau, Rogers & Passingham, 2006; Nachev, Rees, Parton, Kennard & Husain, 2005; Nachev, Wydell, O'Neill, Husain & Kennard, 2007) and that pre-SMA and rostral SMA were more strongly activated when participants performed self-initiated movements compared to externally cued ones (Deiber et al., 1999; Jenkins, et al., 2000). Others found activation differences for this contrast in the adjacent rostral cingulate zone (RCZ) (Forstmann et al., 2006; Mueller, Brass, Waszak & Prinz, 2007) or showed that RCZ and pre-SMA differed only in the onset of activation (Cunnington et al., 2002). Lau and colleagues tried to isolate the neural basis of intention and execution for self-initiated movements and demonstrated that focusing on the onset of upcoming *intentions* activated the pre-SMA (Lau et al., 2004a) while attending to the motor *execution* activated the RCZ, especially the CMA (Lau, Rogers, Ramnani & Passingham, 2004b). None of these studies, however, directly demonstrated the *encoding* of specific decision outcomes in these brain regions as achievable with multivariate decoding methods (see Haggard & Eimer, 1999, for a different approach to infer causality via detecting co-variations of potential causes and effects). Hence, these regions might play a role in free decision making different to computing the final decision outcome.

Other prefrontal areas that have been reported to be important for free decision making were located anterior to the pre-SMA. Fronto-polar cortex (FPC) has been suggested to represent the most abstract contents of cognition (Badre & D'Esposito, 2009). This region was discussed to be involved in keeping a pending action plan active while another task is performed (Koechlin et al., 2000; Koechlin & Hyafil, 2007). Recent decoding studies linked medial FPC to the generation of free decisions (motor decisions as well as abstract ones), long before these reach awareness (Bode et al., in prep.; Soon et al., 2008; in prep.). Medial

prefrontal cortex (PFC), anterior to the pre-SMA, was found to be activated during veto-decisions (Brass & Haggard, 2007). Outcomes of free decisions maintained in memory during a delay period could also successfully be read out from medial PFC (Haynes et al., 2007).

Medial PFC, however, was not found to encode participants' guesses in the previous study; decoding was only possible from a region in the precuneus extending to lateral posterior parietal cortex (PPC). Accordingly, others have shown that the pre-SMA might be more involved in the *timing* of free decisions rather than encoding the decision contents (Soon et al., 2008; in prep.). Studies using single-cell recordings in monkeys recently linked neurons in PPC to movement intentions (Quiroga, Snyder, Batista, Cui & Andersen, 2006). Correlations of spikes and local field potentials (LFPs) between neurons in PPC and dorsal pre-motor cortex were found to be greater when monkeys made free decisions compared to simply following instructions (Pesaran et al., 2008). The authors concluded that this could mirror an exchange of information between these regions when making a free decision. A recent study impressively demonstrated that direct electrical stimulation of human PPC evoked strong motor intentions (Desmurget et al., 2009). Stimulation of pre-motor cortex, however, did not lead to any urge to move; instead it directly triggered movements, which were not even perceived by the participants.

Taken together, recent findings suggest that a network involving medial prefrontal cortex as well as posterior parietal cortex (extending to the precuneus and posterior cingulate cortex), is involved in the generation of voluntary decisions. While precuneus / PPC were also found to encode guessing decisions in the previous study, medial prefrontal regions were not. On the one hand, this suggests a shared neural mechanism for guesses and free decisions located in parietal cortex; on the other hand, there might be important differences, reflected by differential involvement of medial PFC. However, no study so far has directly compared the *encoding* of guesses and intended free decisions using the same paradigm.

5.1.3 The goal of the present study

Many studies investigated internal, free decisions rather indirectly by asking participants to *think* about their intentions instead of acting according to them and also did not control for

the influences of personal values and attitudes (e.g. Blakemore, den Ouden, Choundhury & Frith, 2007; den Ouden et al., 2005). Others successfully implemented tasks in which the choice alternatives did not differ in value, but differed from the previous study in terms of stimuli and decision timing (Mueller et al., 2007; Soon et al., 2008; in prep.). The present study was conducted to achieve a *direct comparison* between random guessing decisions (in a perceptual decision making task as implemented in the previous study) with intended free decisions using the same visual stimulation and decision entities. This allowed the investigation of whether the same neural mechanisms were recruited in both situations.

A second shortcoming of the previous masking study was that residual visibility of strongly masked objects could not be fully ruled out. Many studies have demonstrated that visual information can be processed in the brain without reaching awareness and can influence behavior (e.g. Dehaene et al., 1998; Fahrenfort, Scholte & Lamme, 2007; Fang & He, 2005; Haynes & Rees, 2005a; Van Gaal, Ridderinkhof, Fahrenfort, Scholte & Lamme, 2008; Vorberg et al., 2003; for classical masking studies see also Bachmann, 2006; Breitmeyer & Stoerig, 2006). The involvement of the precuneus / PPC could in principle be explained by weak visual input to object selective regions in the dorsal visual cortex (Konen & Kastner, 2008). To rule out this possibility, in the present study the low visibility condition was replaced by a true *invisibility* condition: only strongly masked non-meaningful noise images were presented, which served as neutral stimuli while all decisions were made between two object categories (pianos and chairs) as choice alternatives.

The experimental procedure was similar to the previous study. Participants either had to categorise the presented object as being a piano or a chair (as in the previous study) or they had to spontaneously and freely choose one of the two categories. The same multivariate decoding techniques (Haynes & Rees, 2005; Haynes et al., 2007; Kriegeskorte et al., 2006; Norman et al., 2005) were applied as for the previous study to analyse the data. It was tested whether decision outcomes from guessing and intended free decisions were based on the same neural patterns in any brain region, as hypothesised for the precuneus / PPC. For this, an unbiased whole-brain cross-condition pattern analysis was performed, testing for *inter-changeability* between the underlying activation patterns for guesses and free decision outcomes. Finally, the results of the pattern classification analyses were compared to those obtained from the previous study.

5.2 Methods

Participants

Sixteen participants took part in the fMRI experiment and gave written informed consent to the test procedure. The experiment was approved by the local ethics committee and was conducted according to the Declaration of Helsinki. Participants were told that they took part in a decision making experiment but were not informed about the crucial experimental manipulations. All participants were right-handed and had normal or corrected to normal visual acuity. No participant indicated to be a professional musician (or to be directly related to one) or took part in earlier experiments presented here. Only one participant indicated after the experiment that he had noticed the missing object images and therefore stopped doing the task as instructed. This participant's data was excluded from all analyses. The remaining sample consisted of seven female and eight male participants (mean age 25 years; range 21-28).

Stimuli

The experiment implemented a categorisation task with either highly visible objects or poorly visible ones (similar to the previous study) and a free decision task. For the categorisation task under high visibility (hereafter referred to as HV), the target stimuli needed to be fairly visible objects. For this, standardised grey-scaled images of pianos and chairs were presented using a moderate masking procedure (see Appendix C, Behavioural pre-tests). For the categorisation task using invisible objects (IV) and the free decision task (FD), the target stimuli must not contain meaningful object stimuli, yet should nonetheless create the illusion that there is indeed an object stimulus presented, albeit difficult to detect. For this, a phase-randomised scrambled pattern image was presented, which was additionally strongly masked. It was therefore guaranteed that participants' decisions were pure guesses or free choices, without any residual visibility influencing either process. The object stimuli (HV) were taken from the previous study (see Experimental Section II for details) and were comprised of 24 grey-scaled images per object category, displaying pianos and chairs with

natural backgrounds (size 400 x 400 pixels). Stimuli were again masked using scrambled checkerboard-masks of 10 x 10 cells (each 40 x 40 pixels) for strong sandwich-masking (see Experimental Section II and Appendix B for details). To create the non-object target stimuli used for the IV and FD conditions, one of the (non-meaningful) mask images was randomly chosen for a two-dimensional Fourier transformation. The phase map was scrambled by adding a random value of $\pm 1.75 * \pi$ to each phase angle. Subsequently, the new phase map was read back using an inverse Fourier transformation with the unchanged power spectrum. The resulting image was completely neutral in terms of its shape and visual information while serving as an average stimulus image for all IV and FD trials.

Experimental procedure

The paradigm operated with three experimental conditions (Figure 5-1b):

1. Perceptual categorisation task for moderately masked, highly visible object images of pianos and chairs (HV);
2. Perceptual categorisation task for strongly masked (invisible) noise images, which also had to be categorised as pianos or chairs (IV) and
3. Free decision task (FD) in which participants had to choose object categories (pianos or chairs) independent from the visual presentation (of invisible noise images)

Participants were given written instructions first, explaining that they had to perform *two* different kinds of task, a *perceptual categorisation task* and a *free decision task*, indicated by the colour of the fixation cross in the beginning of each trial. Participants were *not* told that in the IV and FD trials, no object images were actually presented. During the FD task, participants were instructed to look at the screen but to spontaneously choose the category that first came to their mind, independent of the presentation. Additionally, it was pointed out that they should not change their mind during the trial but always keep the first decision. Most importantly, participants were instructed to initially choose an *object category* instead of a response button to ensure that a semantic rather than a motor decision was made, comparable to the perceptual categorisation task. The instruction text balanced the words “piano” and “chair” with respect to the order of their appearance to avoid biasing participants’ subsequent decisions. Participants were asked to make a decision on each trial, even

if uncertain (in categorisation trials). In this case, they were instructed to follow their intuition or to guess rather than not to respond.

Participants were told to fixate upon a cross in the middle of the screen throughout the entire experiment. Upon the start of each trial, the task cue was given via a change in the colour of the fixation cross during the 750 ms before target presentation. A red cross indicated a perceptual categorisation trial; a green cross indicated a free decision trial. In half of the perceptual categorisation trials, moderately masked images of pianos and chairs were presented (HV); in the other half, only the neutral noise image was presented between strongly effective masks (IV). In free decision trials (FD), again only the strongly masked noise image was presented. Depending on the condition, the target image was presented for 66.7 ms (highly visible object) or 16.7 ms (noise image). The sequence of masked images (pre-mask 150 ms; post-mask 500 ms minus target image duration) was shown three times in direct succession to prolong the overall visual stimulation and to enhance the ratio of signal to noise in fMRI data acquisition (compare Experimental Section II). Finally, a response-mapping screen was presented for 2000 ms. Participants indicated their choice by pressing one of two response buttons (right index- and middle finger), corresponding to the position of the letters “K” (“Klavier”, German for piano) and “S” (“Stuhl”, German for chair), which were presented on either side of the fixation cross. The position of letters was pseudo-randomised on a trial-by-trial basis, separately for all three experimental conditions. This procedure ensured that motor responses were dissociated from participants’ category choices. Following the response mapping screen, the white fixation cross was shown again, completing a jittered delay of seven, eight or nine seconds per trial (Figure 5-1a). Participants performed six runs (72 trials per run). In each run, 48 trials were perceptual categorisation trials (24 HV, 24 IV) and 24 were free decision (FD) trials. The same object image was never shown twice in one run. The order of object images was individually randomised for each participant. Participants first performed ten practice trials outside the scanner. During the fMRI experiment, stimuli were presented using a projector (resolution 1024x768 pixel) placed at the head-end of the scanner, projecting onto a screen behind the participants’ heads. They watched the stimulation via a mirror attached to the head coil. The estimated angle of vision was $\alpha \sim 7.2^\circ$ for all stimuli. The presentation was carried out using the Cogent toolbox for MATLAB 7.0 (The MathWorks, Inc.). After the experiment, participants completed a self-administered questionnaire, which asked about the levels of perceived difficulty, adherence to the instructions and the timing of the free decisions (Appendix C, Figure C-3).

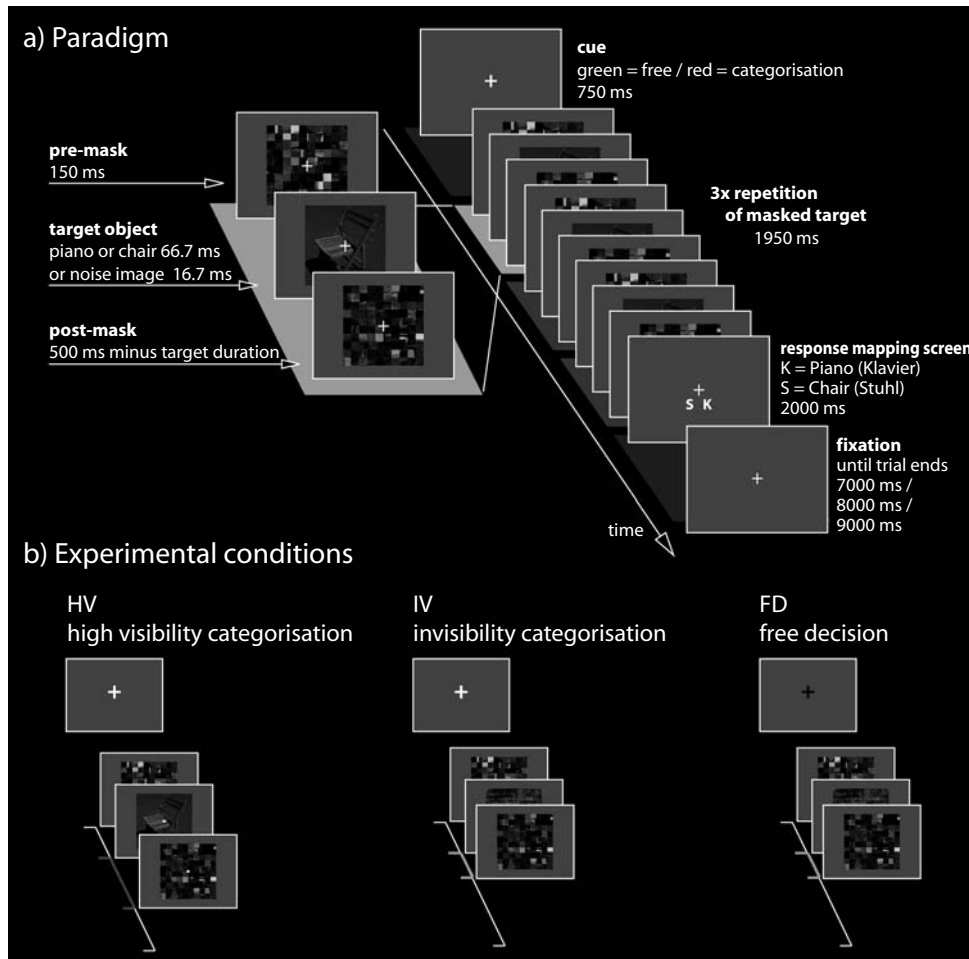


Figure 5-1: Experimental paradigm and conditions. a) Left side: A pre-mask was presented for 150 ms, immediately followed by the target image, which was a piano or a chair (HV) or a noise image (IV, FD). Noise images were shown for 16.7 ms, object images were shown for 66.7 ms. The post-mask duration was calculated as 500 ms minus target duration, resulting in 433.3 ms (HV) and 483.3 ms (IV, FD). The sequence of masked images was directly repeated three times (right side), preceded by a cue (750 ms), indicating the task via the colour of the fixation cross (red = categorisation task; green = free decision). The total duration of this presentation was 2700 ms, followed by a response-mapping screen (2000 ms) with a pseudo-randomised order of response buttons associated with the categories to avoid confounds of categorical choices with motor responses. Subsequently, a fixation period completed the total trial duration of 7000 ms, 8000 ms or 9000 ms. b) Each experimental condition was presented in one third of all 72 trials per run: 24 trials categorisation task with highly visible object images of pianos and chairs (HV; left side); 24 trials categorisation task with strongly masked noise images, which also had to be classified as pianos or chairs (IV; middle); 24 trials free decision task with strongly masked noise images (FD; right side) in which participants had to choose object categories independent from the presentation. Participants were not aware that in IV and FD trials no objects were shown. Trials from all three conditions were presented in an individually randomised order for each run and each participant.

Functional Imaging

Functional imaging was conducted on a Siemens TIMTrio 3T scanner (Erlangen, Germany) with a standard head coil. Gradient-echo EPI functional MRI volumes of the whole brain were acquired (42 axial slices, TR = 2800 ms, echo time TE = 30 ms, matrix size 64 x 64, voxel size 3 x 3 x 2 mm³ with 1 mm interslice gap). In each of the six functional runs, 208 EPI volumes were acquired for each participant. The first two images of each run were excluded from the analyses to allow for magnetic saturation effects.

Data Analyses

Data analysis was performed separately for the HV, the IV and the FD condition. In order to minimise the risk of false positive results, there were restrictions as to which participants and which runs were included in the following analyses. Participants were only included in the analysis for the HV condition if their performance exceeded 80% for both object categories in more than half of the runs (> 3 runs) and did not drop below 70% overall for any category. Participants' data were excluded from the analyses for the IV and FD condition if they were less balanced than a decision ratio of 20:80 (or 80:20; meaning less than 4 trials per condition) in a single run. The estimation of decision-related signals was likely to be distorted in these cases. If a participant's choices were strongly imbalanced in more than half of the runs, that participant's data were completely excluded from the analysis.

The first stage of data processing involved motion correction to the first image of the first run using SPM2. No additional normalisation or smoothing were performed at that stage in order to maximize the sensitivity for information encoded in the fine-grained voxel patterns (Haynes & Rees, 2005a,b, 2006; Kamitani & Tong, 2005). A GLM was then used for statistical analyses on the individual subject level. A high-pass filter with a cut-off of 128 s removed low frequency drifts in the time series at each voxel. Trials were assigned to piano- and chair-trials according to the participants' *choices*. For each run, the onsets of events, beginning with the presentation of the cue, were convolved with the canonical haemodynamic response function (HRF), to obtain the regressors. The parameter estimates for each voxel position formed the basis of subsequent decoding analyses.

Again, a searchlight approach of multivariate pattern classification was used to analyse the data (Haynes & Rees, 2006; Haynes et al., 2007; Kriegeskorte et al., 2006; Norman et al., 2006). The analysis sought to find brain regions that allowed *participants' categorical choices* to be decoded from fine-grained patterns of activity as measured by the BOLD signal. In all cases, participants' choices were limited to pianos and chairs and chance level for correct prediction was 50%. Trials with no decisions were excluded from the analyses. The procedure was largely identical to the previous study but will be briefly described again. Starting with an arbitrary voxel position v_i , a spherical cluster with the radius of four voxels was defined around v_i containing N surrounding voxels ($v_{i...N}$). The parameter estimates of these voxels were transformed into two pattern vectors, one for decisions for pianos, the other one for decisions for chairs. For each run of each participant, such a pair of pattern vectors was extracted. Five of the six pairs of pattern vectors served as the "training data set" for multivariate pattern classification analysis. A linear support vector classifier (Müller et al., 2001) was used with a fixed regularisation parameter $C = 1$ (LIBSVM implementation). The remaining pair of pattern vectors served as the "test data set" and was to be classified based on the classification hyperplane estimated from the training data. The procedure was reiterated six times, each time leaving out the pattern vectors of a different run to be used as a "test data set". Thus, a six-fold cross-validation was achieved by averaging the decoding accuracies from all six iterations (or if less than six runs were used, then a five- or four-fold cross-validation was performed). The average decoding accuracy [%] was then assigned to the central voxel of the searchlight cluster. This procedure was repeated using the next searchlight cluster, constructed around the next spatial position at voxel v_j . After performing this procedure for every voxel in the brain, a three-dimensional map of decoding accuracies could be created for each condition (HV, IV, FD) and each participant. In a last step, these maps were spatially normalised to a standard stereotaxic space (MNI EPI template), re-sampled to an isotropic spatial resolution of $3 \times 3 \times 3 \text{ mm}^3$ and smoothed with a Gaussian kernel of 6 mm FWHM using SPM2. The resulting maps were fed into random-effects group analyses in order to identify common regions across all participants that displayed high decoding accuracies and therefore contained relevant *information* about the participants' choices.

In addition, a cross-condition classification analysis was performed for the IV condition and the FD condition. This analysis aimed at searching for regions in which the activity patterns used for the prediction of decision outcomes were *interchangeable* between the two conditions. This means, the patterns from one condition should allow the prediction of

decision outcomes from the other condition and vice versa. Finding such a brain region would speak for a similar neural code for both decisions. The pattern vectors from one condition (e.g. IV) were therefore used to train the classifier to distinguish between choices for pianos and chairs, but instead of predicting this participants' choices in the one remaining run within the same condition, it was used to predict decision outcomes from the left-out run for the *other* condition (e.g. FD) (Figure 5-2). Note that this procedure maximised the comparability with the previous analyses (e.g. the number of training and test data sets and cross-validation steps) (Walther et al., 2009). Both ways of cross-condition classification were performed and an ANOVA was applied followed by an inclusive conjunction test for both contrasts. Only regions that successfully supported both directions of cross-condition classification were considered significant.

Additional control decoding analyses were conducted. The first analysis decoded the *category of objects presented* during HV trials. Trials were therefore modelled according to the object category presented. The regressors for pianos and chairs were then used for subsequent decoding, which was identical to the analysis described above. A second control analysis was conducted to decode the *motor responses* by using different regressors for left and right button positions, i.e., the finger that operated the response box on each trial. This served to ensure that the decoding of choices was not confounded with the decoding of motor responses. The minimal statistical threshold used for all analyses was $p < .0001$ (uncorrected) with a voxel-threshold of ten voxels.

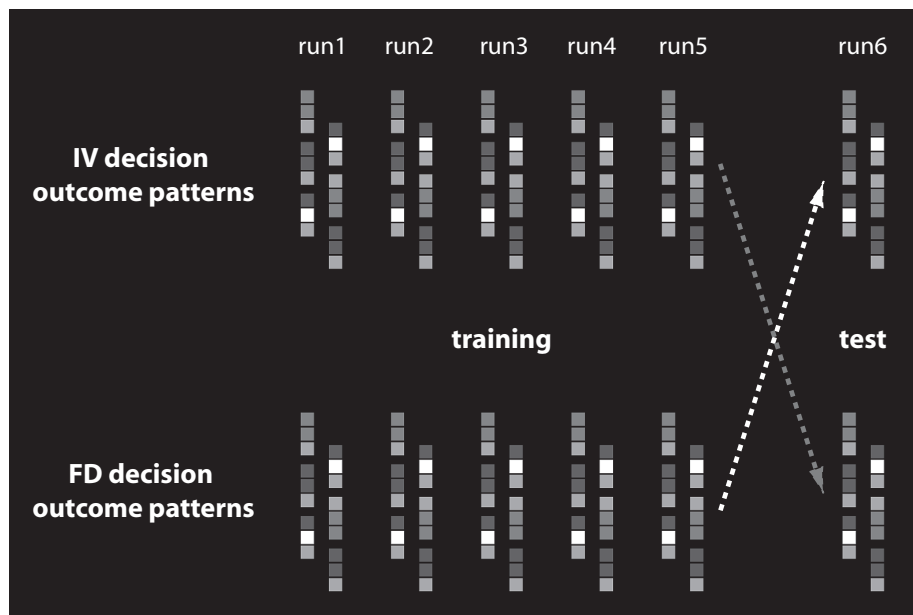


Figure 5-2: Cross-condition pattern classification. The pattern vectors from decision outcomes of the IV condition were used to train the classifier to distinguish between choices for pianos and chairs. Instead of predicting participants' choices from the one remaining run within the same condition, it was used to predict decision outcomes from the left-out run for the FD condition. The same six-fold cross-validation procedure was used as described for the regular decoding analysis and an average decoding accuracy was calculated for the respective cluster. Again, a searchlight algorithm was applied; the analysis was repeated for all possible searchlight clusters ($r = 4$ voxels). The resulting 3-dimensional decoding accuracy maps quantified how good patterns from IV predicted FD decision outcomes. The same analysis was repeated training the classifier with activation patterns from the FD condition and predicting the decision outcomes from the IV condition. The normalised decoding accuracy maps from each participant were used for an ANOVA. Only those regions were considered significant in which both ways of cross-condition classification were successful above chance level.

5.3 Results

The hit rate for both highly visible pianos (85.2%) and chairs (91.5%) was very high. Only one participant showed an exceptionally bad performance for one category ($s_{15} < 70\%$ for chairs; see Figure 5-3) and was excluded from the decoding analysis for the HV condition. Another two participants were excluded from this analysis because they did not reach the criterion of $>80\%$ hit rate for both categories in at least half of the runs, which challenged the assumption that they identified enough objects in these runs.

In total, participants were highly balanced in their choices for pianos and chairs in the IV as well as in the FD condition (Figure 5-4). In the IV condition, participants chose pianos in 47.5% and chairs in 51.2% of all trials (1.3% trials missed). In the FD condition, participants chose pianos in 51.5% and chairs in 47.5% of all trials (1.0% trials missed). The individual performance varied more strongly, but only one participant had to be excluded from the analyses due to being too imbalanced in more than half of all runs (s_5 was above the ratio of 80:20 for pianos – chairs). For four participants, data from one run was excluded due to imbalanced decisions; for two participants, data from two runs were excluded (for details see Appendix C, Choice ratios fMRI experiment).

The analysis of the questionnaire indicated that participants adhered to the instructions. The perceptual categorisation task was perceived as being rather difficult and most participants believed there to be 3 to 4 levels of difficulty. This confirmed that participants did not notice the two-fold scaling of visibility. As instructed, free decisions were reported to be made with regard to categories instead of randomly pressing response buttons. Free decisions were made early in the trial, which was similar in perceptual decisions. Additionally, in almost all cases, participants reported that they did not recognise any (illusory) objects during free decisions. These results confirmed that the experimental manipulations were successful. For details see Appendix C (Figure C-4).

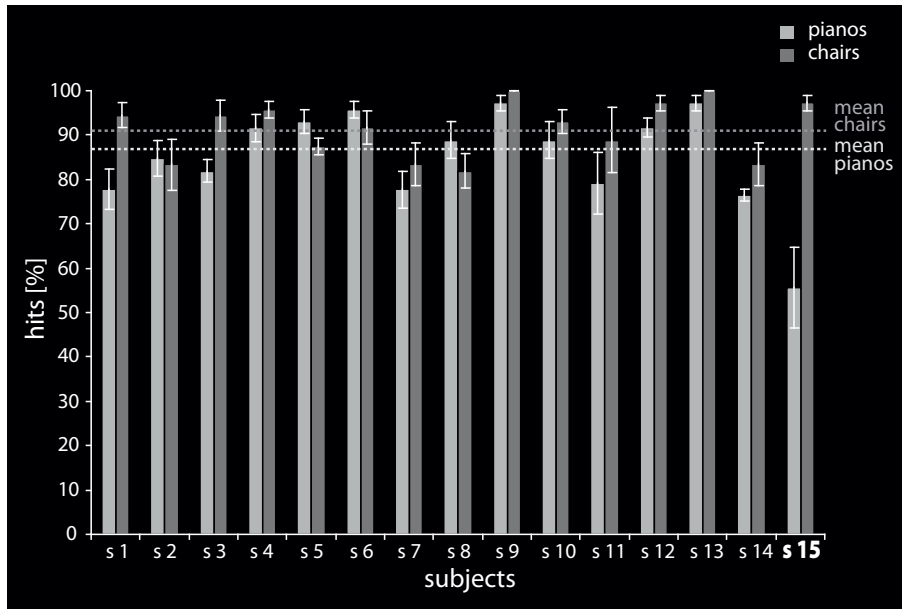


Figure 5-3: Mean hit rates for each participant. Hit rates for pianos and chairs are shown for the high visibility categorisation condition (HV). Displayed are the mean values [%] across all runs and the standard errors (SE) for each participant individually. After excluding participant s15 because of an exceptionally low hit rate, the average hit rates were 87.3% for pianos (bright dashed line) and 91.1% for chairs (dark dashed line). The difference was not significant (t-test, $p > .05$).

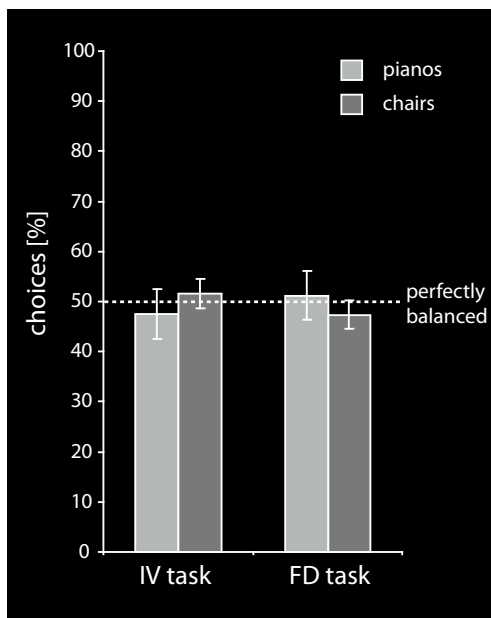


Figure 5-4: Average choice ratios. Ratio of choices for pianos and chairs in the invisible categorisation condition (IV; left) and the free decision condition (FD; right) for all $n = 15$ participants. In the categorisation task, participants chose pianos in 47.5% and chairs in 51.2% off all trials (1.3% trials missed). In the free decision task, participants chose pianos in 51.5% and chairs in 47.5% off all trials (1.0% trials missed). The overall ratios were close to being perfectly balanced (50%) in both conditions. See Appendix C for individual results.

Multivariate decoding was used to investigate which brain areas encoded decision outcomes in perceptual decision making when object images were highly visible (HV). Decoding accuracies significantly above chance level (50%) could be found bilaterally in extrastriate visual cortex extending to the LOC, with greater extensions on the right side (see Figure 5-5a). The peak decoding accuracy was 61%. The same method was applied to decode decision outcomes for perceptual decision making without visible objects (IV). The analysis revealed two regions in medial parietal cortex, mainly located in the right precuneus, which encoded the choices (Figure 5-5b). The first cluster extended to the posterior cingulate cortex (PCC) and peaked with 61% accuracy. The second cluster was located more dorsally and posteriorly, extending laterally to the inferior parietal lobe, peaking with 60% decoding accuracy. Decoding accuracies in left precuneus were just below threshold. No further regions could be found displaying above chance accuracies for the IV condition.

Decoding free decisions (FD) revealed an extended cluster in bilateral anterior medial prefrontal cortex. The peak decoding accuracy was 64% (Figure 5-5c). The regions in the precuneus, which encoded the decision outcome for the IV task (as well as clusters nearby) also displayed high decoding accuracies for free decisions. These, however, missed the strict statistical threshold due to higher variance (peaking at 58% accuracy; $p = .009$). When inclusively masked with decoding results from the IV condition ($p < .0001$ uncorrected), a cluster in the precuneus remained marginally significant for free decisions ($p < .01$ uncorrected). For an illustration of the overlap see Figure 5-6.

To further investigate the relationship between guessing in perceptual decision making (IV) and free decisions (FD), cross-condition classification analyses were conducted. Pattern vectors associated with decision outcomes from the IV task were taken to predict those from the FD task and vice versa. Two clusters, again bilaterally located in the precuneus, were found to allow bi-directional interchange between the predictive activation patterns (Figure 5-5d). One cluster was located close to the second precuneus cluster that encoded decision outcomes in the IV task. The other cluster was located more ventral at the border between precuneus, cuneus and inferior parietal lobe. The peak decoding accuracy was 57%. When the threshold was relaxed, a greater extent of regions in the precuneus could be revealed (see Figure 5-6). No other combination of cross-classification analyses showed any significant clusters.

Conventional univariate analyses could not reveal any differences between decisions for pianos and chairs for any of the conditions, confirming that information about decision

outcomes was only present in the fine-grained patterns of activation rather than reflected in differences in the average amplitude of the BOLD signal (see Appendix C, Univariate analyses).

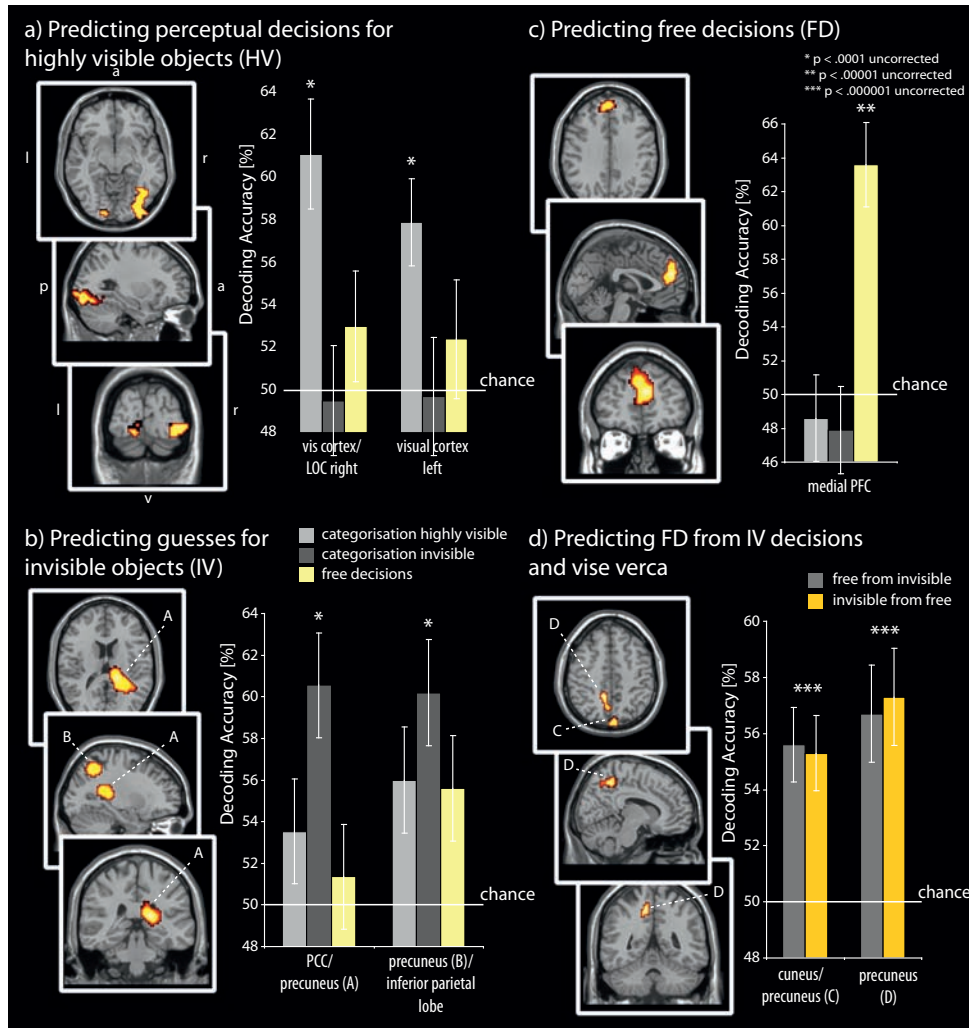


Figure 5-5: Decoding decision outcomes using multivariate pattern classification. Displayed are mean decoding accuracies and standard errors (SE) for searchlight decoding with a radius of $r = 4$ voxels. All results are displayed using a threshold of $p < .001$ (uncorrected) for better visualisation. a) Results for perceptual decision making for highly visible objects (HV). Decoding accuracies significantly above chance level (50%) were found bilaterally in extrastriate visual cortex extending to the LOC with a greater extension on the right side (MNI [36 -66 -15]; peak accuracy 61%; $p < .0001$ uncorrected; MNI [-12 -93 -3]; peak accuracy 58%; $p < .0001$ uncorrected). Decoding IV and FD decisions outcomes did not significantly exceed chance level in these regions. b) Results for guessing decisions in perceptual decision making (IV). Decoding was possible from two

clusters located in (A) precuneus extending to posterior cingulate cortex (PCC) (MNI [24 -45 18]; peak accuracy 61%; $p < .0001$ uncorrected) and (B) precuneus extending to inferior parietal cortex (MNI [30 -63 51]; peak accuracy 60%; $p < .0001$ uncorrected). No other regions were found to encode the decision outcomes in this condition. c) Results for decoding free decisions (FD). Decoding accuracies significantly above chance level were found in bilateral medial prefrontal cortex (MNI [3 48 21]; peak accuracy 64%; $p < .00001$ uncorrected). Additionally, high decoding accuracies could be found in precuneus but were just short of the statistical threshold (MNI [33 -60 45]; peak accuracy 57%; $p = .009$ uncorrected; not displayed). d) Results for cross-condition classification for IV and FD. Pattern vectors from decision outcomes for the IV task were taken to predict decision outcomes for the FD task and vice versa. Two clusters in bilateral precuneus allowed both directions of decoding (C: MNI [6 -84 39]; peak accuracy 55%; $p < .000001$ uncorrected; D: MNI [-12 -48 48]; peak accuracy 57%; $p < .000001$ uncorrected).

A further decoding analysis was performed to predict the *true categories* of the objects during HV trials. The category of the presented object and participants' choices were highly correlated under high visibility, hence, not surprisingly the strongest cluster was located in bilateral visual cortex and LOC, peaking with 61% accuracy in right visual cortex / LOC (Appendix C, Figure C-6b). To confirm that the decoding results for participants' choices were not confounded with the motor responses, an additional decoding analysis for button presses was performed. It could be shown that for all conditions, the motor responses were encoded in left motor cortex only (58% peak accuracy; Appendix C, Figure C-6a) and not in any choice-related areas. Further analyses were conducted to ensure that individual differences in decision balance did not drive the decoding effects. For this, the data was divided into groups of more balanced and less balanced participants by median-split and decoding accuracies were compared between groups. No differences in decoding accuracies could be found between more or less balanced participants. This held true for the whole brain volume as well as for the best searchlight clusters in precuneus and MPFC (for details see Appendix C, Choice ratios and decoding accuracies).

5.3.1 Comparison to results of the previous study

One remaining question was if the precuneus regions in the present study were comparable to the ones found to encode perceptual choices about poorly visible objects in the previous study. There was no complete overlap between the single clusters revealed by the respective analyses, even when a similar task was performed (Figure 5-6; compare blue: previous study

– poorly visible object categorisation, and red: present study – invisible object categorisation). One cluster in precuneus / PCC appeared to be spatially isolated, yet it was not functionally dissociable. All clusters, however, strongly overlapped with the cluster revealed by Soon et al., 2008 (see Figure B-16). The precuneus regions revealed by the previous study overlapped with the precuneus regions in the present study that encoded the IV decision outcomes and allowed IV and FD cross-classification. Additional analyses using ROIs derived from the previous study to analyse the present data yielded a similar *encoding profile* compared to the whole-brain searchlight decoding analyses (for details see Appendix C, ROI analyses). Taken together, these parietal clusters formed a patchy but connected ensemble, reaching from bilateral medial ventral parts of the precuneus to the lateral surface of the inferior parietal lobe.

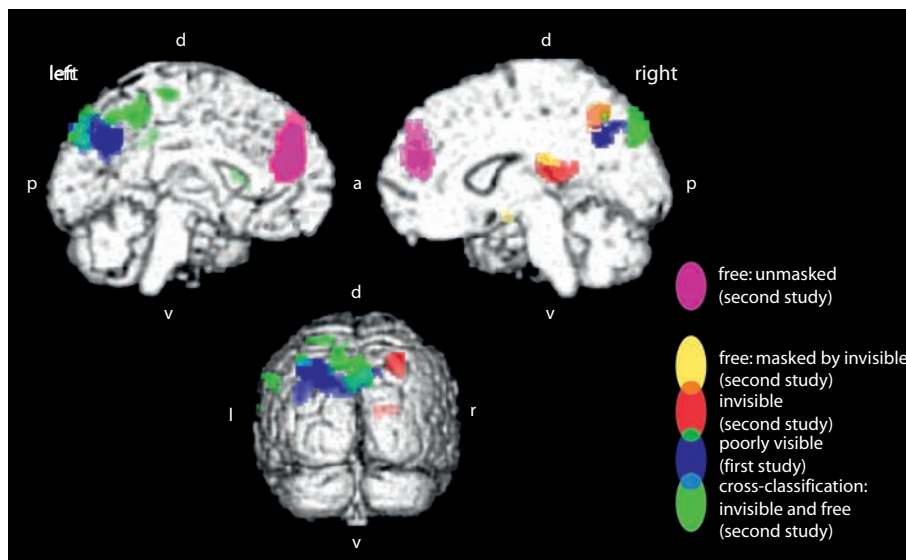


Figure 5-6: Direct comparison of precuneus regions between both masking studies. Medial views of the left and right hemisphere and one posterior view. Blue: The cluster from the previous study that allowed decoding of decisions about poorly visible objects. Red: The comparable cluster from the present study (IV). Green: Areas from which cross-condition classification between free decisions (FD) and IV was successful. Yellow: Clusters from which FD could be decoded in the present study ($p < .01$, uncorrected; after inclusively masking with IV decoding with $p < .0001$, uncorrected). Orange: Overlap of cross-condition classification (IV and FD) with IV. Pink: Clusters encoding only FD in the present study. In summary, searchlight decoding analyses from both studies revealed a rather widespread, overlapping cluster in precuneus / PPC. A second, functionally not dissociable cluster was located in anterior precuneus / PCC. Both clusters overlapped with the regions encoding free decisions revealed by Soon et al. (2008). For comparison a threshold of $p < .0001$ was used for displaying all clusters. p = posterior; a = anterior; l = left; r = right; d = dorsal; v = ventral. Peak coordinates and accuracies can be found in the text.

5.4 Discussion

This study demonstrated that the precuneus was informative about outcomes of guessed decisions when visual information was insufficient for a perceptually based decision. It therefore confirmed the previous findings by excluding the possibility of any residual visibility driving the effects, because a non-meaningful strongly masked noise image replaced the object images in a true invisibility condition. A direct comparison between guessing and free decisions was achieved by using the same visual stimulation while participants freely and spontaneously chose object categories. Decoding accuracies for free decision outcomes were highest in aMPFC, but activation patterns in the precuneus were found to be interchangeable for guesses and free decisions. Comparing the different clusters in precuneus / PPC from which decoding was possible between these studies, a wide-spread medial posterior parietal network emerged. This might be the neural substrate of a symmetry-breaking network, which is engaged during different kinds of internal decision making.

5.4.1 Two modes of perceptual decision making

The present results entirely confirmed the findings about perceptual decision making from the previous study. Again, a double dissociation emerged between the LOC, which encoded decision outcomes for highly visible objects only, and the precuneus, which encoded the decisions made with insufficient visual input (or guessing). These findings are at odds with models derived from SDT (Parker & Newsome, 1998; Swets, 1961), which predict that decisions about perceptual entities, even invisible ones, should still be based on the same dimension of sensory signal. If stimuli cannot be discriminated, trial-by-trial noise fluctuations in the same system should determine the choices. Several studies using single cells recordings in monkeys to quantify choice probabilities (e.g. Britten et al., 1996; Huk & Shadlen, 2005; Kim & Shadlen, 1999; Purushothaman & Bradley, 2005; Romo et al., 2002; 2004; Salinas et al., 2000; Shadlen & Newsome, 2001; Uka et al., 2005) suggested the existence of neural systems for perceptual decisions for all levels of stimulus ambiguity. The present findings, however, strongly support the conclusion of the previous study: that a switch from perceptual decisions to internal, stimulus-independent decisions occurs in the

absence of sufficient stimulus information. It was argued above that, contrary to classical studies using random dot motion or flutter discrimination, extended temporal integration was not required with the present paradigm due to the use of static images (Philiastides & Sajda, 2006). Under these circumstances a switch to a decision mechanism that generated random outcomes was facilitated. This switch could be explained within a framework that assumes that if a perceptual decision cannot be made because of insufficient sensory input or equal evidence for each alternative, a *symmetry-breaking network* becomes important. A similar model has been proposed for the generation of free decisions (Haggard, 2008). For a detailed discussion, see Experimental Section II.

The finding that decision outcomes about highly visible objects was encoded in the LOC is in line with previous studies linking the LOC to the processing of object categories (e.g. Grill-Spector & Malach, 2004; Haushofer et al., 2008a; Haxby et al., 2001; Op de Beeck et al., 2008a,b; Reddy & Kanwisher, 2006; Schwarzlose et al., 2008), but more importantly, to successful object categorisation performance (Grill-Spector et al., 2000; Williams et al., 2007). The LOC was more strongly activated for the categorisation task with highly visible objects than for the free decision task. Interestingly, there was no difference in activation when categorisation of highly visible objects and invisible objects were compared. Given that these differences in activation were clearly present in the previous study, the reduced number of trials per condition and therefore the limited statistical power might be an explanation. Accordingly, the decoding effects found in LOC were also much weaker compared to the previous study. It might be that the lower proportion of highly visible objects presented (50% in the previous study; 33% in the present study) also made perceptual decisions seem more difficult in general and reduced effects in the LOC.

The motor responses were again dissociated from the decisions by using randomised response mapping screens. Additional decoding analyses confirmed that motor responses were encoded in left motor cortex only, confirming that participants' decisions were always related to the objects, not the buttons. In the previous study, one concern was that residual visibility might have been responsible for decision encoding in the precuneus, as it has been shown that PPC can process object information (Konen & Kastner, 2008). This issue was addressed in the present study by using strongly masked neutral noise images instead of real objects to ensure that target images were truly invisible and only the *illusion* was created that objects were presented. By taking this path, it was possible to avoid any concerns regarding d' values greater than zero (Schmidt & Vorberg, 2006). Results from the detailed post-experimental questionnaire confirmed the effectiveness of this manipulation. It can be

concluded that neither the present nor the previous results were significantly influenced by residual visibility effecting encoding in the precuneus. The subjective experience of not having perceived the objects might be the important factor that triggered the involvement of the symmetry-breaking network.

There is still the possibility, however, that other mechanisms drove the encoding of decisions in precuneus. For instance, the precuneus might have been engaged in mental imagery (Fletcher et al., 1995; Knauff et al., 2003) of objects as interpretations of the (non-meaningful) visual stimulation, driven by the expectation of object images. A comparison with images from memory could also have involved the precuneus (Tulving et al., 1994). These alternatives cannot be fully dismissed by the present study. However, as argued before, mental imagery should have involved the respective sensory systems instead of excluding those (Kosslyn et al., 2001). The results therefore strongly support the alternative hypothesis: The precuneus might have been engaged in the generation of *internal decisions* with random outcomes rather than in the creation of mental object images.

The crucial analysis involved identifying brain activity patterns common to guesses and free decisions. It was found that a rather wide-spread area in precuneus / PPC demonstrated those pattern similarity effects and cross-predicted the decision outcomes in both directions, even though statistically significant decoding was only possible for guessing but barely for free decisions alone. It has to be noted, however, that decoding of free decisions from the precuneus was statistically weak but still possible with comparable accuracy. Second, the results suggest that the classification hyperplane related to free decisions was estimated based on comparably less stable patterns in the precuneus (as demonstrated by the classifier's performance). This hyperplane, if also meaningful for the "guessing" condition, should nonetheless be able to successfully classify the more stable patterns underlying guesses, which was demonstrated. Likewise, if stable and informative patterns were used to estimate a meaningful and applicable classification hyperplane, it should also successfully classify weaker patterns in most cases, as observed for classifying free decisions based on the hyperplane estimated from guessing. The cross-classification analysis therefore validly tested the similarity of the underlying activation patterns.

Overlaid with the findings from the previous study, a wide-spread, patchy and partially overlapping cluster in precuneus / PPC emerged for internal decisions (see Figure 5-6). These clusters also strongly overlapped with the cluster revealed for free decisions by Soon and colleagues (2008) (also see Appendix B, Figure B-16). As argued before, the

precuneus / PPC has also been found to be involved in other internal tasks such as focussing on one's own intentions (e.g. Blakemore et al., 2007; Cavanna & Trimble, 2006; den Ouden et al., 2005), internally selecting actions (Jenkins et al., 2000; Mueller et al., 2007), intentional self-processing (Kircher et al., 2000; 2002), prospective memory (Burgess, et al., 2001), the formation of free decisions (Cui & Andersen, 2007; Pesaran et al., 2008) and movement intentions (Desmurget et al., 2009). Taken together, these findings strongly support the hypothesis that regions in medial and posterior parietal cortex play a dominant role in the generation of intentions and free decisions. The specific role of the precuneus could be to break the symmetry of choice alternatives in the absence of sufficient perceptual input (guessing) or in the absence of a more valuable alternative (free decisions) in order to make an internal decision.

5.4.2 From guessing to free decisions

So far, it was argued that guessing resembles internal free decisions, both resulting from the same process of symmetry-breaking. It is obvious, however, that there are also differences between these two kinds of decisions: first, a network of brain regions was differentially activated during guessing and free decisions. Second, when looking at free decisions, the highest decoding accuracies were found in anterior medial PFC while outcomes of guesses could not be decoded from this region. Medial prefrontal cortex was also found to be activated by the voluntary selection of task-sets (Forstmann et al., 2005; 2006), by self-related motivational factors (Kouneiher, Charron & Koechlin, 2009) and by reflecting on one's own mental states (Amodio & Frith, 2006). More importantly, widespread anterior medial PFC regions have been found to directly encode freely formed intentions to perform arithmetic tasks (Haynes et al., 2007). The most anterior part of medial PFC was found by Soon and colleagues (2008; in prep.) to encode free decision outcomes. Note that differences in precise localization might be due to the temporal aspects of the generation of decisions, which cannot be resolved here due to differences in modelling. Some processes might only be visible when using unconstrained models (such as FIR) and monitoring decisions or intentions across longer timescales.

In the present study, cross-condition classification clearly points towards a shared neural substrate of guessing and free decisions in the precuneus. Only free decisions,

however, were additionally encoded in medial PFC. These differences might be explained by what made the tasks distinct. Both tasks required the generation of an *internal* decision, which was not based on external input and had equally valuable outcomes. Only in the free decision task, however, did participants *intend* to make an internal decision. Likewise, in most studies that demonstrated medial PFC involvement, participants were instructed to make free decisions and therefore *intended* to do so (Deiber et al., 1999; Forstmann et al., 2005; 2006; Haynes et al., 2007; Jenkins, et al., 2000; Mueller et al., 2007; Soon et al., 2008; in prep.). Contrarily, the categorisation task required participants to select their response on the basis of (insufficient) visual input. Even though the behavioural outcome was indistinguishable compared to the free decision task, the initial intention was not to produce a random decision outcome. A likely interpretation for the differences in encoding is therefore that the precuneus is always involved in symmetry-breaking when no choice alternative is more likely or valuable. However, only when participants consciously *intend* to make a free decision, the aMPFC is additionally involved in generating the decision outcome. This might then be achieved in concert with the precuneus / PPC (Soon et al., 2008), hypothetically by strong synchronisation between neural activity in medial PFC and (medial) PPC (Pesaran et al., 2008).

5.4.3 Posterior parietal cortex and high self-reference

The exact nature of the suggested symmetry-breaking network has to be further investigated in the future. It can only be speculated that central noise fluctuations might play an important role (Deco & Romo, 2008; Faisal et al., 2008). The extraordinary role of the precuneus in self-processing (Cavanna & Trimble, 2006) suggests that *self-reference* could be the linking element. Self-reference is usually given by values and preferences; these mechanisms would guide decisions in most situations. One's self-related value system is therefore a candidate for a symmetry-breaker: in the absence of more or less valuable outcomes, noise fluctuations in one part of this system could guide the decisions.

There is little direct evidence for this hypothesis, even though it has been suggested that random value-based decisions could exploit the same mechanisms as random perceptual decisions (Gold & Shadlen, 2007). In support of this, some studies have shown that a network including the precuneus, posterior cingulate cortex and posterior parietal cortex is

involved in value- and reward-based decision making as well as in decision making under uncertainty (Hsu, Bhatt, Adolphs, Tranel & Camerer, 2005; Huettel, Song & McCarthy, 2005; Luhmann, Chun, Yi, Lee & Wang, 2008; McCoy & Platt, 2005; Platt & Huettel, 2008; Volz, Schubotz & von Cramon, 2004). It has been pointed out that the integration of sensory evidence and prior knowledge (e.g. about reward and values) in parietal cortex might unify decision systems at their final stages (Glimcher, 2001). PPC might therefore be critical for many kinds of value judgements (Kable & Glimcher, 2007; Platt & Huettel, 2008). For example, Platt and Glimcher (1999) found that a population of neurons in monkey parietal cortex was directly involved in decision processes if the outcomes were valuable for the decision maker. Using fMRI in humans, Vickery and Jiang (2009) showed that a region in the inferior parietal lobe was modulated by feedback in a simple decision task, suggesting a role in constructing a *value representation* under uncertainty. A recent study demonstrated that while participants were passively watching images of different cars of varying value, subsequent purchase intentions could be decoded from regions including PPC / precuneus (Tusche et al., 2010). One interpretation linking these studies and the present findings is that the precuneus and PPC are part of an internal decision making system that is usually engaged in value- and preference-based decisions. Normally, the attachment of values determines the decision outcome. In cases of equal or no valence, for instance in guessing as well as equivocal free decisions, noise fluctuations in this system could determine a decision and serve as a symmetry-breaker. These interpretations, however, remain speculative and have to be addressed in future studies.

5.4.4 Conclusion

The present study confirmed that the precuneus encoded guessing when no relevant perceptual input was provided. Activation patterns in this region were found to be interchangeable with those for free decisions. These findings strengthened the interpretation that the precuneus might be part of the neural substrate of a symmetry-breaking network, involved in the generation of different kinds of *internal* decisions with random outcomes. Anterior medial prefrontal cortex, however, only encoded decision outcomes when decisions were *intended* to be internal and might therefore be exclusively important for free decisions made without an external frame of reference.

6 General Discussion

This final part discusses the implications of the three experimental sections beyond the respective discussion sections. A short interpretation of the present results will be given within a broader framework of the transformation of stimuli to responses. Methodological problems will be addressed, in particular, the limitations and challenges faced by multivariate pattern classification as implemented here. Ideas for future experiments, which resulted from the present findings, will then be discussed, finally followed by a concluding summary. Of course, there would also be much to say about limitations of fMRI in general; some problems were already addressed in the introduction to methods section. For more detailed discussions, which go beyond the scope of the present work, refer to e.g. Logothetis (2002; 2003; 2008), Logothetis and Wandell (2004) as well as Heeger and Ress (2002) for the relationship of the BOLD signal and activity of single neurons, Raichle and Mintun (2006) for details about brain metabolism, Brett and colleagues (2002) for the problem of localisation, Haynes and Rees (2006) as well as Norman and colleagues (2006) for the limits of univariate analyses and Kriegeskorte and colleagues (2009) for potential problems with circularity in fMRI data analyses in general.

6.1 Information encoding in parietal cortex

The present studies made use of new decoding approaches to investigate different aspects of the transformation of stimuli to motor responses. In the first study, the temporal decoding of task-sets demonstrated that abstract rules were encoded in the IPS even before a target stimulus was presented. Interestingly, rule information in parietal cortex preceded the build-up of rule information in VLPFC. These findings stressed that a *network* of prefrontal and

parietal regions is involved in rule-guided behavior (e.g. Brass & von Cramon, 2004a; 2004b; Bunge, 2004; Bunge & Wallis, 2007). This finding clearly goes beyond the current literature by demonstrating in humans that parietal cortex also encoded rules in an abstract form. So far, this has only been shown in monkeys (e.g. Gail & Andersen, 2006; Stoet & Snyder, 2004; 2007). The role of parietal cortex in rule-guided behavior has been hitherto undervalued. Many authors focused on PFC (Sakai, 2008) or neglected the similarity of activation profiles in parietal cortex compared to PFC (e.g. Brass & von Cramon, 2004a; 2004b; Crone et al., 2008); only few acknowledged the lack of differences between these regions (e.g. Rowe et al., 2008). The present findings also ruled out that the role of parietal cortex is restricted just to attention or the preparation of motor responses following abstract rule processing in PFC (Bunge, 2004). Parietal cortex might therefore function as an active information integration hub, as proposed recently (Singh-Curry & Husain, 2009). The second study adds to this view by demonstrating that decision outcomes of guessing in perceptual decision making were encoded in posterior parietal cortex, spanning from the medial part of the precuneus to the inferior parietal lobe. It was concluded that medial PPC is involved in symmetry-breaking and the generation of internal decisions in the absence of sufficient perceptual information. This is in line with recent findings demonstrating precuneus and PPC involvement in the generation of free decisions (Cui & Andersen, 2007; Pesaran et al., 2008; Soon et al., 2008). The third study supported and extended this view by showing that local activity patterns in PPC from intended free decisions could be used to predict guessing decisions and vice versa. Decoding from medial PFC, however, was limited to intended free decisions. Taken together, it was possible to decode different kinds of internal decisions from a network of mostly medial posterior parietal regions, which were hypothesised to be important for symmetry-breaking. A highly self-referential decision context could be the linking element between the PPC and internal decisions, given the involvement of PPC in reward- and value-based decisions (Platt & Glimcher, 1999; Platt & Huettel, 2008; Vickery & Jiang, 2009). It is therefore believed that the present studies provided strong new evidence for the importance of posterior parietal regions in high-level cognitive processes related to internal decision making.

How could PPC (including precuneus) guide the formation of internal decisions? It was already speculated that in the absence of values and reward to guide internal decision making, noise fluctuations in a part of one's internal value system, located in PPC, could produce random decision outcomes. As mentioned before, a recent study used electrical cortical stimulation in human patients and demonstrated that it was possible to evoke the

desire or “will” to move different body parts - or even the strong illusion that a movement had actually been carried out - by stimulating posterior parietal cortex (Desmurget et al., 2009). The authors concluded that motor intentions and the awareness of one’s own intentions could arise from activity in parietal cortex. This activity might be used to predict and anticipate the consequences of intended actions beforehand, being the basis for the perception of subsequent movements and the self-ascription of authorship (Desmurget et al., 2009; for a related framework see also Wegner, 2003). In a similar manner, the anticipation (and understanding) of *others’* actions has been suggested to be based on re-enactment of those actions in one’s own neural system (Prinz, 2006b). The relevance of Desmurget’s study for the interpretation of the present work might be criticised since it only demonstrated the induction of motor intentions and not decisions about object categories or other abstract decisions. It is, however, unlikely that an arbitrary abstract intention or decision could be elicited by cortical stimulation without additionally providing meaningful *context*. It would be reasonable to assume that cortical maps for motor intentions in parietal cortex (Andersen & Buneo, 2002) also form the basis for more abstract intentions, which, in turn, could be the basis for other internal decisions. These maps might be flexibly re-codable depending on the context, similar to what was suggested for PFC (Duncan, 2001). This is likely since, first, most abstract decisions are usually closely linked to motor action; and second, it has been proposed that evolutionary newer functions coming with cultural development could make use of already existing neural structures evolved for more basic but similar functions (Dehaene & Cohen, 2007). This explanation would prevent the unlikely assumption that single neurons (or EINs) are hardwired for all possible abstract intentions and decisions. Importantly, this idea is also supported by recent findings pointing towards the same neural substrate for simple motor intentions and abstract intentions in prefrontal cortex and precuneus (Soon et al., in prep.).

In conclusion, the presented studies have illustrated that parietal cortex plays a far more active and important role in human decision making and cognitive control than assumed so far. This was achieved by providing direct evidence for information encoding in parietal cortex. At the same time, recent views stressed capacity limitations of high-level prefrontal areas for cognitive control (Koechlin & Hyafil, 2007). The interaction between PFC and PPC in controlling behavior (Quintana & Fuster, 1999) might therefore be less dominated by PFC than previously thought. While earlier work on PPC predominantly focussed on its role in attention and visual processing (e.g. Milner & Goodale, 1993; Rizzolatti & Matelli, 2003; Ungerleider & Mishkin, 1982), recent research (mostly from

studies with monkeys) revealed its role in a variety of high-level cognitive functions such as the integration of sensory and cognitive information and motor actions (Gottlieb, 2007), the formation of highly abstract intentions and movement plans (Andersen & Buneo, 2002; Quiroga et al., 2006), as well as keeping the cognitive system goal-focussed and in a task-engaged state (Singh-Curry & Hussain, 2009). Of course, the present work does not question the well established findings of the tremendous importance of PFC in cognitive control and executive functions (e.g. Badre & D’Esposito, 2009; Duncan, 2001; Goldman-Rakic, 1996; Miller & Cohen, 2001; Passingham, 1995). However, it hopefully inspires future research focused on the fronto-parietal *network* and the fast interplay of processes within this system (Pesaran et al., 2008). The application of information decoding approaches for human fMRI data could serve as a key tool for the dissociation of different aspects and functional specialisation within this network. Specifically, it could be useful when the correspondence between monkey and human sub-regions is not clear (Orban et al., 2006; Singh-Curry & Hussain, 2009), when the cognitive functions under investigation are uniquely human, or when detailed verbal reports about perception or decision certainty are required. As demonstrated here, multivariate pattern classification can be used to go beyond classical fMRI studies and reveal a direct link between the outcome of high level cognitive processes and the underlying neural substrate. This might show important contributions of brain areas that would have been overlooked otherwise. On the other hand, it is important to be aware of some limitations of this approach that will be discussed in the next paragraph.

6.2 General methodological considerations

The studies presented in this work demonstrated that information about higher cognitive processes such as rules and decisions can be decoded from brain activity using a combination of fMRI and multivariate pattern classification (Haynes & Rees, 2006; Norman et al., 2006). While the spatial resolution of fMRI is still far from the scale of single neurons or EINs, this method breaks through the conventional limits imposed by voxel size (Haynes & Rees, 2005a; Kamitani & Tong, 2005; Kamitani & Sawahata, 2010; Swisher et al., 2010). Nevertheless, it still can only reveal what was called “a lower bound of information” (Haynes, 2009); it must be considered that *more* information might be encoded in brain regions that cannot be revealed with the current methods. An illustration of this limitation is

when too few neurons (or local EINs) coded for single properties, such that the signals from these voxels were dominated by noise. Too little is known about the proportion of information-carrying and information-irrelevant neurons that is sufficient for the build-up of stable detectable biases within a voxel, given the restrictions on measurement repetitions in a typical fMRI experiment. Another problem arises if a population of neurons flexibly changed their coding with changes of the task context, as suggested for prefrontal cortex (Duncan, 2001). For example, it has been demonstrated that task rules were encoded by the same neural populations in PFC, but in orthogonal patterns during different phases of task preparation (Sigala et al., 2008). Averaging *across* activation patterns from different task phases – wrongly assuming pattern stability – would make it impossible to decode anything. These problems could potentially also arise for parietal cortex in which neurons are tuned for several properties in a wide range of sub-regions (e.g. Quintana & Fuster, 1999).

While pattern classification algorithms are usually trained on the individual activation patterns of non-normalised images, they are not susceptible to differences in micro-level brain anatomy between participants. Individual spatial variation of informative clusters in the range of several centimetres, however, could be problematic. These individual “hot spots” could get lost due to averaging in group-level analyses. A potential solution (as performed in the present studies) is to additionally smooth the resulting decoding accuracy maps, similar to conventional univariate analyses that face the same localisation problems (Brett et al, 2002). So far, searchlight decoding still lacks a comprehensive systematic methodological assessment (for a first approach see Kriegeskorte et al., 2006) including effects of different smoothing kernels, searchlight radii and potential decoupling of information and activation, as observed in the present studies and others (Harrison & Tong, 2009; Serences et al., 2009; Soon et al., 2008). It would also be useful to address whether *individual* decoding accuracy maps (of prefrontal and parietal regions) might reveal distinguishable *sub-types* of information-to-anatomy mapping on a macro-anatomical level.

The usability of pattern classification might also depend on the functional organisation of the brain region under investigation. Such a region-specific lack of power, however, could also be problem in conventional univariate analysis and should be less problematic in multivariate analyses because many voxels contribute to the activation patterns. Wherever multivariate pattern classification is able to successfully decode *distinct* mental states, it can be concluded that information was indeed encoded in this region. Of course, this does not imply that the brain *uses* this information in the same way as it was read out by the rather

crude classification algorithm (Haynes & Rees, 2005a), nor does it suggest a higher-level ‘interpretation area’ elsewhere in the brain, as pointed out before (Haynes, 2009).

Decoding information from a brain region is still no evidence that the respective region plays a *causal* role in the underlying process; the nature of the relationship between activity patterns and mental states is – however tempting it is to assume otherwise – still a correlative one. In principle, information encoding could be an epi-phenomenon of a non-detected process elsewhere in the brain. Nevertheless, local activation during decision making and decoding of decision outcomes differ in terms of the *quality* of the revealed information. Stronger activation could more easily be explained by general preparatory activity, unrelated to the formation of the specific decision outcome. The possibility of spill-over effects of activation between highly interconnected regions makes it hard to pinpoint the role of single areas in cognitive processes, even if cortical excitability is probed directly (Bode, Koenecke & Jäncke, 2007; Eisenegger, Herwig & Jäncke, 2005). Demonstrating *information encoding* in a brain region, however, is more likely to point to a causal involvement in the cognitive process. Of course, this argument does not exclude the presence of unused information. On the other hand, the latter assumption needs an even greater amount of explanation compared to the causal interpretation. Information decoding is therefore *conceptually* closer to the underlying function than detecting activation differences with univariate methods. However, an experiment should always be referable back to psychological models instead of just demonstrating encoding. This can be done by demonstrating the behavioural relevance of information (Walther et al., 2009; Williams et al., 2008), the generalisation of information (Eger et al., 2009; Harrison & Tong, 2009: Experimental section III) or temporal or functional double-dissociations (Experimental Section I, II, III).

6.3 Future directions

The present studies are believed to provide new insights into the mechanisms of task preparation and decision making. They also provide starting points for further research by raising new questions. The first study revealed *lateral* regions in prefrontal and parietal cortex encoding task rules in cued task conditions, whereas the last study revealed *medial* PFC and predominantly *medial* parietal regions encoding participants’ internal decisions.

The existence of a general lateral-medial gradient for increased degrees of freedom in decision making could be hypothesised, as suggested for PFC (Forstmann et al., 2005; Haynes et al., 2007; Kounieher et al., 2009). One direct test for this assumption is to extend the paradigm used for the first study to free decisions. A *task cue* could either indicate a cued decision trial or a free decision trial. In cued decision trials, a second *rule cue* would indicate the rule. In free decision trials, participants would freely choose the rule to apply. After each trial, an additional response screen could explicitly ask for the rule used to avoid random responses. Additional *switch cues* could instruct the participants to immediately switch from the rule in mind (cued or freely chosen) to the alternative rule during some trials. Pattern classification could then be applied to test the following hypothesis: If an instructed switch occurred from a *cued* task rule to the alternative rule, it should be possible to decode both rules from the *same* regions in prefrontal and parietal cortex, which were found in the first study. If the switch occurred from a *freely* chosen rule, however, it could be assumed that the instructed switch transforms an initially free choice into a forced choice and the switch cue would have the same power as a rule cue. This should be accompanied by a shift of task rule information from regions encoding free decision outcomes (medial PFC / medial PPC) to regions encoding cued choices (lateral PFC / IPS). Taken together, the proposed experiment would elegantly close the loop between the first and the third study presented here.

Another question resulting from experiments two and three is whether encoding of choices in the LOC and the precuneus / PPC depends on trial-by-trial *variations in stimulus detection* or *decision certainty*. This could be directly tested by establishing a fixed medium level of visibility, as done by Williams and colleagues (2007). Unlike this study, however, it would be of interest to decode the *decision outcomes* as in the studies presented here. One hypothesis, as predicted by SDT (Swets, 1961), would be that LOC encodes the decision outcomes regardless of decision correctness, given that it receives the same (and sufficient) amount of visual information on each trial. Fluctuations in choice should then be due to trial-by-trial fluctuations in the underlying sensory signal. An alternative hypothesis would be that a constant switch occurs between information encoding in LOC and precuneus / PPC, depending on the *subjective decision certainty* (as opposed to its correctness), which could involve symmetry-breaking. This could also be investigated using e.g. magnetoencephalography (MEG) or EEG, which have a better temporal resolution than fMRI. Additionally, the strength of functional coupling between LOC and precuneus could be addressed under different conditions of correctness and certainty (Haynes et al., 2005). It

would also be worth investigating the possibility of a *transition* of information encoding from one region to the other by using several levels of visibility (e.g. Grill-Spector et al., 2000). To date, however, it is not clear how *relative* values of decoding accuracy (reflecting information) could be satisfactorily compared. The interpretation of significant differences in decoding accuracies *between* regions is not valid because the underlying BOLD signal can differ for various reasons (Logothetis, 2008). Interpreting accuracy differences *within* a region between tasks might also be problematic, because of imperfect reliabilities of the measurement tools (for imaging and for pattern classification). Also, to interpret a significant drop of information encoding (e.g. as 20% less information) within a region would require the unproven assumption of a linear relationship between decoding accuracy and neural information encoding. Local information encoding above a certain threshold might be sufficient for successful task performance, rendering differences above the threshold meaningless. At the moment, in most cases the safest interpretation is to regard information decoding as a dichotomy: information is either present or not.

Finally, it would be worth testing if the involvement of a symmetry-breaking network generalises to other task contexts. A perceptual decision making task could be used for other modalities. Sound discrimination (Binder, Liebenthal, Possing, Medler & Ward, 2004) or flutter discrimination (Romo & Salinas, 2003), however, also operate with prolonged temporal integration due to long stimulus exposure, rendering these approaches problematic for the investigation of the symmetry-breaking network. One radically different approach would be to investigate symmetry situations in logical deductive decisions. For example, a task similar to the Raven standard progressive matrix test (Raven, 1996) could be implemented. Participants would be asked to select the abstract symbol that is logically the missing element in a matrix or an array. In many trials, however, none of the given alternative solutions would be correct such that participants would have to guess. So long as it is possible to keep the choice alternatives the same such that they could be used for decoding, this paradigm could test whether the precuneus still encodes guessed decisions, which would then point to a *general* symmetry-breaking mechanism. Again, *believing* that one did not find the correct solution might moderate the involvement of the proposed mechanism.

These examples illustrate ways to directly test and extend the interpretations drawn from the present studies. These approaches could be combined with enhanced methods such as optimised classification algorithms, surface-based decoding techniques (Chen et al., in prep.), high-resolution fMRI (Kriegeskorte & Bandettini, 2007a; Swisher et al., 2010) or ultra high-field fMRI (Bode et al., in prep.; Swisher et al., 2010), which might allow the

investigation of finer-scaled activation patterns. The present studies are believed to provide an exciting basis for even more sophisticated investigations of human decision making in the future.

6.4 Concluding summary

Every day, humans are constantly required to react on stimuli in their dynamically changing environment. Appropriate behaviour therefore requires the ability to flexibly adapt to environmental cues, such as road signs or traffic lights. Different contexts might require different reactions to the same cues; a competence that has been conceptualised as task switching. An even more basic precondition is the ability to successfully detect sensory stimuli in a noisy environment, i.e., perceptual decision making. Only if, for instance, a stop sign can be identified as such in a hail storm, can the car driver act according to the associated rule. In a series of functional magnetic resonance imaging (fMRI) experiments, this present work investigated the neural encoding of these different components along the chain from sensory stimuli to behavioural responses.

The first experiment tracked the information flow from the encoding of task rules (or *task-sets*) to stimuli and to motor responses using a task switching paradigm. Each trial, participants were presented with one of two rule cues. These cues defined the joystick movement that had to be performed in response to one of two distinct visual target stimuli. Multivariate pattern classification (Haynes & Rees, 2006; Norman et al., 2006) was applied to the fMRI data in order to search for brain regions that encoded the *temporal sequence* of information about rules, target stimuli and motor responses. It was found that first, preceding the presentation of the target stimulus, the left intraparietal sulcus (IPS) encoded the abstract rule. Subsequently, a build-up of rule information was observed in left posterior ventrolateral prefrontal cortex (VLPFC), coinciding with the presentation of the target stimulus. The target stimuli were encoded in visual cortex. Motor responses could be decoded from supplementary and primary motor regions as soon as they could be prepared. Around the time of the motor response, the rules and target stimuli were encoded in left anterior VLPFC. A comparison with fixed task rules confirmed that information about stimuli and rules was encoded in prefrontal and parietal cortex only if a flexible switch between task rules was required. This study emphasised the importance of parietal cortex in

establishing abstract task-sets during early stages of rule-guided task preparation (Stoet & Snyder, 2004; Gail & Andersen, 2006). This function has often been linked to prefrontal cortex alone (Bunge, 2004). In the present work, however, encoding in parietal cortex even preceded encoding in prefrontal cortex.

The second project investigated whether networks for perceptual decision making depend on stimulus visibility. Participants categorised images of masked objects as belonging to one of three categories. Multivariate pattern classification of spatial patterns of fMRI signals was used to decode the participants' category choices, separately under high and low visibility conditions. From signals in the lateral-occipital complex (LOC), it was possible to decode participants' category choices about highly visible objects but not about poorly visible ones. This finding could be replicated using decoding from individual object-selective regions of interest (ROIs) derived from independent localizer scans. When participants were trying to classify poorly visible objects, however, the signals encoding the choices shifted to the precuneus / posterior parietal cortex (PPC). These findings suggest that participants' choices were based on different neural population signals for low and high visibility conditions, challenging current signal detection theory models of perceptual decision making (Swets, 1961). These predict that decisions about perceptual entities should be based on the same dimension of sensory signal for all levels of visibility. If stimuli cannot be discriminated, trial-by-trial noise fluctuations in the same system should determine the choices. Contrarily, the present findings suggest that the precuneus / PPC might be involved in "symmetry-breaking" when no alternative is more supported by sensory evidence. These regions could act as the brain's generator of internal random decisions in the absence of useful perceptual information, comparable to the generation of free decisions (Soon et al., 2008).

Finally, the last project directly tested whether decisions made by guessing (in perceptual decision making with insufficient visual input) resemble intended free decisions about the same object categories, using the same visual stimulation. Participants had to either guess the presented category or freely select a category, independent from the presentation. In order to avoid biasing participants' decisions, the target stimuli did not contain meaningful object images, yet nonetheless created the illusion that there was an object stimulus presented, albeit difficult to detect. Pattern classification was used to search for regions in which activation patterns were interchangeable between the two decision conditions. It could be shown that a network in precuneus / PPC encoded outcomes of guessing and to some extent free decisions. The activation patterns for both kinds of decisions were

demonstrated to be interchangeable in this region. The precuneus therefore appears to be generally involved in the generation of random internal decisions. Anterior medial prefrontal cortex, however, only encoded decision outcomes when decisions were *intended* to be free and might therefore be exclusively important for free decisions made without an external frame of reference.

Taken together, the present work highlights the importance of parietal cortex in controlling both rule-guided (IPS) and self-determined (precuneus / PPC) behavior in humans. Parietal cortex functions might be best described within a framework that acknowledges its capacity for multi-modal information integration and processing of intentions with high self-reference. The present findings also confirmed the power of multivariate pattern classification for decoding high-level mental states from brain activity. By directly decoding the *content* of rules and decisions, it is believed that these studies revealed a closer link between cognitive mechanisms and brain functions than has been shown in humans before. It is hoped that the present work thereby prepared the ground for a deeper understanding of flexible human behavior in the future.

6.5 Zusammenfassung

Wir leben in einer dynamischen, sich ständig verändernden Welt. Zu jeder Zeit die richtigen Entscheidungen zu treffen und Handlungen auszuwählen, hängt maßgeblich von der Fähigkeit ab, flexibel auf Hinweisreize (wie z.B. Straßenschilder oder Ampeln) reagieren zu können. Verschiedene Situationen verlangen dabei oft unterschiedliche Reaktionen auf identische Reize. So erfordert das Überqueren der Straße in Deutschland zuerst die Beachtung des Verkehrs von links, in Australien hingegen Aufmerksamkeit für Verkehr von rechts. Situationsabhängiges flexibles Reagieren auf identische Reize kann im Rahmen des Aufgaben-Wechsel Paradigmas (engl. *task-switching*) konzeptualisiert und untersucht werden. Noch grundlegender ist die Fähigkeit, einen Hinweisreiz unter sub-optimalen Bedingungen überhaupt als solchen zu erkennen. So erfordert beispielsweise die Wahrnehmung eines Stop-Schildes während eines schweren Hagelsturms zuerst eine *perzeptuelle Entscheidung*. In der vorliegenden Arbeit wurden in einer Serie von Experimenten die neuronalen Grundlagen der Verarbeitung dieser Komponenten, von

visuellen Reizen bis hin zu adäquaten Reaktionen, mittels funktioneller Magnetresonanztomographie (fMRT) untersucht.

Das erste Experiment bediente sich des Aufgaben-Wechsel Paradigmas, um den Informationsfluss, beginnend mit der neuronalen Enkodierung abstrakter Regeln, über visuelle Reize, bis hin zur Auswahl einer motorischen Reaktion, zu untersuchen. In jedem Durchgang wurde den Teilnehmern eine von zwei Regeln (*task-sets*; Sakai, 2008) vorgegeben, welche die Zuordnung zwischen zwei möglichen visuellen Reizen und zwei motorischen Antworten festlegte. Im Folgenden wurde einer der visuellen Reize präsentiert, auf den gemäß der aktiven Regel reagiert werden sollte. Mittels Anwendung von multivariaten Mustererkennungsverfahren auf die fMRT Daten (Haynes & Rees, 2005; Norman et al., 2005) wurde nach Hirnregionen gesucht, welche die *zeitliche Abfolge der Informationen* über Regeln, visuelle Stimulation und motorische Antwort enkodierten. Der intraparietale Sulcus (IPS) enkodierte dabei die abstrakte Regel noch bevor der relevante visuelle Stimulus präsentiert wurde. Mit dessen Präsentation konnte im Folgenden ein Zuwachs an Regel-Information im linken posterioren ventrolateral Präfrontalkortex (VLPFC) beobachtet werden. Die visuellen Stimulationsreize selbst waren im visuellen Kortex enkodiert. Die motorischen Antworten waren, sobald sie geplant werden konnten, aus dem supplementären und primären motorischen Kortex dekodierbar. Während der motorischen Antwort schließlich konnten die visuellen Reize und abstrakten Regeln aus dem linken anterioren VLPFC ausgelesen werden. Ein Vergleich mit festen Zuordnungen von Reizen und Antworten bestätigte, dass Informationen über visuelle Reize und Regeln nur dann im präfrontalen und parietalen Kortex enkodiert waren, wenn flexibel zwischen Regeln gewechselt werden musste. Diese Studie demonstriert vor allem die Relevanz des parietalen Kortex für die Verarbeitung abstrakter Regelinformationen, welche bisher oft ausschließlich dem präfrontalen Kortex zugeschrieben wurde (Bunge, 2004). Diese Funktionen beziehen jedoch den parietalen Kortex mit ein (Stoet & Snyder, 2004; Gail & Andersen, 2006), dessen Beteiligung jener des präfrontalen Kortex sogar vorausgehen kann.

Das zweite fMRT Projekt untersuchte die Abhängigkeit neuronaler Netzwerke für perzeptuelle Entscheidungen von der Sichtbarkeit visueller Reize. Dafür wurden Probanden Bilder von stark oder schwach maskierten Objekten dargeboten, die bezüglich ihrer Kategoriezugehörigkeit zu klassifizieren waren. Erneut wurden Mustererkennungsverfahren benutzt, um aus lokalen Aktivitätsmustern im Gehirn für beide Sichtbarkeitsbedingungen einzeln vorherzusagen, für welche Kategorie sich die Probanden entschieden. Signale aus dem ventralen temporalen und okzipitalen Kortex (*lateral-occipital complex, LOC*)

erlaubten dabei die Vorhersage von Entscheidungen für gut sichtbare Objekte, aber nicht für stark maskierte Objekte. Es war jedoch möglich, die kategoriellen Entscheidungen für stark maskierte Objekte aus dem Precuneus im medialen dorsalen Parietalkortex vorherzusagen. Diese doppelte Dissoziation legt nahe, dass Entscheidungen unter hoher und niedriger Sichtbarkeit auf verschiedenen Mechanismen (Netzwerken) beruhen. Dieser Befund widerspricht Annahmen der Signal-Entdeckungs-Theorie (Swets, 1961), die vorhersagt, dass auch unter niedriger Sichtbarkeit Signale in den gleichen sensorischen Arealen für perzeptuelle Entscheidungen ausschlaggebend sein sollten. In diesem Falle sollten zufällige Fluktuationen („noise“) im perzeptuellen System statt echtem sensorischen Input eine zufällige Entscheidung produzieren. Die Befunde der vorliegenden Studie sprechen jedoch dafür, dass im Falle des Vorliegens unzureichender sensorischer Information (sensorische „Symmetrie“ bezüglich der relevanten Alternativen) der Precuneus in die Generierung einer zufälligen internalen Entscheidung involviert ist. Diese Interpretation wird durch eine kürzlich publizierte Studie gestützt, die zeigen konnte, dass der Precuneus auch an der Bildung freier zufälliger Entscheidungen beteiligt ist (Soon et al., 2008).

Basierend auf den vorausgegangenen Befunden, testete das dritte fMRT Experiment direkt die Vergleichbarkeit der neuronalen Mechanismen von Rate-Entscheidungen, die bei unzureichendem visuellem Input getroffen werden, mit geplant freien Entscheidungen (Haggard, 2008). Hierzu wurden Probanden in einem Teil der Durchgänge gebeten, erneut eine perzeptuelle Entscheidung über die Kategorie von stark maskierten Bildern zu treffen. Alternativ sollte in anderen Durchgängen bei gleicher visueller Stimulation eine spontane, freie Entscheidung für eine der Kategorien getroffen werden. Um jegliche Beeinflussung der Entscheidungen durch das verwendete Stimulusmaterial zu vermeiden, wurde in beiden Fällen kein echtes Objekt präsentiert, was durch die kurze Darbietungszeit und starke Maskierung jedoch nicht wahrnehmbar war. Erneut wurden Mustererkennungsverfahren auf die fMRT-Daten angewendet und Hirnregionen identifiziert, deren Muster regionaler Signale die Vorhersage der Entscheidungen erlaubten. Insbesondere war von Interesse, nach Regionen zu suchen, in denen die Aktivierungsmuster der Rate-Entscheidungen und der freien Entscheidungen eine so große Ähnlichkeit aufwiesen, dass sie austauschbar waren. Es konnte erneut gezeigt werden, dass der Precuneus, bis hin zum lateralen posterioren Parietalkortex, Rate-Entscheidungen enkodierte, wenn diese unter Abwesenheit nützlicher visueller Information getroffen werden mussten. Zusätzlich war diese Region auch begrenzt prädiktiv für freie Entscheidungen und zeigte außerdem die vermutete Austauschbarkeit neuronaler Muster für Rate-Entscheidungen und freie Entscheidungen. Aus

Aktivierungsmustern des medialen anterioren Präfrontalkortex hingegen, konnten wie in früheren Experimenten (Haynes et al., 2007; Soon et al., 2008), nur geplant freie Entscheidungen vorhergesagt werden.

Die vorliegende Arbeit demonstriert die große Bedeutung parietaler Areale, sowohl für die Kontrolle regelgeleiteter Handlungen (IPS) als auch für selbstbestimmte Handlungen (Precuneus / posteriorer Parietalkortex). Der parietale Kortex kann somit als multi-modales Integrations-Modul verstanden werden, das an der Auswahl von Handlungszielen und dessen Aufrechterhaltung, sowie an der Verarbeitung von Intentionen und Entscheidungen mit hohem Selbstbezug beteiligt ist. Vom methodischen Standpunkt betrachtet, bestätigt die vorliegende Arbeit den großen Nutzen multivariater Mustererkennungsverfahren für die Analyse von fMRT Daten, auch für höhere kognitive Prozesse. In den hier vorgestellten Studien konnte die Möglichkeit demonstriert werden, aus lokalen Aktivierungsmustern direkt vorherzusagen, welche abstrakte Regel ein Proband gerade anwendet oder welche Entscheidung er trifft. Diese Methodik erlaubt es, einen stärkeren Bezug zwischen menschlichen kognitiven Funktionen und den ihnen zugrunde liegenden neuronalen Netzwerken herzustellen als es mit nicht-invasiven bildgebenden Verfahren bislang möglich war. Damit verknüpft ist die Hoffnung, dass die vorliegende Arbeit auch eine Grundlage für weitere Forschung und ein tieferes Verständnis der neuronalen Basis von flexiblen menschlichen Entscheidungsprozessen für die Zukunft darstellen kann.

References

- Addis, D.R., McIntosh, A.R., Moscovitch, M., Crawley, A.P. & McAndrews, M.P. (2004). Characterizing spatial and temporal features of autobiographical memory retrieval networks: a partial least squares approach. *Neuroimage*, 23(4), 1460-1471.
- Aguire, G.K. (2006). Experimental design and data analysis for fMRI. In S.H. Faro & F.B. Mohamed (Eds.), *Functional MRI. Basic principles and clinical applications* (pp. 58-74). New York, NY, USA: Springer.
- Aguirre, G.K., Zarahn, E. & D'Esposito, M. (1998). An area within human ventral cortex sensitive to 'building' stimuli: evidence and implications. *Neuron*, 21(2), 373-383.
- Allport, A., Styles, E.A. & Hsieh, S. (1994). Shifting intentional set: exploring the dynamic control of tasks. In: C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421-452). Cambridge, MA, USA: MIT Press.
- Allport, A. & Wylie, G. (1999). Task switching: positive and negative priming of task set. In G.W. Humphreys, J. Duncan & A. Treisman (Eds.), *Attention, space, and action: studies in cognitive neuroscience* (pp.273-296). Oxford, UK: Oxford University Press.
- Amemori, K. & Sawaguchi, T. (2006). Rule-dependent shifting of sensorimotor representation in the primate prefrontal cortex. *European Journal of Neuroscience*, 23(7), 1895-1909.
- Amodio, D.M. & Frith, C.D. (2006). Meeting of minds: the medial prefrontal cortex and social cognition. *Nature Reviews Neuroscience*, 7(4), 268-277.
- an der Heiden, U. & Schneider, H. (2007). *Hat der Mensch einen freien Willen? Die Antworten der großen Philosophen*. Ditzingen, Germany: Reclam.
- Andersen, R.A. & Buneo, C.A. (2002). Intentional maps in posterior parietal cortex. *Annual Review of Neuroscience*, 25, 189-220.
- Asaad, W.F., Rainer, G. & Miller, E.K. (1998). Neural activity in the primate prefrontal cortex during associative learning. *Neuron*, 21(6), 1399-1407.
- Asaad, W.F., Rainer, G. & Miller, E.K. (2000). Task-specific neural activity in the primate prefrontal cortex. *Journal of Neurophysiology*, 84(1), 451-459.
- Asari, T., Konishi, S., Jimura, K. & Miyashita, Y. (2005). Multiple components of lateral posterior parietal activation associated with cognitive set shifting. *Neuroimage*, 26(3), 694-702.
- Averbeck, B.B., Sohn, J.W. & Lee, D. (2006). Activity in prefrontal cortex during dynamic selection of action sequences. *Nature Neuroscience*, 9(2), 276-282.
- Bachmann, T. (2006). Microgenesis of perception: Conceptual, psychophysical and neurobiological aspects. In: B.G. Breitmeyer & H. Ögmen (Eds.), *The first half second. The microgenesis and temporal dynamics of unconscious and conscious visual processes*. Cambridge, USA: MIT Press.

- Badre, D. & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience*, *19*(12), 2082-2099.
- Badre, D. & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews Neuroscience*, *10*(9), 659-669.
- Badre, D., Hoffman, J., Cooney, J.W. & D'Esposito, M. (2009). Hierarchical cognitive control deficits following damage to the human frontal lobe. *Nature Neuroscience*, *12*(4), 515-522.
- Badre, D. & Wagner, A.D. (2005). Frontal lobe mechanisms that resolve proactive interference. *Cerebral Cortex*, *15*(12), 2003-2012.
- Badre, D. & Wagner, A.D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, *45*(13), 2883-2901.
- Ball, T., Schreiber, A., Feige, B., Wagner, M., Lücking, C.H. & Kristeva-Feige, R. (1999). The role of higher-order motor areas in voluntary movement as revealed by high-resolution EEG and fMRI. *Neuroimage*, *10*(6), 682-694.
- Bandettini, P.A., Wong, E.C., Hinks, R.S., Tikofski, R.S. & Hyde, J.S. (1992). Time course EPI of human brain function during task activation. *Magnetic Resonance in Medicine*, *25*(2), 390-397.
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*(8), 617-629.
- Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmid, A.M., Dale, A.M., Hämäläinen, M.S., Marinkovic, K., Schacter, D.L., Rosen, B.R. & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 449-454.
- Blakemore, S.J., den Ouden, H., Choudhury, S. & Frith, C. (2007). Adolescent development of the neural circuitry for thinking about intentions. *Social Cognitive and Affective Neuroscience*, *2*(2), 130-139.
- Blackmore, S. (2006). *Conversations on consciousness: what the best minds think about the brain, free will, and what it means to be human*. New York, NY, USA: Oxford University Press.
- Binder, J.R., Liebenthal, E., Possing, E.T., Medler, D.A. & Ward, B.D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nature Neuroscience*, *7*(3), 295-301.
- Bode, S. & Haynes, J.D. (2009). Decoding sequential stages of task preparation in the human brain. *Neuroimage*, *45*(2), 606-613.
- Bode, S., Bogler, C., Soon, C.S. & Haynes, J.D. (under review). Similar neural mechanisms for guesses and free choices.
- Bode, S., He, A.H., Soon, C.S., Trampel, R., Turner, R. & Haynes, J.D. (in prep.). Tracking the unconscious generation of free decisions using ultra-high field fMRI.
- Bode, S., Koenecke, S. & Jäncke, L. (2007). Different strategies do not moderate primary motor cortex involvement in mental rotation: a TMS study. *Behavioral and Brain Functions*, *3*:38.
- Bogler, C., Bode, S. & Haynes, J.D. (in prep.). Multivariate decoding reveals successive computational stages of saliency processing.
- Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S. & Cohen, J.D. Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624-652.
- Bowers, J.S. & Jones, K.W. (2008). Detecting objects is easier than categorizing them. *The Quarterly Journal of Experimental Psychology*, *61*(4), 552-557.
- Bradley, D.C., Chang, G.C. & Andersen, R.A. (1998). Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature*, *392*(6677), 714-717.

- Brass, M., Derrfuss, J., Forstmann, B. & von Cramon, D.Y. (2005a). The role of the inferior frontal junction area in cognitive control. *Trends in Cognitive Sciences*, 9(7), 314-316.
- Brass, M. & Haggard, P. (2007). To do or not to do: the neural signature of self-control. *Journal of Neuroscience*, 27(34), 9141-9145.
- Brass, M., Ruge, H., Meiran, N., Rubin, O., Koch, I., Zysset, S., Prinz, W. & von Cramon, D.Y. (2003). When the same response has different meanings: recoding the response meaning in the lateral prefrontal cortex. *Neuroimage*, 20(2), 1026-1031.
- Brass, M., Ullsperger, M., Knoesche, T.R., von Cramon, D.Y. & Phillips, N.A. (2005b). Who comes first? The role of the prefrontal and parietal cortex in cognitive control. *Journal of Cognitive Neuroscience*, 17(9), 1367-1375.
- Brass, M. & von Cramon, D.Y. (2002). The role of the frontal cortex in task preparation. *Cerebral Cortex*, 12(9), 908-914.
- Brass, M. & von Cramon, D.Y. (2004a). Decomposing components of task preparation with functional magnetic resonance imaging. *Journal of Cognitive Neuroscience*, 16(4), 609-620.
- Brass, M. & von Cramon, D.Y. (2004b). Selection for cognitive control: a functional magnetic resonance imaging study on the selection of task-relevant information. *Journal of Neuroscience*, 24(40), 8847-8852.
- Braver, T.S., Reynolds, J.R. & Donaldson, D.I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron*, 39(4), 713-726.
- Breitmeyer, B.G. & Stoerig, P. (2007). Neural correlates and levels of conscious and unconscious vision. In: B.G. Breitmeyer & H. Ögmen (Eds.), *The first half second. The microgenesis and temporal dynamics of unconscious and conscious visual processes*. Cambridge, USA: MIT Press.
- Brett, M., Johnsrude, I.S. & Owen, A.M. (2002). The problem of functional localization in the human brain. *Nature Reviews Neuroscience*, 3(3), 243-249.
- Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S. & Movshon, J.A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience*, 13(1), 87-100.
- Britten, K.H., Shadlen, M.N., Newsome, W.T. & Movshon, J.A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, 12(12), 4745-4765.
- Britten, K.H. & van Wezel, R.J.A. (1998). Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nature Neuroscience*, 1(1), 59-63.
- Brouwer, G.J. & Heeger, G.J. (2009). Decoding and reconstructing color from responses in human visual cortex. *Journal of Neuroscience*, 29(44), 13992-14003.
- Bunge, S.A. (2004). How we use rules to select actions: a review of evidence from cognitive neuroscience. *Cognitive, Affective, & Behavioral Neuroscience*, 4(4), 564-679.
- Bunge, S.A., Hazeltine, E., Scanlon, M.D., Rosen, A.C. & Gabrieli, J.D. (2002). Dissociable contributions of prefrontal and parietal cortices to response selection. *Neuroimage*, 17(3), 1562-1571.
- Bunge, S.A., Kahn, I., Wallis, J.D., Miller, E.K. & Wagner, A.D. (2003). Neural circuits subserving the retrieval and maintenance of abstract rules. *Journal of Neurophysiology*, 90(5), 3419-3428.
- Bunge, S.A. & Wallis, J.D. (2007). *Neuroscience of rule-guided behavior*. New York, USA: Oxford University Press.
- Bunge, S.A., Wallis, J.D., Parker, A., Brass, M., Crone, E.A., Hoshi, E. & Sakai, K. (2005). Neural circuitry underlying rule use in humans and nonhuman primates. *Journal of Neuroscience*, 25(45), 10347-10350.
- Burgess, P.W., Quayle, A. & Frith, C.D. (2001). Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia*, 39(6), 545-555.

- Buschman, T.J. & Miller, E.K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, 315(5820), 1860-1862.
- Buxton, R.B. (2002). *Introduction to functional magnetic resonance imaging: Principles and techniques*. Cambridge: Cambridge University Press.
- Cavanna, A.E. & Trimble, M.R. (2006). The precuneus: a review of its functional anatomy and its behavioural correlates. *Brain*, 129(3), 564-583.
- Chen, Y., Namburi, P., Elliott, L.T., Heinzle, J., Soon, C.S., Chee, M.W. & Haynes, J.D. (in prep.). Cortical surface-based searchlight decoding.
- Chiang, T. (2005). What's expected of us. *Nature*, 436(7047), 150.
- Chiu, Y.C. & Yantis, S. (2009). A domain-independent source of cognitive control for task sets: shifting spatial attention and switching categorization rules. *Journal of Neuroscience*, 29(12), 3930-3938.
- Cisek, P., Puskas, P.A. & El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. *Journal of Neuroscience*, 29(37), 11560-11571.
- Coltheart, M. (2006). What has functional neuroimaging told us about the mind (so far)? *Cortex*, 42(3), 323-331.
- Constantinidis, C., Franowicz, M.N. & Goldman-Rakic, P.S. (2001a). The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nature Neuroscience*, 4(3), 311-316.
- Constantinidis, C., Franowicz, M.N. & Goldman-Rakic, P.S. (2001b). Coding specificity in cortical microcircuits: a multiple-electrode analysis of primate prefrontal cortex. *Journal of Neuroscience*, 21(10), 3646-3655.
- Constantinidis, C., Williams, G.V. & Goldman-Rakic, P.S. (2002). A role for inhibition in shaping the temporal flow of information in prefrontal cortex. *Nature Neuroscience*, 5(2), 175-180.
- Cox, D.D. & Savoy, R.L. (2003). Functional magnetic resonance imaging fMRI "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage*, 19(2 pt 1), 261-270.
- Crone, E.A., Wendelken, C., Donohue, S.E. & Bunge, S.A. (2006). Neural evidence for dissociable components of task-switching. *Cerebral Cortex*, 16(4), 475-486.
- Cui, H. & Andersen, R.A. (2007). Posterior parietal cortex encodes autonomously selected motor plans. *Neuron*, 56(3), 552-559.
- Cunnington, R., Windischberger C., Deecke, L. & Moser E. (2002). The preparation and execution of self-initiated and externally-triggered movement: a study of event-related fMRI. *Neuroimage*, 15(2), 373-385.
- Cunnington, R., Windischberger C. & Moser E. (2005). Premovement activity of the pre-supplementary motor area and the readiness for action: studies of time-resolved event-related functional MRI. *Human Movement Science*, 24(5-6), 644-656.
- Davatzikos, C., Ruparel, K., Fan, Y., Shen, D.G., Acharyya, M., Loughhead, J.W., Gur, R.C. & Langleben, D.D. (2005). Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *Neuroimage*, 28(3), 663-668.
- Deecke, L., Lang, W., Heller, H.J., Hufnagl, M. & Kornhuber, H.H. (1987). Bereitschaftspotential in patients with unilateral lesions in supplementary motor area. *Journal of Neurology, Neurosurgery, and Psychiatry*, 50(11), 1430-1434.
- Deco, G., Pérez-Sanagustín, M., de Lafuente, V. & Romo, R. (2007). Perceptual detection as a dynamical bistability phenomenon: a neurocomputational correlate of sensation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(50), 20073-20077.
- Deco, G. & Romo, R. (2008). The role of fluctuations in perception. *Trends in Neuroscience*, 31(11), 591-598.

- Dehaene, S. & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56(2), 384-398.
- Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P.F. & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, 359(6702), 597-600.
- Deiber, M.P., Honda, M., Ibañez, V., Sadato, N. & Hallett, M. (1999). Mesial motor areas in self-initiated versus externally triggered movements examined with fMRI: effect of movement type and rate. *Journal of Neurophysiology*, 81(6), 3065-3077.
- De Jong, R. (2000). An intention-activation account of residual switch costs. In: S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: control of cognitive processes* (pp. 357-376). Cambridge, MA, USA: MIT Press.
- de Lafuente, V. & Romo, R. (2005). Neuronal correlates of subjective sensory experience. *Nature Neuroscience*, 8(12), 1698-1703.
- den Ouden, H.E., Frith, U., Frith, C. & Blakemore, S.J. (2005). Thinking about intentions. *Neuroimage*, 28(4), 787-796.
- Dennett, D.C. (1991). *Consciousness explained*. Boston, MA, USA: Little, Brown & Co.
- Dennett, D.C. (2003). *Freedom evolves*. New York, NY, USA: Viking Press.
- Derrfuss, J., Brass, M., Neumann, J. & von Cramon, D.Y. (2005). Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. *Human Brain Mapping*, 25(1), 22-34.
- Derrfuss, J., Brass, M., & von Cramon, D.Y. (2004). Cognitive control in the posterior frontolateral cortex: evidence from common activations in task coordination, interference control, and working memory. *Neuroimage*, 23(2), 604-612.
- Desimone, R. & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193-222.
- Desmurget, M., Reilly, K.T., Richard, N., Szathmari, A., Mottolese, C. & Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science*, 324(5928), 811-813.
- Ditterich, J. (2006). Stochastic models of decisions about motion direction: behavior and physiology. *Neural Networks*, 19, 981-1012.
- Donohue, S.E., Wendelken, C., Crone, E.A. & Bunge, S.A. (2005). Retrieving rules for behavior from long-term memory. *Neuroimage*, 26(4), 1140-1149.
- Dosenbach, N.U., Visscher, K.M., Palmer, E.D., Miezin, F.M., Wenger, K.K., Kang, H.C., Burgund, E.D., Grimes, A.L., Schlaggar, B.L. & Petersen, S.E. (2006). A core system for the implementation of task sets. *Neuron*, 50(5), 799-812.
- Dove, A., Pollmann, S., Schubert, T., Wiggins, C.J. & von Cramon, D.Y. (2000). Prefrontal cortex activation in task switching: an event-related fMRI study. *Brain Research. Cognitive Brain Research*, 9(1), 103-109.
- Downing, P.E., Jiang, Y., Shuman, M. & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473.
- Downing, P.E., Wiggett, A.J. & Peelen, M.V. (2007). Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. *Journal of Neuroscience*, 27(1), 226-233.
- Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience*, 2(11), 820-829.
- Edelman, S., Grill-Spector, K., Kushnir, T. & Malach, R. (1998). Towards direct visualization of the internal shape representation space by fMRI. *Psychobiology*, 26(4), 309-321.

- Eger, E., Ashburner, J., Haynes, J.D., Dolan, R.J. & Rees, G. (2008). fMRI activity patterns in human LOC carry information about object exemplars within category. *Journal of Cognitive Neuroscience*, 20(2), 356-370.
- Eger, E., Michel, V., Thirion, B., Amadon, A., Dehaene, S. & Kleinschmidt, A. (2009). Deciphering cortical number coding from human brain activity patterns. *Current Biology*, 19(19), 1608-1615.
- Eisenegger, C., Herwig, U. & Jäncke, L. (2005). The involvement of primary motor cortex in mental rotation revealed by transcranial magnetic stimulation. *European Journal of Neuroscience*, 25(4), 1240-1244.
- Elliott, R., Rees, G. & Dolan, R.J. (1999). Ventromedial prefrontal cortex mediates guessing. *Neuropsychologia*, 37(4), 403-411.
- Epstein, R., Harris, A., Stanley, D. & Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding? *Neuron*, 23(1), 115-125.
- Epstein, R. & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676), 598-601.
- Fahrenfort, J.J., Scholte, H.S. & Lamme, V.A. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience*, 19(9), 1488-1497.
- Faisal, A.A., Selen, L.P. & Wolpert, D.M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4), 292-303.
- Fang, F. & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience*, 8(10), 1380-1385.
- Faro, S.H. & Mohamed, F.B. (2006). *Functional MRI. Basic principles and clinical applications*. New York, NY, USA: Springer.
- Fletcher, P.C., Frith, C.D., Baker, S.C., Shallice, T., Frackowiak, R.S. & Dolan, R.J. (1995). The mind's eye – precuneus activation in memory-related imagery. *Neuroimage*, 2(3), 195-200.
- Fletcher, P.C., Shallice, T., Frith, C.D., Frackowiak, R.S. & Dolan, R.J. (1996). Brain activity during memory retrieval. The influence of imagery and semantic cueing. *Brain*, 119(Pt 5), 1587-1596.
- Formisano E., De Martino F. & Valente, G. (2008). Multivariate analysis of fMRI time series: classification and regression of brain responses using machine learning. *Magnetic Resonance Imaging*, 26(7), 921-34.
- Formisano, E. & Goebel, R. (2003). Tracking cognitive processes with functional MRI mental chronometry. *Current Opinion in Neurobiology*, 13(2), 174-181.
- Forstmann, B.U., Brass, M., Koch, I. & von Cramon, D.Y. (2005). Internally generated and directly cued task sets: an investigation with fMRI. *Neuropsychologia*, 43(6), 943-952.
- Forstmann, B.U., Brass, M., Koch, I. & von Cramon, D.Y. (2006). Voluntary selection of task sets revealed by functional magnetic resonance imaging. *Journal of Cognitive Neuroscience*, 18(3), 388-398.
- Forstmann, B.U., Ridderinkhof, K.R., Kaiser, J. & Bledowski, C. (2007). At your own peril: an ERP study of voluntary task set selection processes in the medial frontal cortex. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4), 286-296.
- Freedman, D.J. & Assad, J.A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature*, 443(7107), 85-88.
- Friston, K.J., Ashburner, J., Kiebel, S., Nichols, T.E. & Penny, W.D. (2006). *Statistical Parametric Mapping: The analysis of functional brain images*. London, UK: Elsevier.
- Friston, K.J., Holmes, A.P., Poline, J.P., Grasby, P.J., Williams, S.C., Frackowiak, R.S. & Turner, R. (1995a). Analysis of fMRI time-series revisited. *Neuroimage*, 2(1), 45-53.

- Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D. & Frackowiak, R.S. (1995b). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2, 189-210.
- Funahashi, S., Chafee, M.V. & Goldman-Rakic, P.S. (1993). Prefrontal neural activity in rhesus monkeys performing a delayed anti-saccade task. *Nature*, 365(6448), 753-756.
- Gail, A. & Andersen, R.A. (2006). Neural dynamics in monkey parietal reach region reflect context-specific sensorimotor transformations. *Journal of Neuroscience*, 26(37), 9376-9384.
- Gauthier, I., Skudlarski, P., Gore, J.C. & Anderson, A.W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3(2), 191-197.
- Gauthier, I., Tarr, M.J., Anderson, A.W., Skudlarski, P. & Gore, J.C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience*, 2(6), 568-573.
- Glimcher, P.W. (2001). Making choices: The neurophysiology of visual-saccadic decision making. *Trends in Neuroscience*, 24(11), 654-659
- Gold, J.I. & Shadlen, M.N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature*, 404(6776), 390-394.
- Gold, J.I. & Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, 5(1), 10-16.
- Gold, J.I. & Shadlen, M.N. (2003). The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *Journal of Neuroscience*, 23(2), 632-651.
- Gold, J.I. & Shadlen, M.N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30, 535-374.
- Goldman-Rakic, P.S. (1996). The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 351(1346), 1445-1453.
- Gottlieb, J. (2007). From thought to action: the parietal cortex as a bridge between perception, action, and cognition. *Neuron*, 53(1), 9-16.
- Green, D.M & Swets, J.A. (1966). *Signal detection theory and psychophysics*. New York, NY, USA: John Wiley and Sons, Inc.
- Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology*, 13(2), 159-166.
- Grill-Spector, K. & Kanwisher, N. (2005). Visual recognition: as soon as you know it is there, you know what it is. *Psychological Science*, 16(2), 152-160.
- Grill-Spector, K., Kushnir, T., Hendler, T. & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience*, 3(8), 837-843.
- Grill-Spector, K. & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, 27, 649-677.
- Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nature Reviews Neuroscience*, 9(12), 934-946.
- Haggard, P. (2009). Neuroscience. The sources of human volition. *Science*, 324(5928), 731-733.
- Haggard, P. & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, 126(1), 128-133.
- Hampton, A.N. & O'Doherty, J.P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, 104(4), 1377-1382.

- Harrison, S.A. & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238), 632-635.
- Hassabis, D., Chu, C., Rees, G., Weiskopf, N., Molyneux, P.D. & Maguire, E.A. (2009). Decoding neuronal ensembles in the human hippocampus. *Current Biology*, 19(7), 546-554.
- Haushofer, J., Baker, C.I., Livingstone, M.S. & Kanwisher, N. (2008a). Privileged coding of convex shapes in human object-selective cortex. *Journal of Neurophysiology*, 100(2), 753-762.
- Haushofer, J., Livingstone, M.S. & Kanwisher, N. (2008b). Multivariate patterns in object-selective cortex dissociate perceptual and physical shape similarity. *PLoS Biology*, 6(7):e187.
- Haynes, J.D. (2009). Decoding visual consciousness from human brain signals. *Trends in Cognitive Sciences*, 13(5), 194-202.
- Haynes, J.D. (submitted). Beyond Libet: long-term prediction of free choices from neuroimaging signals.
- Haynes, J.D., Driver, J. & Rees, G. (2005). Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. *Neuron*, 46(5), 811-821.
- Haynes, J.D. & Rees, G. (2005a). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, 8(5), 686-691.
- Haynes, J.D. & Rees, G. (2005b). Predicting the stream of consciousness from activity in human visual cortex. *Current Biology*, 15(14), 1301-1307.
- Haynes, J.D. & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523-534.
- Haynes, J.D., Sakai, K., Rees, G., Gibert, S., Frith, C. & Passingham, R.E. (2007). Reading hidden intentions in the human brain. *Current Biology*, 17(4), 323-328.
- Hayworth, K.J. & Biederman, I. (2006). Neural evidence for intermediate representations in object recognition. *Vision Research*, 46(23), 4024-4031.
- Haxby, J.V., Hoffman, E.A. & Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223-233.
- Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L. & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293(5539), 2425-2430.
- Heeger, D.J. & Ress, D. (2002). What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience*, 3(2), 142-151.
- Heekeren, H.R., Marrett, S., Bandettini, P.A. & Ungerleider, L.G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature*, 431(7010), 859-862.
- Heekeren, H.R., Marrett, S. & Ungerleider, L.G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience*, 9(6), 467-479.
- Heisenberg, M. (2009). Is free will an illusion? *Nature*, 459(7244), 164-165.
- Hernández, A., Zainos, A. & Romo, R. (2000). Neural correlates of sensory discrimination in the somatosensory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 97(11), 6191-6196.
- Hernández, A., Zainos, A. & Romo, R. (2002). Temporal evolution of a decision-making process in medial premotor cortex. *Neuron*, 33(6), 959-972.
- Ho, T.C., Brown, S. & Serences, J.T. (2009). Domain general mechanisms of perceptual decision making in human cortex. *Journal of Neuroscience*, 29(27), 8675-8687.

- Hoshi, E., Shima, K. & Tanji, J. (1998). Task-dependent selectivity of movement-related neuronal activity in the primate prefrontal cortex. *Journal of Neurophysiology*, *80*(6), 3392-3397.
- Hoshi, E., Shima, K. & Tanji, J. (2000). Neuronal activity in the primate prefrontal cortex in the process of motor selection based on two behavioral rules. *Journal of Neurophysiology*, *83*(4), 2355-2373.
- Hoshi, E. & Tanji, J. (2004). Area-selective neuronal activity in the dorsolateral prefrontal cortex for information retrieval and action planning. *Journal of Neurophysiology*, *91*(6), 2707-2722.
- Howard, J.D., Plailly, J., Grueschow, M., Haynes, J.D. & Gottfried, J.A. (2009). Odor quality coding and categorization in human posterior piriform cortex. *Nature Neuroscience*, *12*(7), 932-938.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D. & Camerer, C.F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, *310*(1680), 1680-1683.
- Huettel, S.A. & McCarthy, G. (2001). Regional differences in the refractory period of the hemodynamic response: an event-related fMRI study. *Neuroimage*, *14*(5), 967-976.
- Huettel, S.A., Song, A.W. & McCarthy, G. (2005). Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *Journal of Neuroscience*, *25*(13), 3304-3311.
- Huk, A.C. & Shadlen, M.N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *Journal of Neuroscience*, *25*(45), 10420-10436.
- Humphreys, G.W., Price, C.J. & Riddoch, M.J. (1999). From objects to names: a cognitive neuroscience approach. *Psychological Research*, *62*(2-3), 118-130.
- Ishai, A., Ungerleider, L.G., Martin, A., Schouten, J.L. & Haxby, J.V. (1999). Distributed representation of objects in the ventral visual pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(16), 9379-9384.
- Jazayeri, M. & Movshon, J.A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, *9*(5), 690-696.
- Jazayeri, M. & Movshon, J.A. (2007). Integration of sensory evidence in motion discrimination. *Journal of Vision*, *7*(12):7, 1-7.
- Jenkins, I.H., Jahanshahi, M., Jueptner, M., Passingham, R.E. & Brooks, D.J. (2000). Self-initiated versus externally triggered movements: II. The effect of movement predictability on regional cerebral blood flow. *Brain*, *123*(pt 6), 1216-1228.
- Jersild, A.T. (1927). Mental set and shift. *Archives of Psychology*, *14*, 5-82.
- Jezzard, P., Matthews, P.M. & Smith, S.M. (2003). *Functional MRI: An introduction to methods*. Oxford, UK: Oxford University Press.
- Johnston, K., Levin, H.M., Koval, M.J. & Everling, S. (2007). Top-down control-signal dynamics in anterior cingulate and prefrontal cortex neurons following task switching. *Neuron*, *53*(3), 453-462.
- Kable, J.W. & Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, *10*(12), 1625-1633.
- Kamitani, Y. & Sawahata, Y. (2010). Spatial smoothing hurts localization but not information: Pitfalls for brain mappers. *Neuroimage*, *49*(3), 1949-1952.
- Kamitani, Y. & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*(5), 679-685.
- Kandel, E.R., Schwartz, J.H. & Jessell, T.M. (2000). *Principles of neural science*. New York, NY, USA: McGraw-Hill.

- Kanwisher, N., McDermott, J. & Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302-4311.
- Kay, K.N., Naselaris, T., Prenger, R.J. & Gallant, J.L. (2008). Identifying natural images from human brain activity. *Nature*, *452*(7185), 352-356.
- Kiani, R. & Shadlen, M.N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, *324*(5928), 759-764.
- Kim, J.N. & Shadlen, M.N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience*, *2*(2), 176-185.
- Kimberg, D.Y., Aguirre, G.K. & D'Esposito, M. (2000). Modulation of task-related neural activity in task-switching: an fMRI study. *Brain Research. Cognitive Brain Research*, *10*(1-2), 189-196.
- Kircher, T.T., Brammer, M., Bullmore, E.T., Simmons, A., Bartels, M. & David, A.S. (2002). The neural correlates of intentional and incidental self-processing. *Neuropsychologia*, *40*(6), 683-692.
- Kircher, T.T., Senior, C., Phillips, M.L., Benson, P.J., Bullmore, E.T., Brammer, M., Simmons, A., Williams, S.C., Bartels, M. & David, A.S. (2000). Towards a functional neuroanatomy of self-processing: effects of faces and words. *Brain Research. Cognitive Brain Research*, *10*(1-2), 133-144.
- Knauff, M., Fangmeier, T., Ruff, C.C. & Johnson-Laird, P.N. (2003). Reasoning, models, and images: behavioral measures and cortical activity. *Journal of Cognitive Neuroscience*, *15*(4), 559-573.
- Koechlin, E., Corrado, G., Pietrini, P. & Grafman, J. (2000). Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(13), 7651-7656.
- Koechlin, E. & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science*, *318*(5850), 594-598.
- Konen, C.S. & Kastner, S. (2008). Two hierarchically organized neural systems for object information in human visual cortex. *Nature Neuroscience*, *11*(2), 224-231.
- Kornhuber, H.H. & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Archiv für die gesamte Physiologie der Menschen und Tiere*, *284*, 1-17.
- Kosslyn, S.M., Ganis, G. & Thompson, W.L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*(9), 635-642.
- Kouneiher, F., Charron, S. & Koechlin, E. (2009). Motivation and cognitive control in the human prefrontal cortex. *Nature Neuroscience*, *12*(7), 939-945.
- Kriegeskorte, N. & Bandettini, P. (2007a). Analyzing for information, not activation, to exploit high-resolution fMRI. *Neuroimage*, *38*(4), 649-662.
- Kriegeskorte, N. & Bandettini, P. (2007b). Combining the tools: activation- and information-based fMRI analysis. *Neuroimage*, *38*(4), 666-668.
- Kriegeskorte, N., Formisano, E., Sorger, B. & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(51), 20600-20605.
- Kriegeskorte, N., Goebel, R. & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3863-3868.
- Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K. & Bandettini, P.A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, *60*(6), 1126-1141.

- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S. & Baker, C.I. (2009). Circular analysis in systems neuroscience: the danger of double dipping. *Nature Neuroscience*, 12(5), 535-540.
- Kunimoto, C., Miller, J. & Pashler, H. (2001). Confidence and accuracy of near-threshold discrimination responses. *Consciousness and Cognition*, 10(3), 294-340.
- Kusunoki, M., Sigala, N., Gaffan, D. & Duncan, J. (2009). Detection of fixed and variable targets in the monkey prefrontal cortex. *Cerebral Cortex*, 19(11), 2522-2534.
- Kwong, K.K., Belliveau, J.W., Chesler, D.A., Goldberg, I.E., Weisskoff, R.M., Poncelet, B.P., Kennedy, D.N., Hoppel, B.E., Cohen, M.S. & Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences of the United States of America*, 89(12), 5675-5679.
- Lamme, V.A. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10(11), 494-501.
- Lau, H.C., Rogers, R.D., Haggard, P. & Passingham, R.E. (2004a). Attention to intention. *Science*, 303(5661), 1208-1210.
- Lau, H.C., Rogers, R.D. & Passingham, R.E. (2006). Dissociating response selection and conflict in the medial frontal surface. *Neuroimage*, 29(2), 446-451.
- Lau, H.C., Rogers, R.D. & Passingham, R.E. (2007). Manipulating the experienced onset of intention after action execution. *Journal of Cognitive Neuroscience*, 19(1), 81-90.
- Lau, H.C., Rogers, R.D., Ramnani, N. & Passingham, R.E. (2004b). Willed action and attention to the selection of action. *Neuroimage*, 21(4), 1407-1415.
- Leon, M.I. & Shadlen, M.N. (1998). Exploring the neurophysiology of decisions. *Neuron*, 21(4), 669-672.
- Li, S., Ostwald, D., Giese, M. & Kourtzi, Z. (2007). Flexible coding for categorical decisions in the human brain. *Journal of Neuroscience*, 27(45), 12321-12330.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8, 529-566.
- Libet, B. (2004). *Mind time: The temporal factor in consciousness*. Cambridge, MA, USA: Harvard University Press.
- Libet, B., Gleason, C.A., Wright, E.W. & Pearl, D.K. (1983). Time of conscious intention to act in relation to onset of cerebral activities (readiness-potential): the unconscious initiation of a freely voluntary act. *Brain*, 106(Pt 3), 623-642.
- Logothetis, N.K. (2002). The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 357(1424), 1003-1037.
- Logothetis, N.K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. *Journal of Neuroscience*, 23(10), 3963-3971.
- Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. *Nature*, 453(7197), 869-878.
- Logothetis, N.K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843), 150-157.
- Logothetis, N.K. & Pfeuffer, J. (2004). On the nature of the BOLD fMRI contrast mechanism. *Magnetic Resonance Imaging*, 22(10), 1517-1531.
- Logothetis, N.K. & Wandell, B.A. (2004). Interpreting the BOLD signal. *Annual Review of Physiology*, 66, 735-769.

- Luhmann, C.C., Chun, M.M., Yi, D.J., Lee, D. & Wang, X.J. (2008). Neural dissociation of delay and uncertainty in intertemporal choice. *Journal of Neuroscience*, 28(53), 14459-14466.
- Lundstrom, B.N., Ingvar, M. & Petersson, K.M. (2005). The role of precuneus and left inferior prefrontal cortex during source memory episodic retrieval. *Neuroimage*, 27(4), 824-834.
- Lundstrom, B.N., Petersson, K.M., Andersson, J., Johansson, M., Fransson, P. & Ingvar, M. (2003). Isolating the retrieval of imagined pictures during episode memory: activation of the left precuneus and left prefrontal cortex. *Neuroimage*, 20(4), 1934-1943.
- Mack, M.L., Gauthier, I., Sadr, J. & Palmeri, T.J. (2008). Object detection and basic-level categorization: sometimes you know it is there before you know what it is. *Psychonomic Bulletin & Review*, 15(1), 28-35.
- Macknik, S.L. & Livingstone, M.S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience*, 1(2), 144-149.
- Malach, R., Levy, I. & Hasson, U. (2002). The topography of high-order human object areas. *Trends in Cognitive Sciences*, 6(4), 176-184.
- Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R. & Tootell, R.B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 92(18), 8135-8139.
- Mazurek, M.E., Roitman, J.D., Ditterich, J. & Shadlen, M.N. (2003). A role for neural integrators in perceptual decision making. *Cerebral Cortex*, 13(11), 1257-1269.
- McCoy, A.N. & Platt, M.L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. *Nature Neuroscience*, 8(9), 1220-1227.
- Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22(6), 1423-1442.
- Miller, E.K. & Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167-202.
- Milner, A.D. & Goodale, M.A. (1993). Visual pathways to perception and action. *Progress in Brain Research*, 95, 317-337.
- Mitchell, T.M., Hutchinson, R., Niculescu, R.S., Pereira, F., Wang, X., Just, M. & Newman, S. (2004). Learning to decode cognitive states from brain images. *Machine Learning*, 57(1-2), 145-175.
- Monsell, S. (1996). Control of mental processes. In: V. Bruce (Ed.), *Unsolved mysteries of the mind* (pp. 93-148). Danvers, MA, USA: Erlbaum Taylor and Francis.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134-140.
- Mormann, F., Kornblith, S., Quiroga, R.Q., Kraskov, A., Cerf, M., Fried, I. & Koch, C. (2008). Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *Journal of Neuroscience*, 28(36), 8865-8872.
- Mountcastle, V.B., Talbot, W.H., Sakata, H. & Hyvärinen, J. (1969). Cortical neuronal mechanisms in flutter-vibration studied in unanesthetized monkeys. Neuronal periodicity and frequency discrimination. *Journal of Neurophysiology*, 32(3), 452-484.
- Mourão-Miranda, J., Bokde, A.L.W., Born, C., Hampel, H. & Stetter, M. (2005). Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data. *Neuroimage*, 28(4), 980-995.
- Mourão-Miranda, J., Friston, K.J. & Brammer, M. (2007). Dynamic discrimination analysis: a spatial-temporal SVM. *Neuroimage*, 36(1), 88-99.

- Mourão-Miranda, J., Reynaud, E., McGlone, F., Calvert, G. & Brammer, M. (2006). The impact of temporal compression and space selection on SVM analysis of single-subject and multi-subject fMRI data. *Neuroimage*, 33(4), 1055-1065.
- Müller, K.R., Mika, S., Rätsch, G., Tsuda, K. & Schölkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, 12(2), 181-201.
- Mueller, V.A., Brass, M., Waszak, F. & Prinz, W. (2007). The role of the preSMA and the rostral cingulate zone in initially selected actions. *Neuroimage*, 37(4), 1354-1361.
- Mur, M., Bandettini, P.A. & Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI – an introductory guide. *Social Cognitive and Affective Neuroscience*, 4(1), 101-109.
- Murphey, D.K. & Maunsell, J.H.R. (2007). Behavioral detection of electric microstimulation in different cortical visual areas. *Current Biology*, 17(10), 862-867.
- Murphey, D.K., Maunsell, J.H.R., Beauchamp, M.S. & Yeshor, D. (2009). Perceiving electrical stimulation of identified human visual areas. *Proceedings of the National Academy of Sciences of the United States of America*, 106(13), 5389-5393.
- Nachev, P., Kennard, C. & Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nature Reviews Neuroscience*, 9(11), 856-869.
- Nachev, P., Rees, G., Parton, A., Kennard, C. & Husain, M. (2005). Volition and conflict in human medial frontal cortex. *Current Biology*, 15(2), 122-128.
- Nachev, P., Wydell, H., O'Neill, K., Husain, M. & Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *Neuroimage*, 36 Suppl. 2, T155-T163.
- Newsome, W.T., Britten, K.H. & Movshon, J.A. (1989). Neuronal correlates of a perceptual decision. *Nature*, 341(6237), 52-54.
- Nienborg, H. & Cumming, B.G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature*, 459(7243), 89-92.
- Norman, K.A., Polyn, S.M., Detre, G.J. & Haxby, J.V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424-430.
- O'Craven, K.M. & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience*, 12(6), 1013-1023.
- Ogawa, S., Lee, T.M., Kay, A.R. & Tank, D.W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, 87(24), 9868-9872.
- Ogawa, S., Lee, T.M., Nayak, A.S. & Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine*, 14(1), 68-78.
- Ogawa, S., Tank, D.W., Menon, R., Ellermann, J.M., Kim, S.G., Merkle, H. & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the United States of America*, 89(13), 5951-5955.
- Op de Beeck, H.P. (2010). Against hyperacuity in brain reading: Spatial smoothing does not hurt multivariate fMRI analyses? *Neuroimage*, 49(3), 1943-1948.
- Op de Beeck, H.P., Haushofer, J. & Kanwisher, N.G. (2008a). Interpreting fMRI data: maps, modules and dimensions. *Nature Reviews Neuroscience*, 9(2), 123-135.
- Op de Beeck, H.P., Torfs, K. & Wagemans, J. (2008b). Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway. *Journal of Neuroscience*, 28(40), 10111-10123.

- Orban, G.A., Claeys, K., Nelissen, K., Smans, R., Sunaert, S., Todd, J.T., Wardak, C., Durand, J.B. & Vanduffel, W. (2006). Mapping the parietal cortex of human and non-human primates. *Neuropsychologia*, *44*(13), 2647-2667.
- Oristaglio, J., Schneider, D.M., Balan, P.F. & Gottlieb, J. (2006). Integration of visuospatial and effector information during symbolically cued limb movements in monkey lateral intraparietal area. *Journal of Neuroscience*, *26*(32), 8310-8319.
- O'Toole, A.J., Jiang, F., Abdi, H. & Haxby, J.V. (2005). Partially distributed representations of objects and faces in ventral temporal cortex. *Journal of Cognitive Neuroscience*, *17*(4), 580-590.
- Parker, A.J. & Newsome, W.T. (1998). Sense and the single neuron: probing the physiology of perception. *Annual Review of Neuroscience*, *21*, 227-277.
- Passingham, R.E. (1973). Anatomical differences between the neocortex of man and other primates. *Brain, Behavior and Evolution*, *7*(5), 337-359.
- Passingham, R.E. (1995). *The frontal lobes and voluntary action*. Oxford: Oxford University Press.
- Passingham, R.E., Toni, I. & Rushworth, M.F. (2000). Specialisation within the prefrontal cortex: the ventral prefrontal cortex and associative learning. *Experimental Brain Research*, *133*(1), 103-113.
- Peelen, M.V., & Downing, P.E. (2007). Using multi-voxel pattern analysis of fMRI data to interpret overlapping functional activations [comment]. *Trends in Cognitive Sciences*, *11*(1), 4-5.
- Peelen, M.V., Wiggett, A.J. & Downing, P.E. (2006). Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron*, *49*(6), 815-822.
- Pereira, F., Mitchell, T. & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage*, *45*(1 Supplement), S199-209.
- Pesaran, B., Nelson, M.J. & Andersen, R.A. (2008). Free choice activates a decision circuit between frontal and parietal cortex. *Nature*, *453*(7193), 406-409.
- Pessoa, L. & Padmala, S. (2007). Decoding near-threshold perception of fear from distributed single-trial brain activation. *Cerebral Cortex*, *17*(3), 691-701.
- Philiastides, M.G., Ratcliff, R. & Sajda, P. (2006). Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *Journal of Neuroscience*, *26*(35), 8965-8975.
- Philiastides, M.G. & Sajda, P. (2006). Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cerebral Cortex*, *16*(4), 509-518.
- Philiastides, M.G. & Sajda, P. (2007). EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *Journal of Neuroscience*, *27*(48), 13082-13091.
- Picard, N. & Strick, P.L. (2001). Imaging the premotor areas. *Current Opinion in Neurobiology*, *11*(6), 663-672.
- Platt, M.L. & Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, *400*(6741), 233-238.
- Platt, M.L. & Huettel, S.A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. *Nature Reviews Neuroscience*, *11*(4), 398-403.
- Polyn, S.M., Natu, V.S., Cohen, J.D. & Norman, K.A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science*, *310*(5756), 1963-1966.
- Prinz, W. (2003). Emerging selves: representational foundations of subjectivity. *Consciousness and Cognition*, *12*(4), 512-528.
- Prinz, W. (2006a). Free will as a social institution. In S. Pockett, W. P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior?* (pp. 257-276). Cambridge, MA: MIT Press.

- Prinz, W. (2006b). What re-enactment earns us. *Cortex*, 42(4), 515-517.
- Purushothaman, G. & Bradley, D.C. (2005). Neural population code for fine perceptual decisions in area MT. *Nature Neuroscience*, 8(1), 99-106.
- Quintana, J. & Fuster, J.M. (1999). From perception to action: temporal integrative functions of prefrontal and parietal neurons. *Cerebral Cortex*, 9(3), 213-221.
- Quiroga, R.Q., Mukamel, R., Isham, E.A., Malach, R. & Fried, I. (2008). Human single-neuron responses at the threshold of conscious recognition. *Proceedings of the National Academy of Sciences of the United States of America*, 105(9), 3599-3604.
- Quiroga, R.Q., Reddy, L., Koch, C. & Fried, I. (2007). Decoding visual inputs from multiple neurons in the human temporal lobe. *Journal of Neurophysiology*, 98(4), 1997-2007.
- Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C. & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, 435(7045), 1102-1107.
- Quiroga, R.Q., Snyder, L.H., Batista, A.P., Cui, H. & Andersen, R.A. (2006). Movement intention is better predicted than attention in the posterior parietal cortex. *Journal of Neuroscience*, 26(13), 3615-3620.
- Raichle, M.E. and Mintun, M.A. (2006). Brain work and brain imaging. *Annual Review of Neuroscience*, 29, 449-476.
- Raven, J. (1996). *Progressive Matrices: A Perceptual test of Intelligence — Individual Form*. Oxford, UK: Oxford Psychologists Press.
- Reddy, L. & Kanwisher, N. (2006). Coding of visual objects in the ventral stream. *Current Opinion in Neurobiology*, 16(4), 408-414.
- Ress, D. & Heeger, D.J. (2003). Neuronal correlates of perception in early visual cortex. *Nature Neuroscience*, 6(4), 414-420.
- Resulaj, A., Kiani, R., Wolpert, D.M. & Shadlen, M.N. (2009). Changes of mind in decision-making. *Nature*, 461(7261), 263-266.
- Riddoch, M.J. & Humphreys, G.W. (2003). Visual agnosia. *Neurologic Clinic*, 21(2), 501-520.
- Rizzolatti, G. & Matelli, M (2003). Two different streams form the dorsal visual system: anatomy and functions. *Experimental Brain Research*, 153(2), 146-157.
- Rogers, R.D. & Monsell, S.D. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 207-231.
- Roitman, J.D. & Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, 22(21), 9475-9489.
- Romo, R., Hernández, A., Salinas, E., Brody, C.D., Zainos, A., Lemus, L., de Lafuente, V. & Luna, R. (2002). From sensation to action. *Behavioural Brain Research*, 135(1-2), 105-118.
- Romo, R., Hernández, A. & Zainos, A. (2004). Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron*, 41(1), 165-173.
- Romo, R. & Salinas, E. (2003). Flutter discrimination: neural codes, perception, memory and decision making. *Nature Reviews Neuroscience*, 4(3), 203-218.
- Rowe, J., Hughes, L., Eckstein, D. & Owen, A.M. (2008). Rule-selection and action-selection have a shared neuroanatomical basis in the human prefrontal and parietal cortex. *Cerebral Cortex*, 18(10), 2275-2285.
- Rorie, A.E. & Newsome, W.T. (2008). A general mechanism for decision-making in the human brain? *Trends in Cognitive Sciences*, 9(2), 41-43.

- Rubinstein, J.S., Meyer, D.E. & Evans, J.E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 27(4), 763-797.
- Rushworth, M.F., Hadland, K.A., Paus, T. & Sipila, P.K. (2002a). Role of the human medial frontal cortex in task switching: a combined fMRI and TMS study. *Journal of Neurophysiology*, 87(5), 2577-2592.
- Rushworth, M.F., Passingham, R. & Nobre, A.C. (2002b). Components of switching intentional set. *Journal of Cognitive Neuroscience*, 14(8), 1139-1150.
- Rushworth, M.F., Paus, T. & Sipila, P.K. (2001). Attention systems and the organization of the human parietal cortex. *Journal of Neuroscience*, 21(14), 5262-5271.
- Sakai, K. (2008). Task set and prefrontal cortex. *Annual Review of Neuroscience*, 31, 219-245.
- Sakai, K. & Passingham, R.E. (2003). Prefrontal interactions reflect future task operations. *Nature Neuroscience*, 6(1), 75-81.
- Sakai, K. & Passingham, R.E. (2006). Prefrontal set activity predicts rule-specific neural processing during subsequent cognitive performance. *Journal of Neuroscience*, 26(4), 1211-1218.
- Salinas, E., Hernández, A., Zainos, A. & Romo, R. (2000). Periodicity and firing rate as candidate neural codes for the frequency of vibrotactile stimuli. *Journal of Neuroscience*, 20(14), 5503-5515.
- Salzman, C.D., Britten, K.H. & Newsome, W.T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. *Nature*, 346(6280), 174-177.
- Salzman, C.D., Murasugi, C.M., Britten, K.H. & Newsome, W.T. (1992). Microstimulation in visual area MT: effects on direction discrimination performance. *Journal of Neuroscience*, 12(6), 2331-2355.
- Schmidt, T. & Vorberg, D. (2006). Criteria for unconscious cognition: three types of dissociation. *Perception & Psychophysics*, 68(3), 489-504.
- Schwarzlose, R.F., Swisher, J.D., Dang, S. & Kanwisher, N. (2008). The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 105(11), 4447-4452.
- Searle, J.R. (2004). *Mind: a brief introduction*. Oxford, UK: Oxford University Press.
- Serences, J.T. & Boynton, G.M. (2007). The representation of behavioural choice for motion in human visual cortex. *Journal of Neuroscience*, 27(47), 12893-12899.
- Serences, J.T., Ester, E.F., Vogel, E.K. & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, 20(2), 207-214.
- Sereno, A.B. & Maunsell, J.H. (1998). Shape selectivity in primate lateral intraparietal cortex. *Nature*, 395(6701), 500-503.
- Shadlen, M.N., Britten, K.H., Newsome, W.T. & Movshon, J.A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *Journal of Neuroscience*, 16(4), 1486-1501.
- Shadlen, M.N., Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, 86(4), 1916-1936.
- Shinkareva, S.V., Mason, R.A., Malave, V.L., Wang, W., Mitchell, T.M. & Just, M.A. (2008). Using fMRI brain activation to identify cognitive states associated with perception of tools and dwellings. *PLoS One*, 3(1):e1394.
- Shulman, G.L., Ollinger, J.M., Linenweber, M., Petersen, S.E. & Corbetta, M. (2001). Multiple neural correlates of detection in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 98(1), 313-318.

- Sigala, N., Kusunoki, M., Nimmo-Smith, I., Gaffan, D. & Duncan, J. (2008). Hierarchical coding for sequential task events in the monkey prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(33), 11969-11974.
- Simmons, W.K. & Martin, A. (2009). The anterior temporal lobes and the functional architecture of semantic memory. *Journal of the International Neuropsychological Society*, *15*(5), 645-649.
- Singh-Curry, V. & Husain, M. (2009). The functional role of the inferior parietal lobe in the dorsal and ventral stream dichotomy. *Neuropsychologia*, *47*(6), 1434-1448.
- Smith, P.L. & Ratcliff, R. (2004). Psychology and Neurobiology of simple decisions. *Trends in Neurosciences*, *27*(3), 161-168.
- Smith, P.L. & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review*, *116*(2), 283-317.
- Smith, S.M. (2004). Overview of fMRI analysis. *The British Journal of Radiology*, *77*, 167-175.
- Sohn, M.H., Ursu, S., Anderson, J.R., Stenger, V.A. & Carter, C.S. (2000). The role of prefrontal cortex and posterior parietal cortex in task switching. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(24), 13448-13453.
- Soon, C.S., Brass, M., Heinze, H.J. & Haynes, J.D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*(5), 543-545.
- Soon, C.S., He, A.H., Bode, S. & Haynes, J.D. (in prep.). Unconscious formation of voluntary abstract intentions.
- Spector, A. & Biederman, I. (1976). Mental set and shift revisited. *American Journal of Psychology*, *89*(4), 669-679.
- Spiridon, M., Fischl, B. & Kanwisher, N. (2006). Location and spatial profile of category-specific regions in human extrastriate cortex. *Human Brain Mapping*, *27*(1), 77-89.
- Spiridon, M. & Kanwisher, N. (2002). How distributed is visual category information in human occipito-temporal cortex? An fMRI study. *Neuron*, *35*(6), 1157-1165.
- Sterzer, P., Haynes, J.D. & Rees, G. (2008). Fine-scale activity patterns in high-level visual areas encode the category of invisible objects. *Journal of Vision*, *8*(15):10, 1-12.
- Stoet, G. & Snyder, L.H. (2004). Single neurons in posterior parietal cortex of monkeys encode cognitive set. *Neuron*, *42*(6), 1003-1012.
- Stoet, G. & Snyder, L.H. (2007). Correlates of stimulus-response congruence in the posterior parietal cortex. *Journal of Cognitive Neuroscience*, *19*(2), 194-203.
- Striedter, G.F. (2004). Brain Evolution. In: G. Paxinos and J.K. Mai (Eds.), *The human nervous system* (pp. 3-21). Burlington, MA, USA: Elsevier.
- Süper, H., Spekreijse, H. & Lamme, V.A. (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nature Neuroscience*, *4*(3), 304-310.
- Swets, J.A. (1961). Is there a sensory threshold? *Science*, *134*, 168-177.
- Swisher, J.D., Gatenby, J.C., Gore, J.C., Wolfe, B.A., Moon, C.H., Kim, S.G. & Tong, F. (2010). Multiscale pattern analysis of orientation-selective activity in the primary visual cortex. *Journal of Neuroscience*, *30*(1), 325-330.
- Talairach, J. & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York, NY, USA: Thieme Medical.

- Tarr, M.J. & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, 3(8), 764-769.
- Thompson, G.K. & Schall, J.D. (1999). The detection of visual signals by macaque frontal eye field during masking. *Nature Neuroscience*, 2(3), 283-288.
- Thorndike, E.L. (1911). *Animal intelligence: experimental studies*. New York, NY, USA: The Macmillan Company.
- Toni, I., Schluter, N.D., Josephs, O., Friston, K. & Passingham, R.E. (1999). Signal-, set- and movement-related activity in the human brain: an event-related fMRI study. *Cerebral Cortex*, 9(1), 35-49.
- Tosoni, A., Galati, G., Romani, G.L. & Corbetta, M. (2008). Sensory-motor mechanisms in human parietal cortex underlie arbitrary visual decisions. *Nature Neuroscience*, 11(12), 1446-1453.
- Toth, L.J. & Assad, J.A. (2002). Dynamic coding of behaviourally relevant stimuli in parietal cortex. *Nature*, 415(6868), 165-168.
- Tulving, E., Kapur, S., Markovitsch, H.J., Craik, F.I.M., Habib, R. & Houle, S. (1994). Neuroanatomical correlates of retrieval in episodic memory: auditory sentence recognition. *Proceedings of the National Academy of Sciences of the United States of America*, 91(6), 2012-2015.
- Turner, R., Le Bihan, D., Moonen, C.T., Despres, D. & Frank, J. (1991). Echo-planar time course MRI of cat brain oxygenation changes. *Magnetic Resonance in Medicine*, 22(1), 159-166.
- Tusche, A., Bode, S. & Haynes, J.D. (2010). Neural responses to unattended products predict later consumer choices. *Journal of Neuroscience*, 30(23), 8024-8031.
- Uka, T. & DeAngelis, G.C. (2003). Contribution of middle temporal area to coarse depth discrimination: comparison of neuronal and psychophysical sensitivity. *Journal of Neuroscience*, 23(8), 3515-3530.
- Uka, T. & DeAngelis, G.C. (2004). Contribution of area MT to stereoscopic depth perception: choice related response modulations reflect task strategy. *Neuron*, 42(2), 297-310.
- Uka, T., Tanabe, S., Watanabe, M. & Fujita, I. (2005). Neural correlates of fine depth discrimination in monkey inferior temporal cortex. *Journal of Neuroscience*, 25(46), 10796-10802.
- Ungerleider, L.G. & Mishkin, M. (1982). Two cortical visual systems. In: D.J. Ingle, M.A. Goodale & R.J. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549-586). Cambridge, MA, USA: MIT Press.
- Van Gaal, S., Ridderinkhof, K.R., Fahrenfort, J.J., Scholte, H.S. & Lamme, V.A. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *Journal of Neuroscience*, 28(32), 8053-8062.
- VanRullen, R. & Koch, C. (2003). Visual selective behavior can be triggered by a feed-forward process. *Journal of Cognitive Neuroscience*, 15(2), 209-217.
- Vickery, T.J. & Jiang, Y.V. (2009). Inferior parietal lobule supports decision making under uncertainty in humans. *Cerebral Cortex*, 19(4), 916-925.
- Volz, K.G., Schubotz, R.I. & von Cramon, D.Y. (2004). Why am I unsure? Internal and external attributions of uncertainty dissociated by fMRI. *Neuroimage*, 21(3), 848-857.
- Vorberg, D., Mattler, U., Heinecke, A., Schmidt, T. & Schwarzbach, J. (2003). Different time courses for visual perception and action priming. *Proceedings of the National Academy of Sciences of the United States of America*, 100(10), 6275-6280.
- Wallis, J.D., Anderson, K.C. & Miller, E.K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, 411(6840), 953-956.
- Wallis, J.D. & Miller, E.K. (2003). From rule to response: neuronal processes in the premotor and prefrontal cortex. *Journal of Neurophysiology*, 90(3), 1790-1806.

- Walther, D.B., Caddigan, E., Fei-Fei, L. & Beck, D.M. (2009). Natural scene categories revealed in distributed patterns of activity in the human brain. *Journal of Neuroscience*, *29*(34), 10573-10581.
- Waszak, F., Wascher, E., Keller, P., Koch, I., Aschersleben, G., Rosenbaum, D.A. & Prinz, W. (2004). Intention-based and stimulus-based mechanisms in action selection. *Experimental Brain Research*, *162*(3), 346-356.
- Weber, M., Thomson-Schill, S.L., Osherson, D., Haxby, J. & Parsons, L. (2009). Predicting judged similarity of natural categories from their neural representations. *Neuropsychologia*, *47*(3), 859-868.
- Wegner, D.M. (2003). The Mind's best trick: how we experience conscious will. *Trends in Cognitive Sciences*, *7*(2), 65-69.
- Williams, M.A., Baker, C.I., Op de Beeck, H.P., Shim, W.M., Dang, S., Triantafyllou, C. & Kanwisher, N. (2008). Feedback of visual object information to foveal retinotopic cortex. *Nature Neuroscience*, *11*(12), 1439-1445.
- Williams, M.A., Dang, S. & Kanwisher, N.G. (2007). Only some spatial patterns of fMRI response are read out in task performance. *Nature Neuroscience*, *10*(6), 685-686.
- Yeung, N., Nystrom, L.E., Aronson, J.A. & Cohen, J.D. (2006). Between-task competition and cognitive control in task switching. *Journal of Neuroscience*, *26*(5), 1429-1438.
- Zilles, K., Armstrong, E., Moser, K.H., Schleicher, A. & Stephan, H. (1989). Gyrification in the cerebral cortex of primates. *Brain, Behavior and Evolution*, *34*(3), 143-150.

List of Figures

Experimental Section I

3-1	Experimental paradigm	34
3-2	Multivariate pattern classification	37
3-3	Decoding the task-sets, target stimuli and motor responses over time	40
3-4	Information time courses for task-set decoding	41
3-5	Averaged decoding accuracies and FIR contrast estimates	43
3-6	Decoding fixed task-sets in the control experiment	44
<hr/>		
A-1	Joystick and custom-made fMRI compatible plexiglass table	ii
A-2	Decoding target stimuli	vi
A-3	Decoding motor responses	vii
A-4	Unconstrained decoding of rule-target combinations	ix

Experimental Section II

4-1	Experimental paradigm and examples of stimuli and masks	64
4-2	Stimuli used in the localizer experiment	68
4-3	Visibility of stimuli	70
4-4	Whole-brain searchlight decoding results for categorical choices	72
4-5	Whole-brain searchlight decoding results for object categories	73
4-6	Whole-brain searchlight decoding results for motor responses	74
4-7	BOLD contrast estimation profiles for all conditions	75
<hr/>		
B-1	Visibility characteristics of categories and masking procedures	xii
B-2	Visibility of object categories using repetitive masking	xiv

B-3	Images of pianos used for fMRI masking experiment	xvii
B-4	Images of chairs used for fMRI masking experiment	xvii
B-5	Phase randomised noise images used for fMRI masking experiment	xviii
B-6	Confusion matrix in perceptual decision making	xix
B-7	Mean hit rates for object categories	xx
B-8	Activation contrasts for visual stimulation	xxiii
B-9	Results of univariate error analyses for low visibility condition	xxiii
B-10	Correlation analysis of confusion matrices	xxvii
B-11	Individual searchlight patterns in LOC	xxviii
B-12	Individual searchlight patterns in the precuneus	xxix
B-13	Decision confidence and visibility	xxx
B-14	Decoding category choices from LOC based on localizer-defined ROIs	xxxi
B-15	Decoding presented categories from LOC based on localizer-defined ROIs	xxxii
B-16	Comparison of precuneus regions	xxxiii

Experimental Section III

5-1	Experimental paradigm and conditions	95
5-2	Cross-condition pattern classification	99
5-3	Mean hit rates for each participant	101
5-4	Average choice ratios	101
5-5	Decoding decision outcomes using multivariate pattern classification	103
5-6	Direct comparison of precuneus regions between both masking studies	105

C-1	Results of behavioural pre-test	xxxv
C-2	Comparison of questionnaire results between paradigm variants	xxxvi
C-3	Post-experimental questionnaire	xxxvii
C-4	Results from post-experimental questionnaire for fMRI experiment	xxxviii
C-5	Individual choice ratios from fMRI experiment	xxxix
C-6	Decoding of motor responses and true object categories	xliii

List of Tables

Experimental Section I		
A-1	General peak activation for cue presentation	iv
A-2	General peak activation for target stimuli	iv
A-3	General peak activation for motor responses	v
A-4	Decoding rule-target combinations from parietal and prefrontal regions	ix

Experimental Section II		
4-1	Decoding from LOC ROIs	76
B-1	Descriptive statistics (hits) for criteria-based selection of target objects	xvi
B-2	Table of activation contrasts for univariate analysis	xxii
B-3	Results of two-class decoding analyses for object categories	xxv

Experimental Section III		
C-1	Table of activation contrasts for univariate analysis	xlii

List of Abbreviations

ACC	Anterior cingulate cortex
ANOVA	Analysis of variance
ATP	Adenosine-5'-triphosphate
BOLD	Blood oxygen level dependent
BP	Bereitschaftspotential (readiness-potential)
CMA	Cingulate motor area
d'	d-prime
dHb	Deoxyhaemoglobin
CP	Choice probability
DLPFC	Dorsolateral prefrontal cortex
DV	Decision variable
EEG	Electroencephalography
EIN	excitation-inhibition network
EPI	Echo-planar imaging
FD	Free decision task
FFA	Fusiform face area
FIR	Finite impulse response
fMRI	Functional magnetic resonance imaging
FPC	Frontopolar cortex
FWE	Family-wise error
FWHM	Full width at half maximum
GE	Gradient echo
GLM	General linear model
Hb	Oxyhaemoglobin
hMT+	human medial temporal complex
HRF	Haemodynamic response function
HV	Highly visible object categorisation task
IFJ	Inferior frontal junction
IPL	Inferior parietal lobe
IPS	Intraparietal sulcus
IT	Inferior temporal
IV	Invisible object categorisation task

MEG	Magnetoencephalography
MNI	Montreal Neurological Institute
MPFC	Medial prefrontal cortex
MRI	Magnetic resonance imaging
MTL	Medial temporal lobe
MVPA	Multivariate pattern analysis
LDA	Linear discriminant analysis
LFP	Local field potential
LIP	Lateral intraparietal area
LOC	Lateral-occipital complex
PCA	Principal components analysis
PCC	Posterior cingulate cortex
PET	Positron emission tomography
PFC	Prefrontal cortex
PPA	Parahippocampal place area
PPC	Posterior parietal cortex
pre-SMA	Pre-supplementary motor area
RCZ	Rostral cingulate zone
RF	Radio frequency
ROI	Region of interest
SE	Standard error
SD	Standard deviation
SDT	Signal detection theory
SMA	Supplementary motor area
S-R	Stimulus-response
SVM	Support vector machine
TE	Echo time
TR	Repetition time
SNR	Signal-to-noise ratio
V1	Primary visual cortex
V3	Tertiary visual cortex
VLPFC	Ventrolateral prefrontal cortex
WCST	Wisconsin card sorting test

Curriculum Vitae

Stefan Bode (nee Bilgenroth)

born 26. April 1979 in Celle, Germany

Education

1998	Abitur, Ganztagsgymnasium Barsinghausen, Germany
1998-1999	Civilian Service: Werner-Dicke School and boarding house for physically disabled children (Annastift e.V.), Hannover, Germany
1999-2002	Undergraduate Psychology (Vordiplom), University of Göttingen, Germany
2002-2005	Diploma of Psychology (Dipl.-Psych.), University of Göttingen, Germany
2005-2006	Scientific assistant: Dept. Medical Psychology and Medical Sociology, University Clinic Göttingen, Germany
2006-2010	PhD student: Dept. Attention & Awareness, Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany / Dept. of Neurology, Otto-von-Guericke University Magdeburg, Germany

Appendices

Appendices	i
Appendix A.....	ii
Joystick and custom-made plexiglass table for fMRI	ii
Univariate fMRI data analysis.....	iii
Decoding of target stimuli	vi
Decoding of motor responses	vii
Decoding of stimulus-response combinations.....	viii
Appendix B	x
Behavioural pre-tests I: Selection of categories and masking procedure	x
Behavioural pre-tests II: Repetitive masking	xiii
Behavioural pre-test III: Selection of target stimuli	xv
Confusion matrix in perceptual decision making.....	xix
Mean hit rates for fMRI masking experiment	xx
Univariate fMRI data analysis.....	xxi
Two-class pattern classification for presented objects	xxiv
Correlation analysis of confusion matrices	xxvi
Individual searchlight patterns for decoding of choices.....	xxviii
Behavioural control experiment: Decision confidence and hit rates... ..	xxx
Decoding results from conventional LOC ROIs	xxxii
Comparison of precuneus regions with Soon et al. (2008)	xxxiii
Appendix C	xxxiv
Behavioural pre-test	xxxiv
Results questionnaire fMRI experiment	xxxviii
Choice ratios fMRI experiment	xxxix
Choice ratios and decoding accuracy	xl
Univariate fMRI data analysis.....	xli
Decoding motor responses and visual stimulation	xliii
ROI analyses.....	xliv

Appendix A

Joystick and custom-made plexiglass table for fMRI

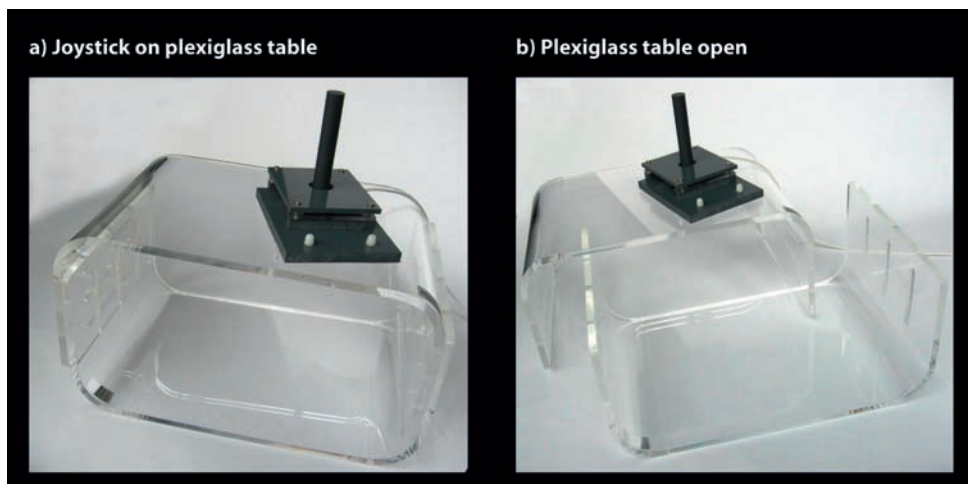


Figure A-1: Joystick and custom-made fMRI compatible plexiglass table. a) The joystick was fixed on the plexiglass construction using plastic screws. Only left and right movements could be performed. Movements in both directions required the same amount of muscle strength, adjustable by using different in-built bridge layers b) The plexiglass table comprised two similar parts, each of which featuring striations on the side. Being perpendicular to their counterparts, these striations allowed fixing the upper part of the table in every desired tilt angle using plastic screws. This assured that the participants' arms rested in a comfortable position throughout the entire experiment. The lower part of the plexiglass table was placed below the scanner mattress, such that participants' weight prohibited any movement of the table (photo: Stephan Liebig).

Univariate fMRI data analysis

Only trials with correct responses were included in the univariate data analysis. The fMRI data were motion corrected, spatially normalised to a standard stereotaxic space (Montreal Neurological Institute EPI template) and re-sampled to an isotropic spatial resolution of $3 \times 3 \times 3 \text{ mm}^3$ in SPM2. Data were smoothed with a Gaussian kernel of 8 mm FWHM to account for anatomical variability and to satisfy the assumption of Gaussian random field theory. For statistical analyses, a FWE corrected threshold of $p < .05$ was used. The analysis aimed to find brain regions that were significantly activated during processing of the cue (task-set), the perception of the target stimulus and the execution of the motor response. Furthermore, the contrasts between the different task-sets, between target stimuli and between responses were calculated. A GLM as implemented in SPM2 was used for the analyses. It consisted of six boxcar regressors, each convolved with a canonical HRF. The first two regressors modelled cue “A” and cue “B”, the third and fourth regressors modelled target 1 and target 2, and the last two regressors modelled the responses (left and right).

The analysis searched for regions with an overall increase in activation during the different stages of information processing. During processing of the rule beginning with the presentation of the cue, increased activation was found in visual cortex, intraparietal sulcus (IPS), the medial frontal gyrus corresponding to supplementary motor cortex (SMA) and pre-SMA (Table A-1). During presentation of the target, the visual cortex was also activated, along with several regions of prefrontal cortex, including the left insula, premotor cortex (predominantly left) and left primary motor cortex. This presumably indicates that motor preparation was already taking place during the early stages of target stimulus presentation (Table A-2). During response execution, activation was increased primarily in motor-related areas including primary motor cortex, premotor cortex, basal ganglia and cerebellum as well as in temporal cortex (Table A-3). No brain region could be identified that was more activated during the presentation of cue “A” compared to cue “B” nor was there conversely a brain region that was more activated during the presentation of cue “B” compared to cue “A”. The same held true for the comparison of target stimuli as well as comparing left and right motor responses (note that both, left and right responses with the joystick were performed with the right hand only).

Table A-1: General peak activation for cue presentation.

Anatomical area	L/R	T-value	x	y	z
Middle occipital gyrus	L	8.85	-27	-84	-6
	R	7.32	30	-84	0
Calcarine sulcus	L	7.86	-15	-93	-6
	R	7.82	15	-90	-9
Intraparietal sulcus	L	7.14	-30	-54	54
Medial frontal gyrus	L	6.29	-6	15	51
	L	6.01	-27	-3	54

Note: The coordinates are given according to MNI space with their T-values. L = left hemisphere, R = right hemisphere; all FWE-corrected ($p < .05$; voxel threshold = 10 voxels).

Table A-2: General peak activation for target stimuli.

Anatomical area	L/R	T-value	x	y	z
Calcarine sulcus	L	9.63	-15	-90	-9
Medial occipital gyrus	R	8.94	30	-75	-12
Inferior occipital gyrus	R	9.05	36	-79	3
Cuneus	R	7.20	9	-72	12
	L	5.68	-12	-75	12
Medial superior frontal sulcus	L	8.04	-6	-18	51
Medial superior frontal gyrus	L	7.20	-6	0	48
Precentral gyrus	L	6.32	-9	-30	54
	L	7.58	-33	-15	42
Central sulcus	L	6.75	-27	-27	48
Postcentral gyrus	R	5.65	63	-15	39
Insula	L	6.40	-42	3	3

Note: The coordinates are given according to MNI space with their T-values. L = left hemisphere, R = right hemisphere; all FWE-corrected ($p < .05$; voxel threshold = 10 voxels).

Table A-3: General peak activation for motor responses.

Anatomical area	L/R	T-value	x	y	z
Cerebellum	L	14.85	-48	-69	-27
	R	14.41	51	-57	-30
Basal ganglia	R	14.81	18	0	27
	L	6.19	-18	0	-9
Precentral gyrus	R	10.73	6	-27	78
Postcentral gyrus	R	8.26	3	-51	72
	L	7.88	-3	-51	72
Superior temporal gyrus	L	7.42	-51	-6	6
Medial temporal gyrus	L	5.88	-66	-48	-3

Note: The coordinates are given according to MNI space with their T-values. L = left hemisphere, R = right hemisphere; all FWE-corrected ($p < .05$; voxel threshold = 10 voxels).

Decoding of target stimuli

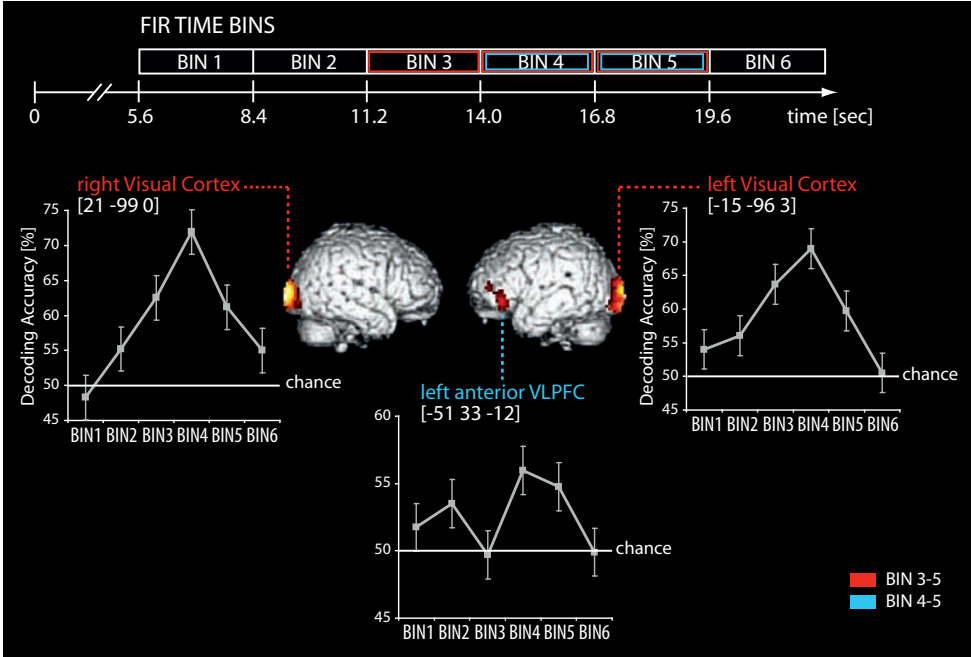


Figure A-2: Decoding target stimuli. The figure shows decoding accuracy [%] of target stimuli collapsed across both task-sets, which made it possible to dissociate encoding of stimuli from encoding of motor responses (radius = 4 voxels). The graphs show mean decoding accuracies and standard errors over time. Regions displayed showed decoding above chance level (50%; voxel threshold = 10 voxels) in time bins 3-5. The cue was presented at the beginning of each trial (onset 0 sec), followed by the target stimulus (onset 4.2 sec) and the motor response (onset 8.4 sec). It was necessary to account for the temporal delay of the BOLD signal such that time bins were shifted by 2 volumes (i.e. time bin 1 is the earliest that could reflect cue related activity). Visual cortex encoded the *target stimulus* from time bin 3 on, peaking in bin 4 ($p < .05$ FWE corrected). Anterior ventrolateral prefrontal cortex showed significant decoding accuracy in time bins 4 and 5, peaking at 56% ($p < .001$ uncorrected). No other regions were predictive above chance for target stimuli in any other time bin. Coordinates displayed are MNI coordinates.

Decoding of motor responses

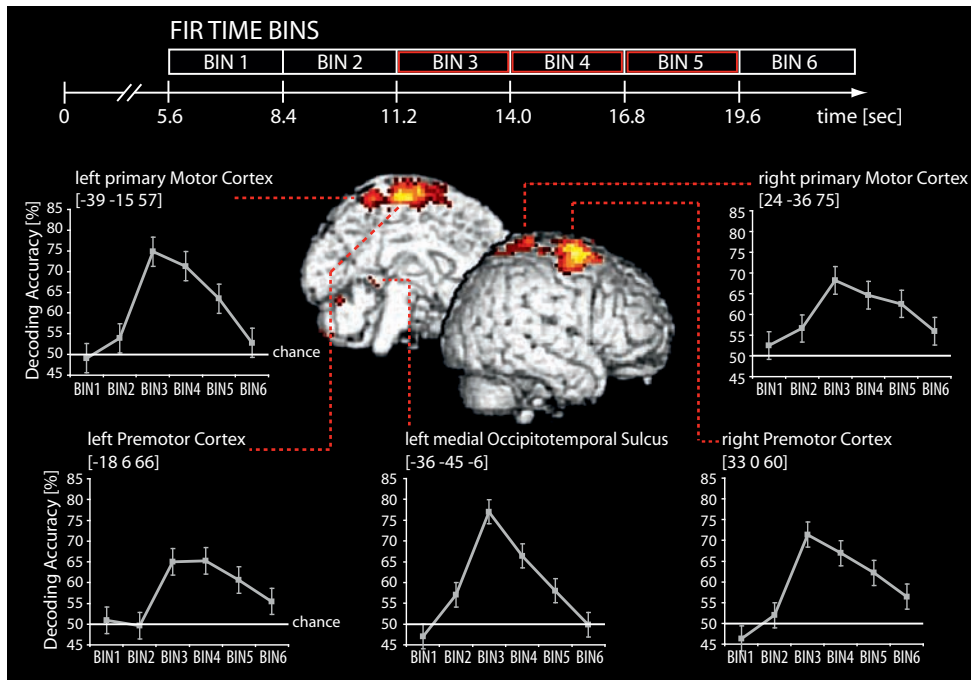


Figure A-3: Decoding motor responses. The figure shows decoding accuracy [%] of motor responses collapsed across both task-sets, which makes it possible to dissociate encoding of motor responses from encoding of target stimuli (radius = 4 voxels). The graphs show mean decoding accuracies and standard errors over time. The cue was presented at the beginning of each trial (onset 0 sec), followed by the target stimulus (onset 4.2 sec) and the motor response (onset 8.4 sec). It was necessary to account for the temporal delay of the BOLD signal such that time bins were shifted by 2 volumes (i.e. time bin 1 is the earliest that could reflect cue related activity). Regions displayed allowed decoding of motor responses above chance level (50%; voxel threshold = 10 voxels) after FWE correction ($p < .05$) in time bins 3-5. Bilateral premotor cortex, bilateral primary motor cortex and left medial occipito-temporal sulcus encoded the motor response. All regions showed similar information time courses, reaching their maxima in the third or fourth time bin (peak of 77%). No other regions were predictive above chance for motor responses in any other time bin. Coordinates displayed are MNI coordinates.

Decoding of stimulus-response combinations

In an additional analysis, the question was addressed whether the four individual *combinations* of rules and target stimuli were also encoded in prefrontal and parietal cortex. Asaad and colleagues (1998) demonstrated in monkeys that lateral PFC can encode combinations of stimuli and rules. It should be noted that the conditions defined by rule-stimulus combinations can also be regarded as stimulus-response combinations. This analysis therefore cannot disentangle target stimuli from motor responses. For this analysis, the fMRI data were motion corrected and re-sampled to an isotropic spatial resolution of $3 \times 3 \times 3 \text{ mm}^3$ in SPM2 as described for the main analyses. Again, a GLM approach as implemented in SPM2 was used, consisting of four boxcar regressors modelling the trials according to the rule-stimulus combination involved. Each regressor was convolved with a canonical HRF. Finally, a searchlight pattern classification analysis was implemented (as described in methods in Experimental Section I) to decode the four individual combinations of rules and target stimuli.

First, the searchlight clusters in pVLPFC, aVLPFC and IPS that encoded the abstract rules and target stimuli were analysed (Table A-4). All regions encoded the individual rule-stimulus combinations with decoding accuracies between 7 to 12% above chance level (25%). Nevertheless, this trend did not reach statistical significance using the same threshold as used for the other analyses reported here (all $p > .001$ uncorrected). Statistically significant accuracies for decoding individual rule-stimulus combinations could only be found in visual cortex and pre-motor as well as motor areas (Figure A-4). These findings confirm that combinations of stimuli and rules are also simply combinations of stimuli and responses, hence, encoded in the respective areas. Further regions, which could also encode these combinations, might not have been detectable because of the decrease in statistical power due to the increase in distinct conditions (four instead of two as in the main analyses) and the decrease of trials per conditions.

Table A-4: Decoding rule-target combinations from parietal and prefrontal regions.

Anatomical area	L/R	Decoding accuracy	SE	T-value	x	y	z
Intraparietal sulcus	L	32 %	2.1	2.86	-24	-45	42
Posterior prefrontal cortex	L	32 %	2.7	2.05	-48	12	21
Anterior prefrontal cortex (I)	L	36 %	4.9	1.97	-51	33	3
Anterior prefrontal cortex (II)	L	37 %	4.1	2.85	-51	33	-12

Note: Listed are only results for the searchlight clusters in IPS, pVLPFC and aVLPFC that showed high decoding accuracies for task-sets and target stimuli in the main analyses. Chance level was 25% for four rule-target combinations. The coordinates are given according to MNI space with their corresponding T-values. L = left hemisphere, R = right hemisphere. None of the decoding accuracies significantly exceeded chance level (all $p > .001$, uncorrected).

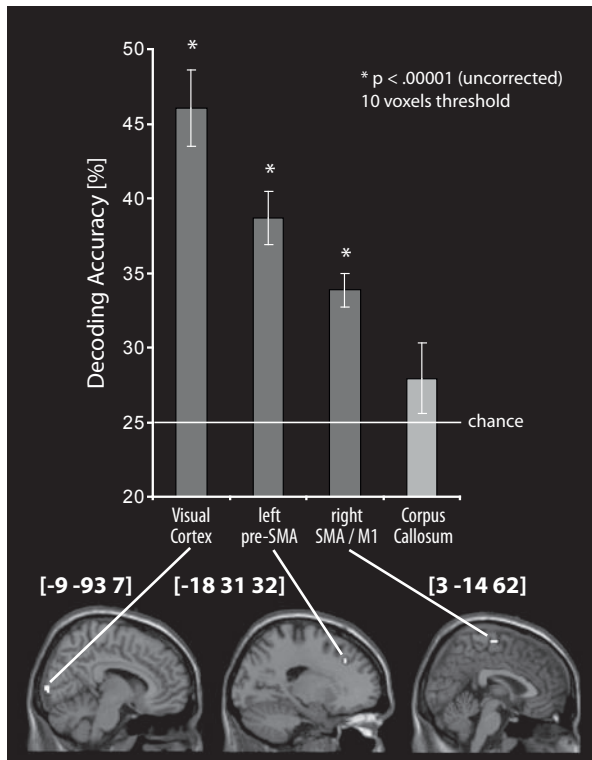


Figure A-4: Unconstrained decoding of rule-target combinations.

Decoding was performed using a moving searchlight (radius = 4 voxels). Displayed are mean decoding accuracies and standard errors in regions showing the highest accuracies above chance (25%) as well as corpus callosum as a non-significant control region. A statistical threshold of $p < .00001$ (uncorrected) and a voxel threshold of 10 voxels were used. Significant accuracies could be found in visual cortex (peak accuracy 46%; MNI [-9 -93 7]), pre-SMA (peak 38%; MNI [-18 31 32]) and SMA/M1 (peak 34%; MNI [3 -14 62]). Combinations of stimuli and rules are also combinations of stimuli and responses, explaining the high accuracies in visual and motor areas.

Appendix B

Behavioural pre-tests I: Selection of categories and masking procedure

Participants and procedure

The pre-test aimed at selecting optimal object categories and a masking procedure for the fMRI experiment. Ten participants (seven female, mean 24 years, SD = 2.31) took part in this experiment. None of them took part in any other pre-test or the final fMRI study. All participants were right-handed and had normal or corrected to normal visual acuity. Participants were paid 7 € for their participation. Stimuli were grey-scaled object images (400 x 400 pixels) from four different categories (pianos, chairs, shoes, cars) in natural backgrounds. Their construction is described in the main text (see Methods). Twelve images (out of thirty) from each category were used for the experiment. Two different types of masking were tested: *backward-masking* (Grill-Spector et al., 2000) and *sandwich-masking*, which used an additional pre-mask. As characteristics of the mask may be critical determinants of the degree of invisibility that can be achieved, differences between two mask compositions were additionally examined. For this purpose, every target picture was scrambled into random square tiles of either 40 x 40 pixels or 50 x 50 pixels in size (see Methods). The size of the tiles in the masks served as an additional experimental condition. For the sandwich-masking procedure, two different masks with the respective tile size were used. The backward masking procedure required only one mask. In sum, masking conditions tested were *backward-* and *sandwich-masking* with masks consisting of randomised but non-meaningful picture parts, namely *10 x 10 (small)* or *8 x 8 (big) tiles*.

Different masking strengths were achieved by varying the target duration with respect to the post-mask duration in all masking conditions. The pre-mask was kept constant with 167 ms for sandwich-masking. The four presentation durations of the target stimulus were a) 16.7 ms (very short; visible for 1 frame given a monitor refresh rate of 60 Hz) with a post-mask duration of 483.3 ms, b) 33.3 ms (short; 2 frames) with a post-mask duration of 466.7 ms, c) 50.0 ms (long; 3 frames) and a post-mask duration of 450.0 ms and d) 66.7 ms (very long; 4 frames) resulting in a post-mask duration of 433.3 ms. The masking methods (backward vs. sandwich) were used in distinct experimental blocks. The order of

blocks was pseudo-randomised between participants. In every block, each of the twelve objects from the four different categories was shown once in every visibility condition. This equated to a total of 192 trials performed per block. Each trial was followed by a two second delay, allowing time for responses. All stimuli were shown on a 17" TFT screen using a Barebone Shuttle Pentium4 personal computer (Megware Computers); the presentation software used was the Cogent toolbox for MATLAB 7.0 (The MathWorks, Inc.). Comparable to the fMRI study the estimated angle of vision of $\alpha \sim 0.72^\circ$ for all stimuli. The background was kept a constant dark-grey. Participants had to fixate on the white cross that was always superimposed. On each trial, they were instructed to report the category they believed the presented object to have by pressing one of four buttons on a response device, each randomly assigned to one of the four categories, with their right and left middle- and index fingers. In the case that participants could not make a decision or were unsure, they were instructed to guess. The data was analysed by calculating a d-prime (d') visibility index for each condition as described for the main experiment (see Methods and caption of Figure B-1). This procedure allowed a direct comparison of distributions of visibility characteristics between the experimental conditions for each category.

Results

The distribution of visibility indices for the different categories in each visibility condition are shown in Figure B-1. Because of the low number of participants and the explorative nature of the pre-test, no statistical tests were performed. From visual inspection it became obvious that sandwich-masking led to stronger masking and less visibility in all conditions. Sandwich-masking showed a gradual increase in visibility with increasing target duration and seemed to be the more appropriate procedure. In general, cars and shoes were detected more easily compared to pianos and chairs, even under strong masking conditions. Participants' self-reports also led to the conclusion that the characteristic shapes of cars and shoes were detected too easily. Chairs and pianos demonstrated better and highly similar visibility characteristics and were selected for the fMRI experiment. Target durations of 16.7 ms and 66.7 ms were selected for the fMRI experiment because they facilitated a maximum difference in object visibility. Since the difference in tile sizes of the masks did not lead to substantial differences, the one tending to yield slightly better masking properties was selected, namely the mask comprised of 40 x 40 pixel tiles.

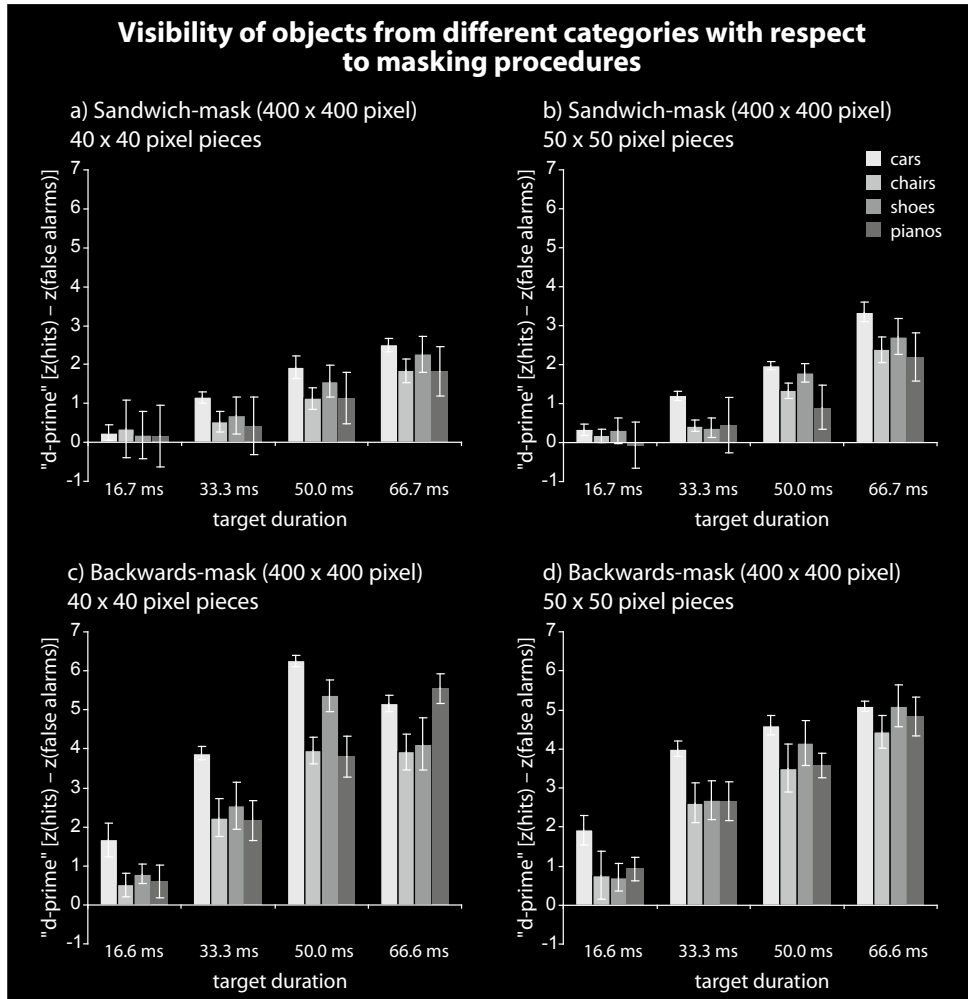


Figure B-1: Visibility characteristics of categories and masking procedures. a) Sandwich-masking using 400 x 400 pixel masks comprised of 10 x 10 tiles (each 40 x 40 pixels), randomly chosen from different images. 12 target images from four categories (cars, chairs, shoes, pianos) were tested in a perceptual category detection task. Displayed are mean d' values [$d' = z(\text{hits}) - z(\text{false alarms})$] and standard errors for four visibility conditions (target durations of 16.7 ms, 33.3 ms, 50.0 ms, 67.7 ms; post-mask duration 500 ms minus target duration; pre-mask duration was kept constant at 167 ms) for $N=10$ participants. b) Sandwich-masking as described for (a) but using masks comprised of 8 x 8 scrambled image tiles (50 x 50 pixel). c) Backward-masking using a 400 x 400 pixel mask comprised of 10 x 10 scrambled image tiles (40 x 40 pixel) shown directly after target image presentation. d) Backward masking as described for (c) using a 400 x 400 pixel mask comprised of 8 x 8 squares of scrambled image pieces (50 x 50 pixel). Since the gradual increase of visibility was best in sandwich-masking and target durations of 16.7 ms and 66.7 ms gave best separation between conditions, these were selected for the fMRI experiment. Pianos and chairs were found to be most comparable and hardest to detect; these were also chosen for the fMRI experiment.

Behavioural pre-tests II: Repetitive masking

Participants and procedure

Additional behavioural pre-tests were conducted to test a four-time repetition of target and masks in rapid succession to extend the total visual stimulation duration without substantially changing the visibility characteristics (Macknik & Livingstone, 1998).

The sample consisted of ten right-handed participants with normal visual acuity (seven female, $M = 23.9$, $SD = 2.69$). None of the participants took part in any other pre-test or the final fMRI study. Participants were paid 7 € for their participation. Sixteen masked stimuli from three object categories (chairs, pianos and noise) were presented using repetitive backward-masking and repetitive sandwich-masking in separate blocks (masks 10 x 10 scrambled tiles). Two visibility conditions (high: target 66.7 ms, post-mask 433.3 ms; low: target 16.7 ms, post-mask 483.3 ms; chosen from pre-test I) were used, resulting in 96 trials in total per block. Every presentation of masked targets was repeated four times in direct succession before the participants were allowed to respond (three seconds response delay). The task and setting was the same as for the first pre-test (see above). Responses (left middle and index fingers, right index finger) were randomly assigned to the three categories on a trial-by-trial basis.

Results

Statistical analyses were not performed due to the small pre-test sample; however, the general pattern of visibility characteristics using the different masking procedures was similar to those obtained in the first pre-test. While the visibilities achieved using sandwich masking were very low and uniform for all categories, the indices for backward masking were already high using 66.7 ms target duration (Figure B-2). Visibility indices might have plateaued and showed a ceiling effect using backward masking. Values around $d' = 5$ can be considered nearly perfect, at which point the masking can be assumed ineffective. In summary, optimal masking could only be achieved using repetitive sandwich-masking. Additionally, considering the more balanced distribution of mean d' -values between categories using repetitive sandwich-masking, this approach was confirmed to be the most suitable method for the final fMRI experiment.

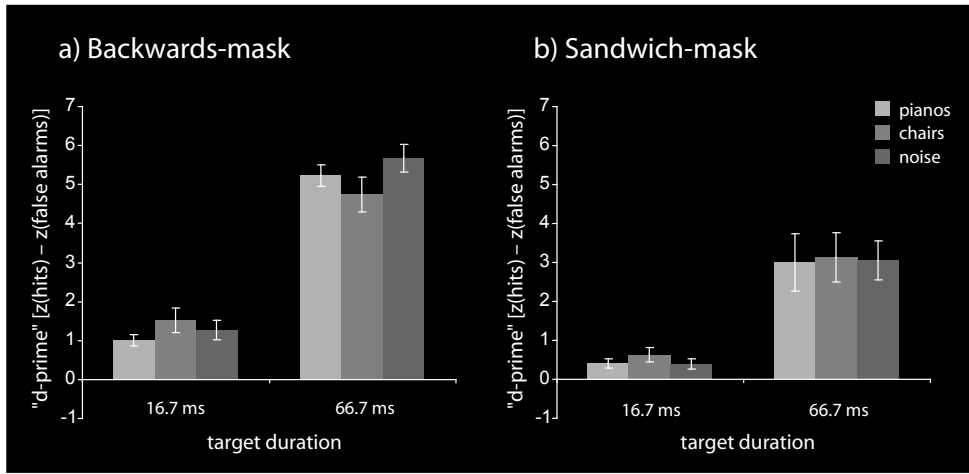


Figure B-2: Visibility of object categories using repetitive masking. Shown are the d' values (for $n=10$ participants) for a) backward-masking (only post-mask, duration of 500 ms minus target duration) and b) sandwich-masking (duration of 167 ms for the pre-mask and 500 ms minus target duration for the post-mask), for two different target durations (16.7 ms and 66.7 ms). While the visibility achieved using sandwich-masking was still reduced for 66.7 ms target duration (right side), visibility showed a ceiling effect using backward-masking (left side). Optimal masking could only be achieved in the condition of 16.7 ms target presentation with repetitive sandwich-masking.

Behavioural pre-test III: Selection of target stimuli

The purpose of the third pre-test was to select *object images* from a pool of images for each category that showed distinct visibility characteristics under the final two masking conditions, such that the highest degree of comparability was maintained across categories.

Participants and procedure

An independent sample of nine participants (5 female, $M = 26$, $SD = 3.5$) participated in this pre-test. They were also exclusively right-handed, had normal or corrected to normal visual acuity and were paid 7 € for their participation. The general procedure and material was exactly the same as in the second pre-test but using the repetitive sandwich-masking approach only (see pre-test II). The final two visibility conditions used target durations of 16.7 ms (low visibility: post-mask duration of 483.3 ms) and 66.7 ms (high visibility: post-mask duration of 433.3 ms). 30 target objects of each category were presented once per run in both visibility conditions, resulting in 180 trials. The order of trials was individually randomised for each participant and each run. All participants performed two experimental runs.

The overall percentage of hits was calculated for each object in all three categories. Note that for single-object analyses, false alarms were impossible to calculate. Only objects that showed substantially *different* visibility characteristics under the two visibility conditions were selected for the fMRI experiment while objects showing exceptional visibility characteristics were discarded. More specifically, objects had to be classified correctly in more than 55% of all cases under high visibility and in more than 25% but not more than 75% under low visibility. Objects that were misclassified even more often were considered systematically misleading and were therefore discarded. Some objects were dismissed because of too characteristic picture properties (according to subjects' reports).

Results

A complete overview of the visibility characteristics is given in Table B-1. The best 24 object images were chosen for the final fMRI experiment (Figures B-3, B-4 and B-5).

Table B-1: Descriptive statistics (hits) for criteria-based selection of target objects.

Object number	Pianos		Chairs		Noise	
	low visibility	high visibility	low visibility	high visibility	low visibility	high visibility
1	44.44	100.00	61.11	100.00	27.78	77.78
2	38.89	72.22	61.11	100.00	22.22	55.56
3	50.00	100.00	61.11	100.00	27.78	38.89
4	50.00	77.78	77.78	100.00	25.22	66.67
5	27.78	55.56*	50.00	100.00	50.00	61.11
6	44.44	100.00	66.67	100.00	33.33	88.89
7	22.22	94.44	38.89	100.00	50.00	83.33
8	61.11	100.00	50.00	100.00	44.44	83.33
9	50.00	94.44	66.67	88.89	27.78	88.89
10	22.22	94.44	66.67	100.00	38.89	83.33
11	38.89	77.78	88.89	100.00	61.11	72.22
12	55.56	100.00	38.89	100.00	27.78	88.89
13	50.00	100.00	44.44	88.89	33.33	72.22
14	44.44	100.00	27.78	94.44	38.89	61.11
15	38.89	100.00	55.56	100.00	27.78	55.56
16	38.89	94.44	50.00	94.44	33.33	77.78
17	55.56	88.89	83.33	100.00	44.44	88.89
18	38.89	100.00	72.22	100.00	27.78	22.22
19	27.78	100.00	61.11	100.00	61.11	66.67
20	66.67*	88.89	22.22	72.22	38.89	77.78
21	44.44	94.44	33.33	66.67	55.56	72.22
22	38.89	100.00	44.44	83.33	16.67	50.00
23	50.00	100.00	55.56	100.00	38.89	88.89
24	61.11	88.89	72.22	100.00	38.89	88.89
25	55.56*	100.00	77.78	100.00	27.78	72.22
26	50.00	94.44	38.89	100.00	22.22	33.33
27	27.78	100.00	33.33	50.00	22.22	66.67
28	50.00	83.33	38.89	77.78	50.00	77.78
29	33.33	50.00	61.11	100.00	38.89	55.56
30	55.56	94.44	38.89	100.00	33.33	77.78

Note: Mean hit rates [%] are reported for all possible target images from the selected categories under two visibility conditions (sample of n = 9 participants). Target durations were 16.7 ms (low visibility) and 66.7 ms (high visibility). Post-mask duration was calculated 500 ms minus target duration, pre-mask duration was always 167 ms. The selection criteria are given in the text. Some objects were dismissed because of too characteristic picture properties (*). Those images whose visibility characteristics did not satisfy the criteria (red) are highlighted in grey. The selected images are can be found in Figures B-3 (pianos), B-4 (chairs) and B-5 (noise images).

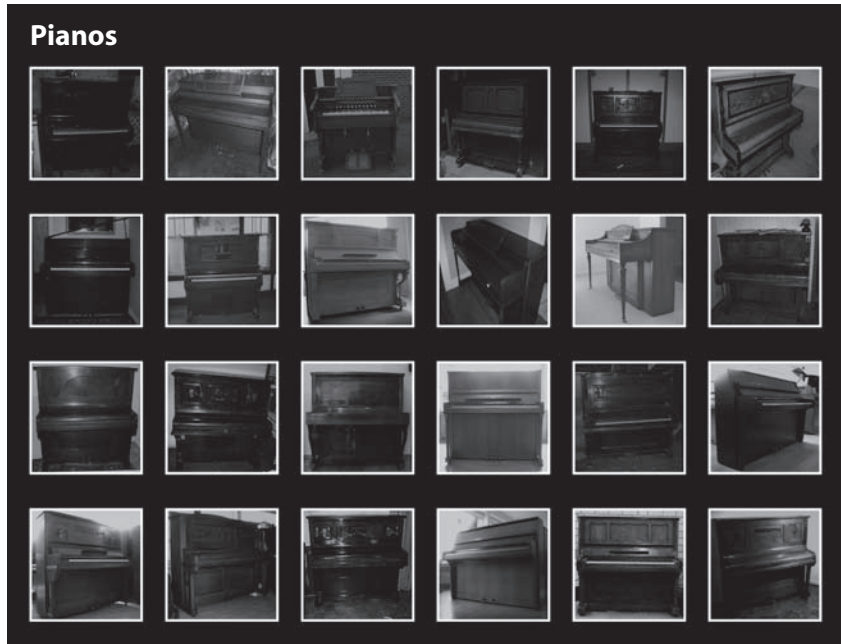


Figure B-3: Images of pianos used for fMRI masking experiment. The figure displays the final selection of 24 different images (400 x 400 pixels) of pianos (for selection procedure see above).



Figure B-4: Images of chairs used for fMRI masking experiment. The figure displays the final selection of 24 different images (400 x 400 pixels) of chairs (for selection procedure see above).

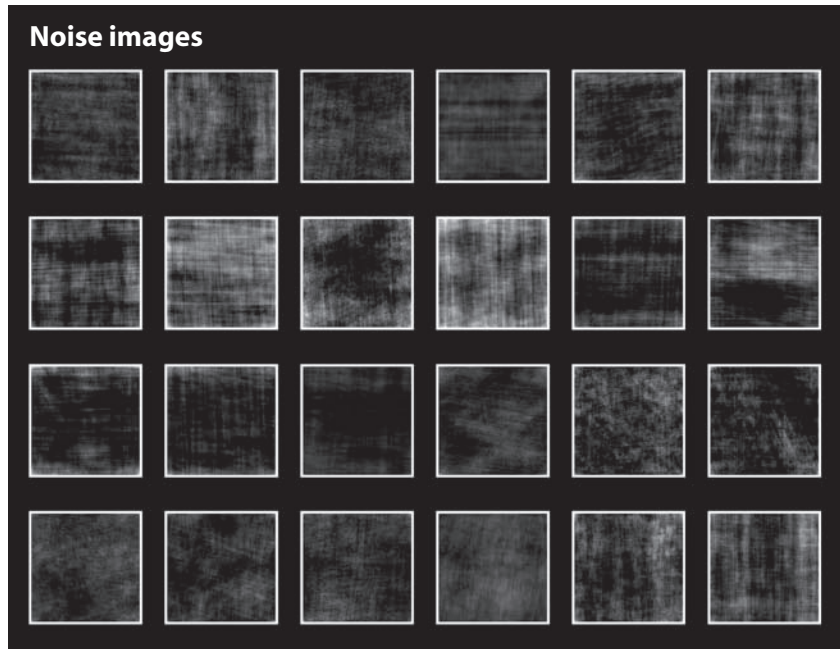


Figure B-5: Phase-randomised noise images used for fMRI masking experiment. The figure displays the final selection of 24 different noise images (400 x 400 pixels; for selection procedure see above). The images were created from all piano and chair images using a phase-randomisation procedure following a 2-D Fourier transformation (for details see Methods).

Confusion matrix in perceptual decision making

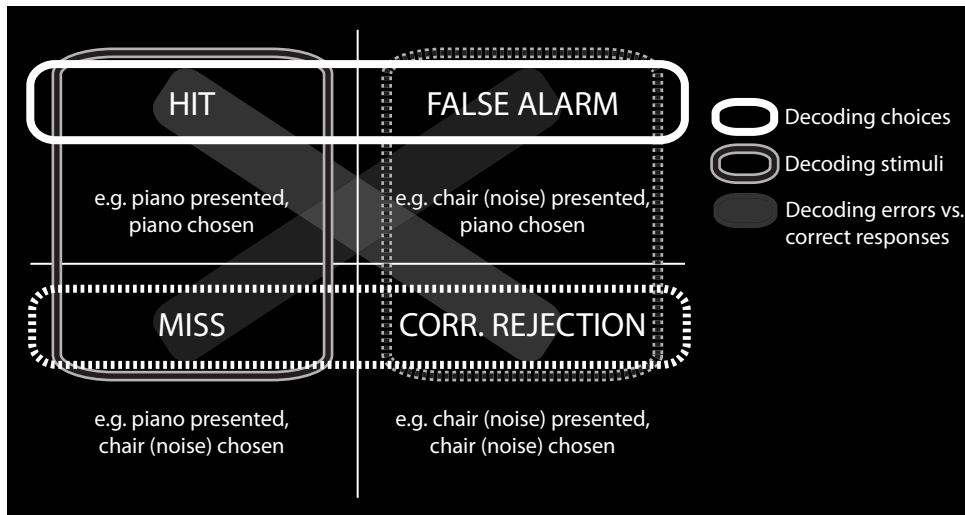


Figure B-6: Confusion matrix in perceptual decision making. For each category (here illustrated for pianos as the target category), four cases are possible when objects have to be classified. “Hit” refers to correct classification of an object from the target category; “false alarm” refers to the incorrect classification of the object as belonging to the target category while a different category was presented; “correct rejection” refers to the classification as a non-target category object while a non-target category object was presented; “miss” refers to the classification as a non-target category object while the target category was presented. Note that in the present study, there were always two non-target categories such that object categories have to be sorted individually. The decoding analyses averaged across each combination of the two cells (choices, independent from real stimuli; stimuli, independent from choices; errors vs. correct responses). Cell-wise analyses could not be performed due to the small number (and uneven distribution) of trials per cell.

Mean hit rates for fMRI masking experiment

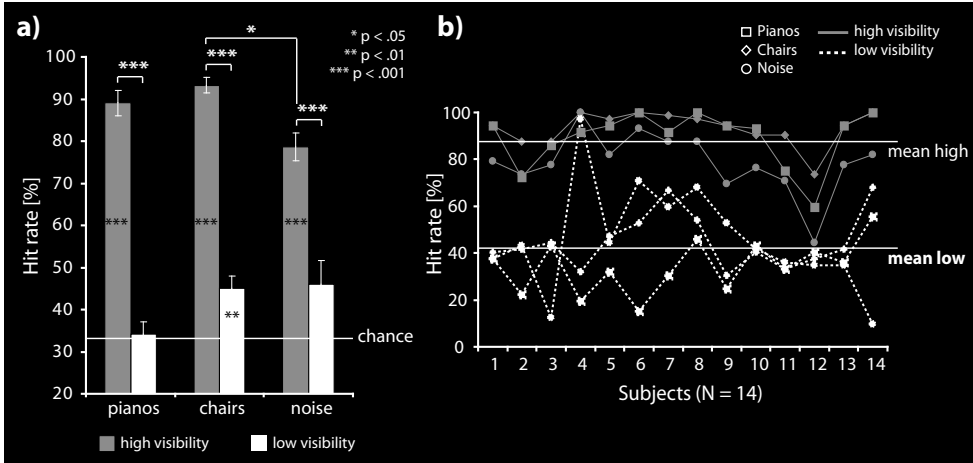


Figure B-7: Mean hit rates for object categories. a) Behavioural data was analysed by calculating hit rates as a visibility index for each condition (high visibility = dark bars; low visibility = bright bars). Displayed are mean hit rates and standard errors for the final sample (N=14). The statistical results were highly comparable to the analyses by means of d' values. Following an ANOVA for repeated measures [$F(5,65) = 57.71$; $\text{Eta}^2 = 0.75$; $p < .001$], post-hoc Sheffé-tests correcting for multiple comparisons revealed significant differences between high and low visibility for all object categories ($p < .001$ for all comparisons; pianos: high visibility 89%, SD = 2.99; low visibility 34%, SD = 2.79; chairs: high visibility 93%, SD = 1.80; low visibility: 45%, SD = 2.93; noise: high visibility 79%, SD = 3.24; low visibility 46%, SD = 5.74). Highly visible objects achieved significant hit rates above chance level ($p < .001$ for all). In the low-visibility condition, only the hit rate for chairs slightly exceeded chance level (pianos: $p > .05$; chairs: $p < .01$; noise: $p > .05$). b) Individual visibility indices. The graph displays individual hit rates for each participant for both visibility conditions and all three categories. Dark lines: categories in “high visibility” condition (mean 87.0%; SE 2.68); bright lines: categories in “low visibility” condition (mean 41.7%; SE 3.82). Nearly all participants showed a clear separation between visibility conditions (exception participant 4 in one case). The same category was never perceived better under “low visibility” compared to “high visibility” by any of the participants.

Univariate fMRI data analysis

Conventional univariate analyses were used to identify brain regions where the overall level of activity was significantly increased during the presentation of objects shown in different visibility conditions; it was also used to contrast processing of real object images (pianos and chairs) with noise images in both visibility conditions. For these analyses the data of the main experiment underwent motion correction, spatial normalisation to a standard stereotaxic space (Montreal Neurological Institute EPI template) and re-sampling to an isotropic spatial resolution of $3 \times 3 \times 3 \text{ mm}^3$ with SPM2. Additionally, data were smoothed with a Gaussian kernel of 8 mm FWHM to account for anatomical variability and to satisfy the assumption of Gaussian random field theory. A general linear model (GLM) as implemented in SPM2 was used. It consisted of six boxcar regressors, each convolved with a canonical haemodynamic response function (HRF). These modelled the presentation of pianos, chairs and noise images (individually for “high visibility” and “low visibility”).

For the “high visibility” condition, stronger activations for *real objects compared to noise images* were found exclusively in left LOC and in a region in right anterior temporal cortex, which has been linked to semantic memory and conceptual knowledge about socially used stimuli (Simmons & Martin, 2009). No region showed stronger activations for real objects compared to noise images under low visibility. Also, no region could be found that showed the inverse effect of stronger activations for noise images than for real object images, neither in the “high visibility” nor in the “low visibility” condition. Next, the contrast of “high visibility” versus “low visibility” was calculated *across all categories*. Only the same region in right anterior temporal cortex showed stronger activation for “high visibility” than for “low visibility”. Conversely, a region in dorso-medial frontal cortex in the rostral cingulate zone (RCZ) was stronger activated in the “low visibility” condition compared to the “high visibility” condition for all categories (all results are reported in Table B-2, Figure B-8). The RCZ and adjacent regions have been linked to different aspects of decision uncertainty (Nachev et al., 2008) and guessing (Elliott, Rees & Dolan, 1999).

Next, it was investigated if any region was more activated for error trials compared to correct trials and vice versa. Again, a GLM was used as described above. It consisted of three boxcar regressors, each convolved with a canonical HRF. These modelled (i) the *correct trials* for the low visibility condition, (ii) the *error trials* for the low visibility

condition as well as (iii) *all trials* for the high visibility condition (there were not sufficient errors under high visibility to allow an independent estimation of correct / error). Only the first two regressors were of interest. A group level analysis was conducted as described above. No region showed a difference between correct trials and error trials for the low visibility condition, even when a very low threshold of $p < .01$ (uncorrected; 10 voxels threshold) was applied. Additionally, the analyses were restricted to three regions of interest (ROIs), namely the left LOC [-21 -51 6], the right LOC [27 -39 -12] and the precuneus [-9 -63 30] defined by the results of the independent pattern classification analysis for participants' choices (applying a threshold of $p < .0001$ for ROI definition). Again, no brain region showed such a difference (Figure B-9), supporting the interpretation that the images have not been processed at the object-level when strong masking was used.

Table B-2: Table of activation contrasts for univariate analysis.

Contrast	Anatomical area	L/R	T	p <	x	y	z
Objects > Noise (all)	---						
Noise > Objects (all)	---						
Objects > Noise (HighVis)	LOC	L	3.45	.001	-18	-39	-12
Noise > Objects (HighVis)	---						
Objects > Noise (LowVis)	---						
Noise > Objects (LowVis)	---						
Low > HighVis (all)	medial frontal / RCZ	L/R	5.52	.05 *	-6	21	45
High > LowVis (all)	anterior temporal	R	5.34	.05 *	21	6	-18
Low > HighVis (Objects)	medial frontal / RCZ	L/R	5.59	.05 *	-6	21	42
High > LowVis (Objects)	anterior temporal	R	4.10	.001	21	6	-18
	LOC	L	4.07	.001	-24	-48	-9
Low > HighVis (noise)	---						
High > LowVis (noise)	---						

Note: * FWE corrected; LowVis = "low visibility" condition; HighVis = "high visibility" condition; LOC = lateral-occipital complex; RCZ = rostral cingulate zone; Objects = images of pianos and chairs; Noise = phase-randomised noise images; voxel threshold = 5 voxels in all analysis; coordinates refer to the MNI coordinate system.

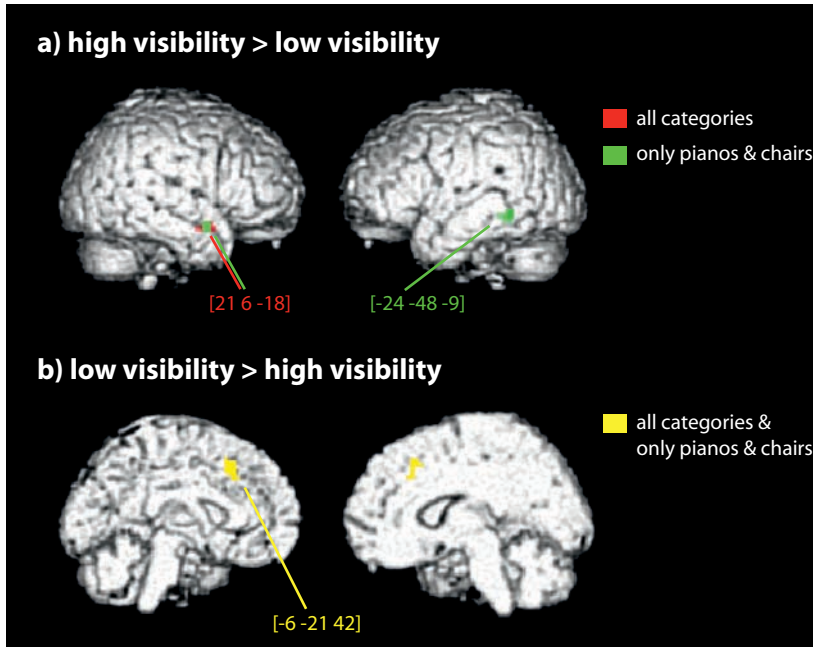


Figure B-8: Activation contrasts for visual stimulation. a) Contrast of activations related to stimuli presented under high visibility versus low visibility. When all categories were included, a region in right anterior temporal cortex showed stronger activation for high visibility compared to low visibility (red; MNI [21 6 -18]; $p < .05$ FWE-corrected). When only the *real* object categories (pianos and chairs) were used for the same contrast, this region again demonstrated stronger activation but at a lower statistical threshold, along with a cluster in left LOC (green; MNI [-24 -48 -9]; $p < .001$ uncorrected) b) Areas more strongly activated under low visibility than under high visibility. Regardless of whether all categories or only real object categories were included in the analysis, only one area located in medial frontal cortex / rostral cingulate zone (RCZ) was more strongly activated (yellow; MNI [-6 -21 42]; $p < .05$ FWE-corrected). Coordinates are MNI coordinates.

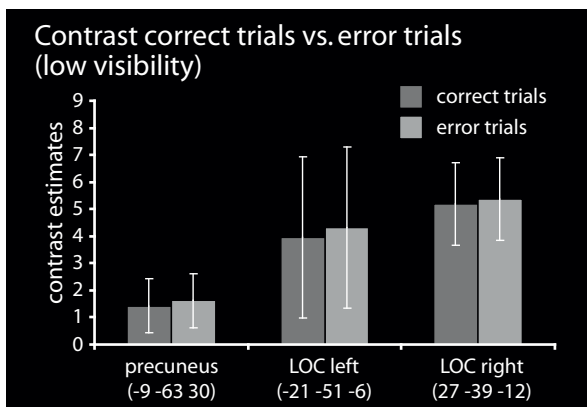


Figure B-9: Results of univariate error analyses for low visibility condition. a) Contrast estimates for correct trials (dark) and incorrect trials (light) for peak voxels of searchlight-clusters from which decoding of participants' choices was possible. None of these regions displayed any differences between correct and incorrect trials ($p > .01$ uncorrected). No other region could be found displaying such a difference.

Two-class pattern classification for presented objects

Decoding participants' categorical choices in the "high visibility" condition was possible from regions in bilateral LOC while decoding the *presented* object categories was possible from a wider range of visual areas (compare Figures 4-4 and 4-5). Additional multivariate decoding analyses were therefore performed to test the hypothesis that the wide spread of regions in the second case was caused by greater differences in low-level features between the real objects (pianos and chairs) compared to the noise images. In these analyses, only two categories of objects were used for classification at a time, thus, analyses for all three combinations of paired categories were conducted separately. The process of classification followed the same searchlight approach as described above (see Methods). The only difference was that by using two instead of three classes at a time, the chance level for decoding accuracy was 50%. The resulting brain maps of decoding accuracies of the two analyses using one real object category and the noise category were averaged (both based on the same number of images per category), preventing the use of uneven distributions in pattern vectors for single decoding analyses. Hence, the analyses resulted in two decoding accuracy maps: "*decoding objects*" and "*decoding objects from noise*".

Using only real object categories (pianos and chairs), decoding was possible for the "high visibility" condition from left LOC (see Figure 4-5b). Decoding object from noise images under high visibility was possible from a wider range of bilateral visual regions and LOC (Table B-3), highly overlapping with the regions found for three-class searchlight decoding. Again, under "low visibility" conditions no region allowed decoding. Taken together, these results confirmed that patterns in several visual areas could distinguish between visible objects and non-objects, pointing to the importance of differences in low-level features of real objects and noise images (Kay et al., 2008; Walther et al., 2009). The LOC, on the other hand, was the only region revealing information encoding about distinct object categories, underlining its exclusive role in processing object category information in the ventral visual pathway.

Table B-3: Results of two-class decoding analyses for object categories.

Decoding	Anatomical area	L/R	Accuracy		T	p <	x	y	z
			M	SE					
Obj only (LowVis)	---								
Obj only (HighVis)	LOC	L	58	1.6	5.26	.00001	-27	-60	-6
Obj – Noise (LowVis)	---								
Obj – Noise (HighVis)	VisCortex	L	66	1.8	8.93	.05*	-36	-87	12
		R	60	1.3	7.78	.05*	45	-84	-6
	dorsal VisC	L	62	1.6	7.75	.05*	-36	-54	42
		R	63	1.6	7.77	.05*	3	-81	30
	LOC	L	67	2.0	8.42	.05*	-36	-45	-6
		R	71	2.5	8.39	.05*	33	-45	-24

Note: * FWE corrected; LowVis = “low visibility” condition; HighVis = “high visibility” condition; LOC = lateral-occipital complex; VisCortex = visual cortex; Obj = images of pianos and chairs; Noise = phase-randomised noise images; Voxel threshold = 5 voxels in all analysis; in the second part of the table, averaged decoding accuracies from two analyses (piano and noise images / chair and noise images) are reported. Coordinates refer to the MNI coordinate system.

Correlation analysis of confusion matrices

Using an additional approach, the detailed *confusion matrix* of classification results produced by the support vector machine classifier was correlated with the overall confusion matrices of the participants' categorical choices. In other words, the analysis searched for regions in which the overall classifier performance (the errors of the classifier) and overall behavioural performance (the errors of the participants) showed a similar pattern, again separately for each visibility condition (for a similar approach see Walther et al., 2009). First, a confusion matrix was created for each condition, containing the percentage of how often the classifier predicted *the chosen category* to be piano, chair or noise *with respect to the presentation of each category*. Such a matrix was compiled for the classifier's results of each searchlight cluster at each voxel position in the brain. The same was done for the *participants' behavioural choices* with respect to the presented objects for the whole experiment, separately for both visibility conditions. For each searchlight cluster, the classifier's and the participants' confusion matrices were transformed into two-dimensional vectors and correlated. A Fisher's Z-transformation was performed on the correlation coefficients. The Z-values were reconstructed into 3-dimensional brain maps for each participant. These were again normalised to the MNI EPI template and smoothed with a Gaussian kernel of 6 mm FWHM, allowing for group-level voxel-wise statistical analyses.

This approach, however, only allowed a general correlation analysis between the classifier's and participants' performance across the entire experiment. Nevertheless, it was able to mirror the *similarity of performance including the classification errors*, which could systematically vary between presented object categories and is not reflected in the global decoding accuracy value. Regions displaying a high similarity between the pattern classifier's performance and the participants' performance across all participants were located in bilateral visual cortex, the LOC and left orbito-frontal cortex when objects were highly visible (see Figure B-10b, green regions). Orbito-frontal cortex has been linked to object recognition before (Bar, 2004). For poorly visible objects, a high correlation could be found in the precuneus, located slightly dorsal to the precuneus cluster found for searchlight decoding. A second region was located on medial frontal gyrus (Figure B-10b, red regions). The analyses again revealed no overlap between the high and low visibility conditions.

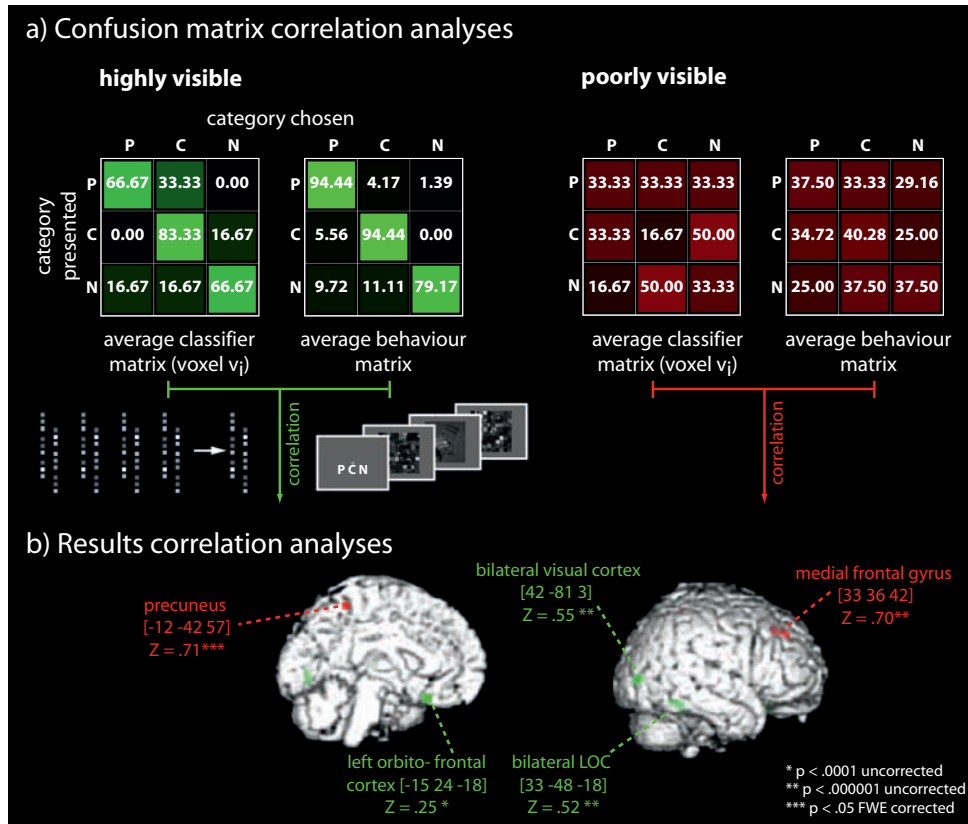


Figure B-10: Correlation analysis of confusion matrices. a) Illustration of the confusion matrix approach by means of one participant's data (classifier data taken from arbitrary searchlight clusters in the LOC and the precuneus). The pattern classifier's performance for each searchlight cluster was transformed into a confusion matrix, displaying the classification errors for the entire experiment. The same transformation was done with participants' categorical choices relative to the presented object. For each participant, one confusion matrix was conducted for the high visibility condition (left side in green) and one for the low visibility condition (right side in red). Since participants' choices closely matched the presented categories for highly visible objects, both matrices had to depict high values in their diagonal cells in order to be highly correlated (left side). Additionally, similarities in tendency to systematic classification errors also added to a high correlation between the matrices. For poorly visible objects high correlations depended on similarities in overall performance of classifier and participant, hence similar errors (right side). b) Results: High correlations could be found for highly visible objects in bilateral visual cortex (peak: MNI [42 -81 3]; Z = .55; p < .000001, uncorrected), bilateral LOC (peak: MNI [33 -48 -18]; Z = .52; p < .000001, uncorrected) and left orbito-frontal cortex (peak: MNI [-15 24 -18]; Z = .25; p < .0001, uncorrected). For the low visibility condition, high correlations could be found in the precuneus (peak: MNI [-12 -42 57]; Z = .71; p < .05, FWE corrected) and right medial frontal gyrus (peak: MNI [33 36 42]; Z = .70; p < .05, FWE corrected).

Individual searchlight patterns for decoding of choices

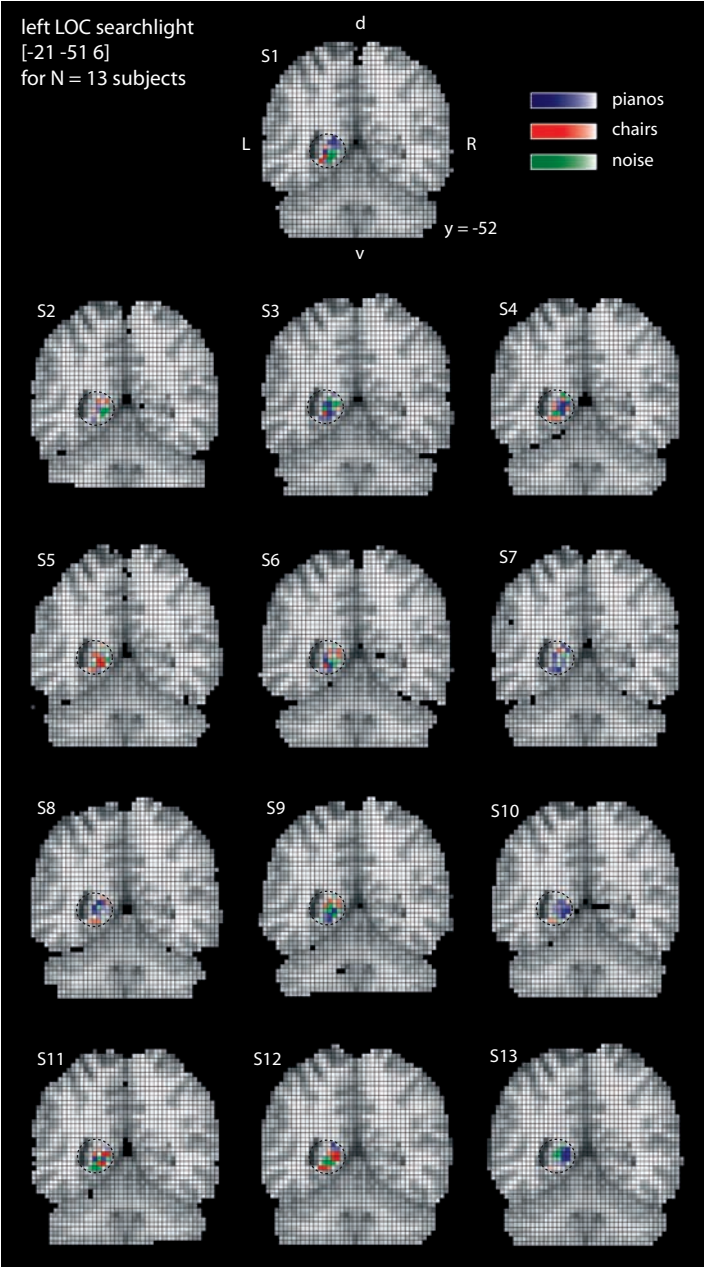


Figure B-11: Individual searchlight patterns in LOC. The individual searchlight patterns are rendered onto the coronal slice containing the cluster from which decoding of participants' category choices for highly visible objects was most robust. The displayed cluster belongs to the medial temporal cortex, extending to the LOC. Grey-matter voxels are colour-coded according to the category choice that elicited the strongest BOLD response (blue = pianos; red = chairs; green = noise). Activation values are re-scaled and transformed into colour values; strong saturation represents a stronger activation for a single decision. The patterns were stable within participants but unique for each participant. One participant was excluded from the analysis because the noise category was never selected in one run.

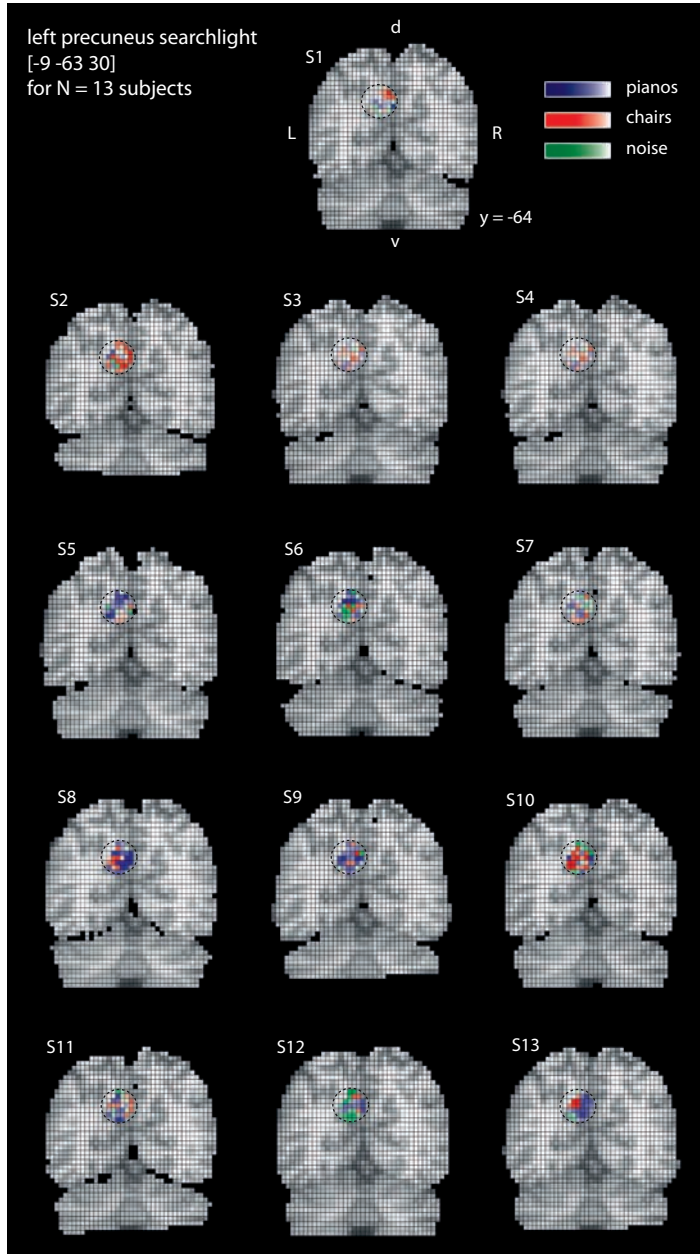


Figure B-12: Individual searchlight patterns in the precuneus. The individual searchlight patterns are rendered onto the coronal slice containing the cluster from which decoding of participants' choices for poorly visible objects was most robust. The displayed cluster belongs to the precuneus. Grey-matter voxels are colour-coded according to the category choice that elicited the strongest BOLD response (blue = pianos; red = chairs; green = noise). Activation values are re-scaled and transformed into colour values; strong saturation represents a stronger activation for a single decision. The patterns were stable within participants but unique for each participant. One participant was excluded from the analysis because the noise category was never selected in one run.

Behavioural control experiment: Decision confidence and hit rates

A behavioural control experiment was conducted outside the scanner with a different sample of 11 participants (6 female, mean age 27.0; range 18-31). All of them were right-handed, had normal or corrected to normal visual acuity, did not take part in any of the other experiments and were paid 7 € for their participation. Participants performed three runs of the main experiment (see Methods) under similar viewing conditions (see pre-test I for setup and material). Participants additionally indicated their decision confidence on a five-point scale (1 = “no confidence” to 5 = “very confident”) after each trial. Responses were given via keyboard (buttons 1-3 for categories; buttons 1-5 for confidence rating). Visibility was strongly associated with decision confidence (for details see Figure B-13).

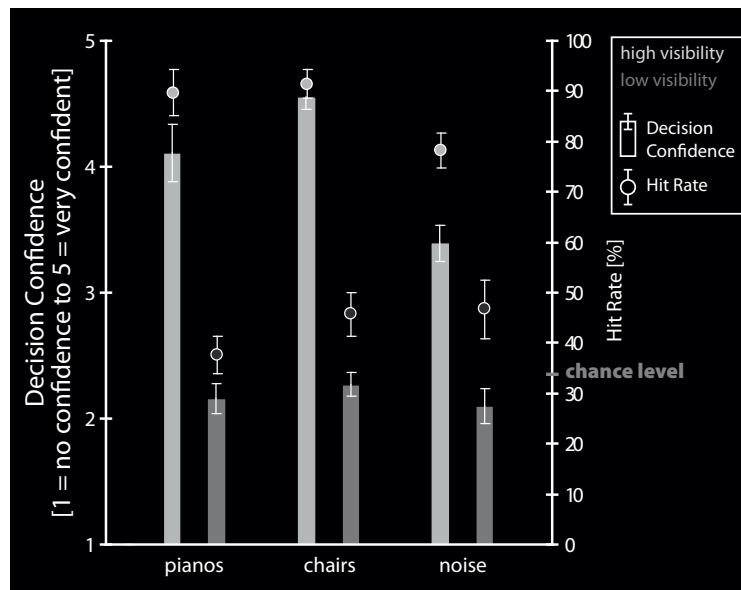


Figure B-13: Decision confidence and visibility. Eleven different participants were presented with 3 runs of the main experiment outside the scanner. They additionally indicated their decision confidence (1 = “no confidence” to 5 = “very confident”) after each trial. The graph displays mean confidence ratings and hit rates with standard errors (SE). Confidence ratings mirrored hit rates with significantly higher decision confidence for highly visible objects compared to poorly visible objects for each category (two-tailed t-tests: pianos: $t(10) = 9.51$, $p < .001$; chairs: $t(10) = 18.05$, $p < .001$; noise: $t(10) = 11.06$, $p < .001$). The correlation (Pearson’s r) between hit rates and decision confidence was highly significant for each category (pianos: $r = .89$, $p < .001$; chairs: $r = .90$, $p < .001$; noise: $r = .76$, $p < .001$).

Decoding results from conventional LOC ROIs

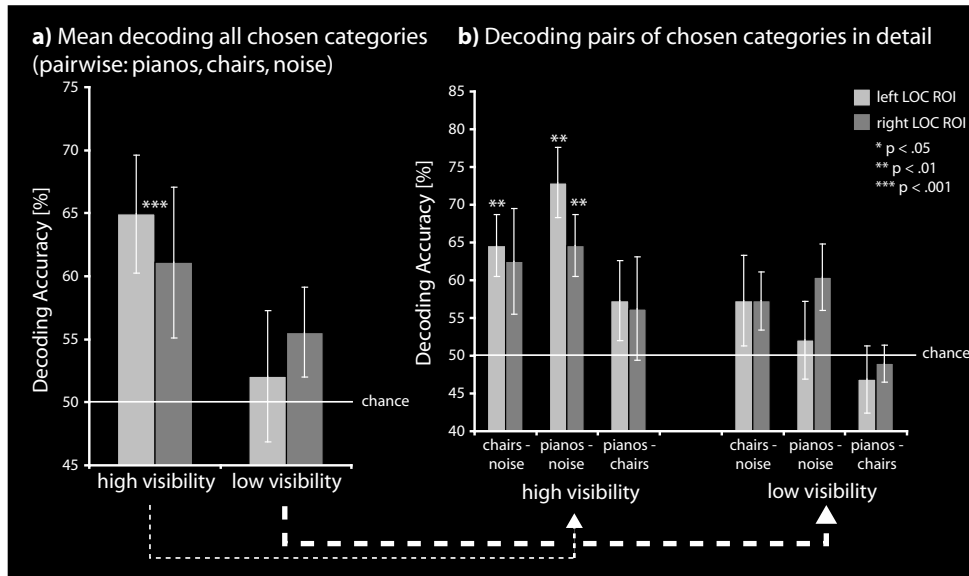


Figure B-14: Decoding category choices from LOC based on localizer-defined ROIs. a) Decoding accuracies are averaged for the three pair-wise classification processes of the three categories (chance level always 50%) for $n = 8$ participants. Bars are standard errors (SE). Left LOC showed significant decoding accuracy [65%; $t(7) = 5.94$; $p < .001$] and right LOC just missed the threshold, still displaying a high accuracy of 61% [$t(7) = 2.10$; $p = .07$] for the “high visibility” condition. For the “low visibility” condition, left LOC [52%; $t(7) = 0.61$; $p = .56$] as well as right LOC [56%; $t(7) = 1.68$; $p = .13$] did not encode the categorical choices. b) Detailed analyses of the single pair-wise classification processes: The analyses revealed that for the “high visibility” condition, above chance classification was possible from left LOC for chairs and noise at 65% [$t(7) = 3.56$; $p < .01$] and pianos and noise at 73% [$t(7) = 4.92$; $p < .01$] but barely for pianos and chairs [57%; $t(7) = 1.37$; $p = .21$]. The right LOC consistently showed lower values: 65% accuracy for pianos and noise [$t(7) = 3.56$; $p < .01$] and non-significant accuracies for chairs and noise [63%; $t(7) = 1.77$; $p = .12$] as well as pianos and chairs [56%; $t(7) = 0.91$; $p = .39$]. For the “low visibility” condition, neither the left LOC nor the right LOC was predictive above threshold for any category. In sum, the left LOC was especially predictive for participants’ choices of categories in the “high visibility” condition while LOC was not predictive for choices in the “low visibility” condition. The statistical power was reduced compared to the main analyses due to the smaller sample size. These findings, however, generally confirm the results from searchlight decoding in the LOC regions.

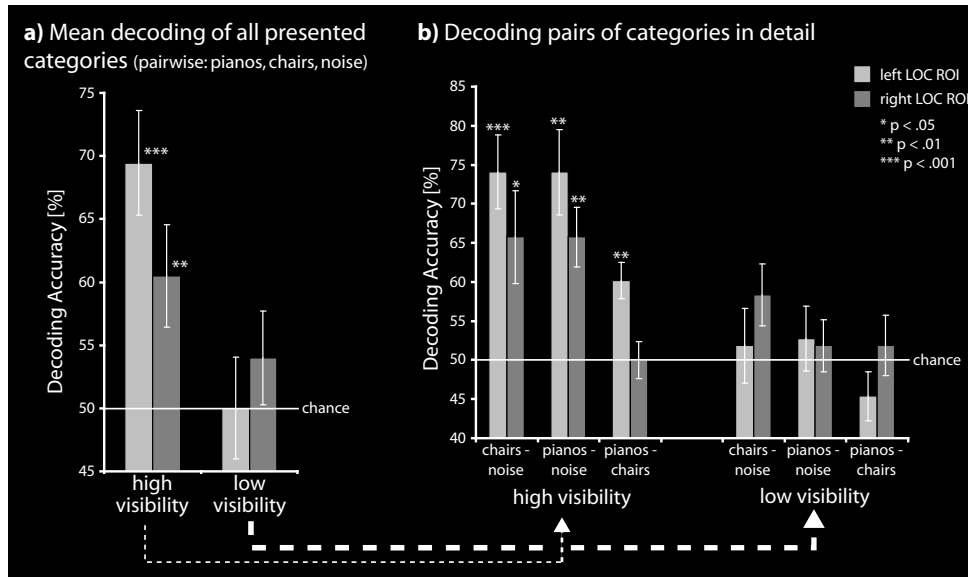


Figure B-15: Decoding *presented categories* from LOC based on localizer-defined ROIs. a) Decoding accuracies are averaged for the three pair-wise classification processes of the three categories (chance level 50%) for $n = 9$ participants. Bars are standard errors (SE). The statistical power was reduced compared to the main analyses due to the smaller sample size. Significant decoding accuracies above chance, however, were found for both the left LOC [69%; $t(8) = 7.67$; $p < .001$] and the right LOC [60%; $t(8) = 3.42$; $p < .01$] for the “high visibility” condition. For the “low visibility” condition, decoding accuracies for left LOC [50%; $t(8) = .01$; $p = .99$] and right LOC [54%; $t(8) = 1.55$; $p = .16$] did not exceed chance level. Post hoc tests revealed that the difference in decoding accuracy between the “high visibility” and the “low visibility” condition was significant ($p < .001$) for the left LOC. b) For the “high visibility” condition, the single pair-wise analyses revealed that left LOC was always more informative than the right LOC, confirming the results from the searchlight decoding analysis. Separate t-tests revealed that above chance classification was possible from left LOC for chairs and noise at 74% [$t(8) = 5.12$; $p < .001$], pianos and noise at 74% accuracy [$t(8) = 4.41$; $p < .01$] as well as pianos and chairs at 60% [$t(8) = 4.40$; $p < .01$] for the “high visibility” condition. The right LOC consistently showed lower accuracies: 66% for chairs and noise [$t(8) = 2.64$; $p < .05$], 66% accuracy for pianos and noise [$t(8) = 4.15$; $p < .01$] and 50% for pianos and chairs, which did not reach statistical significance [$t(8) = 0.01$; $p = .99$]. For the “low visibility” condition, neither the left LOC nor the right LOC was predictive above threshold in any category analysis.

Comparison of precuneus regions with Soon et al. (2008)

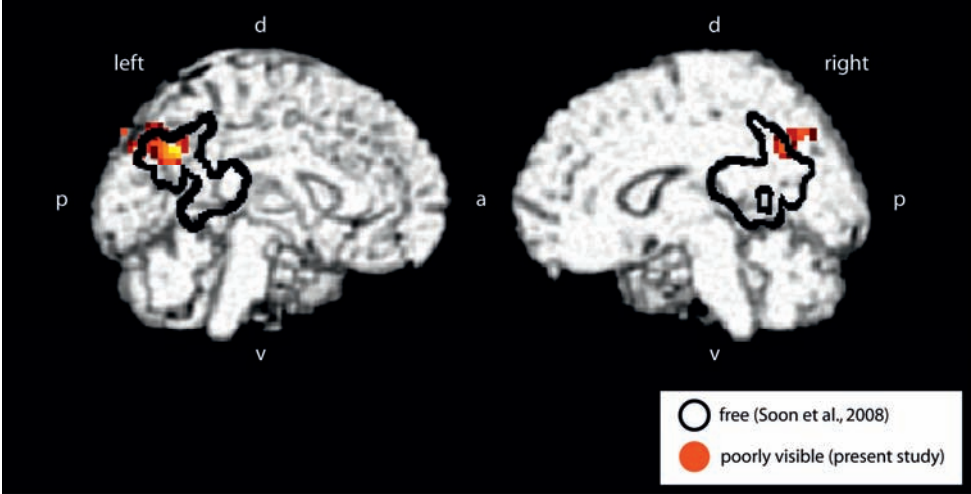


Figure B-16: Comparison of precuneus regions. In red: Medial view of precuneus region that encoded perceptual choices made with insufficient visual input in the present study. Black line: Medial view of precuneus region that encoded free decisions in Soon et al. (2008). Data from the present study and from Soon et al. (2008) are both displayed using a threshold of $p < .0001$ (uncorrected) for better visualisation. The regions display substantial overlap but the cluster from Soon et al. (2008) extends more anterior to posterior cingulate cortex (PPC). Differences might also be due to differences in modelling and the analysed time periods.

Appendix C

Behavioural pre-test

Behavioural pre-tests were conducted in order to assess the suitability of the experimental stimuli. First, they must create the illusion that hard-to-detect object stimuli were presented in all trials. Also, if participants experienced the two-fold scaling of visibility they would have given up on hard categorisation trials and only performed on easy categorisation trials. Second, the stimuli must not induce a systematic bias in participants' choices towards any particular category. A sample of nine independent participants (5 female; mean age 24.8; range 20-30; right handed, normal or corrected to normal visual acuity; 7 Euros for participation) took part in the experiment outside the scanner. The setup was the same as for the pre-tests of the previous study (17" TFT monitor, refresh rate of 60 Hz, resolution 1024 x 768 pixel screen; presentation using the Cogent toolbox for MATLAB 7.0; estimated angle of vision $\alpha \sim 7.2^\circ$ comparable to the fMRI experiment). The paradigm was the same as for the fMRI experiment (see Methods), using *noise images* for invisible categorisation (IV) and free decisions (FD). Participants consecutively performed two experimental blocks, which tested two variants of the paradigm (short: 66.7 ms target durations vs. long: 83.3 ms target duration for HV, while IV and FD were unchanged in both variations). The order of variants was randomised between participants. Participants were not informed about these differences. After performing each block, a self-conducted questionnaire was used to assess details about their impressions (Figure C-3). Finally, they were informed about the experimental manipulation and asked if they noticed any missing object stimuli.

Hit rates for the HV condition under both variants were comparably high. Participants were nearly equally well balanced in for the IV and the FD condition, close to 50% ($p > .05$ for all t-tests) (see Figure C-1). The shorter variant was perceived as harder by trend and was associated with slightly earlier decisions for the FD task (see Figure C-2). The illusion achieved by the noise stimuli was successful: participants made decisions for object categories in free decision trials and did not notice the missing objects. Given the slightly more challenging nature of the short duration variant and better comparability with the former study, the short duration variant was used for the fMRI study.

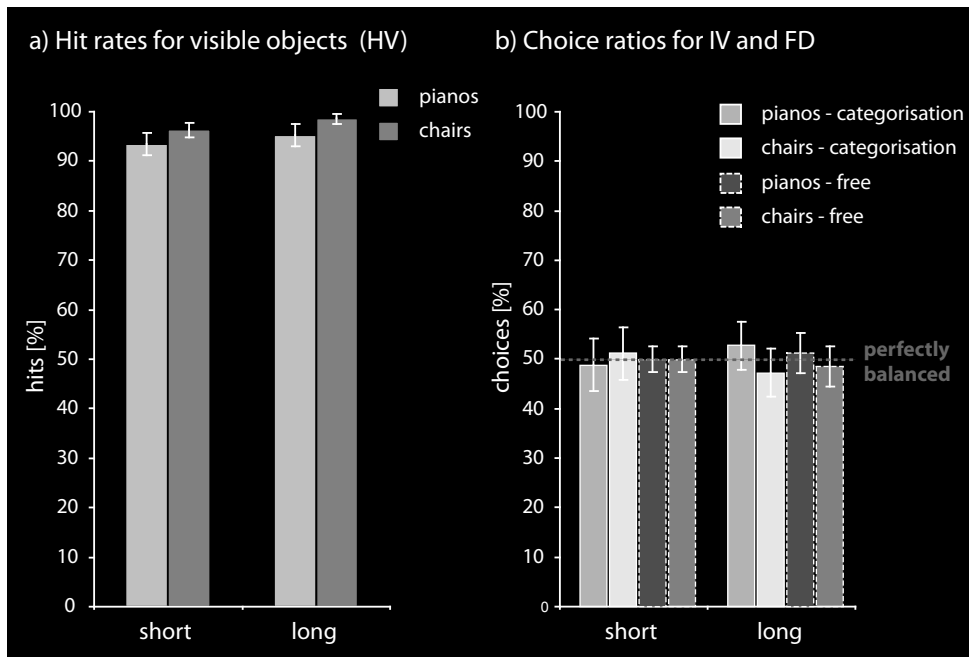


Figure C-1: Results of behavioural pre-test. a) Hit rates in the categorisation task for highly visible objects (HV). Displayed are the mean hit rates and standard errors for all participants ($n = 9$) using both variants of the paradigm (left side: “short”, 66.7 ms target duration and 433.3 ms post-mask duration; right side: “long”, 83.3 ms target duration and 417.7 ms post-mask duration). Hit rates were 93.5% for pianos and 96.3% for chairs (short) and 95.4% for pianos and 98.6% for chairs (long), indicating that the “short” variant was slightly more challenging, however not significant. b) Choice ratios for invisible detection task (IV) and the free decision (FD) task for both variants of the paradigm. The variants did not differ in stimulation for these conditions. Participants were close to perfectly balanced on average (dotted line) in all conditions for both variants (short: 49% pianos and 51% chairs for IV task, 50% pianos and 50% chairs for FD task; long: 53% pianos and 47% chairs for IV task, 51% pianos and 49% chairs for FD task).

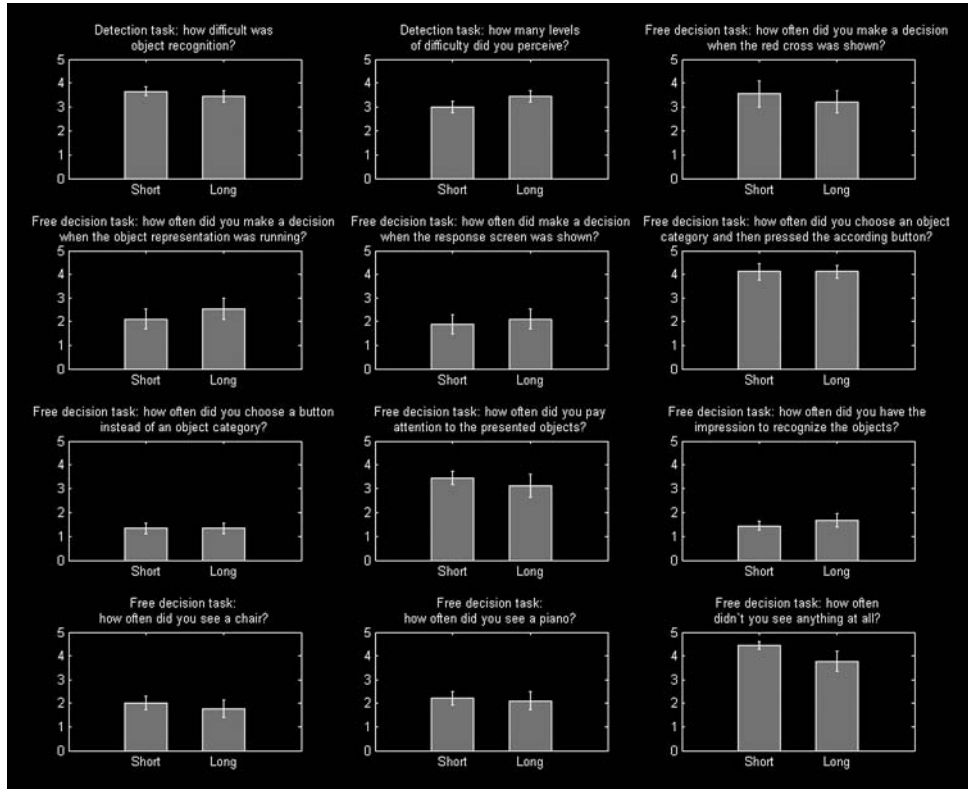


Figure C-2: Comparison of questionnaire results between paradigm variants. Given are the mean values for participants' approval to the question on a five-point scale (from "1 = never" ranging to "5 = always") after the pre-test. Questions 1-12 are sorted from left to right. The mean values for the questions are: question 1 (short: $M = 3.7$, $SD = 0.5$; long: $M = 3.4$, $SD = 0.7$), question 2 (short: $M = 3.0$, $SD = 0.7$; long: $M = 3.4$, $SD = 0.7$), question 3 (short: $M = 3.6$, $SD = 1.7$; long: $M = 3.2$, $SD = 1.4$), question 4 (short: $M = 2.1$, $SD = 1.3$; long: $M = 2.6$, $SD = 1.3$), question 5 (short: $M = 1.9$, $SD = 1.3$; long: $M = 2.1$, $SD = 1.3$), question 6 (short: $M = 4.1$, $SD = 1.1$; long: $M = 4.1$, $SD = 0.8$), question 7 (short: $M = 1.3$, $SD = 0.7$; long: $M = 1.3$, $SD = 0.7$), question 8 (short: $M = 3.4$, $SD = 0.9$; long: $M = 3.1$, $SD = 1.5$), question 9 (short: $M = 1.4$, $SD = 0.5$; long: $M = 1.6$, $SD = 0.9$), question 10 (short: $M = 2.0$, $SD = 0.9$; long: $M = 1.8$, $SD = 1.1$), question 11 (short: $M = 2.2$, $SD = 0.8$; long: $M = 2.1$, $SD = 1.2$), question 12 (short: $M = 4.4$, $SD = 0.5$; long: $M = 3.8$, $SD = 1.3$).

Results questionnaire fMRI experiment

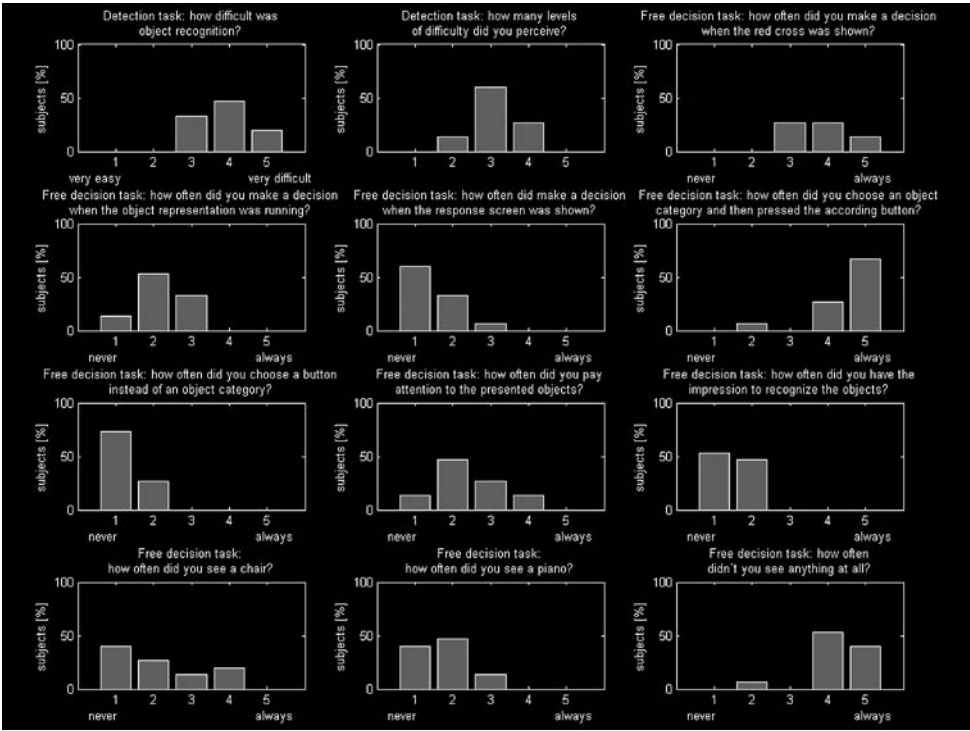


Figure C-4: Results from post-experimental questionnaire for fMRI experiment. The questionnaire was the same as for the pre-test. It asked for participants' rating on a five-point scale ("1 = never" to "5 = always"). Questions are sorted from left to right. The graphs show the distribution of participants' responses in percent (n = 15 = 100%). The categorisation/detection task was perceived as rather difficult (question 1: 3 = 33.3%; 4 = 46.7%; 5 = 20%). Most participants thought that at least three levels of difficulty were used, confirming the successful illusion (question 2: 2 = 13.3%; 3 = 60%; 4 = 26.7%). Free decisions were made rather early in the trial (question 3: free decisions during cue presentation: 3 = 26.7%; 4 = 26.7%; 5 = 13.3%; question 4: free decisions during target presentation: 1 = 13.3%; 2 = 53.3%; 3 = 33.3%; question 5: free decisions during response mapping presentation: 1 = 60%; 2 = 33.3%; 3 = 6.7%). Free decisions were nearly always made for object categories instead of response buttons, as instructed (question 6: free decisions for object categories: 2 = 6.7%; 4 = 26.7%; 5 = 66.7%; question 7: free decisions for buttons instead of categories: 1 = 73.3%; 2 = 26.7%). Participants did not pay much attention to the presented objects during free decision trials (question 8; 1 = 13.3%; 2 = 46.7%; 3 = 26.7%; 4 = 13.3%) and did not believe they identified any presented objects (question 9: 1 = 53.3%; 2 = 46.7%). The majority seldom believed that they identified a chair during free decisions (question 10: 1 = 40%; 2 = 26.7%; 3 = 13.3%; 4 = 20%) or a piano (question 11: 1 = 40%; 2 = 46.7%; 3 = 13.3%). Instead, the majority indicated that they had not seen anything (question 12: 2 = 6.7%; 4 = 53.3%; 5 = 40%).

Choice ratios fMRI experiment

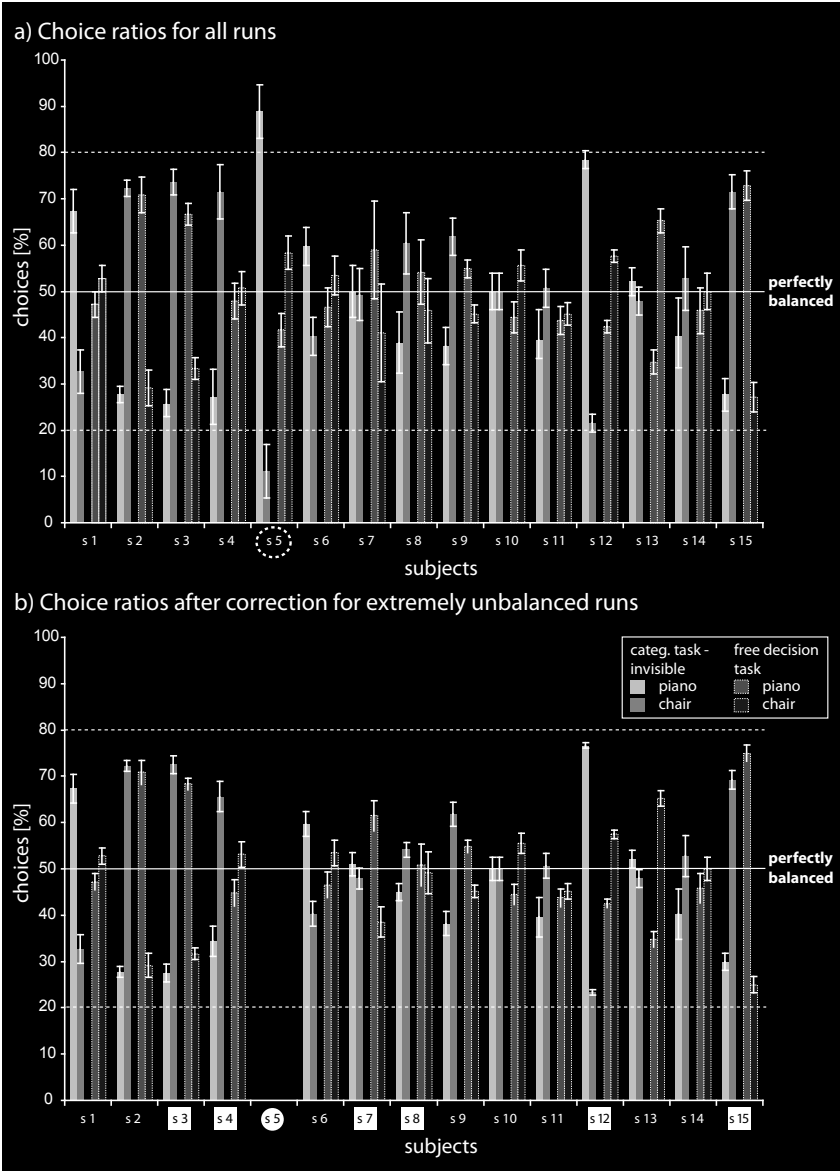


Figure C-5: Individual choice ratios from fMRI experiment. Choice ratios for pianos (bright) and chairs (dark) for invisible categorisation (IV; no frame) and free decisions (FD; white frame). a) Ratios for all participants including all runs. One participant was extremely imbalanced (s5: > 80:20 for pianos-chairs in IV, dotted line) and was excluded completely. b) Ratios after excluding s5 (circle) and individual unacceptably imbalanced runs from six participants (boxes; 2 runs excluded for s4, s7; 1 run excluded for s3, s8, s12, s15). Runs were considered unacceptably imbalanced if the decision ratio was more extreme than 20:80 or 80:20 for one category. Individual decision ratios were closer to optimal decision ratio of 50:50 (solid line) after correction.

Choice ratios and decoding accuracy

After excluding imbalanced individual participants and single runs, there were still differences in decision ratios between participants (see Figure C-5). These differences are a logical consequence of the guessing and free decision task and speak *for* randomness in decisions rather than against it. There was no trend towards one choice option across participants. Nevertheless, stronger imbalance might have influenced the decoding accuracy. To control for this possibility, additional analyses were conducted. The 14 participants were divided into two groups with respect to their decision balance. The absolute difference in percentage of decisions (averaged across IV and FD) was taken as a balance indicator (BI):

$$BI = (| \text{pianos}(\%) - \text{chairs}(\%) |)$$

Groups (BI_{high} mean = 26.19%; SD = 8.45; BI_{low} mean = 9.23; SD = 4.41) were divided by median split (BI = 10%). Two-tailed t-tests for independent samples were conducted to test whether these groups differed in decoding accuracies for i) average searchlight decoding results for the whole brain volume (pointing to *general* methodological problems caused by decision imbalance) or ii) decoding results for the best searchlight clusters in precuneus and MPFC (pointing to *specific* regional biases caused by decision imbalance).

There were no significant differences in the overall decoding accuracies between the two groups, neither for the IV condition (BI_{high} 51.7%, SD 2.45; BI_{low} 49.6%, SD = 2.71; $t(6) = 1.41$; $p = 0.21$) nor for the FD condition (BI_{high} 51.2%, SD 2.24; BI_{low} 50.6%, SD = 2.40; $t(6) = 0.44$; $p = 0.68$).

There were also no significant differences in decoding accuracies for the specific searchlight clusters for guessing decisions in a) precuneus / PCC (MNI 24 -45 18; BI_{high} 60.5%, SD = 12.60; BI_{low} 57.6%, SD = 4.29; $t(6) = 0.60$; $p = 0.57$) and b) precuneus / PPC (MNI 30 -63 51; BI_{high} 63.2%, SD = 10.37; BI_{low} 58.1%, SD = 10.15; $t(6) = 1.29$; $p = 0.25$). Again, no differences could be found for decoding accuracies for free decisions in MPFC (MNI 3 48 21; BI_{high} 61.3%, SD = 9.21; BI_{low} 65.9%, SD = 8.15; $t(6) = -1.06$; $p = 0.33$). The same picture emerged when BIs were calculated for each condition (IV, FD) separately to compare BI_{high} and BI_{low} groups ($p > .05$ for all comparisons). It can therefore be concluded that differences in decision ratios cannot explain the present findings.

Univariate fMRI data analysis

Conventional univariate analyses were conducted on the fMRI data to investigate whether any areas showed differences in general activation when participants were presented with (or chose) pianos compared to chairs. Additionally, areas should be identified that differed in activation for the HV task, the IV task and the FD task. First, all data was motion corrected to the first volume of the first run using SPM2. All volumes were then normalised to MNI space, re-sampled to a voxel size of $3 \times 3 \times 3 \text{ mm}^3$ and smoothed with a Gaussian kernel of 8 mm FWHM. A GLM was used for statistical analyses on subject level. Modelling was the same as described for the multivariate analyses, resulting in separate regressors for piano and chair trials, individually for each experimental condition and each run. T-tests were used for random effects group statistics and again a statistical threshold of $p < .0001$ (uncorrected) was applied. Contrasts between pianos and chairs were calculated for each condition. The three experimental conditions were contrasted across object categories.

For all experimental conditions, strong activation could be found in large parts of the brain when contrasted against baseline. The overall peak was consistently located in bilateral visual cortex and LOC (see Table C-1). No region was found to show stronger activation related to the presentation (or choice) of pianos compared to chairs or vice versa in any condition. Contrasting FD trials with perceptual decision trials in general (HV as well as IV), a network of brain regions could be revealed that showed stronger activation for FD. This network included bilateral inferior parietal lobe, bilateral precuneus, bilateral ventrolateral prefrontal cortex (VLPFC), bilateral dorsolateral prefrontal cortex (DLPFC) and the left inferior temporal sulcus. These findings are in line with an earlier study using Positron Emission Tomography (PET) that revealed a similar network for self-initiated actions (Jenkins, et al., 2000). Contrasting HV with FD demonstrated stronger activation for HV in bilateral visual cortex and LOC as well as in the right orbito-frontal cortex (see Bar, 2004). No other contrast revealed any significant differences for any of the conditions (for all results see Table C-1). Interestingly, the information-encoding parietal regions were again not strongly activated during guessing but during free decisions. Given that object categorisation was reported to be challenging, it is unlikely that low task difficulty can explain these findings. It rather demonstrates again that guessing and free decisions differed in important aspects and cannot be regarded as being completely identical.

Table C-1: Table of activation contrasts for univariate analysis.

Contrast	Anatomical area	cluster size §	T	p <	x	y	z
HV > IV	-						
IV > HV	-						
HV > FD	R LOC	4	5.03	.00001	36	-39	-12
	L LOC	7	4.43	.0001	-33	-39	-15
	R orbito-frontal	8	4.7	.0001	24	36	-12
FD > HV	R DLPFC	402	7.65	.05*	33	36	45
	L inf parietal lobe	397	7.54	.05*	-60	-48	33
	R inf parietal lobe	302	7.31	.05*	57	-60	39
	bil precuneus	167	7.3	.05*	15	-54	36
	L DLPFC	191	7.07	.05*	-36	24	45
	L inf temp sulcus	140	6.82	.05*	-60	-36	-6
	L VLPFC	92	6.5	.05*	-51	27	-6
	R VLPFC	26	4.77	.0001	57	18	-6
IV > FD	-						
FD > IV	L inf parietal lobe	366	7.32	.05*	-57	-57	30
	L VLPFC	35	6.25	.05*	-51	27	-6
	R inf parietal lobe	198	6.18	.05*	57	-60	39
	L DLPFC	95	6.01	.05*	-33	27	48
	L inf temp sulcus	62	5.83	.05*	-60	-36	-9
	bil precuneus	54	5.6	.05*	12	-54	36
	R DLPFC	42	5.19	.05*	30	39	42
overall peak							
HV > baseline	bil vis cortex/LOC		22.44	.05*	30	-93	3
IV > baseline	bil vis cortex/LOC		20.73	.05*	-27	-96	3
FD > baseline	bil vis cortex/LOC		19.55	.05*	-27	-96	3

Note: HV = perceptual decision task with highly visible objects; IV = perceptual decision task without visible objects (invisible); FD = free decision task; L = left; R = right; inf temp sulcus = inferior temporal sulcus; LOC = lateral-occipital complex; DLPFC = dorsolateral prefrontal cortex; VLPFC = ventrolateral prefrontal cortex; bil vis cortex = bilateral visual cortex; * FWE (family-wise error) correction for multiple comparisons; § cluster size is always given for a statistical threshold of $p < .0001$ uncorrected. Coordinates are MNI coordinates.

Decoding motor responses and visual stimulation

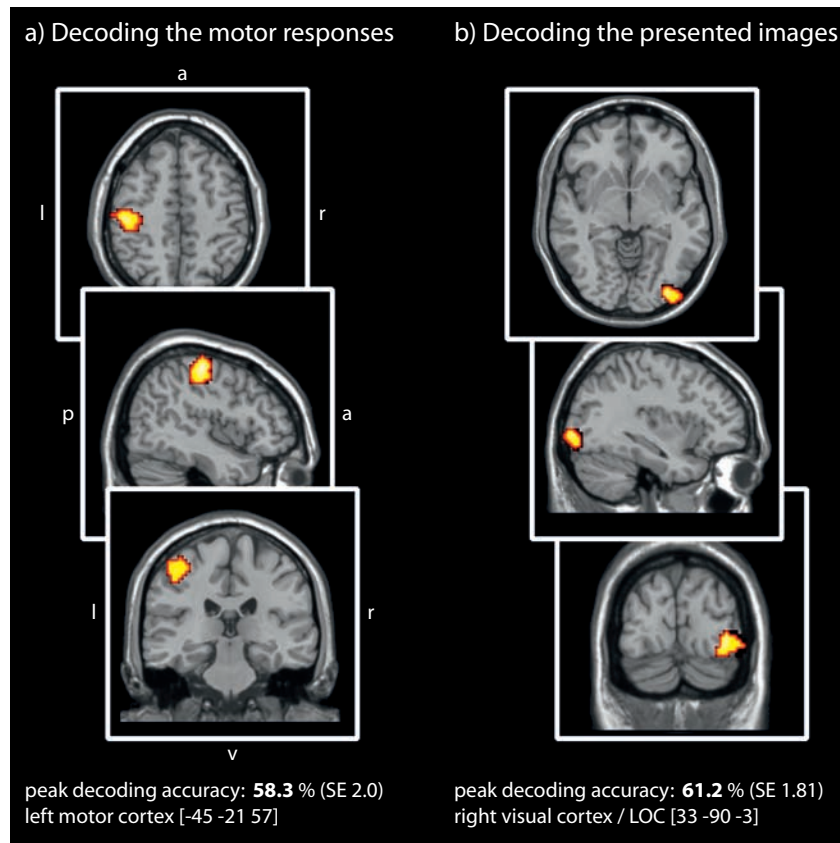


Figure C-6: Decoding of motor responses and true object categories. Displayed are the clusters showing the highest decoding accuracies for the respective analyses using a searchlight pattern classification approach with a radius of $r = 4$ voxels. All results are displayed using a threshold of $p < .001$ (uncorrected) for better visualisation. a) Button presses were performed using the index- and middle finger of the right hand. Motor responses could be decoded from left motor cortex (peak accuracy 58%; SE = 2.0; $p < .0001$ uncorrected; MNI [-45 -21 57]) for all experimental conditions. The analysis confirmed that the participants' categorical choices were made on a category level rather than being motor decisions (which would have been expected for non-adherent response behaviour). b) The peak decoding accuracy for true presented object categories (pianos and chairs) was located in the right visual cortex extending to LOC (also found in the left hemisphere when a more liberal statistical threshold was applied). The peak accuracy was 61% (SE = 1.8; $p < .0001$ uncorrected; MNI [33 -90 -3]). These regions showed a great overlap with visual/LOC regions found to encode perceptual choices for highly visible objects.

ROI analyses

To investigate how similar the present posterior parietal regions were in terms of the *quality* of information encoding compared to the posterior parietal regions found in the previous study, ROI analyses were conducted. Only those voxels were sampled for the analysis of the present data that constituted the significant searchlight clusters in the *previous* study. In order to define the ROIs, an uncorrected threshold of $p < .0001$ (precuneus; decoding choices for poorly visible objects) and $p < .05$ FWE corrected (LOC; decoding choices for highly visible objects) was applied for the searchlight decoding group-analyses from the previous study. The left and right LOC ROIs were combined to a bilateral LOC ROI (238 voxels) in order to be comparable in size with the precuneus ROI (360 voxels). Since the ROIs had to be applied to normalised data (decoding was performed on individual non-normalised data but group analyses were performed on normalised accuracy maps), the ROIs were used to extract searchlight decoding *results* from normalised decoding accuracy maps of the present study. The decoding accuracies from each ROI were averaged and tested against chance level for each condition using ANOVA and t-tests. These averaged accuracies strongly underestimate the information encoded in the ROIs, but it can be used as an approximation for the comparison of searchlight decoding *profiles* from both studies. Note that any effects were expected to be small due to averaging values from a relatively high number of voxels in each ROI. These included mainly voxels that were not significant in the searchlight decoding analyses.

The LOC ROI was found to be informative about choices in the HV condition [$t(11) = 2.66$; $p = .006$], but the precuneus ROI was not [$t(13) = 1.11$; $p = .14$]. For the IV condition, the LOC ROI was not informative [$t(11) = 0.92$; $p = .18$] but the precuneus ROI was indeed weakly informative about decision outcomes, marginally missing the statistical threshold [$t(13) = 1.35$; $p = .09$]. As expected from the searchlight results of the present study, neither ROI encoded decision outcomes for free decisions [LOC ROI: $t(11) = 0.92$; $p = .18$; precuneus ROI: $t(13) = 0.77$; $p = .22$]. Nevertheless, cross-condition classification between FD and IV, incorporating both directions of cross-classification in an ANOVA, was significant for the precuneus ROI [$F(1,12) = 4.01$; $p = .03$] but not for the LOC ROI [$F(1,12) = 0.74$; $p = .49$]. In summary, these results yielded a similar profile of results for ROIs derived from the previous study compared to the whole-brain searchlight decoding results from the present study. Choices for poorly visible and truly invisible objects (and to some extent for free decisions) were similarly encoded in a rather wide-spread region in precuneus / PPC.

List of Publications

Bode, S., Bogler, C., Soon, C.S. & Haynes, J.D. (under review). Similar neural mechanisms for guesses and free choices.

Bode, S. & Haynes, J.D. (2009). Decoding sequential stages of task preparation in the human brain. *Neuroimage*, 45(2), 606-613.

Bode, S., He, A.H., Soon, C.S., Trampel, R., Turner, R. & Haynes, J.D. (in prep.). Tracking the unconscious generation of free decisions using ultra-high field fMRI.

Bode, S., Koeneke, S. & Jäncke, L. (2007). Different strategies do not moderate primary motor cortex involvement in mental rotation: a TMS study. *Behavioral and Brain Functions*, 3:38.

Bogler, C., Bode, S. & Haynes, J.D. (in prep.). Multivariate decoding reveals successive computational stages of saliency processing.

Müller, A., Bode, S., Myer, L., Roux, P. & von Steinbüchel, N. (2008). Electronic measurement of adherence to paediatric antiretroviral therapy in South Africa. *The Paediatric Infectious Disease Journal*, 27(3), 257-262.

Müller, A.D., Bode, S., Myer, L., Stahl, J. & von Steinbüchel, N. (2010). Predictors of adherence to antiretroviral treatment and therapeutic success among children in South Africa. *AIDS Care*, in press.

Soon, C.S., He, A.H., Bode, S. & Haynes, J.D. (in prep.). Unconscious formation of voluntary abstract intentions.

Tusche, A., Bode, S. & Haynes, J.D. (2010). Neural responses to unattended products predict later consumer choices. *Journal of Neuroscience*, 30(23), 8024-8031.

BIBLIOGRAPHISCHE DARSTELLUNG

Bode, Stefan

From stimuli to motor responses: decoding rules and decision mechanisms in the human brain

Fakultät für Biowissenschaften, Pharmazie und Psychologie,

Universität Leipzig

Dissertation

153 Seiten (plus 44 Seiten Appendix), 362 Literaturangaben, 45 Abbildungen, 9 Tabellen

Abstract

In a dynamically changing environment, we are constantly required to flexibly react to stimuli. It is therefore necessary to adapt behaviour to environmental cues, as well as to successfully perceive relevant stimuli. The present work addressed the question of which brain areas form the basis for task preparation and decisions along the processing chain from stimuli to responses. It combined functional magnetic resonance imaging with multivariate pattern classification to search for the encoding of specific *contents* of mental processes. The first study demonstrated, using a task switching paradigm, that task-sets were first encoded in left intraparietal sulcus, preceding left posterior ventrolateral prefrontal cortex. This finding emphasises the importance of parietal cortex in establishing abstract rules in a cued task context. In the second study, the visibility of response-relevant target objects was varied. It was found that the lateral-occipital complex (LOC) only encoded perceptual decisions about highly visible objects. The precuneus, on the other hand, encoded random guessing decisions made with insufficient visual input. Contradicting classical signal detection models, this finding emphasises the notion of two modes for perceptual decision making depending on stimulus visibility. The third study demonstrated a shared neural substrate for random guessing and free decisions in the precuneus, suggesting a general role for the generation of internal decisions. Additionally, anterior medial prefrontal cortex was exclusively engaged when decisions were intended to be made without an external frame of reference. In summary, the present work highlights the importance of parietal cortex in controlling both rule-guided and self-determined behavior in humans. Parietal cortex functions might be best described as related to multi-modal information integration and processing of highly self-referenced intentions.

Selbständigkeitserklärung

Hiermit erkläre ich, dass die vorliegende Arbeit ohne unzulässige Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt wurde und dass die aus fremden Quellen direkt oder indirekt übernommenen Gedanken in der Arbeit als solche kenntlich gemacht worden sind.

Stefan Bode

Leipzig, den 20.02.2010

MPI Series in Human Cognitive and Brain Sciences:

- 1 Anja Hahne
Charakteristika syntaktischer und semantischer Prozesse bei der auditiv Sprachverarbeitung: Evidenz aus ereignis-korrelierten Potentialstudien
- 2 Ricarda Schubotz
Erinnern kurzer Zeitdauern: Behaviorale und neurophysiologische Korrelate einer Arbeitsgedächtnisfunktion
- 3 Volker Bosch
Das Halten von Information im Arbeitsgedächtnis: Dissoziationen langsamer corticaler Potentiale
- 4 Jorge Jovicich
An investigation of the use of Gradient- and Spin-Echo (GRASE) imaging for functional MRI of the human brain
- 5 Rosemary C. Dymond
Spatial Specificity and Temporal Accuracy in Functional Magnetic Resonance Investigations
- 6 Stefan Zysset
Eine experimentalpsychologische Studie zu Gedächtnisabrufprozessen unter Verwendung der funktionellen Magnetresonanztomographie
- 7 Ulrich Hartmann
Ein mechanisches Finite-Elemente-Modell des menschlichen Kopfes
- 8 Bertram Opitz
Funktionelle Neuroanatomie der Verarbeitung einfacher und komplexer akustischer Reize: Integration haemodynamischer und elektrophysiologischer Maße
- 9 Gisela Müller-Plath
Formale Modellierung visueller Suchstrategien mit Anwendungen bei der Lokalisation von Hirnfunktionen und in der Diagnostik von Aufmerksamkeitsstörungen
- 10 Thomas Jacobsen
Characteristics of processing morphological structural and inherent case in language comprehension
- 11 Stefan Kölsch
*Brain and Music
A contribution to the investigation of central auditory processing with a new electrophysiological approach*
- 12 Stefan Frisch
Verb-Argument-Struktur, Kasus und thematische Interpretation beim Sprachverstehen
- 13 Markus Ullsperger
The role of retrieval inhibition in directed forgetting – an event-related brain potential analysis
- 14 Martin Koch
Measurement of the Self-Diffusion Tensor of Water in the Human Brain
- 15 Axel Hutt
Methoden zur Untersuchung der Dynamik raumzeitlicher Signale
- 16 Frithjof Kruggel
Detektion und Quantifizierung von Hirnaktivität mit der funktionellen Magnetresonanztomographie
- 17 Anja Dove
Lokalisierung an internen Kontrollprozessen beteiligter Hirngebiete mithilfe des Aufgabenwechselfaradigmas und der ereigniskorrelierten funktionellen Magnetresonanztomographie
- 18 Karsten Steinhauer
Hirphysiologische Korrelate prosodischer Satzverarbeitung bei gesprochener und geschriebener Sprache
- 19 Silke Urban
Verbinformationen im Satzverstehen
- 20 Katja Werheid
Implizites Sequenzlernen bei Morbus Parkinson
- 21 Doreen Nessler
Is it Memory or Illusion? Electrophysiological Characteristics of True and False Recognition

- 22 Christoph Herrmann
Die Bedeutung von 40-Hz-Oszillationen für kognitive Prozesse
- 23 Christian Fiebach
Working Memory and Syntax during Sentence Processing. A neurocognitive investigation with event-related brain potentials and functional magnetic resonance imaging
- 24 Grit Hein
Lokalisation von Doppelaufgabendefiziten bei gesunden älteren Personen und neurologischen Patienten
- 25 Monica de Filippis
Die visuelle Verarbeitung unbeachteter Wörter. Ein elektrophysiologischer Ansatz
- 26 Ulrich Müller
Die katecholaminerge Modulation präfrontaler kognitiver Funktionen beim Menschen
- 27 Kristina Uhl
Kontrollfunktion des Arbeitsgedächtnisses über interferierende Information
- 28 Ina Bornkessel
The Argument Dependency Model: A Neurocognitive Approach to Incremental Interpretation
- 29 Sonja Lattner
Neurophysiologische Untersuchungen zur auditorischen Verarbeitung von Stimminformationen
- 30 Christin Grünewald
Die Rolle motorischer Schemata bei der Objektrepräsentation: Untersuchungen mit funktioneller Magnetresonanztomographie
- 31 Annett Schirmer
Emotional Speech Perception: Electrophysiological Insights into the Processing of Emotional Prosody and Word Valence in Men and Women
- 32 André J. Szameitat
Die Funktionalität des lateral-präfrontalen Cortex für die Verarbeitung von Doppelaufgaben
- 33 Susanne Wagner
Verbales Arbeitsgedächtnis und die Verarbeitung ambiger Wörter in Wort- und Satzkontexten
- 34 Sophie Manthey
Hirn und Handlung: Untersuchung der Handlungsrepräsentation im ventralen prämotorischen Cortex mit Hilfe der funktionellen Magnet-Resonanz-Tomographie
- 35 Stefan Heim
Towards a Common Neural Network Model of Language Production and Comprehension: fMRI Evidence for the Processing of Phonological and Syntactic Information in Single Words
- 36 Claudia Friedrich
Prosody and spoken word recognition: Behavioral and ERP correlates
- 37 Ulrike Lex
Sprachlateralisierung bei Rechts- und Linkshändern mit funktioneller Magnetresonanztomographie
- 38 Thomas Arnold
Computergestützte Befundung klinischer Elektroenzephalogramme
- 39 Carsten H. Wolters
Influence of Tissue Conductivity Inhomogeneity and Anisotropy on EEG/MEG based Source Localization in the Human Brain
- 40 Ansgar Hantsch
Fisch oder Karpfen? Lexikale Aktivierung von Benennungsalternativen bei der Objektbenennung
- 41 Peggy Bungert
*Zentralnervöse Verarbeitung akustischer Informationen
Signalidentifikation, Signallateralisation und zeitgebundene Informationsverarbeitung bei Patienten mit erworbenen Hirnschädigungen*
- 42 Daniel Senkowski
Neuronal correlates of selective attention: An investigation of electro-physiological brain responses in the EEG and MEG

- 43 Gert Wollny
Analysis of Changes in Temporal Series of Medical Images
- 44 Angelika Wolf
Sprachverstehen mit Cochlea-Implantat: EKP-Studien mit postlingual ertaubten erwachsenen CI-Trägern
- 45 Kirsten G. Volz
Brain correlates of uncertain decisions: Types and degrees of uncertainty
- 46 Hagen Huttner
Magnetresonanztomographische Untersuchungen über die anatomische Variabilität des Frontallappens des menschlichen Großhirns
- 47 Dirk Köster
Morphology and Spoken Word Comprehension: Electrophysiological Investigations of Internal Compound Structure
- 48 Claudia A. Hruska
Einflüsse kontextueller und prosodischer Informationen in der auditorischen Satzverarbeitung: Untersuchungen mit ereigniskorrelierten Hirnpotentialen
- 49 Hannes Ruge
Eine Analyse des raum-zeitlichen Musters neuronaler Aktivierung im Aufgabenwechselparadigma zur Untersuchung handlungssteuernder Prozesse
- 50 Ricarda I. Schubotz
Human premotor cortex: Beyond motor performance
- 51 Clemens von Zerssen
Bewusstes Erinnern und falsches Wiedererkennen: Eine funktionelle MRT Studie neuroanatomischer Gedächtniskorrelate
- 52 Christiane Weber
Rhythm is gonna get you. Electrophysiological markers of rhythmic processing in infants with and without risk for Specific Language Impairment (SLI)
- 53 Marc Schönwiesner
Functional Mapping of Basic Acoustic Parameters in the Human Central Auditory System
- 54 Katja Fiehler
Temporospatial characteristics of error correction
- 55 Britta Stolterfoht
Processing Word Order Variations and Ellipses: The Interplay of Syntax and Information Structure during Sentence Comprehension
- 56 Claudia Danielmeier
Neuronale Grundlagen der Interferenz zwischen Handlung und visueller Wahrnehmung
- 57 Margret Hund-Georgiadis
Die Organisation von Sprache und ihre Reorganisation bei ausgewählten, neurologischen Erkrankungen gemessen mit funktioneller Magnetresonanztomographie – Einflüsse von Händigkeit, Läsion, Performanz und Perfusion
- 58 Jutta L. Mueller
Mechanisms of auditory sentence comprehension in first and second language: An electrophysiological miniature grammar study
- 59 Franziska Biedermann
Auditorische Diskriminationsleistungen nach unilateralen Läsionen im Di- und Telenzephalon
- 60 Shirley-Ann Rüschemeyer
The Processing of Lexical Semantic and Syntactic Information in Spoken Sentences: Neuroimaging and Behavioral Studies of Native and Non-Native Speakers
- 61 Kerstin Leuckefeld
The Development of Argument Processing Mechanisms in German. An Electrophysiological Investigation with School-Aged Children and Adults
- 62 Axel Christian Kühn
Bestimmung der Lateralisierung von Sprachprozessen unter besondere Berücksichtigung des temporalen Cortex, gemessen mit fMRT

- 63 Ann Pannekamp
Prosodische Informationsverarbeitung bei normalsprachlichem und deviantem Satzmaterial: Untersuchungen mit ereigniskorrelierten Hirnpotentialen
- 64 Jan Derrfuß
Functional specialization in the lateral frontal cortex: The role of the inferior frontal junction in cognitive control
- 65 Andrea Mona Philipp
The cognitive representation of tasks – Exploring the role of response modalities using the task-switching paradigm
- 66 Ulrike Toepel
Contrastive Topic and Focus Information in Discourse – Prosodic Realisation and Electrophysiological Brain Correlates
- 67 Karsten Müller
Die Anwendung von Spektral- und Waveletanalyse zur Untersuchung der Dynamik von BOLD-Zeitreihen verschiedener Hirnareale
- 68 Sonja A.Kotz
The role of the basal ganglia in auditory language processing: Evidence from ERP lesion studies and functional neuroimaging
- 69 Sonja Rossi
The role of proficiency in syntactic second language processing: Evidence from event-related brain potentials in German and Italian
- 70 Birte U. Forstmann
Behavioral and neural correlates of endogenous control processes in task switching
- 71 Silke Paulmann
Electrophysiological Evidence on the Processing of Emotional Prosody: Insights from Healthy and Patient Populations
- 72 Matthias L. Schroeter
Enlightening the Brain – Optical Imaging in Cognitive Neuroscience
- 73 Julia Reinholz
Interhemispheric interaction in object- and word-related visual areas
- 74 Evelyn C. Ferstl
The Functional Neuroanatomy of Text Comprehension
- 75 Miriam Gade
Aufgabeneinhibition als Mechanismus der Konfliktreduktion zwischen Aufgabenrepräsentationen
- 76 Juliane Hofmann
Phonological, Morphological, and Semantic Aspects of Grammatical Gender Processing in German
- 77 Petra Augurzky
Attaching Relative Clauses in German – The Role of Implicit and Explicit Prosody in Sentence Processing
- 78 Uta Wolfensteller
Habituelle und arbiträre sensomotorische Verknüpfungen im lateralen prämotorischen Kortex des Menschen
- 79 Päivi Sivonen
Event-related brain activation in speech perception: From sensory to cognitive processes
- 80 Yun Nan
Music phrase structure perception: the neural basis, the effects of acculturation and musical training
- 81 Katrin Schulze
Neural Correlates of Working Memory for Verbal and Tonal Stimuli in Nonmusicians and Musicians With and Without Absolute Pitch
- 82 Korinna Eckstein
Interaktion von Syntax und Prosodie beim Sprachverstehen: Untersuchungen anhand ereigniskorrelierter Hirnpotentiale
- 83 Florian Th. Siebörger
Funktionelle Neuroanatomie des Textverstehens: Kohärenzbildung bei Witzen und anderen ungewöhnlichen Texten

- 84 Diana Böttger
Aktivität im Gamma-Frequenzbereich des EEG: Einfluss demographischer Faktoren und kognitiver Korrelate
- 85 Jörg Bahlmann
Neural correlates of the processing of linear and hierarchical artificial grammar rules: Electrophysiological and neuroimaging studies
- 86 Jan Zwickel
Specific Interference Effects Between Temporally Overlapping Action and Perception
- 87 Markus Ullsperger
Functional Neuroanatomy of Performance Monitoring: fMRI, ERP, and Patient Studies
- 88 Susanne Dietrich
Vom Brüllen zum Wort – MRT-Studien zur kognitiven Verarbeitung emotionaler Vokalisationen
- 89 Maren Schmidt-Kassow
What's Beat got to do with ist? The Influence of Meter on Syntactic Processing: ERP Evidence from Healthy and Patient populations
- 90 Monika Lück
Die Verarbeitung morphologisch komplexer Wörter bei Kindern im Schulalter: Neurophysiologische Korrelate der Entwicklung
- 91 Diana P. Szameitat
Perzeption und akustische Eigenschaften von Emotionen in menschlichem Lachen
- 92 Beate Sabisch
Mechanisms of auditory sentence comprehension in children with specific language impairment and children with developmental dyslexia: A neurophysiological investigation
- 93 Regine Oberecker
Grammatikverarbeitung im Kindesalter: EKP-Studien zum auditorischen Satzverstehen
- 94 Şükrü Barış Demiral
Incremental Argument Interpretation in Turkish Sentence Comprehension
- 95 Henning Holle
The Comprehension of Co-Speech Iconic Gestures: Behavioral, Electrophysiological and Neuroimaging Studies
- 96 Marcel Braß
Das inferior frontale Kreuzungsareal und seine Rolle bei der kognitiven Kontrolle unseres Verhaltens
- 97 Anna S. Hasting
Syntax in a blink: Early and automatic processing of syntactic rules as revealed by event-related brain potentials
- 98 Sebastian Jentschke
Neural Correlates of Processing Syntax in Music and Language – Influences of Development, Musical Training and Language Impairment
- 99 Amelie Mahlstedt
*The Acquisition of Case marking Information as a Cue to Argument Interpretation in German
An Electrophysiological Investigation with Pre-school Children*
- 100 Nikolaus Steinbeis
Investigating the meaning of music using EEG and fMRI
- 101 Tilman A. Klein
Learning from errors: Genetic evidence for a central role of dopamine in human performance monitoring
- 102 Franziska Maria Korb
Die funktionelle Spezialisierung des lateralen präfrontalen Cortex: Untersuchungen mittels funktioneller Magnetresonanztomographie
- 103 Sonja Fleischhauer
Neuronale Verarbeitung emotionaler Prosodie und Syntax: die Rolle des verbalen Arbeitsgedächtnisses
- 104 Friederike Sophie Haupt
The component mapping problem: An investigation of grammatical function reanalysis in differing experimental contexts using event-related brain potentials

- 105 Jens Brauer
Functional development and structural maturation in the brain's neural network underlying language comprehension
- 106 Philipp Kanske
Exploring executive attention in emotion: ERP and fMRI evidence
- 107 Julia Grieser Painter
Music, meaning, and a semantic space for musical sounds
- 108 Daniela Sammler
The Neuroanatomical Overlap of Syntax Processing in Music and Language - Evidence from Lesion and Intracranial ERP Studies
- 109 Norbert Zmyj
Selective Imitation in One-Year-Olds: How a Model's Characteristics Influence Imitation
- 110 Thomas Fritz
Emotion investigated with music of variable valence – neurophysiology and cultural influence
- 111 Stefanie Regel
The comprehension of figurative language: Electrophysiological evidence on the processing of irony
- 112 Miriam Beisert
Transformation Rules in Tool Use
- 113 Veronika Krieghoff
Neural correlates of Intentional Actions
- 114 Andreja Bubić
Violation of expectations in sequence processing
- 115 Claudia Männel
Prosodic processing during language acquisition: Electrophysiological studies on intonational phrase processing
- 116 Konstanze Albrecht
Brain correlates of cognitive processes underlying intertemporal choice for self and other
- 117 Katrin Sakreida
Nicht-motorische Funktionen des prämotorischen Kortex: Patientenstudien und funktionelle Bildgebung
- 118 Susann Wolff
The interplay of free word order and pro-drop in incremental sentence processing: Neurophysiological evidence from Japanese
- 119 Tim Raettig
The Cortical Infrastructure of Language Processing: Evidence from Functional and Anatomical Neuroimaging
- 120 Maria Golde
Premotor cortex contributions to abstract and action-related relational processing
- 121 Daniel S. Margulies
Resting-State Functional Connectivity fMRI: A new approach for assessing functional neuroanatomy in humans with applications to neuroanatomical, developmental and clinical questions
- 122 Franziska Süß
The interplay between attention and syntactic processes in the adult and developing brain: ERP evidences