Visual Consciousness and Corticocortical Connectivity in the Human Brain

vorgelegt von
Diplom Informatikerin
Fatma İmamoglu
aus München

Von der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr. Olaf Hellwich
Gutachter:	Prof. Dr. Klaus Obermayer
Gutachter:	Prof. Dr. John-Dylan Haynes
Gutachter:	Prof. Dr. Christof Koch

Tag der wissenschaftlichen Aussprache: 17.04.2013

Berlin 2013
D 83

For AJ

Biz daima hakikati arayan ve onu buldukça
ve bulduğumuza kani oldukça ifadeye cür'et
gösteren adamlar olmaliyiz.
Mustafa Kemal Atatürk

We must be the men
who always seek for the truth, and
once we found it, and are content to have found it
we should take the liberty to express it.
Mustafa Kemal Atatürk

Abstract

How does neuronal activity generate complex conscious experiences, such as the redness of red, the sound of a blackbird or the smell of a lily? These questions form the basis for one of the most intriguing neuroscientific questions. Once we can understand how subjective experiences emerge from neuronal activity we could potentially model conscious experiences.

The purpose of this thesis was to investigate the underlying brain mechanisms that mediated conscious experience, in particular, conscious visual perception. Previous research has already shown that the underlying mechanisms mediating conscious perception are complex and might involve different regions distributed across the brain. However, brain connectivity as an underlying mechanism has been rarely considered. The focus of this thesis was thus to investigate the effects of brain connectivity on conscious visual perception.

Two functional magnetic resonance imaging (fMRI) experiments with healthy human subjects were conducted to examine the involvement of a distributed network of brain regions in conscious visual perception. Functional coupling between low-level and high-level brain regions was explored using the psychophysiological interaction method. Furthermore, Granger causality analysis was applied to examine the directionality of the functional coupling. Studying conscious perception is challenging. Brain processes related to conscious perception must be dissociated from stimulus-driven processes as well as from processes associated with the behavioral report. To that end, Mooney images were used in combination with two response modalities and a cognitive conjunction analysis in the first fMRI experiment. Mooney's are binary images that hide an object, which is not directly recognized by the subject, yet give rise to abrupt recognition after some time. An automatically generated database of Mooney images (MoonBase) was created prior to the fMRI experiment. In the second fMRI experiment, a visual masking procedure was used to modulate perceptual visibility.

The results indicated that a network of brain regions involving the frontoparietal cortex was contributing to conscious visual perception. Furthermore, the functional coupling between low-level and high-level brain regions reflected conscious visual perception. Moreover, findings from the Granger causality analysis demonstrated stronger feedforward connections than feedback connections when a stimulus was perceived consciously. Based on these findings, I conclude that functional coupling between long-distance brain regions is one important underlying brain

mechanism that give rise to conscious visual experience. These findings shed light on the underlying mechanisms mediating conscious visual perception and can be further utilized to model it.

Zusammenfassung

Wie generiert ein neuronales System komplexe bewusste Wahrnehmungen, beispielsweise die Farbe Rot, den Ton einer singenden Amsel oder den Geruch einer Lilie? Diese Fragen bilden die Basis für eine der interessantesten neurowissenschaftlichen Fragestellungen. Erst wenn wir verstehen, wie aus neuronalen Aktivierungsmustern subjektive Erfahrungen entstehen, können wir möglicherweise Bewusstsein modellieren.

Eines der Ziele dieser Dissertation war es die grundlegenden Gehirnmechanismen zu untersuchen, die zu Bewusstsein, insbesondere bewusster visueller Wahrnehmung, führen. Frühere Forschungsarbeiten haben bereits gezeigt, dass die grundlegenden Mechanismen für die Produktion bewusster Wahrnehmung komplex sind und verschiedene Gehirnregionen, die im ganzen Gehirn verteilt sind, einbezogen werden. Allerdings wurde die tatsächliche Konnektivität zwischen Gehirnregionen als ein grundlegender Mechanismus bei der bewussten Wahrnehmung bislang kaum betrachtet. Daher lag der Schwerpunkt dieser Dissertation auf der Untersuchung der Auswirkungen von Konnektivität auf bewusste visuelle Wahrnehmung.

Hierfür wurde in zwei funktionellen Magnetresonanztomographie (fMRT) Experimenten mit gesunden menschlichen Probanden untersucht, welche Rolle ein verteiltes Netzwerk von Gehirnregionen bei der bewussten visuellen Wahrnehmung spielt. Dabei wurde die funktionelle Kopplung zwischen frühen visuellen und späten präfrontalen Gehirnregionen durch die Methode der Psychophysiological Interactions untersucht. Darüber hinaus wurde eine Granger causality-Analyse angewandt, um die Richtungsabhängigkeiten dieser Interaktion zu untersuchen. Dabei stellen zwei mögliche Nebeneffekte eine Herausforderung für jede Bewusstseinsforschung dar: Gehirnprozesse, die mit bewusster Wahrnehmung in Verbindung gesetzt werden, müssten stets getrennt werden von stimulusabhängigen Prozessen sowie von Prozessen, die mit Verhaltensberichten der Probanden assoziiert sind. Daher wurden im ersten fMRT-Experiment Mooney-Bilder in Kombination mit zwei Antwortmodalitäten und der Cognitive Conjunction-Methode benutzt, um diese Nebeneffekte zu kontrollieren. Mooney's sind binäre Bilder, die ein Objekt verbergen, das von den Probanden nicht direkt erkennbar ist sondern nach einer bestimmten Zeit abrupt erkannt werden kann. Eine automatisch generierte Datenbank aus Mooney-Bildern (MoonBase) wurde dafür eigens vor den fMRT-Experimenten erzeugt. Im zweiten fMRT-Experiment wurde zu diesem Zweck ein visuelles Maskierungsparadigma angewandt, um die Sichtbarkeit des

Zielbildes zu modulieren.

Die Ergebnisse dieser Studien zeigten, dass ein Netzwerk von Gehirnregionen, das den frontoparietalen Kortex beinhaltet, bei der Generierung von bewusster Wahrnehmung involviert ist. Des Weiteren konnte gezeigt werden, dass die funktionellen Wechselwirkungen zwischen frühen visuellen und späten präfrontalen Gehirnregionen die bewusste Wahrnehmung beeinflussen. Zudem zeigten die Ergebnisse der Granger Causality-Analyse stärkere Verbindungen von frühen visuellen zu präfrontalen Arealen (feedforward) als umgekehrt (feedback) während der bewussten Wahrnehmung. Zusammenfassend machen diese Ergebnisse deutlich, dass funktionelle Konnektivität zwischen relativ weit entfernten Gehirnregionen wichtig ist, um bewusste visuelle Erfahrungen zu generieren. Diese Ergebnisse geben Aufschluss über die grundlegenden Gehirnmechanismen der bewussten visuellen Wahrnehmung und können bei der Modellierung dieser Prozesse benutzt werden.

Acknowledgements

I would like to address my gratitude to some people who made my thesis time intellectually fruitful, and mentally peaceful during difficult times.

I would like to start with a wonderful thank to my supervisor John-Dylan Haynes, for his valuable input, support, and helpful comments throughout this thesis. None of these studies could have been done without his contributions.

Christof Koch for being a great advisor and mentor. In our frequent skype meetings, he always greatly selected the essence of my subject in study out of our conversations. He is always keeping a careful eye on my research, for which I am grateful.

Klaus Obermayer for his support, constant engagement, and interest in my thesis.

Olaf Hellwich for his interest in my thesis and that he agreed to be the chair of my committee.

Marianne Märtens for her comments on my Mooney experiment, and her friendly advices about science in general.

Deutsche Forschungsgemeinschaft for their generous funding that made this PhD possible.

Thorsten Kahnt for his collaboration on my first fMRI experiment, his openness to share his knowledge and enthusiasm for science.

Jakob Heinzle for his collaboration on my second fMRI experiment, and all his questions that always brought my understanding for neuroimaging and connectivity methods way further.

Martin Hebart for his help throughout the MoonBase Experiment, and rating the Mooney images.

Adrian Imfeld for collecting the data in the second fMRI experiment.

Smadar Ovadia-Caro for her tremendous support in the last days of my thesis and for all our beautiful conversations that ranges from science, life to baking a brain cake.

My family and friends for their moral support during the last years.

Furthermore, I would like to thank people who took their time to read this thesis with all its inconsistencies and gave me very useful comments: Carsten Bogler, Latifa Omary, Smadar Ovadia-Caro, Henning Schroll, and Lejla Senderovic.

On the last run of my thesis marathon I was reading a book titled “Ignorance - How it drives science”. I am thankful to the author of this book, Stuart Firestein, who with his first ironic quota in his opening chapter saved days of my thesis. He wrote,

> "It is very difficult to find a black cat in a dark room", warns an old proverb.
> "Especially when there is no cat".

Last but not least, I want to thank Axel Johannis. This thesis is dedicated to Axel and his continuous power in backing me up. He always endowed me with tremendous believe in times when I could not see the end.

Today, I feel proud. The feeling to sit in a wonderful library[1] that the Humboldt University offers its students and others, and to be able to write the last sentences of four years hard work, makes me feel proud. Like any other doctorate, it was an exhausting course that requires hard work, flexibility, understanding for and of yourself, your committee, your reviewers, and patience, patience, patience. Though my mentor Christof Koch would always tell me "You should enjoy the way, the path along you are walking through, and not the end", sitting here at the end of my thesis, I can say with a smile on my face that 'the end' does make me feel wonderful.

Through Stuart Firestein's book and with Christof's advice in my mind I am ready to continue my scientific adventure with an open heart and some "healthy" ignorance.

[1] Jacob-und-Wilhelm-Grimm-Zentrum is located near Friedrichstrasse in Berlin and is a must seen library. So if your way crosses Berlin, do not miss this wonderful place, and if you already are in Berlin go and study there.

Contents

List of Figures

List of Tables

Abbreviations

ANOVA Analysis of Variance Analysis
AR Auto-Regression or Auto-Regressive
BOLD fMRI Blood Oxygenation Level Dependent Functional Magnetic Resonance Imaging
DCM Dynamic Causal Modeling
DLPFC Dorsolateral Prefrontal Cortex
DOI Difference of Influence
EC Extrastriate Cortex
EPI Echo-planar Imaging
EVC Early Visual Cortex
FMRI Functional Magnetic Resonance Imaging
FUS Fusiform Gyrus
FWE Familywise Error
GC Granger Causality
GLM General Linear Model
HRF Hemodynamic Response Function
ITG Inferior Temporal Gyrus
IPS Intraparietal Sulcus
LOC Lateral Occipital Cortex
MOG Middle Occipital Gyrus
NCC Neural Correlates of Consciousness
PPI Psychophysiological Interactions
ROI Region of Interest
SFG Superior Frontal Gyrus
SOG Superior Occipital Gyrus
SPC Superior Parietal Cortex
SPM Statistical Parametric Mapping
V1 Early Visual Cortex

I stand at the window and see a house, trees, and
sky. Theoretically, I might say there were 327
brightness and nuances of colour. Do I have "327"?
No. I have sky, house, and trees.
Max Wertheimer

1 Introduction

In our everyday lives, we continuously experience subjective sensations such as happiness, fear, love, the colors red and green, blue skies, objects, faces, men speaking, the sound of the wind, the taste of a falling tear... Some consciously, others unconsciously. The relationship between the objective material world, its corresponding brain state and the subjective conscious experience is the focus of my research. Although philosophers supporting the dualist approach reject that conscious experience can be fully explained by the brain response (Eccles, 1994; Popper and Eccles, 1996), this view is widely accepted by other philosophers (Block, 2005; Chalmers, 2010), as well as consciousness scientists today (Crick, 1995; Haynes and Rees, 2006; Koch, 2005; Rees et al., 2002; Tong, 2003; Tononi and Koch, 2008). The NEURAL CORRELATES OF CONSCIOUSNESS (NCC) are defined as "the minimal neuronal mechanisms jointly necessary to give a unique conscious experience" (Koch, 2005, p. 18). Although NCC does not account for individual differences in subjective consiousness (e.g. it is not possible to explain the exact same experience of my subjective experience of the color red to you), it has been suggested as a scientific, operative approach aimed at gaining further insight into consious experience (Crick and Koch, 2003).

NCC can be used to investigate two aspects of consciousness (please compare with Blumenfeld (2009, p. 18)):

1. The *level* of consciousness: This is the overall level of arousal which a creature, be it a human (Rees et al., 2002; Tong, 2003), an animal (Edelman and Seth, 2009), or a machine

(Koch and Tononi, 2008) has. For example, subjects in states such as wakefulness or rapid eye movement (REM) sleep have a high *level* of consciousness, while patients in vegetative state (VS) or minimally conscious state (MCS) have a low *level* of consciousness (Laureys and Tononi, 2009).

2. The *content* of consciousness: This is the *content* of a particular conscious experience e.g. an image, a sound, or a touch that is perceived consciously. It is not a state of general arousal, but rather being conscious of a specific experience. However, to be conscious of a *content*, the brain must already have a relatively high *level* of consciousness.

In my thesis, I investigate the *content* of conscious experience in humans. Although it is important to measure the *level* of consciousness one has, for example locked-in syndrome patients (Gosseries et al., 2009) are prone to misclassification as unconscious (Laureys et al., 2005), it is beyond the scope of my thesis[1]. The research of the *level* of consciousness is widely discussed elsewhere (Laureys and Tononi, 2009). Here, I study VISUAL CONSCIOUSNESS as one form of *content* consciousness. It is the subjective experience of the outside world through the visual system. The visual system is one of the most studied system in the human and animal brain, thus serves as a suitable candidate to study the neural correlates of conscious experience (Koch, 2005). Furthermore, visual phenomena, such as binocular rivalry, visual masking, or ambiguous figures can be used to precisely manipulate the subjective experience, while keeping the visual input constant. Thus, providing an approach to study consciousness experimentally. Additionally, to borrow the words from Francis Crick (Crick, 1995, p. 20),

> "If we could understand the mechanisms for one aspect [of consciousness], then we hope we will have gone most of the way to understanding them all."

The current literature supports the hypothesis that consciousness is represented by distributed networks of brain regions (Dehaene et al., 2006; Tononi and Koch, 2008; Rees et al., 2002; Haynes et al., 2005b; Lumer and Rees, 1999). It thus follows that dynamical changes in human brain connectivity should correlate with conscious experience. Evidence from human neuroimaging studies showed covariations between changes in human brain connectivity and visual experience (Dehaene et al., 2001; Haynes et al., 2005b; Lumer and Rees, 1999). Furthermore, our recent neuroimaging results support this view and suggest that further investigations into connectivity can improve our understanding of visual consciousness and consciousness in general (Chapter 4 (Imamoglu et al., 2012b) and Chapter 5 (Imamoglu et al., 2012a)).

In this thesis, I used functional magnetic resonance imaging (fMRI) in combination with connectivity measures to study visual consciousness. FMRI is a neuroimaging technique that can

[1] In this place, I can also recommend the memoir "The diving bell and the butterfly" by Bauby (1997), which was written by a locked-in syndrome patient.

Figure 1.0.1: Rubin's vase face illusion is an example of an ambiguous figure (adapted from Rubin, 1921, p. 263).

be used to localize brain functions during different cognitive states (e.g. conscious vs. unconscious). Statistical methods can be further applied to define the functional interactions between activated cortical regions. Because of its non-invasive nature and high spatial resolution, this method is widely used in cognitive experiments with healthy human subjects. However, it is an indirect measurement of neural activity that depends on blood oxygenation level changes (Chapter 2). Furthermore, the smallest measurable fMRI unit (also called voxel) usually incorporate large number of neurons. Therefore, it has been argued that fMRI studies are not suitable to identify the NCC (Chalmers, 2010). In contrast, others argued that the NCC might be "a parameter defined across multiple populations of neurons" (Haynes, 2009a), and fMRI is therefore a useful tool to identify the brain basis of human consciousness. Furthermore, Logothetis (2008) argued that the brain as a "distributed system", can be best explored using a large-scale imaging technique with high spatial resolution, such as fMRI.

Specifically, using fMRI my research questions are:

- Which distributed brain networks are contributing to conscious vision?
- Do changes in connectivity reflect the underlying mechanisms of conscious visual perception?

Visual processing is sensitive to changes in visual stimulus. Therefore, when studying visual consciousness, it is of high importance to distinguish conscious experience from other possible brain correlates that can be induced by changes in the physical stimulus. Ideally, we want to have a visual stimulus that does not change throughout the presentation but gives rise to more than one conscious experience. As an example, please fixate on the cross in the following image

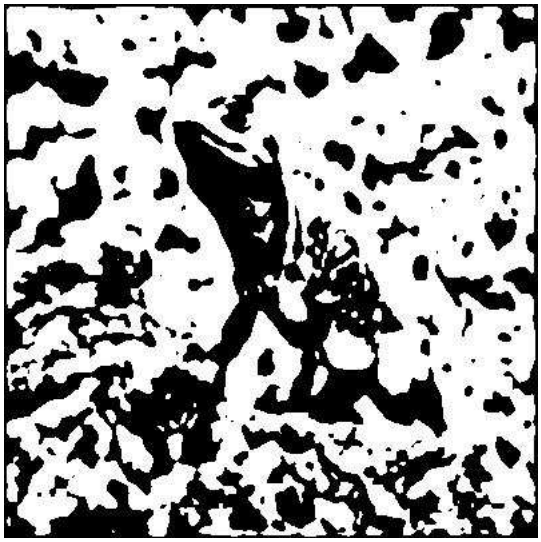


Figure 1.0.2: A two-tone, Mooney image.

(Figure 1.0.1) for ten seconds. At first sight you only perceive a white vase, but when you continue to fixate, the silhouette of two faces is perceived. Such an image is called AMBIGUOUS FIGURE and is an example of a perceptual illusion. During this illusion, the physical stimulus does not change while the perception fluctuates between two representations, i.e., the objective input is constant while subjective experience changes. The transition from one object representation (vase) to the other one (face) is an internal phenomenon that occurs intrinsicaly on average every 2-3 seconds. An observer can only consciously observe one of the two objects at any given moment.

However, to study connectivity changes in combination with fMRI data, we need a stimulus that has longer transition times. Therefore, during my thesis I first created an image database of two-tone, Mooney images (for an example of a Mooney image please see Figure 1.0.2) and designed a behavioral experiment to select suitable stimuli for the fMRI experiment (Experiment 1, described in Chapter 3). A two-tone, Mooney image contains a *hidden* object that is perceived only after a certain presentation time. At a first glance, subjects only perceive some unorganized black and white patches. However, after some internal reorganization, the perception abruptly changes and the subjects recognize the *hidden* object. Similarly to the previously described ambiguous figure, the subjective experience changes even though the physical properties of the image remains constant. However, once the object is perceived, the observer continues to see the object and there is no fluctuation in perception as in the previous example. This feature makes these images ideal to study changes in the functional coupling, i.e., connectivity, during conscious recognition of objects.

In the first fMRI experiment (Experiment 2, described in Chapter 4), I presented a subset of

these two-tone, Mooney images to subjects. In this experiment, I studied the neural correlates of conscious object recognition and the effect of functional coupling between long-distance cortical regions that was correlated with this percept. Although during the Mooney presentation, subjects were always seeing some black and white patches (see Figure 1.0.2), after a certain time, subjects were able to consciously recognize an object and could report the time of recognition (in Figure 1.0.2: a lizard).

In the second fMRI experiment (Experiment 3, described in Chapter 5), I manipulated subjective visibility profiles using a visual masking procedure. VISUAL MASKING enables the researcher to mask a visual stimulus that is projected onto the retina as perceptually invisible. Similarly to the ambiguous figure representation, the retinal input stays constant during the presentation time, but its neural representations, and therefore the subjective perception, changes over the course of the presentation. In this experiment, I studied the brain correlates of subjective visibility and the effects of connectivity on visibility.

In the following sections, I will review the relevant literature including influential electrophysiological as well as human neuroimaging studies that investigated consciousness (Section 1.1.1 and Section 1.1.2). Thereafter, I will present studies that demonstrate the influence of human brain connectivity on consciousness (Section 1.2). In the last section (Section 1.3) I will review some influential hypotheses and theories of consciousness.

1.1 Neural Correlates of Consciousness

There is a debate in the literature concerning the underlying mechanisms of conscious processing. One view is that local cerebral mechanisms such as early visual regions are sufficient to result in a conscious visual percept (Tong, 2003; Zeki, 2007). A contradictory view attributes conscious perception to the interaction of distributed networks of brain regions, specifically between low-level and high-level regions (Crick and Koch, 1995; Rees et al., 2002; Dehaene et al., 2006). As the reader will discover throughout this section, the latter view is emerging as a dominating hypothesis of conscious perception. Recurrent connections are one of the suggested mechanisms that underly this interaction along the visual hierarchy (Lamme and Roelfsema, 2000). Although my thesis concerns visual consciousness, in order to provide a broader understanding of the underlying mechanisms of consciousness, I will review studies investigating different perspectives of consciousness (sensory modalities as well as clinical approaches). Additional reviews can be found in: Rees et al. (2002); Tong (2003); Tononi and Koch (2008); Haynes (2009b), with a focus on clinical aspects in Goldfine and Schiff (2011), and animal consciousness in Edelman and Seth (2009).

1.1.1 Review of animal studies

One of the earlier animal studies in the field of consciousness provided an important piece of evidence about the significance of frontal brain regions in visual perception. It was based on cortical tissue ablation in visual and non-visual areas of 14 macaques performing a visual discrimination task (Nakamura and Mishkin, 1986). The authors reported two important findings that are relevant to the understanding of visual consciousness. First, when the visual cortex (striate, prestriate, medial, and inferior temporal cortex) and non-visual areas (cingulate cortex and ventral frontal area) were left intact, the animals could perform a visually guided behavioral task. However, when the same non-visual area was lesioned, and only the visual cortex was left intact, the animals were unable to perform the task. This means that non-visual areas that were left intact in the first case, but were removed in the second case, contain or process information that is relevant for the visually guided behavior. Though these authors did not show direct evidence of conscious or unconscious perception, this is one of the first studies to convey the importance of non-visual areas in visually guided behavior.

In a study by Cumming and Parker (1997) a random dot stereogram was presented to awake monkeys while neural responses in early visual area V1 were recorded. RANDOM DOT STEREOGRAM is a pair of images containing random dots that produces depth perception when one image is presented to each eye. Depth perception is achieved once the dots displayed to one eye are shifted horizontally such that the dots correlate binocularly. However, when the binocularly correlated dots have an opposite contrast, no depth perception occurs (anticorrelated random dot stereogram by Cumming and Parker (1997)). They demonstrated that V1 neurons responded both when depth perception occured or did not occur. This indicates that V1 neurons respond in an identical way to a stimulus that is differently perceived by the animal. Therefore, early visual areas alone could not account for conscious perception. Gur and Snodderly (1997) used an alternating colored bar stimulus (blue or red), which flickered at different frequencies (15 Hz, 30 Hz, and 60 Hz) while recording from neurons located in V1 and the lateral geniculate nucleus (LGN) of two monkeys. The highest frequency at which heterochromatic color can be perceived by the animals was below the presented frequencies (10 Hz according to Livingstone and Hubel (1987, p. 3422), see also Ives (1923, p. 364)). Although the animals could not perceive the stimuli, the authors found that the firing rate of the neurons located in V1 and LGN followed the fast flickering colored stimuli (also compare with Stockman et al., 1993; Gegenfurtner et al., 1997). Similarly, Macknik and Livingstone (1998) showed that V1 neurons follow the physical properties of perceptually invisible stimuli. These results indicate that neurons located in the early visual hierarchy are representing stimulus properties that are not consciously perceived by the animals. In addition, Logothetis and colleagues (Logothetis and Schall, 1989; Leopold and Logothetis, 1996; Sheinberg and Logothetis, 1997) recorded from visual cortical areas in monkeys while the animals performed a binocular rivalry task. BINOCULAR RIVALRY is

a perceptual illusion that is widely used in consciousness research. One stimulus (e.g. a vertical grating) is presented to the left eye, and another stimulus (e.g. a horizontal grating) to the right eye of the animal. The animal consciously perceives only one stimulus at a time with subsequent alternations of the stimuli (PERCEPTUAL ALTERNATIONS, Blake and Logothetis (2002)). The authors demonstrated that neurons located in early visual area V1 did not correlate with the animals' percept, rather responded only to changes in the physical stimulus. In contrast, neurons located in the inferior temporal (IT) cortex, a high-level visual area, responded mostly to the stimulus that was being reported by the animal. These studies demonstrated that high-level cortical areas are necessary for conscious perception.

Bridgeman (1980) recorded single-unit activity in monkeys' striate cortex. In his study, monkeys' task was to discriminate the brightness of target stimuli that were metacontrast masked. METACONTRAST MASKING is one form of visual masking that is used to render a stimulus perceptually invisible (Breitmeyer and Ogmen, 2006). In visual masking, a briefly presented target stimulus can be masked using different masking techniques (backward[2], forward[3], metacontrast[4], etc.). Depending on the time between the target and the mask (stimulus onset asynchrony or SOA), the perceptual visibility of the target stimulus can be manipulated. In his study, although early single-unit activity was independent of animals' perception (masked target or unmasked target), later neural responses were higher in visible conditions (unmasked target) than in invisible conditions (masked target). This study has one important implication, i.e., neurons in the striate cortex code different information at different time points. Only at a later time point neurons respond to differences in perception (visible or invisible). This result can be viewed as an evidence for feedback signals from high-level areas that entered the striate cortex before the perceptual decision was made. Similarly, Kovacs et al. (1995) recorded from shape selective cells in the IT cortex of monkeys while monkeys were presented with backward masked stimulus. They found that the initial response of these cells did not differ in visible (unmasked target) or invisible conditions (masked target), but the later response significantly decreased during invisible conditions. Furthermore, Lamme and colleagues (Lamme, 1995; Lamme et al., 1999; Zipser et al., 1996; Super et al., 2001) demonstrated using figure-ground texture segmentation experiments and electrophysiological recordings in monkeys' visual cortex that recurrent as well as feedback connections are necessary for conscious perception.

Thompson and Schall (1999) examined the relations between behavioral responses and signals recorded from the frontal eye fields (FEF) of two monkeys under backward masking. They demonstrated that the signals in FEF can predict the behavioral response of the animal. In accordance with this study, others (Libedinsky and Livingstone, 2011) recently showed that FEF neurons predict animals' responses in a motion-induced-blindness experiment. In MOTION-

[2]In backward masking a mask temporally follows the target.
[3]In forward masking a mask temporally precedes the target.
[4]In metacontrast masking target and mask do not overlap spatially.

INDUCED-BLINDNESS, a salient stimulus is rendered invisible by surrounded moving dots (Bonneh et al., 2001). These studies support the notion that the involvement of high-level brain regions are correlates of consciousness.

The electrophysiological studies I reviewed in this section support the hypothesis that early visual areas alone cannot reflect conscious perception. Furthermore, recurrent or feedback connections from high-level brain regions to visual areas are necessary for conscious perception. In the next section, I would like to review some human studies that either confirm or contradict this hypothesis.

1.1.2 Review of human studies

I will first review studies supporting the hypothesis that localized brain regions (such as visual areas) alone can account for conscious perception. In the second half of this section, I will converge to studies that demonstrate the need for distributed cortical mechanisms for conscious perception (Figure 1.1.1).

Tong et al. (1998) conducted a functional magnetic resonance imaging (fMRI) experiment using binocular rivalry to study conscious perception. In the rivalry conditions, they presented an image of a face to one eye and an image of a house to the other eye of the subject. In the nonrivalry condition, they presented subjects alternating face and house stimuli using monocular images of either faces or houses. Subjects' task was to indicate their perceived stimulus via a button-press. Tong et al. reported extrastriate activation (specifically fusiform face area and parahippocampal place area) correlating with subjects' perceptual alternations. From this results, they conclude that conscious perception is resolved in visual processing areas. Similarly, Polonsky et al. (2000) used fMRI and binocular rivalry with gratings to study rivalry-related activity in the human primary visual cortex. They found that early visual area V1 reflected subjects' perceptual alternations. With this experiment, the authors challenged the view that extrastriate or high-level cortical regions are necessary to perceive a visual stimulus consciously. An influential experiment by Tong and Engel (2001) studied perceptual alternations in the blind-spot representation of the visual cortex (BLIND SPOT REGION). The BLIND SPOT or OPTIC DISC is a small region at the back of the retina, where axons of the retinal ganglion cells exit the eye. This region does not contain any photoreceptor cells that convert light into neural signals, which are carried through the OPTIC NERVE to the brain for further processing. The blind spot region in V1 is a monocular representation of the blind spot and receives only signals from the ipsilateral eye. The authors of this study presented a vertical grating stimulus to the ipsilateral eye, and a horizontal grating to the contralateral eye. When subjects reported to perceive the vertical grating (ipsilateral eye) the fMRI signal in the blind spot region in V1 was as expected high, while it was low when the perception switched to the horizontal grating (contralateral eye). This result shows that the signals from the unperceived eye modulate the

fMRI response in the blind spot region. The authors conclude that binocular rivalry is resolved in monocular V1 neurons (neurons receiving signals only from one eye) due to interocular competition. A study by Haynes et al. (2005a) used binocular rivalry with fMRI to study the effects of binocular rivalry in the visual thalamic region called lateral geniculate nucleus and the early visual cortex V1. LATERAL GENICULATE NUCLEUS (LGN) is the visual center in the thalamus that projects to V1 and contains eye-specific (monocular) neurons in specific layers (layers 2, 3 and 5 receives information from the ipsilateral eye; layers 1, 4 and 6 receives information form the contralateral eye (Andrews et al., 1997)). The authors demonstrated that lateral geniculate nucleus (LGN) activity reflects subjects' reported percept rather than the physical inputs from the two eyes. This study challenges the interocular competition idea demonstrated by Tong and Engel (2001) and has one important implication for consciousness research, i.e., perceptual alternations during binocular rivalry are likely to occur via inhibitory feedback connections from V1 to monocular layers of the LGN. Moreover, Haynes and colleguaes suggest modulations by feedback information from high-level cortical regions.

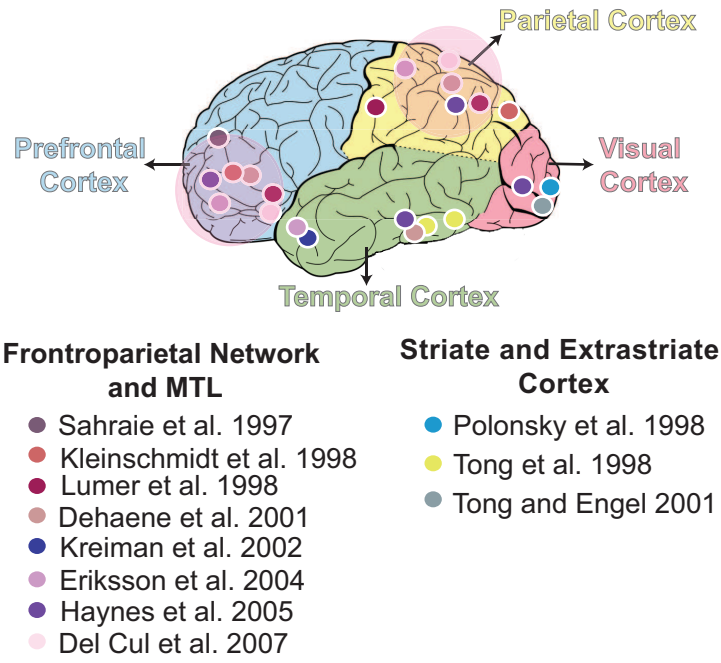


Figure 1.1.1: Brain correlates of human consciousness. The two pink shaded circles represent the regions within the parietal and prefrontal cortex. Each colored circle shows the involvement of different brain regions in different consciousness studies. The corresponding Talairach coordinates for these regions (if mentioned in the original studies) are listed in Appendix A.

The remainder of this section focuses on studies that support the involvement of high-level cortical regions in conscious perception. A study by Sahraie et al. (1997) used functional

magnetic resonance imaging (fMRI) to study a BLINDSIGHT patient (GY) with unilateral V1 lesion. Blindsight patients are blind to their visual field due to V1 lesions, but can perform the right actions (e.g. grasp an object) even though they report not to see the objects that are presented to their blind visual field. Furthermore, they make correct guesses, when they are asked about the form or shape of the object (Weiskrantz, 1990; Stoerig and Cowey, 1997). The patient's task was to discriminate the direction of the motion, which was presented to his blind visual field. The patient sometimes reported to have conscious vision in his blind visual field and sometimes not. In both conditions he could correctly discriminate the direction of the motion above chance performance. They found that the prestriate cortex (also called V2) was active in both conscios and unconscious conditions. In addition, the dorsolateral prefrontal cortex (DLPFC, Figure 1.1.1) was active only when the patient was conscious about the motion direction. This suggests that brain responses in the visual cortex can arise in the absence of conscious perception and high-level brain regions are necessary for conscios perception. An experiment conducted by Rees et al. (2000) used fMRI to study visual extinction. VISUAL EXTINCTION is a neurological disorder that typically occurs following brain damage to the parietal cortex. Patients with this disorder can detect a visual stimulus that is presented to one of the visual fields. However, when two stimuli are simultaneously presented to both visual fields then extinction of perception occurs, and the patient cannot perceive both stimuli (Rees et al., 2000). For example, the patient in this study had a right parietal damage and a left visual extinction. He could detect a visual stimulus that was presented to his left visual field, however, when another stimulus was presented simultaneously to his right visual field, he could not detect the (left visual field) stimulus anymore. Rees and colleguaes observed brain activations in the right early visual cortex (contralateral to the left visual extinction) in the absence of conscious visual perception. This case study further supports the notion that visual regions can be active without consciously perceiving a visual stimulus.

An influential study by Lumer et al. (1998) was one of the first human neuroimaging studies that demonstrated a direct evidence for the involvement of frontoparietal cortex in conscious perception. The authors used binocular rivalry and fMRI to study the effect of perceptual alternations in cortical regions of healthy human subjects. They presented a face stimulus to the one eye of the subjects and a grating stimulus to the other eye, and asked the subjects to press a button whenever their perception switched from a face to a grating, and vice versa (rivalry condition). As a control, the authors presented alternating face and grating stimulus to the subjects without using binocular rivalry (nonrivalry condition). They demonstrated that in both rivalry and nonrivalry conditions the extrastirate areas were active. But prefrontal and parietal (frontoparietal) cortex were active only during perceptual alternations in rivalry conditions. Another binocular rivalry study by Tononi et al. (1998) used magnetoencephalography (MEG) to map the brain activity that occured during perceptual alternations. In rivalry

conditions, they presented a vertical red grating flickering at one frequency to one eye and a horizontal blue grating flickering at a different frequency to the other eye. The specific frequency was selected out of four different frequencies: 7.41 Hz, 8.33 Hz, 9.50 Hz, or 11.12 Hz. In a control condition (nonrivalry conditon) they presented stimulus alternations without binocular rivalry. Similarly to Lumer et al. (1998), the authors showed significant evoked responses in a distributed network of brain regions that extended to anterior areas (prefrontal cortex) during rivalry conditions. Kleinschmidt et al. (1998) used Rubin's vase face illusion (as introduced in Figure 1.0.1) with fMRI to study perceptual alternations using a physically constant stimulus. They showed significant fMRI activation in high-level visual areas, parietal cortex, and prefrontal cortex during perceptual alternations. Furthermore, patients with unilateral prefrontal cortex damage have been shown to have abnormal perceptual transitions compared to controls or patients with posterior damage (Ricci and Blundo, 1990). In addition, Kreiman et al. (2002) recorded single-cell responses of the electrodes implanted in the medial temporal lobe (MTL) of epilepsy patients and demonstrated that these neurons followed perceptual alternations reported by the patients. Similarly, Eriksson et al. (2004) demonstrated MTL and frontoparietal activation when a previously unseen object was abruptly recognized (pop-out) in the image (similar to the Mooney image in Figure 1.0.2). Likewise, I recently reported using two-tone, Mooney stimuli parietal and dorsolateral prefrontal cortex activation when subjects recognized an object in the Mooney image vs. when they did not recognize it (for details see Imamoglu et al. (2012b) and Chapter 4). A prominent study by Dehaene et al. (2001) used visual masking in combination with fMRI and electroencephalography (EEG) using 128 scalp electrodes to map brain activity of visible (unmasked) and invisible (masked) conditions. They manipulated visibility of written words using a combined forward masking (mask preceding the target) and backward masking (mask followed by the target) experiment. Similar to Tononi et al. (1998), they demonstrated a distributed network of brain regions that were active during consious perception (visible conditions). This network included regions such as the extrastriate, parietal, and prefrontal cortex. However, these results might reflect differences in the visual stimuli rather than conscious perception due to the statistical comparison the authors used (masked word vs. blank word for invisible conditions, and unmasked word vs. unmasked blank for visible conditions). Therefore, the authors tested a repetition suppression experiment in which the same masked words preceded visible words. They demonstrated suppression in the fMRI activity in high-level visual areas during repetition suppression, which indicated processing in high-level visual regions in the absence of consciousness. The authors further support these results by analyzing event related potentials (ERPs, brain responses to a cognitive task measured by EEG electrodes). They showed an early positive (~170 ms) response over the occipital electrodes (corresponding to extrastriate activations in fMRI) and two negative (~240 ms and ~470 ms) responses over the left-lateralized electroes (corresponding to left fusiform

gyrus activations in fMRI) during invisible conditions. Similarly, recent EEG and MEG studies showed that early visual activations were preserved during masked conditions (Del Cul et al., 2007; Fahrenfort et al., 2007; Melloni et al., 2007), but activations later in time (Del Cul et al., 2007; Fahrenfort et al., 2007) or oscillatory activations in the gamma frequency band (Melloni et al., 2007) correlated with subjective experience. For example, Del Cul et al. (2007) reported that a backward masked stimulus did not have an effect on early visual areas, but reported modulations in the ERP response (180-250 ms) in high-level visual areas and in the frontoparietal network during conscious report ($>$ 270 ms). The results of these studies demonstrate that a processing stream including early visual and high-level visual areas can occur in the absence of conscious perception. Furthermore, a distributed network including the frontoparietal cortex is a strong correlate of conscious perception and can be seen as part of the NCC.

The studies I reviewed in this section largely highlight a major role of the frontoparietal cortex in consciousness. In the next section, I will review some recent results based on neuroimaging and anatomical studies that not only searched for localized brain activations, but also for global, distributed network correlates of consciousness.

1.2 Consciousness and Connectivity

Influential hypotheses and theories of consciousness (for a detailed description see Section 1.3) have emphasized the importance of brain connectivity in conscious perception. Yet, until recently, addressing these theories empirically was rather difficult (and some of them are still not achievable experimentally (Tononi, 2004)). This has changed since new measurement techniques and statistical methods took the stage in cognitive neuroscience and consciousness research. Here, I will describe studies in which statistical dependencies between different brain regions (functional connectivity) and causal connections between these regions (effective connectivity) are studied (Sporns, 2007). Although functional or effective connectivity do not directly reflect anatomical brain connectivity (Zhang et al., 2008), it has been shown to correlate (Honey et al., 2007).

There are several brain connectivity methods that are useful to infer functional, effective, or causal connectivity in the human brain. Functional connectivity is defined as the temporal correlation between the average fMRI signal in different brain regions (Horwitz, 2003; Fair et al., 2007). psychophysiological interactions (PPI) (Friston et al., 1997) is an additional approach building on functional connectivity. PPI computes the correlation between the neural activity of a priori selected brain region and the rest of the brain, while considering the experimental manipulations (for a detailed explanation see Chapter 2.3). Both approaches (i.e., functional connectivity and PPI) cannot be used to infer causality and/or directionality of connections. Granger causality computes whether the response of one brain region predicts

the response of another brain region at a later time while taking into account the history of both regions (Bressler and Seth, 2011; Roebroeck et al., 2005). Granger causality thus measures the directed functional connectivity without any prior assumption concerning the structural connections. Dynamical causal modelling (DCM), is an effective connectivity method that is also used to infer causal connections between brain regions. It is a model based approach that includes prior assumptions about the structural connectivities (Friston et al., 2003). Effective connectivity is defined as the influence one neural system exerts over another (Valdes-Sosa et al., 2011). In addition to the above mentioned methods, perturbation methods can be further used to infer causality. Transcranial magnetic stimulation (TMS) is a noninvasive technique that uses magnetic field properties to discharge action potentials on the site of the stimulation (Barker et al., 1985). The stimulation perturbs the activity in the stimulated area, hence, creates an experimental condition of transient interference. Using this approach causality between a cognitive task and corresponding brain areas can be studied (Philiastides et al., 2011; Guse et al., 2009; Amassian and Maccabee, 2006).

I will first describe three connectivity studies that support the main intention of this thesis (see Figure 1.2.1). One of the early neuroimaging studies that investigated functional connectivity in humans was conducted by Lumer and Rees (1999). They used binocular rivalry and fMRI to study the covariation between long-distance brain regions. In the rivalry condition, an image of a grating was presented to one eye while an image of a face to the other eye of the subject. In the stable viewing condition, subjects were presented with a constant half-face, half-grating stimulus. The uniqeness of this experiment was that subjects were not asked for their subjective reports (usually via button press) to indicate the binocular alternations in the rivalry condition. The authors rather estimated perceptual alternations using autocorrelations of the fMRI signal. This means that binocular rivalry was studied in the absence of motor response (eliminating a possible motor confound). The authors demonstrated a significant correlation between the prestriate cortex and the inferior frontal cortex, superior parietal cortex, as well as fusiform gyrus in rivalry conditions (Figure 1.2.1, dark purple arrows). Only prestriate cortex and the fusiform gyrus were significantly correlated during stable viewing conditions. This was the first study that showed the importance of frontoparietal-visual cortex interaction in conscious perception. Similarly, Dehaene et al. (2001) showed that various distributed brain regions were functionally connected when a word was consciously perceived. Using word masking paradigm and PPI analysis, the authors found that when words were visible, the left fusiform gyrus increased its connectivity with occipital, high-level visual areas (ventral occipitotemporal), parietal areas (intraparietal sulci), as well as prefrontal areas (dorsolateral prefrontal cortex and inferior frontal cortex) as compared to the invisible condition (Figure 1.2.1, blue arrows). Haynes et al. (2005b) have demonstrated similar results. Using metacontrast masking and a variant of the PPI analysis, they found enhanced functional connectivity between visual area V1 and

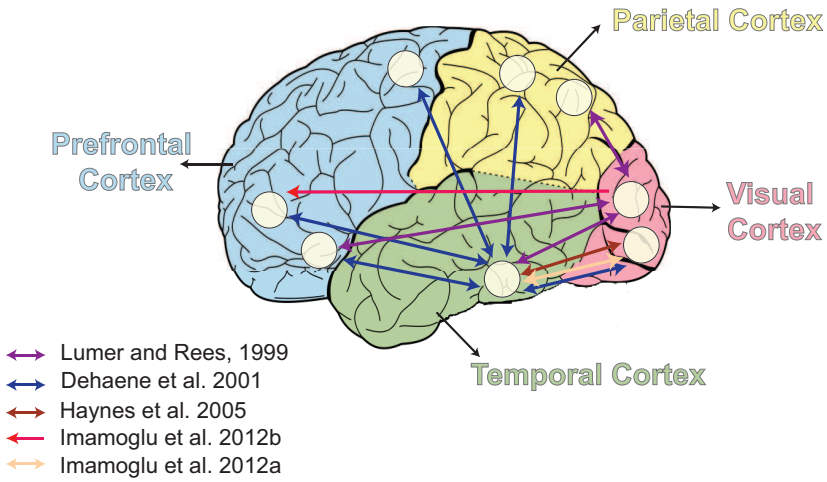


Figure 1.2.1: Brain connectivity of human consciousness. Each colored arrow depicts a different study. The connections between the circles are results of the enhanced functional connectivities during consciousness. Non-directional connectivity results (here illustrated as both sided arrows) are obtained from PPI analysis (Lumer and Rees, 1999; Dehaene et al., 2001; Haynes et al., 2005b; Imamoglu et al., 2012a). Directional connectivity (here illustrated as one sided arrows) result is obtained using Granger causality analysis (Imamoglu et al., 2012b).

bilateral fusiform gyri in high visibility as compared to low visibility conditions (Figure 1.2.1, brown arrow).

Resting state fMRI studies in patients with disorders of consciousness showed frontoparietal functional connectivity reduction in brain-damaged patients as compared to controls (Vanhaudenhuyse et al., 2010). Massimini et al. (2005) used TMS in combination with EEG recordings during awake and NREM (non rapid eye movements) sleep phases in healthy subjects. In the latter case, subjects usually have no or little consciousness. Massimini and his colleagues showed that under wakefulness an initial signal that is caused by the TMS stimulation travels to distant brain areas and reaches the frontal cortex, whereas in NREM phases the effect of the stimulation is limited to the targeted areas. These results further emphasize the importance of the frontoparietal cortex in consciousness.

Resarchers have attempted to investigate whether the interaction between frontal and visual areas is based on feedforward or feedback connections. Gaillard et al. (2009) used backward masking and electrophysiological recordings from electrodes implanted in epileptic patients for surgical purposes. They demonstrated significant changes in long-range connectivity based on Granger causality from occipital to frontal cortex (feedforward) during conscious (unmasked) conditions. In contrast, Boly et al. (2011) observed preserved feedforward connectivity but impared feedback conntectivity (from frontal cortex to temporal cortex) in vegetative state patients using DCM analyis. These patients are intact in their level of arousal yet are unaware

of their environment (Boly et al., 2012; Laureys, 2005). In addition, two recent studies have used binocular rivalry and TMS to study the causal role of parietal cortex in the perception of rivalrous stimulus (Kanai et al., 2010; Carmel et al., 2010). Both reported a causal role of the superior parietal cortex in perceptual alternations. However, when superior parietal cortex was perturbed by TMS an increase in the rate of perceptual alternations was observed (Carmel et al., 2010). In contrast, Kanai et al. (2010) observed a decrease in the rate of perceptual alternations while superior parietal cortex was perturbed. Although these results are contradictory to some extent, both studies demonstrated a causal role of the parietal cortex in these rivalrous alternations. A comment by Clifford (2010) suggests some possible causes of this inconsistency, such as different neural stimulation methods (repetitive TMS vs. theta-burst TMS) that were used in these studies, or different procedures that were used to localise superior parietal cortex. Further studies are needed to clarify whether perturbing the superior parietal cortex by TMS results in increase or decrease in perceptual alternations.

To conclude, based on these studies, frontoparietal network is crucial for conscious perception. Connectivity between frontoparietal regions and other cortical regions (e.g. visual cortex) can be seen as an underlying mechanism for conscious perception. However, the directionality of these connections, i.e., whether feedforward or feedback, is inconclusive given the current literature. The importance of connectivity to the understanding of consciousness is implied in several theories of consciousness (please compare with Tononi and Koch (2008)). It is thus of parmount importance to familiarize oneself with the different theoretical approaches and hypotheses concerning consciousness. Therefore, before we start our journey with the experiments I carried out, I will provide a brief overview of these theoretical aspects in the next section.

1.3 Neural Theories and Hypotheses of Consciousness

Several theories and hypothesis have been proposed to account for consciousness. In the previous section, I have reviewed some studies that support a localized view as the underlying neuronal mechanisms necessary for conscious perception. In contrast, other studies demonstrated a distributed, global network of brain regions and their interaction to be necessary for conscious perception. The theories I will review in this section can be clustered around these two views. I will start with localized views and converge towards the end to global views.

The MICROCONSCIOUSNESS HYPOTHESIS suggested by Zeki and Bartels (1999) is a localist approach to account for the underlying neural mechanisms of conscious perception. According to this hypothesis, several nodes ("microconsciousness") can be active when one is conscious about a particular feature of a stimulus. Furthermore, these nodes can interact with each other in order to integrate or bind perceptual properties together ("macroconsciousness"). This latter processing occurs post consciousness and is not a necessary processing step for conscious

perception. For example, when we hear a sound, our auditory system is active and consciousness of this particular sound arise, or when we perceive a color, the area in the brain that gives rise to color perception (visual area V4) is active, and consciousness of this particular color arise (Zeki, 2007). Therefore, consciousness involves multiple microconsciousnesses, which are processing information independently and can be seen as a NCC (Kouider, 2009). This hypothesis is largely driven from clinical studies that used patients with lesions in their visual cortex (such as blindsight patients). Therefore, it is mainly discussing the *content* of consciosness (specifically in the visual domain) but cannot account for the *level* of consciousness. The visual cortex hypothesis suggested by Tong (2003) can be interpreted as one microconsciousness that explains the underlying mechanisms of visual consciousness. Accordingly, visual consciousness is resolved within the stages of the visual pathway. However, I have previously reviewed a variety of experiments that showed the necessity of high-level visual and high-level brain regions (such as the parietal and prefrontal cortex) for conscious perception (Section 1.1). Although the involvement of early visual area V1 in conscious vision is well accepted (Koch, 2005), this region is not considered as a sufficient processing unit to account for conscious visual experience.

Another localized approach to account for visual consciousness is suggested by Goodale and Milner (1992). In their DUPLEX-VISION HYPOTHESIS, conscious visual perception occurs through information exchange between two distinct visual pathways: (i) The ventral visual pathway is the conscious path, which projects information from visual cortex to the high-level visual areas in the inferior temporal cortex. (ii) The dorsal pathway, which is the fast unconscious path, projects the visual information from visual cortex to parietal and action-related motor areas. This theory was mainly built on observations from blindsight patients. However, there are already evidences that show stream invariant processing of conscious and unconscious information (Hesselmann and Malach, 2011; Hebart and Hesselmann, 2012; Fang and He, 2005).

The REENTRANT HYPOTHESIS proposed first by Edelman (1992), builds upon the categorization problem of visual perception. He suggests that parallel brain signals are passed on to different "cortical maps" in a simultaneous manner. These cortical maps are similar to category specific areas in the brain and are specialized for a specific function (e.g. the fusiform face area, where face related information in the brain is largely processed (Kanwisher et al., 1997)) and can be linked with other maps to perceptually categorize visual inputs (Searle, 1997, p. 39-51). Furthermore, the interaction between these cortical maps is based on parallel reciprocal connections. According to this theory, consciousness arises from these reentrant processing among different cortical maps that are distributed across the brain. Victor Lamme and others used a similar idea for their "local recurrent" (or reentrant) connectivity approach (Lamme et al., 1998). This theory mainly explains visual aspects of conscious perception. In its three processing levels, it suggests that only the neural processes that are locally recurrent can give rise to conscious perception. Here, recurrency refers to local reciprocal connections where the visual

information from the high-level visual areas is fed back to the early visual areas. According to Lamme (2006), the other two processing stages are (i) "feedforward processes", in which fast signals from visual cortex travel to temporal, parietal, and prefrontal cortices in an unconscious manner, and (ii) "long-distance recurrent connections" between visual and frontoparietal cortex, which give rise to "reportable conscious perception". According to Lamme, whether a system is conscious or not should not rely on the behavioral report. Therefore, he distinguishes between unconscious processing that is reflected in the feedforward processing, the conscious processing where local recurrency between the visual areas give rise to a conscious experience, and the access consciousness where long-distance recurrency enables conscious report. It is therefore important to distinguish between "conscious perception" and "reportable conscious perception", and that such distinction is also visible in the neural correlates (local recurrent and long-distance recurrent connections respectively).

In contrast to the above mentioned hypotheses, a global neurobiological account of consciousness is represented in the Global Neuronal Workspace Theory (GWNT, Dehaene et al., 2006, 2011). This theory is the neurobiological extenstion of Baars' (1988) cognitive global workspace theory and associates conscious perception with the frontoparietal network and its top-down processing. According to Dehaene and colleagues, at any given time there are multiple cerebral processes that are active in parallel and being processed automatically (which are not available for consciousness). A distributed neural system (also called WORKSPACE) with long-distance connectivity can connect these multiple processes. Neural processes that are automatic and unconscious are assumed to make their information content "globally available" to other processes through this workspace. Therefore, unconscious processes are only able to engage in consciousness if they are connected to the workspace (Dehaene and Naccache, 2001). Accordingly, attention and top-down processing is a prerequisite of consciousness and is the central unit in this workspace. The workspace is a dynamic set of neurons that mainly share neurons (WORKSPACE NEURONS) in the prefrontal cortex, parietal cortex, as well as anterior cingulate, and act upon unconscious processes in a top-down manner (feedback processes). Furthermore, according to this theory, subjective reports are the key elements to study consciousness. To explain conscious vision the authors define two different stages of non-conscious processing: A "subliminal processing" stage, where only early visual areas are involved, and a "preconscious processing" stage in which, similar to Lamme's recurrent theory, a recurrent occipito-temporal processing takes place, but do not give rise to conscious experience, due to the lack of attentional resources and top-down amplification. What gives rise to conscious vision is the feedback information from the workspace (such as the frontoparietal cortex) (Dehaene et al., 2006). Although others have argued that consciousness and attention are distinct processes, and attentional amplification can also give rise to unconscious perception (Koch and Tsuchiya, 2007; Koch, 2005, p. 167-187), the global workspace theory has been very popular in the field of of consciousness

studies (Kouider, 2009). However, due to its direct link to subjective reportability this theory can only address *content* consciousness.

The Information Integration Theory (IIT), gives a theoretical framework for both *level* and *content* consciousness. According to this theory, consciousness is associated with integration of information and complexity of a cognitive system. The idea behind this theory is that there is no such consciousness network that is involved in conscious processing, such as the global workspace, rather different conscious states can be reflected in alternative combinations of neurons or networks. According to IIT, any system that has consciousness, must first discriminate the conscious *information* from a pool of other states (in informational theoretical language this refers to reduction of uncertainty). In other words, conscious states are highly differentiated. Secondly, the system has to be able to *integrate* this information as a whole, and in a unified manner. For example, we experience a red rectangle in a unified manner and not the shape rectangle separated from its color. Accordingly, consciousness *is* integrated information (Tononi and Balduzzi, 2009), and both differentiated information and its integration are essential for consciousness. The more a system can integrate differentiated information, the more it is conscious. The main idea of this theory lies in an information theoretical measurement that is named ϕ by its authors (Tononi, 2004). Based on bipartitioning the system into subsets, ϕ computes the integrated information of a given system. ϕ has a high value for a system that is complex and whoes neurons are largely interconnected. In contrast, it has a low value for simple systems. Although this measurement in small neural networks has been proven very promising, its computation is very demanding for natural systems (Tononi, 2004; Tononi and Sporns, 2003; Tononi and Balduzzi, 2009), and therefore not directly applicable to neuroimaging or electrophysiological data. However, an approximation of ϕ has already been proposed (Edlund et al., 2011) and it has been recently applied on EEG and ECoG data (Oizumi et al., 2012). The uniqueness of this theory is that it converges with other theories and hypotheses of consciousness, such as the recurrent hypothesis or the global (neuronal) workspace theory, and consciousness is not seen as an all-or-none process, but as a gradually changing phenomenon. Furthermore, IIT is valid for any organism, biological or silicon, and does not only deal with human consciousness.

Until now, I have presented scientific approaches to study consciousness that have been demonstrated over the course of the past several decades. One important conclusion I can draw from the studies I reviewed here is that large scale distributed brain regions are significantly contributing to conscious perception. Furthermore, the connectivity between these regions, which include an interaction between high-level brain regions and low-level brain regions are of particular importance for consciousness. The influence of brain connectivity on consciousness has been mainly studied in the context of *level* of conscioues ness (Boly et al., 2012; Jin and Chung, 2012), but not as extensively in *content* consciousness (but see Lumer

and Rees, 1999; Dehaene et al., 2001; Haynes et al., 2005b; Imamoglu et al., 2012a,b). Therefore, both fMRI experiments that are described in this thesis are designed to study functional connectivity changes between distant brain regions during visual consciousness (Chapter 4 and Chapter 5). Furthermore, the directionality of the functional connectivities was studied using a Granger causality analysis in Experiment 2 (Chapter 4). In addition, studies that contributed to *content* consciousness have largely relied on behavioral responses that were reported via a button press (Dehaene et al., 2001; Haynes et al., 2005b, but see Lumer and Rees, 1999). This can be a possible confound when high-level areas such as the prefrontal cortex are involved in conscious processing. This possible confound is likely to occur due to the anatomical configuration of movement related areas in and near the prefrontal cortex: the premotor cortex, which is responsible for the preparation of a movement (e.g. preparation of the button press) is anterior to the primary motor cortex and is located within the prefrontal cortex. This was ruled out in the design and analysis of my first fMRI experiment (Experiment 2, described in Chapter 4). In addition, in a visual consciousness experiment the stimuli that are presented to the subjects need to be physically constant. The stimuli I used in both my experiments contemplated this important purpose. I will describe these experiments in more detail in Chapter 4 and Chapter 5. The rest of the thesis is structured in the following way: In Chapter 2, I will describe the methods I have used in the fMRI experiments. Chapter 3 focuses on how the Mooney image database (MoonBase) was generated. Finally, in Chapter 6, I will discuss the impact of these studies on our understanding of consciousness and propose some future research directions. The reader should keep in mind that Chapter 4 is adapted from the original published manuscript (Imamoglu et al., 2012b). Chapter 3 and Chapter 5 are adapted from manuscripts that are in preparation (at the time of the submission of this thesis) for submission to peer-reviewed journals.

Measure what is measurable, and make measurable
what is not so.
Galileo Galilei

2

Methods for fMRI Data Analysis

In this chapter I will briefly introduce the fMRI data analysis methods that I used in this thesis. FMRI uses magnetic resonance imaging (MRI) scanners, which generates 3D images of biological tissue using strong magnetic fields. It is a non-invasive imaging technique that is related to activity of large populations of neurons, and is widely used to study the function of the brain as well as in clinic for diagnostics (structural MRI images). Images are created while a subject or patient lies in an MRI scanner with a strong magnetic field strength of at least 1.5 Tesla (the majority of fMRI studies in humans use 3 Tesla scanners, but stronger magnetic fields are possible).

Information transfer between neurons occur at the synaptic level through neurotransmitters. Increased synaptic activity leads to higher oxygen demand in local brain regions, which in turn is accompanied by increased blood flow and/or oxygenation level (Matthews, 2009). In fMRI, using an appropriate pulse sequence (changing magnetic field gradients and oscillating electromagnetic fields) that is sensitive to changes in local blood oxygenation level (BOLD), the MRI machine detects the local magnetization changes in brain regions caused by the blood deoxygenation level (Ogawa et al., 1992). This change is reflected in the hemodynamic response function (HRF) and is related to changes in neuronal activity.

In Chapter 4 and 5, I used an echo-planar imaging (EPI) pulse sequence, which is widely used in fMRI studies because of its speed in acquiring brain images (~30-40 ms for each brain slice and around 20 slices per second). When subjects conduct an experiment while their brain images are acquired, those brain regions that are related to the sensory stimulation or cognitive

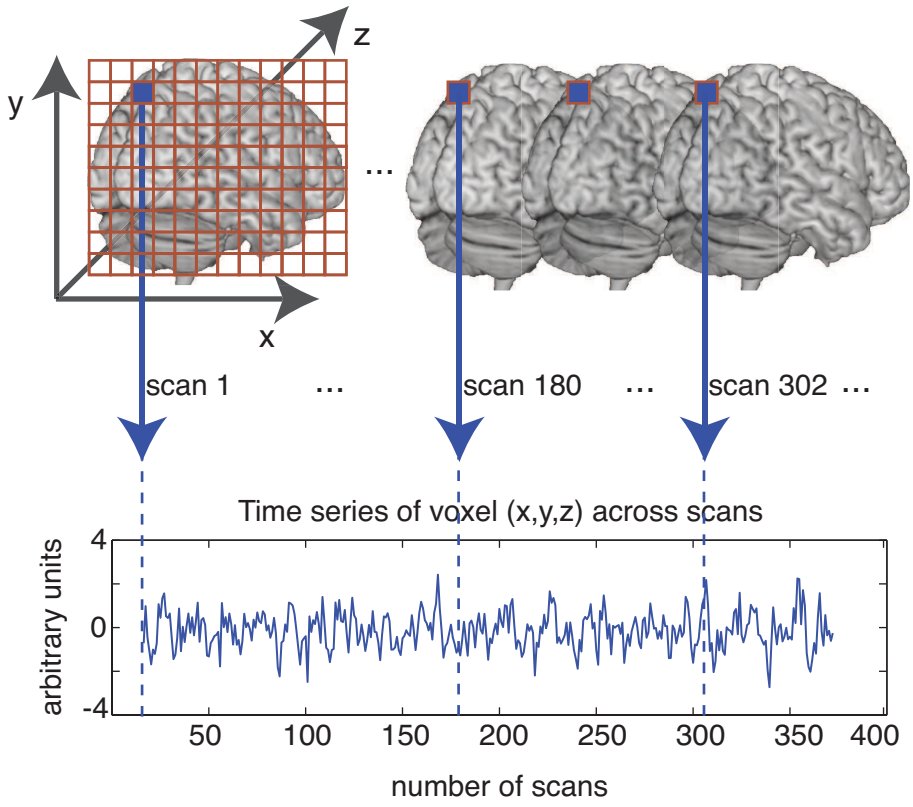


Figure 2.0.1: Examples of fMRI volumes and a voxel time-series.

task will show increased intensity in their brain scans (Smith, 2004). Because fMRI measures BOLD signal, it is a non-direct measurement of neural activity. Furthermore, acquisition of each whole brain image takes around 2 seconds (with a repetition time (TR) of 2 s). This results in poor temporal resolution. In contrast to other neuroimaging techniques such as the electroencephalography (EEG), or magnetoencephalography (MEG) however, the spatial resolution of fMRI is high (a few mm^3).

In a typical fMRI experiment, every few seconds a functional MR volume (also called SCAN or IMAGE) is aquired. During each fMRI run (also called SESSION) several fMRI volumes with thousands of three dimensional volume elements, called voxels are collected. Each voxel can be plotted as a function of time, which represents the voxel intensity changes across different volumes, or acquisition times (Figure 2.0.1).

In this chapter, I tried to capture the main ideas of the methods that I used throughout this thesis and tried to avoid detailed descriptions. These can be found in the following books: Huettel et al. (2008); Ashby (2011); Jezzard et al. (2002)[1]. In Section 2.1, I will briefly describe data analysis steps. In Section 2.2, I will introduce general linear modeling approach. In the last

[1]And websites such as: http://mindhive.mit.edu/ or http://imaging.mrc-cbu.cam.ac.uk/imaging/CbuImaging.

two sections of this chapter, I will describe functional connectivity approaches that go beyond functional localization. Section 2.3, includes the psychophysiological interaction (PPI) method, and Section 2.4, a rather more sophisticated method, called Granger causality that captures causality in the data.

2.1 FMRI Data Analysis Steps

Throughout this thesis I used the Statistical Parametric Mapping (SPM8)[2] software to analyze the fMRI time-series. There are several steps in analyzing an fMRI data-set:

1. **Preprocessing** steps are caried out on the reconstructed brain images to reduce the noise in the data. Following preprocessing steps are usually applied:

   a) **Slice acquisition time correction:** When fMRI volumes are constructed, each slice in each volume is acquired at different time points ($\sim 30 - 40\,ms$ apart). However, it is desirable to have the signal for the whole volume as if it was taken at the same time. Therefore, a temporal interpolation[3] is applied to the time points. At the end of this procedure, an approximate signal is constructed that simulate the signal that would have been measured if we sampled each voxel in the brain at the same time. In both experiments (Chapter 4 and 5), I acquired the fMRI slices in an ascending acquisition order, in which slices were acquired one after the other (e.g. 1-2-3-4-5-...-23-24). We further first spatially realigned the brain images and then applied slice-timing correction (Ashby, 2011).

   b) **Spatial realignment (Motion correction):** This preprocessing step corrects for subjects motion (in particular head movements) in the scanner. Using rigid body registration with six parameters (3 translations and 3 rotations), each slice of the brain volume is aligned to the first volume to correct for this motion (Ashburner and Friston, 2007). It has been shown that the SPM software, can handle head movements up to $10\,mm$ (Ardekani et al., 2001). However, following the suggestions in the literature (Ashby, 2011), in this thesis I excluded subjects that had head movements larger than the voxel size in each direction.

   c) **Normalization:** Images can be normalized to a standard MNI (Montreal Neurological Institute) template to bring brain images acquired from different subjects into a common reference space. In this procedure, each brain scan is matched to the MNI template using registration methods with 12 affine transformation parameters (3 translations, 3 rotations, 3 scalings, and 3 shears). Please note that 6 additional

[2] http://www.fil.ion.ucl.ac.uk/spm/software/spm8/
[3] In SPM8 sinc interpolation is used.

parameters (scaling and shear) are included to account for size and shape differences between individual brains and the MNI template. The standard template used in SPM8 is ICBM152 (International Consortium for Brain Mapping). ICBM152 is generated by taking the average of 152 high resolution MRI scans, which are registered into the original MNI305 template (Collins et al., 1994) using affine transformation (Mazziotta et al., 2001).

In the first fMRI experiment (Chapter 4) I used the standard normalization procedure of SPM8 (Ashburner and Friston, 1999), in which the sum of squared difference between the image to be normalized and a reference image (e.g. the MNI template) is minimized. I first normalized the mean image of the first run to the MNI template and then all images were subsequently normalized to this image.

In the second fMRI experiment (Chapter 5) I used an alternative normalization technique implemented in SPM5 (and higher) called "unified segmentation". This method uses a combination of segmentation, bias correction, and registration. It uses spatially-aligned prior tissue probability maps from other segmentations (Ashburner and Friston, 2005), which describes for each location in an image the prior probability of belonging to a tissue class (white matter, grey matter, and cerebrospinal fluid). Subsequently, Bayes rules are used to estimate the posterior of each measured voxel intensity using these prior probability maps. I first coregistered subject's anatomical image to the mean image of the first run. This anatomical image was then segmented using unified segmentation. The estimated normalization parameters from the unified segmentation are then used to normalize all other images into the same space as the anatomical image, namely the MNI space.

d) **Spatial filtering (Smoothing):** Images can be smoothed using a smoothing kernel to enhance the signal to noise ratio (SNR). However, this comes at a cost of losing spatial resolution. In this procedure, the voxel intensity is replaced by a weighted average of the intensity values in its neighboring voxels (Ashby, 2011, p. 68). The weights are determined by the smoothing kernel. In both experiments (Chapter 4 and 5) a Gaussian smoothing kernel with a $6\,mm$ full width at half maximum (FWHM) was used.

e) **Temporal filtering:** A high pass filter can be applied to the time-series to filter out low frequency noise. SPM includes a set of discrete cosine transform basis functions of different low frequencies into the design matrix to filter out the low frequency noise (Ashby, 2011). The default SPM high pass filter cut-off frequency of $0.008\,Hz$ corresponding to $128\,s$ was used in both experiments.

2. **First level analysis** is done on single subject's data and is the basic statistical analysis

method applied in fMRI studies. The goal of this step is for each subject to determine which voxels are activated by a specific task or sensory stimulation. Throughout this thesis, I used 'univariate analysis methods', in which each voxel's time series is analyzed independently. The standard general linear model (GLM), which I describe in the next section (Section 2.2) is an example of a univariate analysis method. The result of the first level analysis is a statistical parametric map for each subject denoting activated voxels in the brain (Ashby, 2011; Huettel et al., 2008). Other analysis methods such as multi-variate techniques simultaneously consider all voxels in a specific region (a region of interest or searchlight). However, this method would expand the scope of this thesis and are extensively discussed elsewhere (Haynes and Rees, 2006; Kriegeskorte, 2006; Norman et al., 2006).

3. **Second level analysis**, also called group level analysis, is the analysis in which the results from different subjects are integrated to a single statistical parametric map. The attempt is to generalize the results obtained in single subjects to a larger population from which these subjects were drawn. For this procedure it is necessary that single subject images are in the same common space (normalized). Subject's statistical parametric maps are entered into the second level analysis, and standard GLM procedures are applied to investigate the brain regions that are reflecting the stimulation or cognitive task across subjects. At the end of this procedure, the result is a 'group' statistical parametric map and standard inference procedures can be applied to these maps (see Section 2.2.1).

In the next section, I will describe a data analysis technique (the general linear model) that is widely used in fMRI data analysis.

2.2 The General Linear Model

The general linear model (GLM) is a univariate analysis method that is used to identify brain regions that show task-related brain activity. In univariate methods, each variable (in this case each voxel) is analyzed independently. In a typical fMRI experiment we measure the BOLD signal Y_j (RESPONSE VARIABLE) for each observation $j = 1, ..., J$ and voxel. Each observation can be described with a set of EXPLANATORY VARIABLES, x_{jk}, that are associated with specific experimental conditions (where $k = 1, ..., L$). A general linear model then explains the response variable Y_j in terms of a linear combination of the explanatory variables x_{jk}:

$$Y_j = x_{j1}\beta_1 + x_{j2}\beta_2 + ... + x_{jl}\beta_l + ... + x_{jL}\beta_L + \epsilon_j \tag{2.2.1}$$

In this equation, β_k are the (unknown) parameters of the model, and ϵ_j are the normally distributed errors (with zero mean and σ^2 variance). When we expand Equation 2.2.1 for each

observation, we can write it in a matrix form as following:

$$\underbrace{\begin{bmatrix} Y_1 \\ \vdots \\ Y_J \end{bmatrix}}_{BOLD\ signal} = \underbrace{\begin{bmatrix} x_{11} & \cdots & X_{1L} \\ \vdots & \ddots & \vdots \\ x_{J1} & \cdots & X_{JL} \end{bmatrix}}_{Design\ matrix} * \underbrace{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_L \end{bmatrix}}_{Parameter\ Estimates} + \underbrace{\begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_J \end{bmatrix}}_{Residuals}$$

The matrix notation of this is:

$$Y = X * \beta + \epsilon, \tag{2.2.2}$$

where Y is the measured BOLD signal of a particular subject and voxel, with J time points. The design matrix X specifies the timing of the cognitive tasks (conditions) of interest and is a matrix with J time points and L regressors (explanatory variables). β is the (unknown) parameter estimate of the model and contains L parameter weights, ϵ is the idependently, identically, and normally distributed error (with zero mean and σ^2 variance). The residual errors are defined as the difference between the actual measured signal and the estimated model: $\epsilon = Y - X\beta$. Equation 2.2.2 holds independently for each voxel.

Once the fMRI volumes are acquired and preprocessed, the next crucial step is to set up the design matrix. When we design and analyze an fMRI experiment we make the assumption that our experimental conditions (regressors) are contributing to the fMRI signal. These experimental regressors are described in the columns of the design matrix X. Each regressor is a vector with J time points. This vector contains 1's and 0's at each time point to describe the timing of each event that belongs to the condition described in the regressor. These vectors are then convolved with a canonical hemodynamic response function (HRF). In SPM this is defined as the difference of two Gamma functions (Figure 2.2.1). Further regressors can be added into the design matrix, which are not directly related to experimental manipulations but reflect other variations in the data (such as the movement parameters). These regressors are called NUISANCE REGRESSORS and are not convolved with the hemodynamic response function.

After the design matrix is constructed, the parameter estimates are computed according to the ordinary least squares method. This method estimates the unknown parameters such that an optimal parameter estimate, $\hat{\beta}$, minimizes the sum of squared errors (SSE). The sum of squared errors between the measured signal and the fitted model is $SSE = \sum_{j=1}^{J} \epsilon^2 = \epsilon'\epsilon$, where ϵ' is the transpose of ϵ, and can be re-written as:

$$SSE = (Y - X\hat{\beta})'(Y - X\hat{\beta}). \tag{2.2.3}$$

This equation is minimized when we differentiate with respect to $\hat{\beta}$ and solve following equation for $\hat{\beta}$:

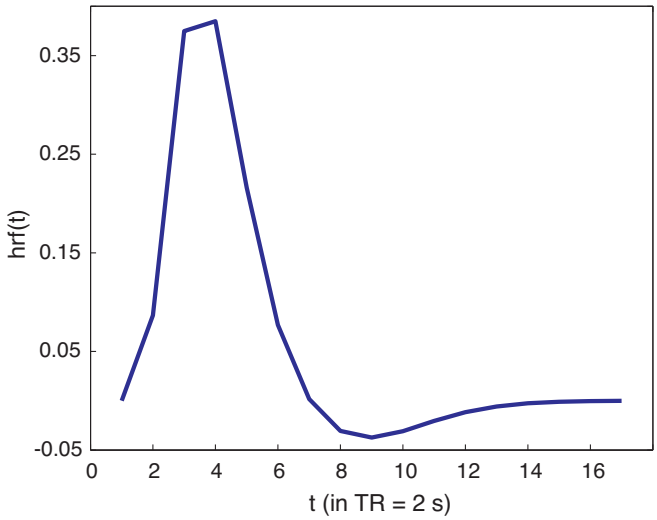


Figure 2.2.1: Hemodynamic response function as modeled in SPM. The difference between two gamma-functions with a TR of 2 s is used to model the HRF. The x-axis illustrates the time in TRs and the y-axis is the intensity weight at a particular time. This filter is used to model the BOLD response.

$$\frac{\partial SSE}{\partial \hat{\beta}} \stackrel{!}{=} 0$$

Solving this equation resolves in:

$$\hat{\beta} = (X'X)^{-1}X'Y, \tag{2.2.4}$$

where $\hat{\beta}$ is the optimal parameter estimate, and X' is the transpose of X. The parameter estimates are then used to make statistical inferences. In the next section, I will describe the statistical tests that I used in this thesis.

2.2.1 Hypothesis testing, inference, and the multiple comparisons problem

In fMRI, a voxel is described as ACTIVATED when a statistical test that is applied to the parameter estimates exceeds a certain threshold. Once the GLM is estimated, a contrast, i.e., the linear combination of parameters, is applied to the parameter estimates. STATISTICAL PARAMETRIC MAPS are generated based on these contrasts, which include the result of the statistical test of a given contrast for every single voxel in the brain. SPM converts the statistical maps to z-scores to display and make inference from the statistics. A z-score is a number from the normal distribution ($\mu = 0$, and $\sigma = 1$) and indicates by how many standard deviations an observation is above or below the mean.

T-contrast tests for differential (condition1 > condition2) activations, and represents whether

one condition effected the change in the BOLD signal more than another condition. This test compares the size of an effect (condition1 > condition2) with the variability of the data. F-contrast on the other hand, tests whether any condition effected the change in the measured BOLD signal. It can be thought as applying an OR operation across several activations. The F-contrast does not tell the experimenter anything about which of the conditions drive the effect in the measured signal. However, it is for example a good measurement if regions of interests need to be selected for further analysis in an unbiased manner (not driven by any condition). In both experiments I applied T- as well as F-contrasts to generate the statistical maps (please compare with Chapter 4 and 5).

After a statistical map is created, the next step is to make inferences based on these statistical maps. In its simplest way, we select a statistical threshold (p threshold) and apply this to every voxel in the statistical map to be able to conclude with some probability whether a task activated a certain voxel or not. To make this judgment, we first choose a false-positive rate, which indicates how many percent of the voxels that are falsely marked as active we are willing to accept (α value). Secondly, from a z-table, we obtain the statistical threshold, or criterion (z-value) that corresponds to the α value. With this we can conclude with some α-probability that a voxel is falsely judged to be active. In order words, the p threshold for a T-statistic gives the probability that the difference between the experimental and real conditions arose by chance.

In a usual whole brain analysis, a statistical map can contain over 50 000 voxels. In this case, with an α value of 0.01 (significance level of $p < 0.01$), we are willing to accept to have 500 voxels that are falsely activated, which is an unacceptably high number. This is called the MULTIPLE COMPARISONS PROBLEM. In order to avoid such high false positives, the chosen p-threshold is *corrected* to account for the number of tests we conduct. Because this type of correction prevents false positives across all voxels (the whole family of tests), it is named to FAMILYWISE ERROR CORRECTION (FWE correction) (Nichols and Hayasaka, 2003). All our results reported in this thesis are FWE corrected results.

In general, a method called BONFERRONI CORRECTION is used to correct for the multiple comparisons problem. In this method, for each voxel the significance level α is divided by the number of voxels V:

$$\alpha_{voxel} = \frac{\alpha}{V} \tag{2.2.5}$$

Hence, for every single voxel the new significance level is set to the new value, α_{voxel}. However, in fMRI, not all voxels are independent from the neighboring voxels (number of independent voxels < number of voxels). Therefore, the Bonferroni correction is considered as a conservative method for fMRI data. A method called GAUSSIAN RANDOM FIELD (GRF) theory is used to set the statistical threshold (Friston et al., 1996; Worsley et al., 1996). This method considers

the correlations between neighboring voxels by creating a Gaussian random field. Under the null hypothesis of no signal, the GRF is a randomly sampled z-map that is smoothed using a Gaussian smoothing kernel. Once the GRF is created, the statistically independent voxels are estimated (the amount of smoothness of the data also called RESOLUTION UNIT) and the euler characteristics for a given threshold T is applied to the GRF ($EC_T = \#of\ connected\ peaks - \#of\ holes$ (Ashby, 2011, p.134)). The probability that a voxel is falsely activated is then computed using the expected value of the euler characteristics for a particular threshold T. Another approach to consider is CLUSTER-LEVEL CORRECTION. This method considers a cluster of voxels rather than each single voxel separately. It computes the probability that a cluster with k and/or more voxels would occur by chance given the GRF (Friston et al., 1993; Forman et al., 1995). I applied cluster-level correction procedure in both experiments, and report the siginificant results in Chapter 4 and 5 with the corresponding cluster extent threshold of k voxels.

2.3 Functional Connectivity Analysis

The methods I described so far were based on the notion that voxels respond in isolation to a particular cognitive task or function. The statistical maps generated using the T- and F-contrasts are based on COACTIVATIONS of different voxels under certain sensory stimulation or cognitive task. Although this tells us that two or more voxels are responding to the same sensory stimulus, it does not help us to understand the underlying mechanisms of a cognitive task. In general, we are interested in explaining how anatomically separate brain regions interact during a particular cognitive task. This means, we want to "explain responses, in one cortical area, in terms of an interaction between the influence of another area and some experimental (sensory or task-related) parameter" (Friston et al., 1997).

Therefore, in this section, I will discuss a functional connectivity method, called PSYCHOPHYSIOLOGICAL INTERACTIONS that computes the correlation between the differential task-related response in one brain region or voxel (SEED REGION) with the signal measured from other brain regions (Friston et al., 1997; Rissman et al., 2004; Macaluso et al., 2000; Haynes et al., 2005b; Imamoglu et al., 2012a,b). In this type of analysis, we are not only computing the correlations between brain activity in different brain regions (or voxels) but also take into acount the experimental manipulations that can have an effect on these correlations. Therefore, it is different from other conventional functional connectivity methods, such as resting state connectivity (Greicius, 2003; Fox and Greicius, 2010; Gusnard et al., 2001).

The psychophysiological interactions is based on three factors:

1. Physiological factor: This is an estimation of the underlying neuronal signal within a seed voxel (deconvolved BOLD signal, Gitelman et al. (2003); David (2011); Friston (2011))

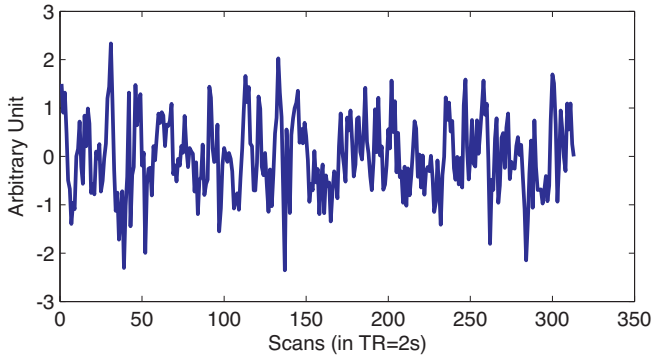


Figure 2.3.1: An example of the physiological factor of one subject and one run with 312 scans. Here, the mean signal of the a region of interest located in the prefrontal cortex (y-axis) is plotted against time (x-axis). This is the deconvolved BOLD signal and is thought to be a good estimation of the underlying neural signal.

and is a vector with n time-points (Figure 2.3.1). If a region of interest rather than a single voxel is chosen as the seed region, then the mean across voxels or a principal component analysis within this region can be used to select the time-series.

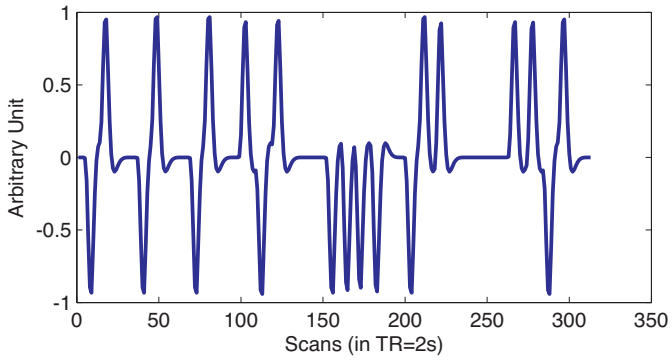


Figure 2.3.2: An example of the psychological factor of one subject and one run with 312 scans. In this example we have two conditions (coded with 1 for condition1 and -1 for condition2) and are interested in the differential activity. The time-course is convolved with the HRF function.

2. Psychological factor: The timings of the sensory stimulation or task during the experiment are depicted in this factor. This is a vector that has the same length as the physiological factor. For example, when we are interested in the effect of a differential activation (condition1 > condition2) in one brain region to explain changes in another brain region, we code in this vector 1's for condition1, and -1's for condition2 (Figure 2.3.2). This vector is then convolved with the hemodynamic response function (Gitelman et al., 2003). Other parametrical models are also possible (Haynes et al. (2005b) and Chapter 5).

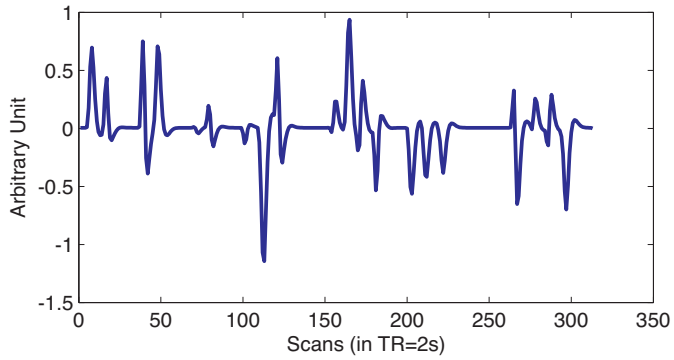


Figure 2.3.3: An example of the interaction factor of one subject and one run with 312 scans. This is the multiplication of the physiological factor with psychological factor.

3. Interaction factor: The element-wise multiplication of these two factors (physiological and psychological) results in the interaction factor that models the neural activity associated with the experimental manipulation (Figure 2.3.3).

The correlation between the condition-specific changes in the seed region and the activation in other voxels in the brain are estimated using a GLM with these three factors as regressors. This can be seen in the following equation:

$$Y = \underbrace{(C_1 - C_2)}_{Psychological\ Factor}\beta_1 + \underbrace{ROI_1}_{Physiological\ Factor}\beta_2 + \underbrace{(C_1 - C_2)ROI_1}_{Interaction\ Factor}\beta_3 + \epsilon$$

For the group analysis, the single subject statistical maps (see also Section 2.2.1) are used to make inference about the task related correlations between distinct brain regions.

2.4 Granger Causality Analysis

The functional connectivity measurement that I described in the last section, can merely refer to correlations of distant brain regions during a particular task. But it cannot tell us anything about the directionality of this connection. To understand whether feedforward (from the back of the brain to the front) or feedback (from the front to the back) connections are mediating the cognitive task, we need more advanced methods.

Therefore in this section, I will describe a directed functional connectiviy method called Granger causality. Granger causality is based on the temporal precedence and predictability of one signal over another. This method has its roots in economics (Granger, 1969; Valdes-Sosa et al., 2011), and has been adapted to neuroscience and functional neuroimaging recently (Goebel et al., 2003; Roebroeck et al., 2005; Bressler and Seth, 2011; Seth, 2007). It is based on linear autoregressive (AR) modeling. Consider two time series X and Y, which e.g. represents

the BOLD signal measured in two brain regions. X "Granger causes" Y, when the history of X contains information that can predict the future of Y over and above the information contained in the history of Y.

Because of the sluggish nature of the BOLD signal, it has been argued that Granger causality in combination with fMRI data "is not a measure of true, physical neural causality" (Huettel et al., 2008). Nevertheless, Granger causality has been applied to resting state (Deshpande et al., 2011; Liao et al., 2010) and cognitive experiments (Cole et al., 2010; Hwang et al., 2010; Imamoglu et al., 2012b), as well as simulated data sets (Deshpande et al., 2010; Roebroeck et al., 2005).

One prerequisite for a valid Granger causality computation is that the time series included into the autoregressive model need to be COVARIANCE STATIONARY. A time series is covariance stationary when the mean and variance do not change over the time course. This can be tested using a unit-root test (Seth, 2010). According to this test, a time-series is **not** covariance stationary when the signal has a unit-root. As an example, a first order autoregressive equation $y_t = \alpha_1 y_{t-1} + \epsilon_t$ is defined to have a unit-root, if $\alpha_1 = 1$. For a given time-series, the ADF tests, whether there is a unit-root, meaning no covariance stationarity. Therefore, once we can reject this null hypotheses, we can conclude that our data set is covariance stationary, and can apply Granger causality. This test is implemented in the Granger Causal Connectivity Toolbox (Seth, 2010).

Furthermore, a MODEL ORDER needs to be selected prior to the Granger causality computation. This indicates how many time points from the history of region X and region Y will be included to the predictive model. The model order can be chosen according to a priori knowledge, using the Akaike information criteria (Akaike, 1974), or a Bayesian information criteria (Schwarz, 1978). In fMRI with a TR = 2 s, it is unlikely to have sensible causalities for more than a model order of $m = 1$. Although temporal smoothing by the HRF is slow in time (around 32 s for a TR = 2 s, see also Figure 2.2.1 on page 26), it is possible to make reasonably accurate predictions about the future of the BOLD signal with a model order $m = 1$ (Roebroeck et al., 2005; Ashby, 2011, p. 223). The model order is equivalently named TIME LAG in this thesis.

To make predictions two influence measurements are computed:

- The influence of X on Y: $F_{x\to y}$
- The influence of Y on X: $F_{y\to x}$

The influence from X to Y ($F_{x\to y}$) is computed using two linear regression models (Roebroeck et al., 2005; Bressler and Seth, 2011):

1. A restricted model:

$$Y(t) = \sum_{k=1}^{m} \alpha_k Y(t-k) + \epsilon_k(t), \tag{2.4.1}$$

where t represents the time points, and m is the time lag (or model order) that indicates how many time points from the past will be included into the model. α_k is the auto-regression coefficient that reflects the regression weights, and ϵ_k is the restricted model's residuals, which is normally distributed with mean 0 and variance σ^2_Y.

2. An unrestricted model :

$$Y(t) = \sum_{k=1}^{m} \left[\alpha_k Y(t-k) + \beta_k X(t-k) + \epsilon_k(t)\right], \tag{2.4.2}$$

where α_k reflects the regression weights from region Y that was observed m time points earlier. β_k reflects the regression weights from region X that was measured m time points earlier. The time lag m indicates, as above, the number of past time points that will be included into the predictive model. ϵ_k is the unrestricted model's residuals, which is normally distributed with mean 0 and variance $\sigma^2_{Y|XY}$.

The second influence term $F_{y\to x}$ is computed in the same way. Keep in mind that these equations (2.4.1 and 2.4.2) are both linear regression equations and the parameters can be estimated as in Equation 2.2.4. At the end of this computation, the variance of the four error terms ($\sigma^2_Y, \sigma^2_{Y|XY}, \sigma^2_X, \sigma^2_{X|YX}$) are used for the quantitative measurement of causality, as these values reflect the accuracy for prediction (Ashby, 2011; Roebroeck et al., 2005).

The influence $F_{X\to Y}$ is then computed as follows (Geweke, 1982; Roebroeck et al., 2005):

$$F_{X\to Y} = ln\left[\frac{\sigma^2_Y}{\sigma^2_{Y|XY}}\right]. \tag{2.4.3}$$

$F_{X\to Y} = 0$ means no causal influence from region X to Y, whereas $F_{X\to Y} > 0$ reflects a causal influence of the brain responses in region X to region Y. Because σ^2_Y can never be less than $\sigma^2_{Y|XY}$ it is assured that $F_{X\to Y} > 0$ is always true. Furthermore, when region X does not have any causal influence over region Y, then we can conclude that $\sigma^2_Y = \sigma^2_{Y|XY}$ and $F_{X\to Y} = 0$ (Ashby, 2011, p. 229-230).

It is known that the Granger causality measurement is sensitive to noise (Roebroeck et al., 2005; Valdes-Sosa et al., 2011; Goebel et al., 2003; David, 2011; Roebroeck et al., 2011a). Furthermore, Ashby (2011, pages 224 and 230) clearly illustrates, when different HRF's are used to model the brain responses in regions X and Y, than the causal influence is lower than when the same HRF would be used in both regions. It has been previously shown that the HRF varies across different brain regions (Aguirre et al., 1998; Handwerker et al., 2004). Therefore, we need to take this possible confound into account. Hence, to make statistical inferences, Roebroeck et al. (2005) suggested to compute not only the individual causalities ($F_{x\to y}$, $F_{y\to x}$), but their differences (difference of influence measure, DOI): $F_{x\to y} - F_{y\to x}$ (Roebroeck et al.,

2005, but see Barnett and Seth, 2011[4]). A consequence of this DOI measurement is that it detects the dominant direction of influence and cannot measure reciprocal connections. Hence, DOI indicates only whether there is comparatively more feedforward than feedback connections (or vice versa).

Because the causal influence measurements lack a statistical distribution, a surrogate statistics is necessary to make statistical inference (Seth, 2010; Roebroeck et al., 2005). Hence, the significance threshold needs to be determined using a permutation test and the bootstrapping procedure (Efron and Tibshirani, 1993).

In the single subject level, three steps are needed to generate the surrogate statistics (empirical null distribution): First, for each subject and ROI, the trial-orders are shuffled (or only the labels of the trials can be rearranged). Second, this procedure is repeated a certain amount of time (e.g. 1000 times) to create surrogate time-series. Third, the Granger causality is applied repeatedly for each of these surrogate time-series. Since the trial-orders were shuffled, an observed causality using these time-series ($DOI > 0$) can only be detected due to chance.

In the group statistics, random samples of subject's surrogate DOI measurements are selected to generate the group null distribution of no influence. Afterwards the DOI measured from the real time-series is tested against the null hypothesis that the difference of influence measure (DOI) is zero.

Before conclusions can be drawn about the directionality of the signal in fMRI, some validation tests need to be carried out (Seth, 2010):

1. Model consistency: Ding's consistency method can be applied to test what portion of the correlation structure in the data is captured by the model (Ding et al., 2000). This method computes the difference between the correlation structure of the real data and the correlation structure of a simulated surrogate data. As a rule of thumb the model consistency should be $> 80\%$.

2. Serially uncorrelated residuals: The Durbin-Watson test (Durbin and Watson, 1950) can be applied to test whether the residuals of the auto-regressive model are serially uncorrelated.

3. Sum of squared errors: The adjusted sum of squared errors of the regression model can be examined as an additional validation. According to Seth (2010), if this value is < 0.3 there is a cause for concern.

Note that the Granger Causal Connectivity Analysis (GCCA) Toolbox (Seth, 2010) is a very useful toolbox implemented in MATLAB to compute Granger causalities with built-in preprocessing and validation tests.

[4]Barnett and Seth present in their recent manuscript that the Granger causalities are invariant to filtering.

The bane of my existence is doing things that I know the computer could do for me.
Dan Connolly

3

Experiment 1: MoonBase - An Automated Two-tone, Mooney Image Database

Figure 3.0.1: A two-tone image created by MoonBase toolbox.

In this chapter, I will introduce our image database (MoonBase) that contains automatically generated two-tone, Mooney images. I will first describe our custom algorithm that creates these images automatically. In the remaining part of this chapter, I will present results of a behavioral experiment that we conducted using these images (Imamoglu et al., 2013).

A Mooney image is a thresholded, binary image that consists a *hidden* object (a foreground) and a background. The *hidden* object is hard to recognize at first sight with recognition times in the second to minute range (Figure 3.0.2). In most of the cases, the recognition is abrupt and gives rise to a feeling as if having solved a difficult problem (aha-feeling or eureka-effect). This abrupt recognition can occur intrinsically (Imamoglu et al., 2012b), after the contour of the object is marked (Tallon-Baudry et al., 1997), or after presenting the subject with the original image (Ludmer et al., 2011; Hsieh et al., 2010; Dolan et al., 1997). In the two latter cases the recognition time is much accelerated.

Such images have already been used to study different visual as well as cognitive phenomena, such as perceptual closure (Mooney, 1957), visual consciousness (Imamoglu et al., 2012b), human memory and insight (Ludmer et al., 2011), object and face recognition (Dolan et al., 1997; Andrews and Schluppeck, 2004; McKeeff and Tong, 2007; Hsieh et al., 2010; Tallon-Baudry and Bertrand, 1999), visual search (Tallon-Baudry et al., 1997), and crowding (Farzin et al., 2009). Most of these studies used manually man-made two-tone images. Therefore, the amount of images that were available in these studies was limited (max. 40 images in Ludmer et al. (2011)). Only recently we were able to use a large amount of unique two-tone images to study human visual consciousness (120 images in Imamoglu et al. (2012b) and see Chapter 4).

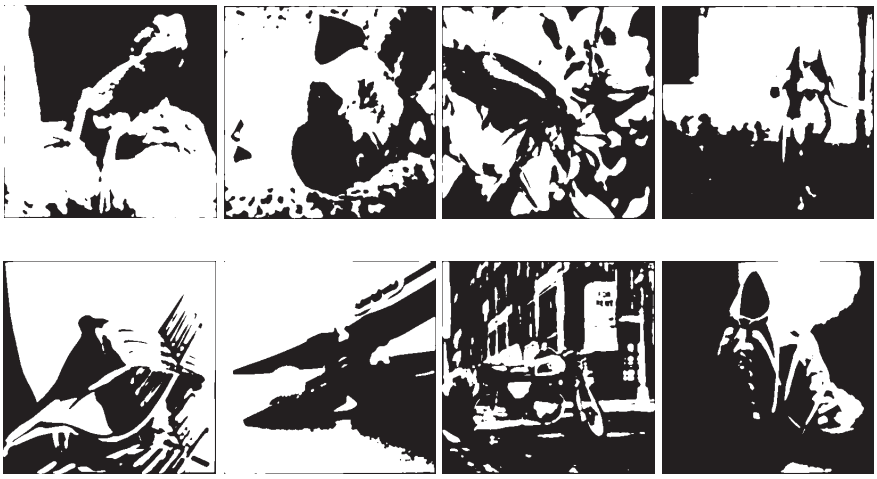


Figure 3.0.2: Examples of two-tone, Mooney images. Upper panel (left to right): A lizard, a dog, a bug, a woman. Lower panel (left to right): Forks, a plier, a motorcycle, a pair of shoes.

I see two reasons, why an image set that contains a large amount of different two-tone images can be useful in vision and cognitive science experiments:

1. Subjects recognize the *hidden* object in a previously recognized Mooney image immediately when they see the same image at another time point. This means that the same image cannot be used twice in the same experiment, when the recognition process is studied. This

characteristic is a drawback for experiments that need many repetitions of the stimulus presentation. However, when there is an image database with many different Mooney images, subjects can be presented with the necessary amount of unique images for an experiment.

2. Subject's chance to make a correct guess is higher, when only a certain number of object categories are used in an experiment. Therefore, it is essential to have a two-tone image database that has a large variety.

In Section 3.1, I will first describe the image collection method, and will introduce the preprocessing and thresholding steps used in this algorithm. In Section 3.2, I will introduce the behavioral experiment and results. In the last section (Section 3.3), I will present future directions that can improve the algorithm.

We named our two-tone image database, "MoonBase" (**Moon**ey Data**base**) as the first two-tone face images were created and used in a study by Craig Mooney (Mooney, 1957).

3.1 Methods

MoonBase is a large two-tone image database created using concrete words. We first selected concrete words[1] from the Medical Research Council Psycholinguistic Database (Wilson, 1988)[2] to ensure a large variety of different objects in our database. We then collected images in an automated fashion using these concrete words as search tags in an online image database. We automatically downloaded 20 images per concrete word. These images were preprocessed and using a thresholding algorithm automatically transformed to two-tone images (Figure 3.1.1).

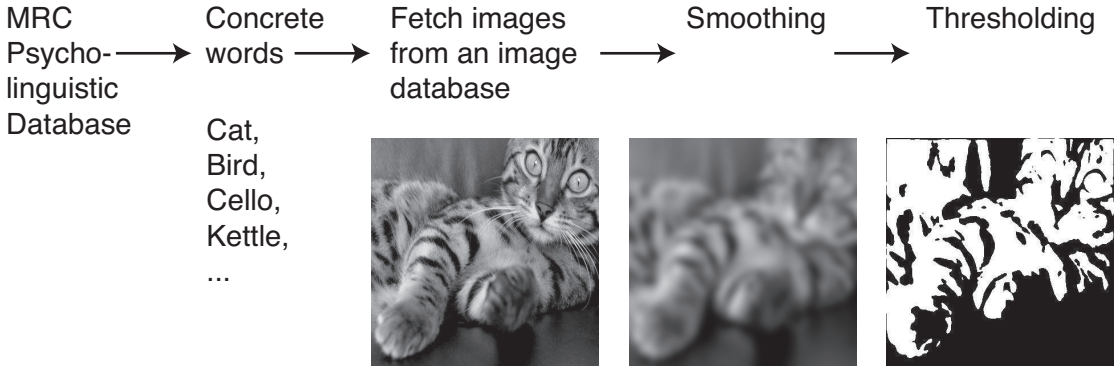


Figure 3.1.1: Automatic generation of Mooney images.

The preprocessing steps were following:

[1] Concreteness and imagability rating range were 550 – 700.
[2] http://www.psych.rl.ac.uk/MRC_Psych_Db.html

1. Converting the colored images to gray scale images
2. Smoothing the images using a 2D smooth operation with a Gaussian kernel ($\sigma = 2$ pixels and full width at half maximum (FWHM) = 5 pixels).
3. Resizing the images to have a size of 400 x 400 pixels. In this case the image is subsampled with a $scale\,factor = new\,size/old\,size$.

After preprocessing, we thresholded the images using a histogram based image thresholding algorithm implemented in Matlab (Otsu, 1979). Otsu's thresholding method assumes that each image has two classes of pixel properties: A foreground and a background. For each possible threshold the algorithm iteratively computes the separability of the two classes and converges when the maximum separability is reached.

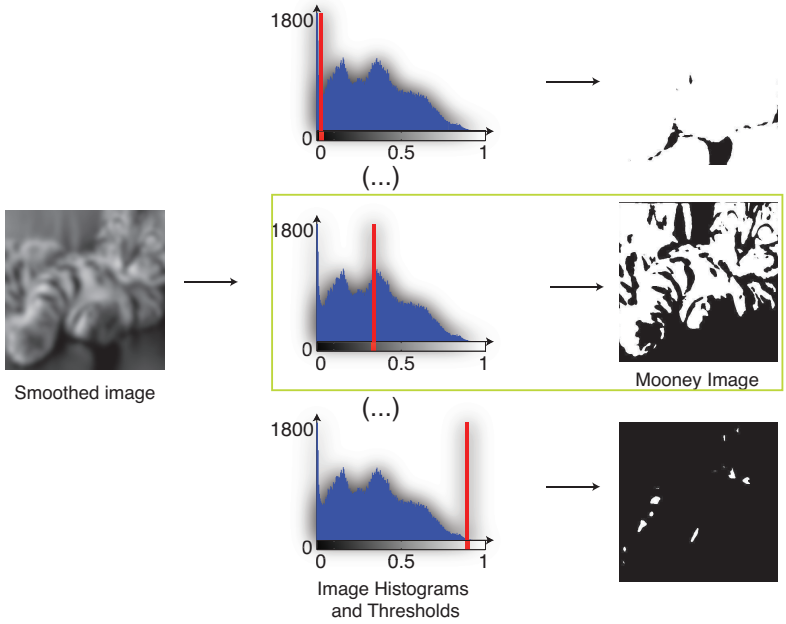


Figure 3.1.2: Threshold selection from the gray scale distribution.

An image can be considered as a two-dimensional matrix with gray scale values $(1, 2, 3..., L) \in [1, 256]$ in every position in the matrix. Otsu's algorithm first creates a distribution of the number of pixels for each gray scale value (image histogram). Then, it computes the probability distribution of the image histogram by computing the probability p_i for each gray scale value:

$$p_i = \frac{n_i}{N}$$

where n_i is the number of pixels that are attributed to gray scale value i, and N the total number of pixels. The algorithm assures the best separability between the foreground and the

background (ideally the object and its background) by maximizing the between class variance, using the following equation (T_{opt}):

$$T_{opt} = argmax\left(\frac{(\mu_{total}\,\omega_k - \mu_k)^2}{\omega_k(1-\omega_k)}\right)$$

where k is a possible threshold (one of the gray scale values in the image), $\mu_{total} = \sum_{i=1}^{L} ip_i$ is the total mean of the gray scale values in the image, $\omega_k = \sum_{i=1}^{k} p_i$ is the cumulative sum of the probabilities of the gray scale values up to gray scale value k, and $\mu_k = \sum_{i=1}^{k} ip_i$ is the mean of the gray scale values up to gray scale value k.

Once the optimal threshold is selected, for each pixel x in image I the thresholding is computed by following:

$$I(x) = \begin{cases} 1 & \text{if } I(x) > T_{opt} \\ 0 & \text{else} \end{cases}, \text{ where } T_{opt} \text{ is the optimal threshold}$$

At the end of this thresholding procedure, we have a binarized image with 0's and 1's for black and white respectively (see also Figure 3.1.2).

Once the images are automatically downloaded and thresholded, a manual clean up session by human subjects needs to be done. This manual cleaning session is necessary because some images that are automatically downloaded from the web may not include the object that corresponds to the search word (e.g. cat) (Deng et al., 2009) and therefore needs to be removed from the image set. This processing step was done manually for our purpose. However, in Section 3.3, I will discuss possibilities to automate this step as suggested by others (Schroff et al., 2007; Deng et al., 2009).

3.2 Behavioral Experiment and Results

In this section I will describe a behavioral study that we conducted to acquire psychophysical measurements for the images in the database. We selected images based on these psychophysical measurements and used these images in Experiment 2 (Chapter 4).

3.2.1 Participants and experimental design

Six subjects (aged 23-31, mean 26 years, 3 females) participated in this experiment. All subjects had normal or corrected to normal vision, were paid for their time, and gave informed consent. We used 330 two-tone, Mooney images (76 concrete words) that we automatically generated as described in the previous Section 3.1.

The experiment started with a Mooney image presentation for 20 seconds (Figure 3.2.1). Subjects' task was to press a button as soon as they recognized an object in the image (T_{recog}).

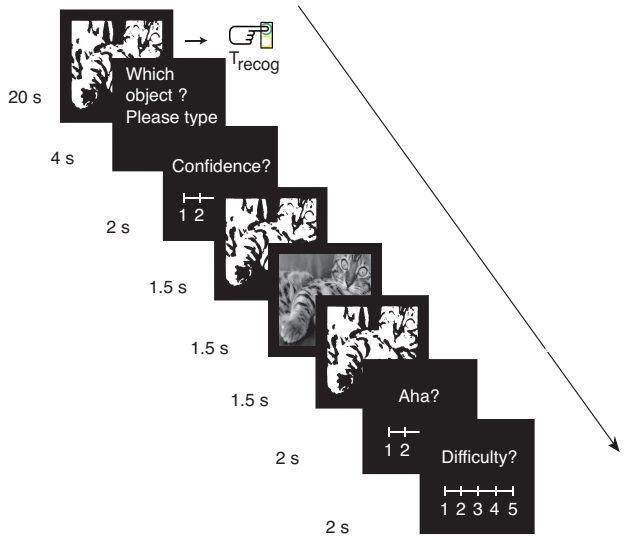


Figure 3.2.1: Experimental design.

When subjects pressed a button, indicating they recognized an object, a screen followed for 4 s, in which they were asked to type the name of the object they recognized. This followed a confidence rating, where subjects were asked to rate how confident they were in their response (2 s). Using the left and right arrow keys on the keyboard, subjects could choose one point on a continuous scale ranging from 1 (not confident) to 5 (very confident). Regardless of recognition (button pressed or not), a continuous stream of Mooney image, original image, and again the same Mooney image was presented to the subjects (each screen lasted for 1.5 s). After the presentation of the original image, two more subjective ratings (2 s each) were collected: (i) Aha rating: Subjects were asked to rate their aha-feeling after they saw the original image on a continuous scale ranging from 1 to 5 (1: no aha-feeling, 5: high aha-feeling). This is at highest when subjects abruptly recognize the previously uninterpretable object in the Mooney image. (ii) Difficulty rating: Subjects were asked to rate how difficult it was to recognize the hidden object in the Mooney image on a continuous scale ranging from 1 to 5 (1: very easy, 5: very difficult).

3.2.2 Results

Recognition times

Average recognition time to all 330 images was $9.36\,s \pm 7.40\,s$ (Figure 3.2.2a). All subjects responded to 113 images and 59 of these images were correctly recognized by all subjects. The mean recognition time for the 113 images was $4.32\,s \pm 3.83\,s$ (button pressed: correct or incorrect responded images). This was $3.18\,s \pm 3.06\,s$ for the 59 correct responded images (Figure 3.2.2b).

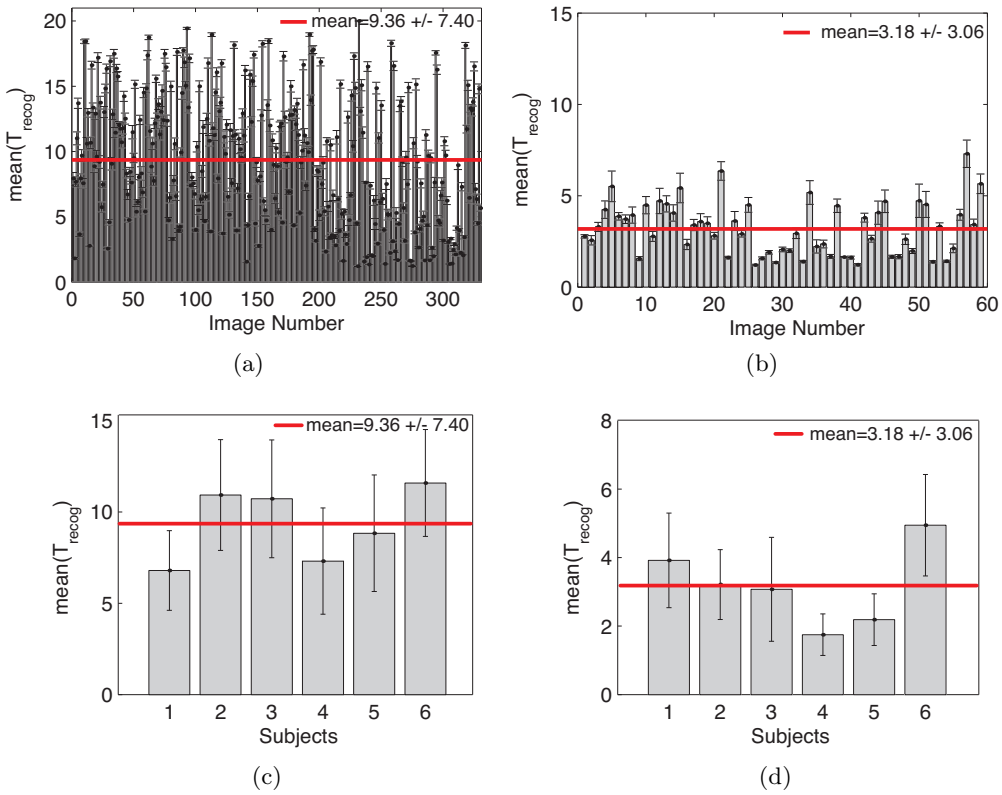
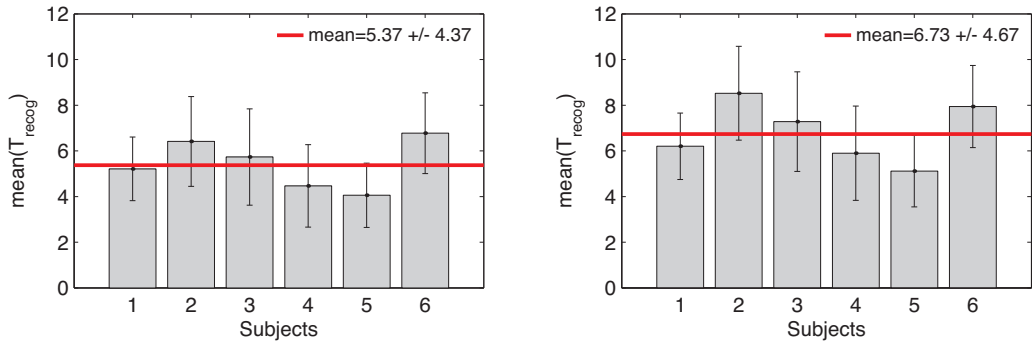


Figure 3.2.2: Average recognition time across subjects and images. Error bars indicate standard error of the mean (s.e.m.). (a) Mean recognition time (across subjects) for all 330 images. (b) Mean recognition time (across subjects) for 59 correct recognized images. (c) Mean recognition time (across images) for all 330 images. (d) Mean recognition time (across images) for 59 correct recognized images.

The mean recognition times to correct responded images were significantly shorter than to incorrectly recognized images ($p < 4.46 \times 10^{-4}$, see also Figure 3.2.3). Indicating that correctly recognized images were easier to recognize.

On average $54.95\% \pm 33.17$ of the subjects recognized the *hidden* objects correctly (Figure 3.2.4a). The fluctuation in percent subjects is apparent in the bar plot in Figure 3.2.4a. This indicates a high inter-subject variability. An object was incorrectly recognized on average by $26.11\% \pm 25.55$ of the subjects. Furthermore, $18.89\% \pm 20.09$ of the subjects did not recognize the *hidden* object in the image.

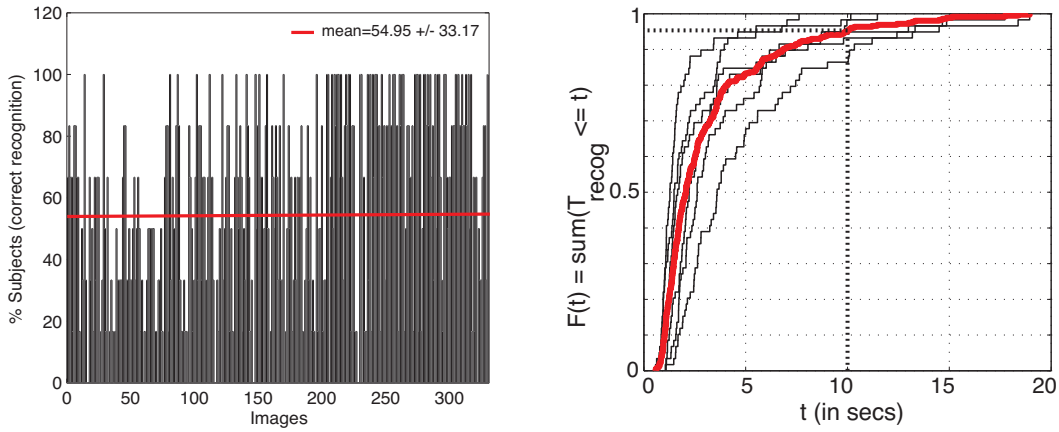
We further computed the percentage of the images that were recognized correctly until a certain time. We computed the cummulative distribution function (CDF) by pooling all T_{recog} across subjects and images. $F(t)$ is the cumulative count of all images less than or equal to a possible recognition time point t. For example if $t = 10\,s$, $F(10\,s)$ is the cummulative count



(a) (b)

Figure 3.2.3: Average recognition time to correctly or incorrectly recognized images. Here, images that were recognized by at least one subject correctly or incorrectly are grouped. Error bars indicate standard error of the mean (s.e.m.). (a) Mean recognition times of subjects to correct responded images. (b) Mean recognition times of subjects to incorrect responded images.

of images that were recognized from $t = 1\,s$ to $t = 10\,s$. We computed the CDF for correctly recognized images and plotted the mean CDF (red solid curve) across six subjects (black solid curves) (Figure 3.2.4b). The dashed black line at $t = 10\,s$ indicates that over 90% of the presented images were recognized until this time point. Therefore, we selected $10\,s$ as the Mooney image presentation duration in the fMRI experiment described in the next Chapter 4.

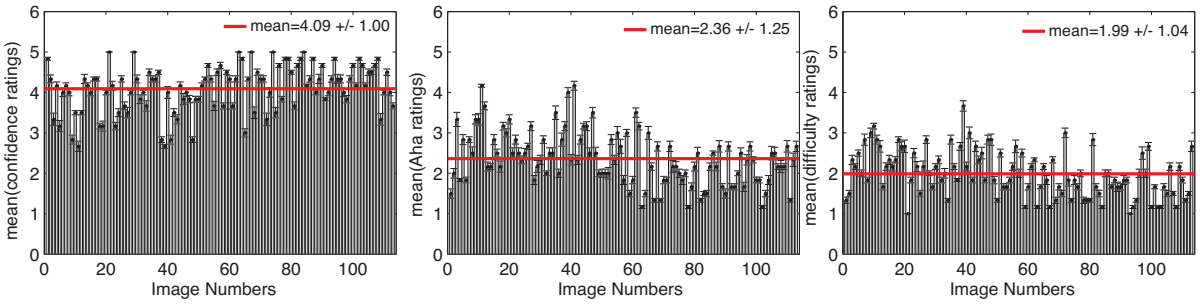


(a) (b)

Figure 3.2.4: Average rate of correct recognition and cummulative distribution function of correct recognized images across subjects. (a) Average rate of correct recognition across subjects. (b) Cummulative distribution function of T_{recog}.

Subjective ratings

In each trial we collected three subjective ratings: confidence ratings, aha ratings, and difficulty ratings. Each rating was selected from a continuous scale ranging from 1 (at least) to 5 (at most). In Figure 3.2.5 the average values of the ratings across subjects are represented for each image. Subjects were highly confident when they indicated that they recognized an object in the image. Average confidence rating for correctly and incorrectly recognized images pooled together (113 images) was 4.09 ± 1.00 (Figure 3.2.5a). This was slightly higher (4.38 ± 0.82) when only correctly recognized images were analyzed (59 images).



(a) (b) (c)

Figure 3.2.5: Mean subjective ratings for recognized images. Error bars indicate standard error of the mean (s.e.m.). (a) Confidence rating (1: not confident at all, 5: very confident). (b) Aha rating (1: no aha-feeling, 5: high aha-feeling). (c) Difficulty rating (1: very easy, 5: very difficult).

Subjects rated their aha-feeling on average with 2.36 ± 1.25 (Figure 3.2.5b). This was lower (1.98 ± 1.05) for only correctly recognized images. The subjective aha-ratings were collected after the subjects saw a continuous stream of Mooney image, original image, and Mooney image (see also 3.2.1). Hence, it is interesting to analyze subjective aha-ratings in not recognized images (no button press). As expected, the aha-ratings were on average higher (3.80 ± 1.26) when only not recognized images were considered. Images with similar characteristic (high aha-feeling only after the exposure of the original image) has been used in other studies (Dolan et al., 1997; Hsieh et al., 2010; Ludmer et al., 2011). Subjects on average did not rate the recognition task as difficult (1.99 ± 1.04, and 1.68 ± 0.82 for correctly recognized images, Figure 3.2.5c).

We were further interested in the correlations between the recognition times and the different ratings. For correctly recognized images, T_{recog} values were positively correlated with subjective aha ratings ($\rho = 0.55$, $p < 10^{-4}$, Figure 3.2.6a) and difficulty ratings ($\rho = 0.46$, $p < 0.0002$, Figure 3.2.6c). Not surprisingly, the fast recognition times were reflected in low aha or difficulty ratings. Furthermore, a negative correlation between T_{recog} and the confidence ratings were observed ($\rho = -0.65$, $p < 10^{-4}$, Figure 3.2.6b). As expected, the more time subjects needed to

recognize the objects the less confident they were. However, notice that the confidence rating does not fall below a value of three, which means subjects were in general confident when they recognized the images correctly.

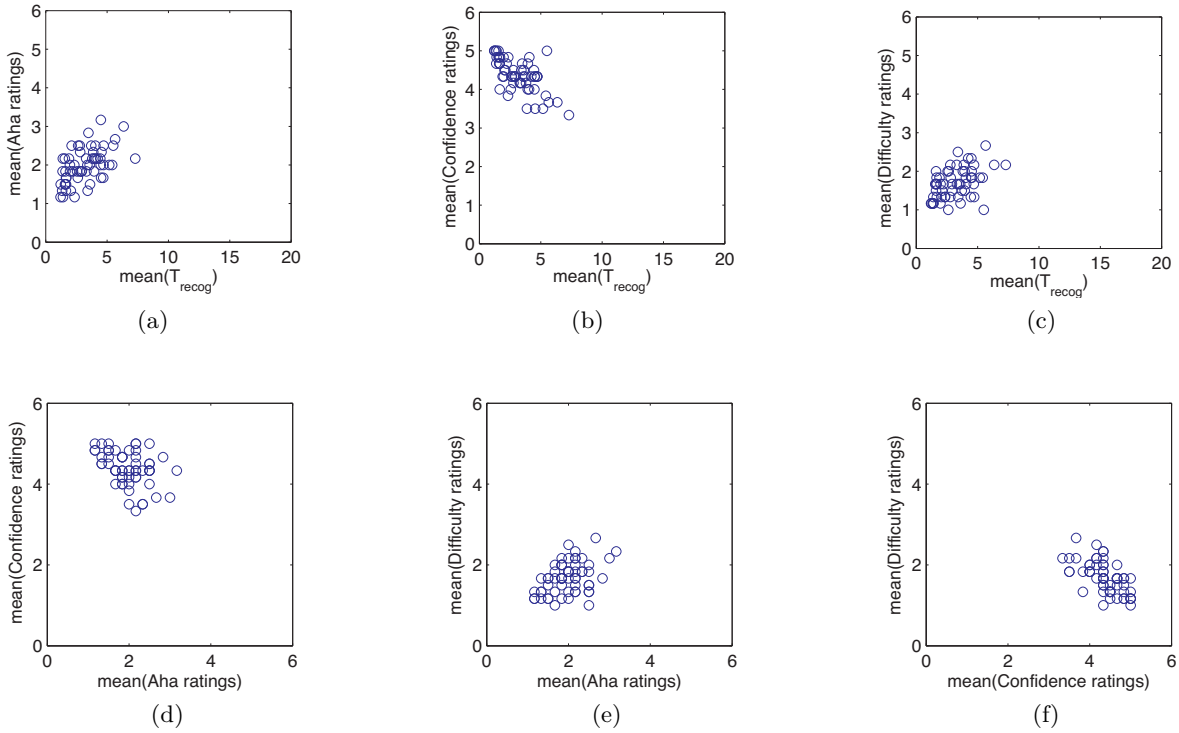


Figure 3.2.6: Correlations between subjective ratings and T_{recog}. Upper panel: Correlations between the subjective ratings and T_{recog} for correct recognized images. (a) T_{recog} vs. Aha. (b) T_{recog} vs. Confidence. (c) T_{recog} vs. Difficulty. Lower panel: Correlations between different subjective ratings for correct recognized images. (d) Confidence vs. Aha ratings. (e) Aha vs. difficulty ratings. (f) Confidence vs. difficulty ratings.

Furthermore, correlations between different subjective ratings are plotted in the lower panels of Figure 3.2.6. As expected, confidence ratings were negatively correlated with aha and difficult ratings ($\rho = -0.38$, $p < 0.003$, and $\rho = -0.62$, $p < 10^{-5}$, respectively. Compare with Figure 3.2.6d and Figure 3.2.6f). This indicates that with increasing confidence, subjects' aha ratings and difficulty rating decrease. Furthermore, a positive correlation was observed between aha and difficulty ratings ($\rho = 0.43$, $p < 0.0006$. See also Figure 3.2.6e).

3.3 Conclusions and Future Directions

In this chapter, I have presented a unique two-tone, Mooney image database with 330 images. Typically, Mooney images need to be painstakingly generated by hand. Here, I presented an approach to automatically generate a two-tone image database, and psychophysical results obtained from six subjects. This algorithm is available as a Matlab toolbox. Using this toolbox, the researcher can add new object names as search words and create new two-tone images easily. This database is, to my knowledge, the largest two-tone image set available to the vision and cognitive science research community[3]. Nevertheless, improvements can be implemented in different stages of the algorithm:

1. Improvement in the psycholinguistic database: We used the MRC Psycholinguistic database to select the concrete words. As an alternative the WordNet linguistic database[4] can be used in further applications. WordNet brings the advantage that the words in this database are categorized by their semantical content (Miller, 1995; Fellbaum, 1998). Furthermore, it can be fully downloaded. Hence, the word selection step can be automatized.

2. Improvement in the gray-scale image selection procedure: The reliability of the images that are automatically downloaded from online image databases using specific search words is quite low. Therefore, candidate images need to be manually selected from a large amount of images that are automatically downloaded. A unique image database called ImageNet uses the properties of WordNet to download images (Deng et al., 2009). In a second cleaning session, it uses Amazon Mechanical Turk agents to outsource the image selection procedure. This database has the advantage that images with low accuracy are filtered and excluded from the database. This could be a very useful approach to include in our toolbox, as it combines two very useful processes: Word selection from WordNet, and image selection from World Wide Web.

3. Improvement in the Mooney image selection procedure: In Section 3.2, we have seen that the recognition times and correct recognition of the objects in the Mooney images vary extremely from subject to subject (high inter-subject variability). Once the two-tone images are automatically generated, Amazon Mechanical Turk (AMT) workers can be used to collect psychophysical properties for each image from a large population of observers. Hence, for different cognitive experiments, the suitable images can be selected in a finer fashion. AMT workers were successfully applied previously in vision experiments (Sorokin and Forsyth, 2008; Deng et al., 2009).

[3] https://sites.google.com/site/hayneslab/links
[4] http://wordnet.princeton.edu/

4. Let the machines learn[5]: Machine learning techniques could be further implemented to learn the properties of two-tone images to create suitable two-tone images in a fully automated fashion.

5. Many-tone images: When we design a cognitive experiment, we would like to control subjects' recognition times. Using an adaptive thresholding experiment, many-tone images can be created to control for inter-subject variability in the recognition times. These images contain an optimal number of tones such that a particular image can be recognized within a certain time range[6]. An algorithm proposed by Liu (2009) could be used to create such images. This algorithm automatically determines the number of thresholds to be found for an input image (it selects different thresholds for different regions). This would allow to create different gray-tones for different object regions.

In the next chapter, I will introduce an fMRI experiment that used a subset of this Mooney image database.

[5] In collaboration with Sebastien Crouzet
[6] Personal communication with Felix Wichmann.

No amount of experimentation can ever prove me right; a single experiment can prove me wrong.
Albert Einstein

4

Experiment 2: Conscious Object Recognition and Functional Connectivity

In this chapter[1], I will describe an fMRI experiment that we conducted using two-tone, Mooney images to investigate whether a distributed network of brain regions including the frontoparietal network was involved in conscious recognition of objects or not. In addition, we investigated functional connectivity changes between high-level brain regions and low-level brain regions during conscious object recognition (Imamoglu et al., 2012b).

The involvement of high-level sensory brain regions and prefrontal cortex in conscious perception of complex stimuli such as faces, words or objects has been studied previously (Beck et al., 2001; Kleinschmidt et al., 1998; Grill-Spector et al., 2000; Dehaene et al., 2001). In addition, it has been suggested that large-scale brain networks extending from primary visual cortex to prefrontal cortex are important for conscious visual perception (Crick and Koch, 1995; Dehaene et al., 2001; Haynes et al., 2005b; Lumer et al., 1998). However, to which degree the functional connectivity between low-level and high-level regions contributes to conscious perception has been rarely studied (Lumer and Rees, 1999; Dehaene et al., 2001; Haynes et al., 2005b). In addition, whether the functional connectivity between distant brain regions is mediated by feedforward or feedback connections is still under debate (Gaillard et al., 2009; Boly et al., 2011).

Two confounds make the study of conscious object recognition challenging:

[1]This chapter is mainly an adaptation of our manuscript (Imamoglu et al., 2012b).

1. Processes related to conscious recognition need to be dissociated from stimulus-driven (only related to physical properties of the stimulus) processes. This can be accomplished by an experiment in which a constant physical stimulus leads to a change in perception, as with Mooney images (see Chapter 3). These images have the characteristic that at first sight, only unorganized black and white patches are perceived. Yet after some time, a distinct object is recognized (Figure 3.0.2 on page 35). At the time of recognition, the constant physical input triggers a mechanism that leads to the conscious recognition of an object where previously only random black-and-white patches were perceived.

2. Processes related to conscious recognition need to be dissociated from the processes associated with the behavioral report. For example, a motor response as a behavioral report can activate the prefrontal cortex, which reflects the preparation of motor action rather than conscious recognition. This confound needs to be excluded.

We used 120 custom generated, two-tone, Mooney images to control for the first confound (Mooney (1957), see also Appendix B.1). In each trial, one of these images was presented to the subjects for ten seconds. Subjects' task was to report the time at which they recognized the object (T_{recog}) and its name (Figure 4.1.1 on page 49). Furthermore, we used two different response modalities ("button" and "memory", see Figure 4.1.1) to account for the second confound: In case of "button" runs subjects pressed a button as soon as they recognized an object in the image. In "memory" runs, a random sequence of single digits was superimposed on the image and subjects had to report the number that temporally coincided with their recognition. We compared the two different perceptual conditions (*recognized* vs. *not recognized*) and searched for modality-independent changes in neural activity using a conjunction analysis (Price and Friston, 1997; Nichols et al., 2005). If recognition is independent of behavioral response, the results in both response modalities will be the same.

In addition, we examined the role of functional coupling between a high-level region, the dorsolateral prefrontal cortex, and the remaining voxels in the brain during conscious recognition of objects using the psychophysiological interaction (PPI) analysis (Friston et al., 1997). The dorsolateral prefrontal cortex (DLPFC) have been previously shown to be necessary for conscious perception (Sahraie et al., 1997; Lau and Passingham, 2006) and is therefore a good candidate to investigate functional connectivity. Furthermore, we applied a Granger causality analysis to understand whether the interaction between the DLPFC and visual brain regions are mediated by feedforward or feedback mechanisms.

In Section 4.1, I will describe the methods we used in this experiment. I will present the results in Section 4.2, and give a conclusion in Section 4.3.

4.1 Methods

4.1.1 Participants

Twenty-eight naïve participants took part in this fMRI experiment (aged 19-34, mean 26 years, 11 females). All participants had normal or corrected to normal vision and were paid for their time. Data from nine participants were excluded either due to poor task performance (80% of the images not recognized correctly; two participants), technical problems (five participants) or strong head movements (two participants). Hence, the results of 19 subjects are reported here. The local ethics committee approved the experimental procedure and all subjects gave informed consent.

4.1.2 Stimuli: Two-tone, Mooney image selection for the fMRI analysis

Stimuli selection was based on a behavioral experiment that we conducted prior to the fMRI experiment (Section 3.2). In this behavioral experiment, Mooney image presentation was 20 seconds and subjects' task was to press a button as soon as they recognized the *hidden* object in the image. When they recognized the image, they were asked to type the name of the object using a keyboard. Overall, we analyzed the data from six subjects. We selected images that were correctly recognized within $3 - 20\,s$ image presentation time by at least two different subjects. Not recognized images, incorrectly recognized images, and images that were recognized too fast (before three seconds) were discarded from the image set. The inter-subject variability of recognition is very high. Therefore, it was challenging to select images that were correctly recognized within the same time interval by more than two subjects.

At the end of the image selection procedure we had 128 images. We used 120 images from these and excluded eight images to have a balanced stimulus set (60 animate, 60 inanimate images). These images can be seen in Appendix B.1. The mean responses of six subjects from the behavioral experiment to these 120 images are listed in Appendix B.2.

The cummulative distribution function across all recognition times in the behavioral experiment showed that in 90% of the cases subjects either recognize the *hidden* object within the first $10\,s$ or they do not recognize it (see also Figure 3.2.4b on page 41). Therefore, the stimulus presentation time for the fMRI experiment was set to $10\,s$.

Stimuli were projected through the rear bore of the scanner onto a translucent screen (screen size: 26.5°, stimulus: 10° of visual angle) and viewed through a mirror mounted on the head coil.

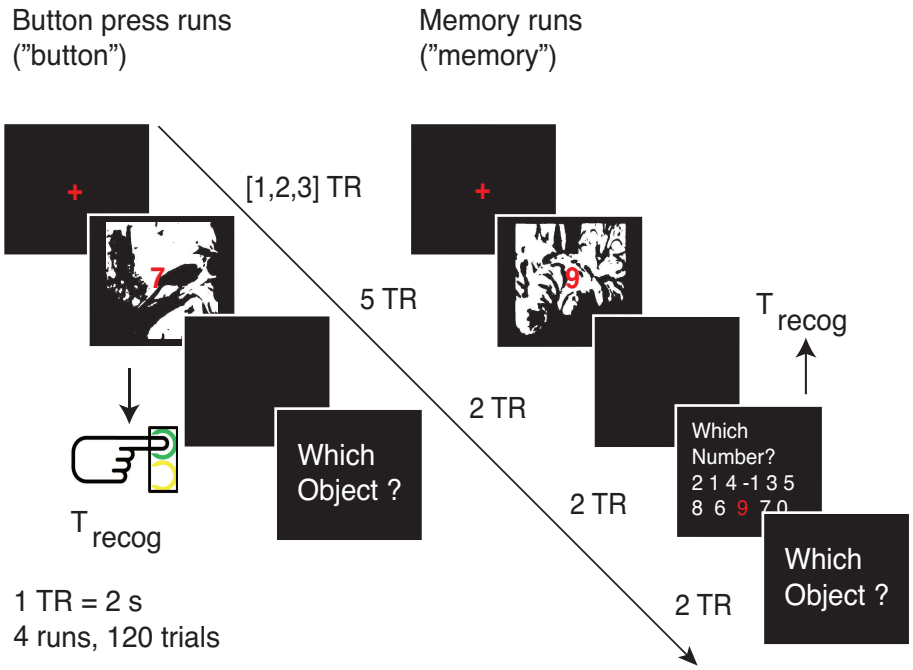


Figure 4.1.1: Experimental design.

4.1.3 Experimental design

There were four experimental runs (Figure 4.1.1 on page 49). Each run consisted of 30 trials, resulting in 120 novel image presentations for the entire experiment. On each trial, a Mooney image was presented with a single, red digit (0-9) superimposed at the center. This number randomly changed every second, resulting in ten different numbers for the 10 s stimulus presentation. We instructed participants to freely view the image and to report (i) the time of recognition (T_{recog}) as soon as they recognized an object, and (ii) which object they recognized.

We used two response modalities and instructed subjects as follows:

1. "button": Press a button with the right index finger as soon as you recognize an object while ignoring the central number (T_{recog} = time of button press).

2. "memory": Look at the number, which is present at the center of the screen just at the time of recognition and report this number at the end of the stimulus presentation (T_{recog} = time of selected number presentation ±0.5 seconds) (Soon et al., 2008).

Each trial started with a fixation cross of variable duration (1, 3, or 5 s), followed by the 10 s stimulus presentation and terminated by a 4 s blank screen. Subsequently, subjects verbally reported the name of the recognized objects into an MR compatible microphone (VisuaStim) (response screen 4 s). When subjects were performing "memory" runs, a number selection screen for 4 s followed the blank screen. On this screen, subjects were asked to select their time

of recognition (numbers were randomized from trial to trial and -1 indicated no recognition). When no object was recognized (no button was pressed or -1 was selected), the next trial started after the blank screen. In both "button" and "memory" runs, responses were recorded during the scanning session and evaluated afterwards. Only correctly recognized and not recognized trials were used for the fMRI analyses.

4.1.4 FMRI data acquisition and preprocessing

A Siemens Trio 3-Tesla scanner equipped with a 12-channel head coil was used to acquire functional MRI volumes. T2*-weighted gradient-echo echo-planar images (EPI) containing 33 axial slices ($4\,mm$ thick, ascending) with a $1\,mm$ inter-slice gap were acquired with the following imaging parameters: repetition time (TR) = $2,000\,ms$, echo time (TE) = $30\,ms$, flip angle = $90°$, matrix size = 64×64 and field of view (FOV) = $192\,mm$ resulting in a voxel size of $3 \times 3 \times 5\,mm^3$. High resolution T1-weighted structural data was collected for anatomical localization, with TR = $1,900\,ms$, TE = $2.52\,ms$, matrix size = 256×256, FOV = $256\,mm$, 192 slices ($1\,mm$ thick), flip angle $9°$.

We used SPM8[2] to preprocess the functional scans and further analyze the data. Functional images were corrected for motion and slice-acquisition time, and were spatially normalized to an EPI template based on the standard Montreal Neurological Institute (MNI) space. Voxels were spatially smoothed with a $6\,mm$ FWHM Gaussian kernel.

4.1.5 FMRI data analysis

The goal of this analysis was to identify the brain regions that were differentially activated in *recognized* vs. *not recognized* trials. For this analysis, we first selected trial periods that were either uniquely *not recognized* or uniquely *recognized*. We first marked valid trials as following: We assigned trials that were correctly recognized within the first four seconds ($T_{recog} < 4\,s$) as *recognized* and trials in which no object was recognized as *not recognized* ($T_{recog} > 10\,s$). Using these valid trials, we modeled a boxcar function of $6\,s$ durations (the same 4–$10\,s$ time interval in *recognized* and *not recognized* trials, see also Figure 4.1.2) separately for each response modality ("button" and "memory"). We excluded incorrect responses from our analyses (on average 4 trials (± 3.8 trials) out of 120, $min = 0$, $max = 14$ trials per subject).

Overall we included four regressors (*recognized* button, *recognized* memory, *not recognized* button, *not recognized* memory) and six movement parameters as nuisance regressors to the general linear model (see Section 2.2). All experimental conditions were convolved with a canonical hemodynamic response function and statistical inference was based on a random effects analysis (Holmes and Friston, 1998; Penny and Holmes, 2007).

[2]Wellcome Department of Imaging Neuroscience, Institute of Neurology, London

Unless I state otherwise, the coordinates of the voxels I report in this chapter are in the MNI standard brain space.

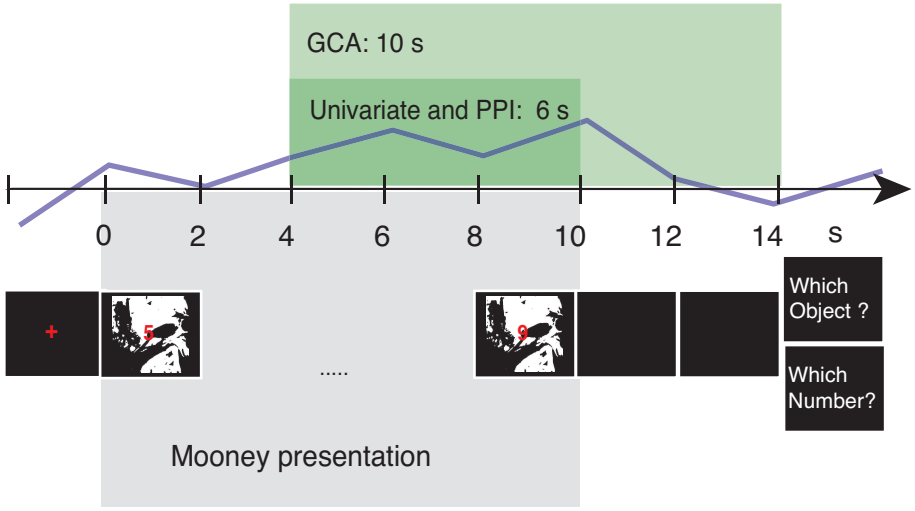


Figure 4.1.2: An example of one trial. For the univariate fMRI analysis and psychophysiological interactions (PPI) we used data points between $4 - 10\,s$ following Mooney image onset. This is modeled as a boxcar function of $6\,s$. For the Granger causality analysis (GCA) we select the data points between $4 - 14\,s$ and model a boxcar function of $10\,s$. The same time-intervals were selected for *recognized* and *not recognized* trials.

Network mask

In this analysis, we were interested in the network that showed modulation to the two conditions *recognized* or *not recognized*, regardless of which condition drives the brain activation. Therefore, we first used a group level F-statistics (see Section 2.2.1) to generate a network mask that represented nonspecific task-related activity. Using this network mask we further analyzed task-related modulations within this network (inclusive mask procedure as in Henson and Friston (2007); Li et al. (2007)). This has the effect that in the group level analysis only those voxels are selected for statistical inference that are within this network mask.

Recognized vs. not recognized: A cognitive conjunction analysis

In the single subject level analysis, a boxcar function of $6\,s$ was used for *recognized* and *not recognized* trials as described previously. We applied for each subject and response modality ("button" or "memory") a contrast *recognized* > *not recognized* to identify the differential brain activations within the subjects. This resulted in two contrast images per subject (one per response modality). In the group level analysis, we used these contrast images to perform a two-

way (2 x 2) ANOVA (Analysis of Variance, Henson and Penny, 2003) with repeated measures including the factors response modalities ("button", "memory") and conditions (*recognized* vs. *not recognized*). ANOVA is a statistical procedure that compares more than two sample means. Hence, it can be seen as a generalization of the statistical t-test. In repeated measures ANOVA (also called within subject ANOVA) the levels and factors are driven from the same subject. We further applied a CONJUNCTION ANALYSIS to identify brain regions, which were active in both response modalities (Price and Friston, 1997; Friston et al., 2005; Nichols et al., 2005). Cognitive conjunction analysis includes two steps: First, the conventional cognitive subtraction between the experimental conditions (*recognized* vs. *not recognized*) is computed for each response modality separately (Price and Friston, 1997). Secondly, a logical AND operation is applied to the difference measurements (Nichols et al., 2005). This means, only those regions that are responsive to the relevant process, in this case *recognition*, will survive the statistical inference. This analysis allowed us to minimize possible confounds that could have emerged when e.g. only one type of response modality ("button" or "memory") would be used.

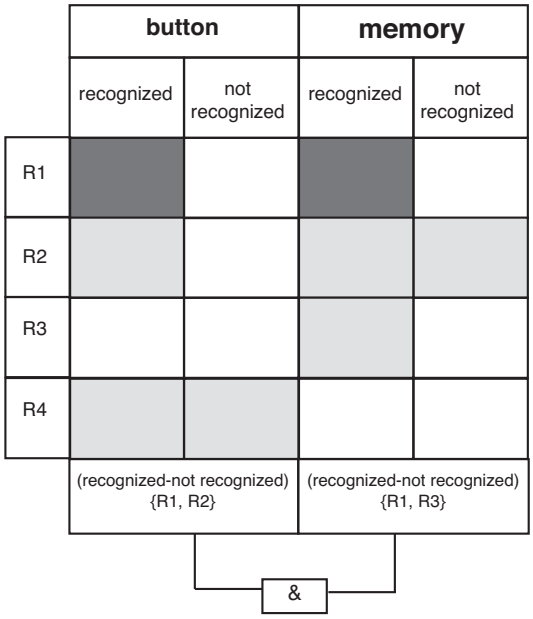


Figure 4.1.3: A toy example of cognitive conjunction. Adapted from Price and Friston (1997).

The cognitive conjunction procedure is illustrated in Figure 4.1.3. In this toy example, the regions R1 and R2 are activated after classical cognitive subtraction in "button". In contrast, regions R1 and R3 are activated when only "memory" modality is used. However, after the conjunction analysis (illustrated with & in Figure 4.1.3), those regions that are not directly related to the recognition process will not survive the statistical inference. Thus, only region R1 will be active as this is the only region that was active in both response modalities.

In addition, we looked for percent signal modulation effects (*recognized* > *not recognized*)

in five anatomically defined regions based on the MNI coordinates in McKeeff and Tong (2007) (see also Table 4.1).

Region of interest	Right hemisphere			Left hemisphere		
	x	y	z	x	y	z
FFA	37	−45	−17	−38	−53	−19
LOC	42	−68	−9	−30	−54	−2
OFA	39	−69	−18	−42	−67	−18
PPA	25	−52	−12	−27	−54	−14
MC	-	-	-	−33	−26	61

Table 4.1: MNI coordinates of the anatomical regions as in McKeeff and Tong (2007) that were used in percent signal modulation analysis.

4.1.6 Functional connectivity analysis

In this analysis, we wanted to identify brain regions that show a dynamical change in the functional coupling related to conscious object recognition. We performed a functional connectivity analysis using a variant of the psychophysiological interaction (PPI) analysis (Friston et al. (1997), also see Section 2.3). Different to Friston et al. (1997), who used a single voxel time-series as a seed region, we selected the mean time course of a region of interest (ROI) as the seed region. We first selected the dorsolateral prefrontal cortex (DLPFC) as the seed region from the nonspecific activity (network mask, Figure 4.2.2 on page 57) by drawing an 8 *mm* spherical cluster around the peak voxel (Figure 4.1.4). We then computed the differential changes in the covariation of this seed region with other voxels in the brain using following factors (please compare with Section 2.3):

1. The physiological factor is extracted from the mean time course of the region of interest (DLPFC).

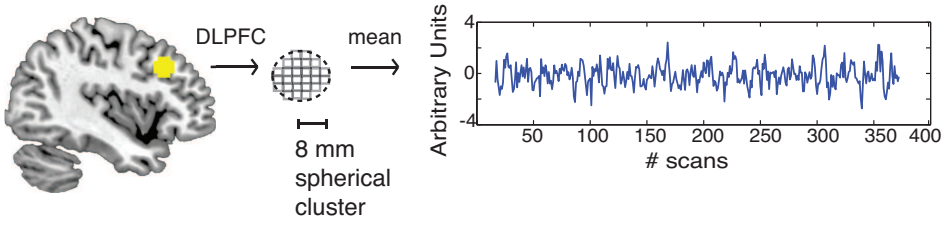


Figure 4.1.4: The seed region used in the functional connectivity analysis. The mean signal in dorsolateral prefrontal cortex (DLPFC, $[-45, 23, 34]$) was selected as the seed region. Here, a time-series with 30 trials is presented for illustration purposes.

2. The psychological factor is a vector with 1's when trials were *recognized* and -1's when trials were *not recognized*. Importantly, the same boxcar function of 6 *s* is used for *recognized*

and *not recognized* trials as described previously in Section 4.1.5 (compare with Figure 4.1.2). Notice that this term has the same length as the physiological factor.

3. The interaction factor is computed by the element-wise vector multiplication of the physiological and the psychological factor.

These three factors were then estimated using a voxel-wise GLM and the individual parameter estimates were used for statistical inference (Macaluso et al., 2000; Haynes et al., 2005b). We also performed this analysis for both response modalities ("button" and "memory") separately.

4.1.7 A trial-by-trial Granger causality analysis

We performed a Granger causality analysis to understand whether feedforward or feedback connections are mediating the increased functional connectivity between extrastriate cortex (EC) and DLPFC during conscious object recognition. The EC was obtained from the results of the PPI analysis. We used autoregressive modeling of the fMRI time-series, and computed a trial-by-trial region of interest (ROI) based Granger causality analysis between the DLPFC and EC by looking one TR ($2\,s$) into the past (Goebel et al., 2003; Roebroeck et al., 2005, 2011b; Bressler and Seth, 2011). This trial-by-trial Granger causality procedure is slightly different than previous studies that used mainly block-designed experiments (Sridharan et al., 2008; Hwang et al., 2010, but see Bressler et al., 2008).

We first generated a boxcar function by coding valid time points with 1's and the rest with 0's separately for each condition (*recognized* and *not recognized*; boxcar function included 4–14 s following Mooney presentation, see also Figure 4.1.2 on page 51). This resulted in a $10\,s$ (5 time points) time window for each valid trial. Trials for the *recognized* and *not recognized* conditions were balanced such that on average 26.63 valid trials ($\pm$ 6.02 trials) and 133.16 time points ($\pm$ 30.10 time points) were used for each condition to estimate the autoregressive model. Table 4.2 lists the number of valid trials and time points for each subject. Subsequently, we multiplied these boxcar functions with the ROI time-series.

We then checked whether these time-series were covariance stationary using the augmented Dickey Fuller test (Hamilton, 1994) as implemented in the Granger Causal Connectivity Analysis Toolbox (Seth, 2010; see also Section 2.4). All time-series that were included into the Granger causality analysis were covariance stationary.

We set the model order, which is the number of past observations included in the regression model to $m = 1\,TR\,(2\,s)$. We created for each ROI (X and Y) and direction ($X \rightarrow Y$ and $Y \rightarrow X$) two vectors according to the restricted and unrestricted model (for a description see Section 2.4). The restricted model contains the regions own past ($Y(t-m)$), whereas, the unrestricted model contains the regions own past ($Y(t-m)$), as well as the second regions past ($X(t-m)$). The same procedure was repeated for X, resulting in four different vectors.

Subject	Length of time series	#Data Points	Recognized	Not Recognized
1	1334	135	27	27
2	1398	140	28	28
3	1319	165	33	33
4	1361	140	28	28
5	1240	100	20	20
6	1342	130	26	26
7	1331	130	26	26
8	1355	95	19	19
9	1317	145	29	29
10	1279	165	33	33
11	1309	195	39	39
12	1379	115	23	23
13	1365	70	14	14
14	1262	170	34	34
15	1276	100	20	20
16	1318	150	30	30
17	1340	135	27	27
18	1292	140	28	28
19	1349	110	22	22

Table 4.2: The number of trials and data points selected for the Granger causality analysis. Each row depicts a different subject.

We excluded all time points, which contained a zero in the vectors and removed the first and last time points of each trial in the vectors before the regression analysis. This procedure concatenates the time-series from multiple trials but at the same time makes sure that the autoregressive model does not pair time points from two different trials.

We computed the following influence measurements between the two ROIs (EC, x and DLPFC, y):

- $F_{x \to y}$: The influence of EC on DLPFC (feedforward).
- $F_{y \to x}$: The influence of DLPFC on EC (feedback).
- $DOI := F_{x \to y} - F_{y \to x}$: The difference of influence measurement (Roebroeck et al. (2005), see also Section 2.4).

It is important to mention, that the DOI measurement only detects the dominant direction of influence (Roebroeck et al., 2005). It cannot measure reciprocal connectivity and only indicates whether there is comparatively more feedforward than feedback connections (or vice versa).

For statistical inference, we first generated 1000 surrogate time-series by shuffling the trial-order independently for each condition, ROI, and subject. We then repeated the Granger

causality analysis using these surrogate time-series. We further contrasted the DOIs of the two conditions ($DOI\,2 = DOI(recognized)$–$DOI(not\ recognized)$) and conducted our statistical inference based on this measurement only.

In the group statistics, we further selected 1000 random samples of subject's surrogate DOI measurements and tested the hypothesis that the difference of influence ($DOI\,2$) measure is larger than zero by setting the significance threshold at $p = 0.05$ (see black lines in Figure 4.2.6c on page 61). When the median $DOI\,2$ (across subjects) measured from the original time-series (see red line in Figure 4.2.6c on page 61) was larger than the significance threshold of the null distribution, the connection was determined to be significant (Roebroeck et al., 2005; Hwang et al., 2010).

4.1.8 Granger causality analysis using Granger causal connectivity analysis (GCCA) toolbox

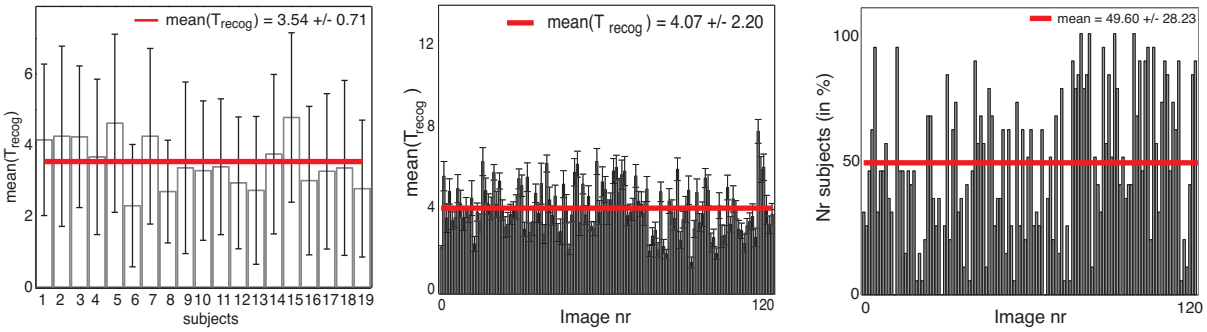
We further applied the multiple realization Granger causality implementation of the Granger causal connectivity analysis (GCCA) toolbox (Seth, 2010). The trial-by-trial Granger causality measurement is an exploratory method that concatenates different trials to create the time-series. Altough we ensured that the covariance stationarity of the data was not affected from this procedure, our implementation remains exploratory. Therefore, to validate our Granger causality results from the previous section (Section 4.1.7) we used the multitrial data implemantation in the GCCA toolbox. This method avoids the concatenation of the trials and allows to make predictions from multitrial data. Each trial is considered as a separate realization of a single underlying stochastic process and is predicted using the Morf algorithm for multivariate auto-regression models (Ding et al., 2006; Seth, 2010). This analysis can be seen as a control to our trial-by-trial Granger causality implementation.

4.2 Results

4.2.1 Behavioral results

Subjects' average recognition time for correct recognized trials was $3.54\,s \pm 0.71\,s$ (Figure 4.2.1a). Average recognition time for images that were correctly recognized within the first four seconds (interval selected for the fMRI analysis) was $2.13\,s\ \pm 0.24\,s$. Average recognition times across subjects are illustrated in Figure 4.2.1b.

On average $49.60\% \pm 28.23\%$ of the subjects recognized the images correctly (Figure 4.2.1c, 9.95% wrong, 37.23% not recognized, 3.22% responses were not understandable). There was no significant difference between correct recognition for animate or inanimate images.



(a) (b) (c)

Figure 4.2.1: Behavioral results. Error bars indicate standard error of the mean. (a) Mean recognition times across images (Correct recognized trials only). (b) Mean recognition times across subjects. (All trials). (c) Average rate of correct recognition per image.

4.2.2 FMRI results

Network mask

We first dentified a network mask that was active in the *recognized* or the *not recognized* conditions (nonspecific task-related activity). This network consists the early visual cortex (EVC, $[30, -97, 10]$, $F_{[4,72]} = 148.72$), the lateral occipital complex (LOC, $[45, -67, -14]$, $F_{[4,72]} = 102.20$), the lateral parietal cortex ($[36, -52, 52]$, $F_{[4,72]} = 28.86$), and the dorsolateral prefrontal cortex (DLPFC, $[-45, 23, 34]$, $F_{[4,72]} = 20.03$) (Figure 4.2.2; $p < 0.01\,FWE$ [Family Wise Error] corrected for multiple comparisons).

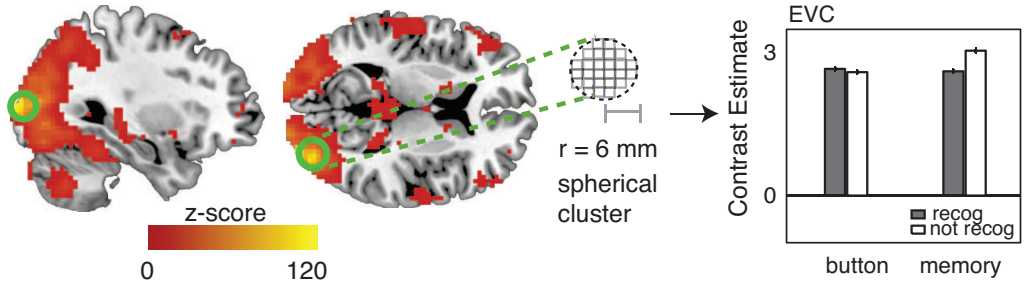
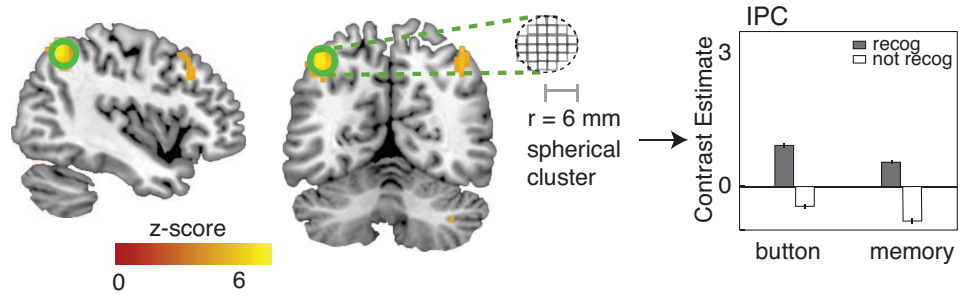


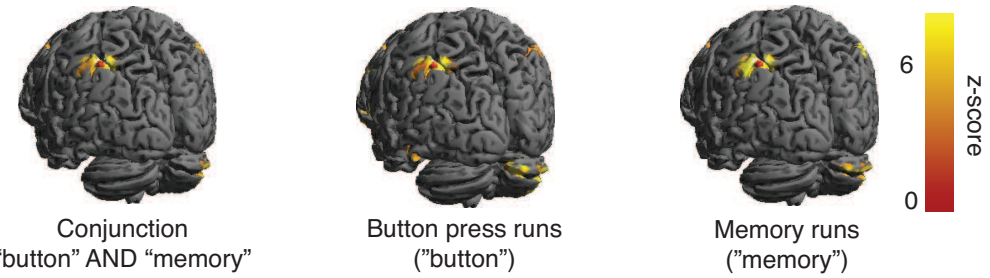
Figure 4.2.2: Network mask. The significant regions are the result of the nonspecific task-related activity ($p < 0.01\,FWE$). Bar plot on the right represents activity in early visual cortex (EVC, $[30, -97, 10]$). No difference between *recognized* (recog) and *not recognized* (not recog) trials is observed.

Recognized vs. not recognized: A cognitive conjunction analysis

We used a conjunction analysis (Price and Friston, 1997; Friston et al., 2005; Nichols and Holmes, 2002) to identify the brain regions that were differentially active in *recognized* vs. *not recognized* conditions (*recognized* > *not recognized*) across both response modalities ("button", "memory").



(a)



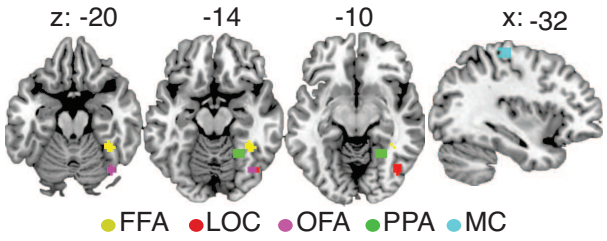
(b)

Figure 4.2.3: FMRI results for *recognized* > *notrecognized*. (a) Significantly higher activity in bilateral inferior parietal cortex (IPC; Left, $[-42, -61, 52]$; Right, $[42, -58, 55]$) revealed by a conjunction analysis across two response modalities ("button", "memory") when subjects consciously recognized the object in the Mooney image correctly versus when they did not recognize it (*recognized* > *notrecognized*). The bar plot on the right shows the mean response of the left parietal region (a sphere at the center of the peak activation, illustrated with the green circle) is computed for button press runs ("button") and memory runs ("memory") separately. Error bars indicate the s.e.m. (b) A lateral surface projection of the results in (a).

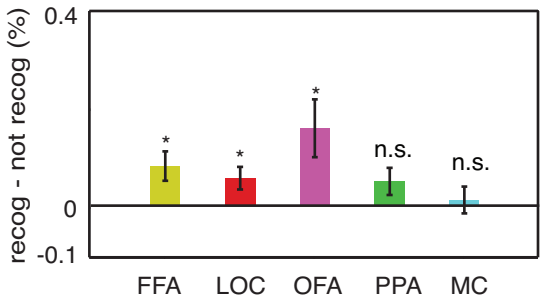
We observed significant activations in bilateral inferior parietal cortex (IPC, Left, $[-42, -61, 52]$, $T^2_{72} = 6.7$; Right, $[42, -58, 55]$, $T^2_{72} = 5.19$), right cerebellum ($[45, -73, -41]$, $T^2_{72} = 5.08$), and left dorsolateral prefrontal cortex (DLPFC, $[-45, 26, 37]$, $T^2_{72} = 4.85$) (Figure 4.2.3; $p < 0.0001$ (uncorrected) with a cluster extent threshold $k = 30$ voxels, surviving a cluster-level correction $p < 0.05\,(FWE)$). In addition, mean BOLD responses in left parietal cortex were significantly higher when subjects correctly recognized the object than when they did not

(*recognized* > *not recognized*), separately in both button press ("button") and memory runs ("memory") (Figure 4.2.3b).

Furthermore, high-level visual areas such as the lateral occipital complex (LOC, $T_{[18]} = 2.43$), fusiform face area (FFA, $T_{[18]} = 2.67$) and occipital face area (OFA, $T_{[18]} = 2.67$) showed significantly higher percent signal modulation for *recognized* > *not recognized* conditions. As expected (McKeeff and Tong, 2007), parahippocampal place area (PPA, $T_{[18]} = 1.79$) and motor cortex (MC, $T_{[18]} = 0.42$) did not show any significant change in activity (Figure 4.2.4). Percent signal modulation was computed using the MarsBar toolbox for SPM (Brett et al., 2002).



(a)



(b)

Figure 4.2.4: Percent signal change modulation in five anatomically defined regions. (a) Anatomical ROIs were LOC; lateral occipital complex, FFA; fusiform face area, OFA; occipital face area, MC; motor cortex. Each region was a 6 *mm* spherical cluster around the central voxel that was based on coordinates in McKeeff and Tong (2007). (b) Percent signal modulation in these ROIs (*recognized* vs. *not recognized).* Asterisks indicate significant effect at the $p < 0.05$ level, n.s. = not significant (recog = recognized and not recog = not recognized).

4.2.3 Functional connectivity results

We used psychophysiological interactions (PPI) to search for differential (*recognized* vs. *not recognized)* covariation between the DLPFC and the rest of the brain. The functional connectivity between left DLPFC and bilateral extrastriate cortex (EC, Left, $[-33, -91, 16], T_{[18]} = 5.35$; Right, $[39, -85, 7], T_{[18]} = 5.67$) was significantly enhanced when subjects recognized the *hidden*

objects than when they did not (Figure 4.2.5a; $p < 0.001\,(uncorrected)$ with a cluster extent threshold $k = 30$ voxels, surviving a cluster-level correction at $p < 0.05\,(FWE)$). Importantly, this connectivity pattern was similar for button press ("button") and memory runs ("memory"), suggesting that this effect is not driven by the specific response modalities of the behavioral report (Figure 4.2.5b).

This result suggests that extrastriate cortex and DLPFC are part of a functional network that supports the conscious recognition of objects by changes in functional connectivity.

Furthermore, we performed a PPI analysis using a seed region that was located within the parietal cortex (MNI : $[36, -52, 52]$). This revealed no significant functional connectivity enhancement ($p > 0.2$ *uncorrected*).

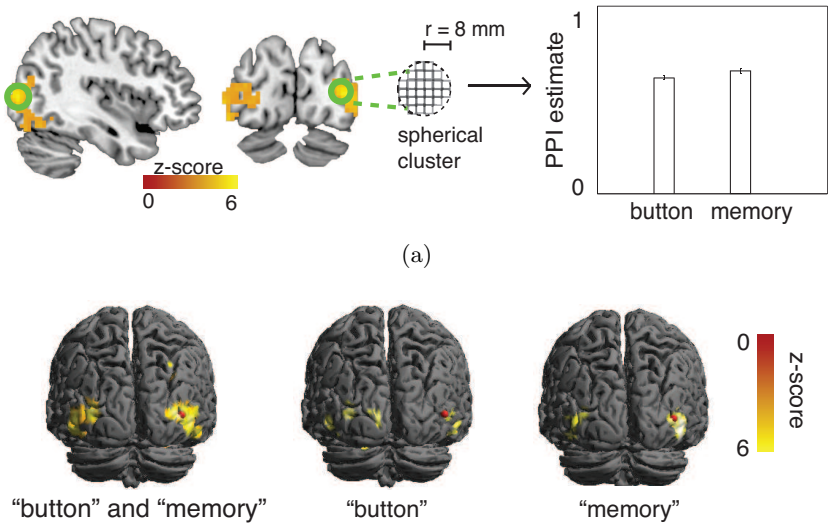


Figure 4.2.5: Functional connectivitiy analysis (psychophysiological interaction, PPI) results. (a) Brain regions showing significant connectivity with DLPFC in *recognized* > *not recognized*. Bar plot on the right is presenting the mean connectivity estimates in the right extrastriate cortex (EC, $[39, -85, 7]$, illustrated with the green circle) for button press runs ("button") and memory runs ("memory") separately. There are no significant differences between button press and memory runs. (b) Lateral surface projection of the results in Figure 4.2.5a shown for different response modalities. 1) Button press runs and memory runs are pooled together; 2) only button press runs are analyzed; 3) only memory runs are analyzed. No significant difference is observed between button press and memory runs.

4.2.4 Granger causality results

We used a trial-by-trial Granger causality analysis (see Section 4.1.7) to identify whether the feedforward or feedback mechanisms mediate the resulting change in functional coupling be-

tween extrastriate cortex (EC) and DLPFC. Using autoregressive modeling the influence of EC on DLPFC ($F_{EC \to DLPFC}$) and the influence of DLPFC on EC ($F_{DLPFC \to EC}$) was computed. Furthermore, a difference of influence measurement ($DOI = F_{EC \to DLPFC} - F_{DLPFC \to EC}$) for each condition was computed to account for differences in the hemodynamic response delay in different brain regions (see also Section 2.4). Furthermore, the differences between the DOI measurements were computed to make statistical inference ($DOI\,2 = DOI(recognized) - DOI(not\,recognized)$).

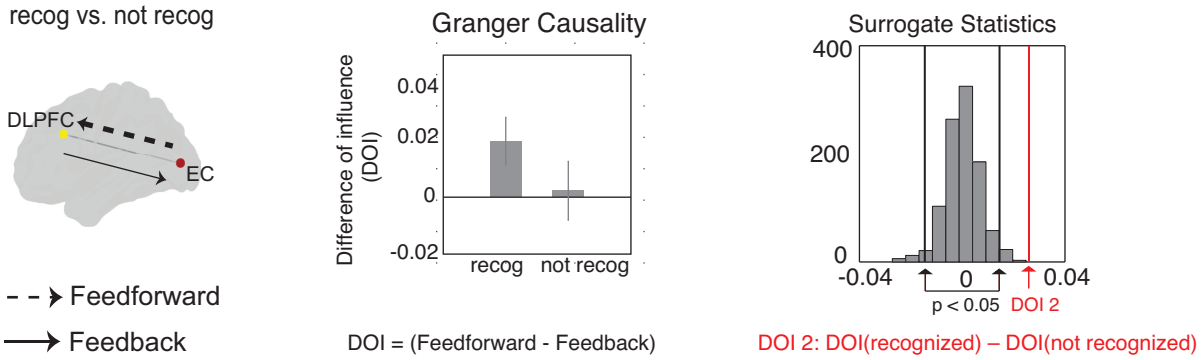


Figure 4.2.6: Granger causality analysis results. (a) Regions of interests used in the Granger causality analysis. EC: Extrastriate cortex, $[39, -85, 7]$. DLPFC: Dorsolateral prefrontal cortex, $[-45, 23, 34]$. (b) Difference of influence (DOI) measurements for recognized ($DOI(recog) = F_{EC \to DLPFC} - F_{DLPFC \to EC}$) and not recognized ($DOI(not\,recog) = F_{EC \to DLPFC} - F_{DLPFC \to EC}$) conditions. (c) Group level histogram of an empirically derived null distribution of no influence. The null distribution was generated by permutation test and the bootstrapping procedure (for details see Section 2.4). Black lines indicate the upper and lower significance threshold at $p = 0.05$. The red line indicates the measured $DOI\,2$ from the original time-series. If the measured $DOI\,2$ from the original time-series is more extreme than the significance threshold, the connection is determined to be significant.

Statistical inference is based on the same procedure as described previously (Section 2.4 and Section 4.1.7). A null distribution of no interest was created using surrogate time-series. The median difference of influence measurement of the original time-series (see red line in Figure 4.2.6c) was larger than the significance threshold ($p = 0.05$, black lines in Figure 4.2.6c) of the null distribution, indicating that the stronger feedforward connectivity during conscious object recognition was significant (Roebroeck et al., 2005; Hwang et al., 2010). We validated this result by using the Durbin-Watson test (Durbin and Watson, 1950) that tests whether the residuals of the autoregressive model were serially correlated or not (see also Section 2.4). In case of a serial correlation, the prediction of the model can be false. We could successfully validate our

model prediction by showing that the residuals was serially uncorrelated.

We have previously seen that the DLPFC was active when the differential activations were tested (*recognized* > *not recognized*). This rises the question whether the Granger causality results are directly linked to the differential activity in DLPFC. More precisely, we need to test whether the differential activity is driving the stronger feedforward connectivity during conscious object recognition, or not. To test this, we computed the correlation between the differential activity (*recognized* > *not recognized*) in DLPFC with the difference of influence (DOI) measurement in the *recognized* conditions for each subject. We did not find any significant correlation between the univariate differential activity and the $DOI(recognized)$ measurement (Figure 4.2.7; Pearson's correlation coefficient $r = -0.05$, $p = 0.83$ for button press runs and $r = 0.02$, $p = 0.93$ for memory runs). This result was a clear indication that the stronger feedforward connectivity was not directly driven by the differential activity in DLPFC.

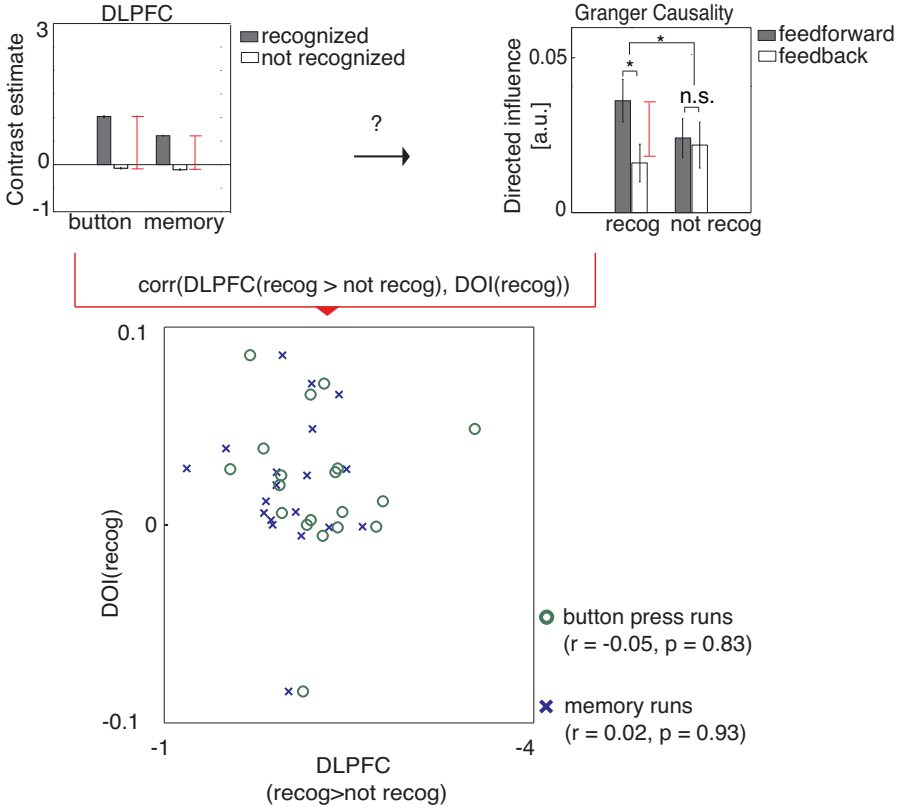


Figure 4.2.7: Correlation between the differential brain response in DLPFC and Granger causality results between DLFC and EC. No significant correlation between the differential activitiy in DLPFC and the Granger causality results were observed (Pearson's correlation coefficient $r = -0.05$, $p = 0.83$ for button press runs and $r = 0.02$, $p = 0.93$ for memory runs).

4.2.5 Granger causality results using the GCCA toolbox

In this analysis, we used the multiple realization Granger causality procedure implemented in the GCCA toolbox to estimate a single model directly from the multiple trial data. In the previous section, I introduced the results of a trial-by-trial Granger causality method, in which trials were concatenated into a long time-series. However, this might result in residual non-stationarities in the data. The multiple trial method implemented in GCCA toobox considers such non-stationarities and allows to make predictions without concatenating different trials.

Statistical inference is based on the same procedure as described previously (Section 2.4 and Section 4.1.7). In a single subject, the multiple trial GCCA is applied to 1000 surrogate time-series to create an empirical null distribution. In the group level, random samples from subjects' surrogate statistics are selected that formed the null distribution of no interest. The median $DOI2$ value measured from the original time-series using this multiple trial analysis (red line in Figure 4.2.8b) was larger than the significance threshold ($p = 0.05$, black lines in Figure 4.2.8b). This indicates a significantly stronger feedforward connection during conscious object recognition.

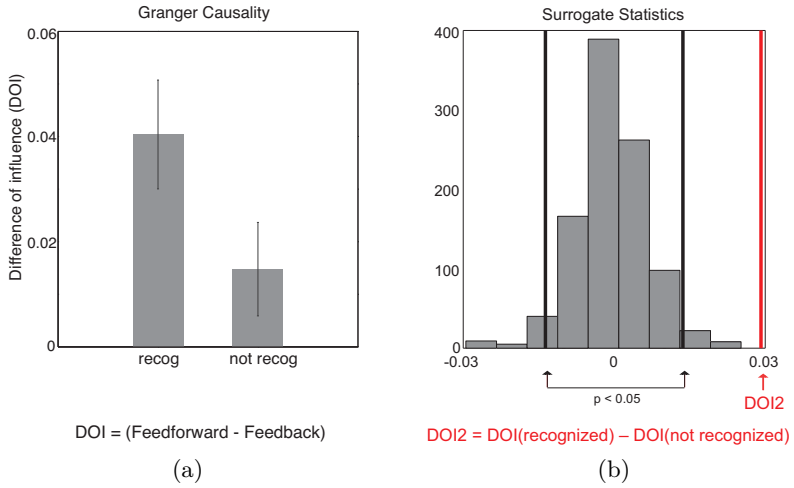


Figure 4.2.8: Granger causality using the multi-trial data implementation of the Granger causal connectivity analysis (GCCA) toolbox. (a) Difference of influence (DOI = Feedforward - Feedback) measurement for recognized (recog) and not recognized (not recog) trials. (b) Group level histogram of an empirically derived null distribution of no influence. The null distribution was generated by the permutation test and the bootstrapping procedure (for details see Materials and Methods). Black lines indicate the upper and lower significance threshold determined by a threshold at $p = 0.05$. The red line indicates the measured $DOI2$ from the original time-series. If the measured $DOI2$ from the original time-series is more extreme than the significance threshold, the connection is determined to be significant.

4.3 Conclusion

I have presented data showing that the frontoparietal network was activated when subjects consciously recognized an object. In addition, dynamic changes in functional connectivity between the extrastriate cortex and the dorsolateral prefrontal cortex were reflecting conscious recognition of objects. These results extend previous conscious perception studies that only measured activity within specific brain regions (Kleinschmidt et al., 1998; Polonsky et al., 2000; Lumer et al., 1998; Portas et al., 2000; Beck et al., 2001; Ress and Heeger, 2003), or looked for whole-brain activity but not for directed functional connectivity between brain regions (Lumer and Rees, 1999; Dehaene et al., 2001; Haynes et al., 2005b).

Two characteristics of our experimental design are significant for the interpretation of our results:

1. We compared two conditions (*recognized* vs. *not recognized*) in the presence of a constant physical stimulus. Hence, subjects' reported perceptual state was not associated with changes in the physical stimulus.

2. We compared two conditions (*recognized* vs. *not recognized*) independent of subjects' behavioral report ("button, "memory"), which is rarely considered for studies investigating conscious perception (but see Lumer and Rees (1999)). Hence, possible confounds in the measured fMRI activity due to behavioral response are controlled.

We have presented significant activations in the frontoparietal network, and functional connectivity changes between visual regions and DLPFC across two different response modalities ("button" and "memory"). Differential activations to *recognized* vs. *not recognized* trials were found in intraparietal cortex, which has been previously associated with attention (Colby and Goldberg, 1999) and visual object recognition (Konen and Kastner, 2008). Furthermore, the right cerebellum was significantly more responding to *recognized* vs. *not recognized* trials. This regions have been previously associated with object naming (Price et al., 1996; Wiggs et al., 1999) and can be related to subjects' overt response. Furthermore, differences in percent signal change during $recognized > not\,recognized$ trials were significant in object related areas, such as LOC. As expected, we did not find any significant difference in the brain activity in early visual areas. Previous studies suggested that cooperative interactions between extrastriate, visual and nonvisual areas are necessary for conscious perception (Lumer and Rees, 1999; Dehaene et al., 2001; Haynes et al., 2005b). Here, we have shown that the functional connectivity between visual and prefrontal areas is contributing to conscious recognition of objects.

We have further presented stronger feedforward connectivity than feedback connectivity during conscious recognition of objects, which has not been investigated previously using fMRI (but see Gaillard et al. (2009)). Our trial-by-trial Granger causality implementation is until now a

rather exploratory causality approach (for block design experiments see Sridharan et al. (2008); Hwang et al. (2010)). Therefore, the multitrial analysis using the GCCA toolbox was applied to validate this method. Using this analysis, we have presented the same stonger feedforward connectivity result (see Section 4.2.5) as in the trial-by-trial Granger causality analysis (compare with Section 4.2.4). Thus, the stronger feedforward connectivity is a reliable result given our data. Similarly, Gaillard et al. (2009) showed in a backward masking experiment that the Granger causality from occipital cortex to prefrontal cortex gets enhanced during conscious unmasked conditions. In contrast, Boly et al. (2011) observed preserved feedforward connectivity but impared conntectivity from frontal cortex to temporal cortex (feedback) in vegetative state patients. The inconsistent results (feedforward connections mediate consciousness vs. feedback connections mediate consciousness) between these studies might be due to the different methods used to infer causality or experimental differences. For example, Gaillard et al. (2009) and we (Imamoglu et al., 2012b) used Granger causality analysis to measure causal connectivity, which is a model-free causality measurement that brings no prior assumptions into the analysis. In contrast, Boly et al. (2011) used a model-based method called Dynamic Causal Modelling (DCM) to infer effective connectivity. Experimental wise, the first two studies searched for content consciousness, however Boly et al. focused on the level of consciousness. Although it is more likely that the different methods used in these studies are causing this discrepant results, we can not directly rule out the second difference as a possible cause. Furthermore, the Granger causality results are based on the DOI measurement that can only detect dominant directionality, but cannot refer to reciprocal connectivity between regions. Therefore, we can only conclude that the feedforward connectivity from EC to DLPFC was more dominant than the feedback connectivity when subjects recognized the object correctly.

To conclude, our results suggest that a distributed network of brain regions (including the frontoparietal network) is involved in conscious object recognition. Additionally, the functional connectivity between visual cortex and prefrontal cortex is an important mechanism that supports conscious object recognition. Furthermore, stronger feedforward connections from extrastriate cortex to dorsolateral prefrontal cortex reflect conscious recognition of objects.

It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong.
Richard P. Feynman

5

Experiment 3: Visual Masking and Functional Connectivity

In this chapter, I will describe a second fMRI experiment that we designed to study the effects of functional connectivity on changes in perceived stimulus visibility (Imamoglu et al., 2012a). We used a common method called visual masking to manipulate perceived stimulus visibility (Breitmeyer and Ogmen, 2006). In visual masking a target stimulus is rendered perceptually invisible by varying the timing of a mask that is followed by a target stimulus. An important advantage of such technique is that the physical appearance of the target stimulus does not change during stimulus presentation. The mask can precede (forward masking) or follow (backward masking) the target stimulus. The aim of this experiment was (i) to identify the brain regions that were reflecting the changes in perceptual visibility when a low-level grating stimulus was used, and (ii) to search for functional connectivity changes between the visual brain regions and the rest of the brain during changes in perceptual visibility.

It has been shown that an overall response including the high-level visual, parietal, and frontal brain regions represented the visibility of masked stimuli (Grill-Spector et al., 2000; Dehaene et al., 2001; Haynes et al., 2005b). Importantly, Grill-Spector et al. (2000) studied brain responses induced by a visible or invisible target stimulus (in this case objects) using fMRI and a backward masking by noise experiment. They have shown that the early visual area V1 was mostly unaffected by the perceived stimulus visibility, but the high-level visual areas such as the lateral occipital and fusiform cortex showed high correlation with subjects'

behavioral responses. In addition, Dehaene et al. (2001) showed a distributed network of brain responses in high-level visual, parietal, and prefrontal regions when words were perceptually visible to the subjects in contrast to when they were not. These authors used a masking technique in which the target words were forward and backward masked (also called sandwich masking; Harris et al., 2011; Kouider et al., 2007). Furthermore, a high-level brain region, the dorsolateral prefrontal cortex, has been shown to respond when stimulus perception was high in contrast to when it was low in a metacontrast masking experiment (Lau and Passingham, 2006). Similiar results in human EEG and animal electrophysiology studies showed that visibility is reflected in extrastriate or high-level brain regions (Bridgeman, 1980, 1988; Kovacs et al., 1995; Dehaene et al., 2001; Del Cul et al., 2007; Lamme et al., 2002; Rolls and Tovee, 1994). These visual masking studies demonstrated that high-level visual areas and the frontoparietal network reflect changes in subjects' perceptual visibility. However, the involvement of high-level brain regions in the subjective visibility of a low-level grating stimulus has not been studied yet (but see Tse et al. (2005) for low-level grating stimuli and the involvement of early visual areas).

Furthermore, it has been shown that the functional coupling between early visual and high-level visual areas were enhanced when stimulus visibility was increased (Dehaene et al., 2001; Haynes et al., 2005b; Lumer and Rees, 1999). For example, during perceptual transitions in binocular rivalry sessions, strong covariations between prefrontal and visual brain regions have been demonstrated (Lumer and Rees, 1999). We recently showed that the dorsolateral prefrontal cortex is functionally more coupled with visual brain regions when subjects' consciously recognized a hidden object in Mooney images (Chapter 4 and Imamoglu et al. (2012b)). Moreover, changes in functional connectivity have been shown to reflect the visibility of a stimulus under metacontrast masking (Haynes et al., 2005b) and sandwich masking (Dehaene et al., 2001). However, changes in functional coupling using low-level stimuli and backward masking have not been investigated yet.

In this study, we used low-level grating stimuli in combination with a backward masking by noise experiment to investigate the involvement of frontoparietal network in perceptual visibility. In combination with fMRI, we searched for voxel-wise neural responses in the human brain that followed subjects' psychometric visibility function. Furthermore, using a parametric functional connectivity approach we studied whether the functional coupling between visual and high-level brain regions were reflecting changes in subjects' visibility function.

In the next section, I will first describe the methods we used in this experiment (Section 5.1). This will follow the results (Section 5.2) and a conclusion section (Section 5.3).

5.1 Methods

5.1.1 Participants

Fourteen healthy subjects (six female, age range 21 to 36 years) participated in this experiments. All subjects had normal or corrected to normal vision and gave written informed consent to participate in the fMRI experiment. Three subjects were excluded from the fMRI analysis: One subject due to systematic motion during the experiment, and two other subjects due to very poor performance in the behavioral task. The experiment was approved by the local ethics review board of the Max Planck Institute for Human Cognitive and Brain Science (Leipzig) and conducted according to the Declaration of Helsinki.

5.1.2 Stimuli

Gratings of two orientations were used in this study (Figure 5.1.1a, right-tilted (45°) and left-tilted (135°) with four different phase-shifts (0°, 90°, 180°, 270°)). The spatial frequency of the gratings was 2 *cpd* (cycles per degree of visual angle). The contrast of the gratings (target stimuli) was computed by dividing the standard deviation of the pixel gray values by the mean pixel gray value, and had a value of 0.3. The grating annulus covered the visual field from 4° to 9° eccentricity, sparing the foveal area to permit better stimulus masking. The noise mask (Figure 5.1.1b) was created by bandpass filtered noise with the same peak spatial frequency as the spatial frequency of the oriented grating without any orientation information. The mask was presented with a higher contrast (0.7) than the target stimulus. Three frames with different random versions of the mask were flashed in succession after the grating stimulus for a powerful masking effect. All parameters of the visual stimuli were selected based on behavioral pre-tests.

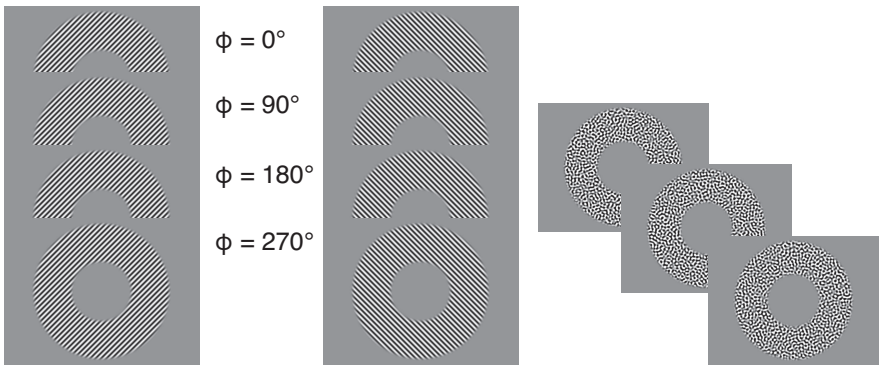


Figure 5.1.1: Example of target and mask stimuli. (a) Left- and right-tilted target stimuli with four different phase-shifts (0°, 90°, 180°, 270°). (b) The noise mask stimuli with three random versions of the mask.

5.1.3 Experimental design

We designed a backward masking by noise experiment to explore visual consciousness by creating different visibility conditions (Macknik and Livingstone, 1998; Breitmeyer and Ogmen, 2006). Visual masking is a widely used procedure to modify the perceived visibility of a stimulus. Backward masking by noise is one masking type in which one briefly presented target stimulus (here a left- or right-tilted grating, see also Figure 5.1.1a) is followed by a temporally succeeding briefly presented mask stimulus (here a random noise mask, see also Figure 5.1.1b). By varying the stimulus onset asynchronies (SOA), that is the delay between the target stimulus and the mask stimulus presentation, perception of the target stimulus can be manipulated (Breitmeyer and Ogmen, 2006). The psychometric visibility function of such masking paradigm is a linear parametric function, which is ascending with increasing visibility level (monotonic Type A masking function, Kolers (1962), see also Figure 5.2.1 on page 74). We used four different SOAs and created a range from perceptually invisible to highly visible stimuli.

Subjects were instructed to fixate on the white cross in the middle of the screen during the whole experiment (Figure 5.1.2). At trial onset, a left- or right-tilted grating stimulus (target) was presented for 17 ms (corresponding to 1 frame at 60 Hz). After a delay of 17, 67, 117 or 217 ms, a mask of three consecutive, different noise frames was flashed for a total duration of 50 ms. The four SOAs created four visibility conditions ranging from perceptually invisible (short SOA, 17 ms) to highly visible (long SOA, 217 ms). Subsequently, 700 ms after the trial onset a response screen containing a left- and right-tilted symbol including dots was presented for 1500 ms.

Subjects' task was to indicate the orientation of the grating they believed they had seen by selecting the corresponding symbol with a left or right button press. Subjects were encouraged to take their best guess if they could not identify the target or were unsure about its orientation. The response screen was pseudo-randomized in order to disambiguate between perceived orientations and button presses. After the response or timeout, a screen with a fixation cross was presented for 800 or 2800 ms. Each trial was completed either 3000 or 5000 ms after trial onset.

The trial order was pseudo-randomized, ensuring 50% left-tilted and 50% right-tilted stimulus presentations. In half of the trials (randomly selected), the symbols in the response mapping screen swapped to decouple stimulus-response mappings. In order to avoid activation differences due potential brightness differences in the stimuli, both orientations contained four phase-shifted grating versions, and thus selectively triggered orientation sensitive cells. The phase-shifted gratings were randomly assigned to the trials.

We conducted ten experimental runs in the scanner, consisting of 80 trials each. Every run contained 10 left-tilted and 10 right-tilted stimulus trials for each visibility condition. The mean trial duration was 4000 ms, half of the trials were 3000 ms, the other half 5000 ms long presented

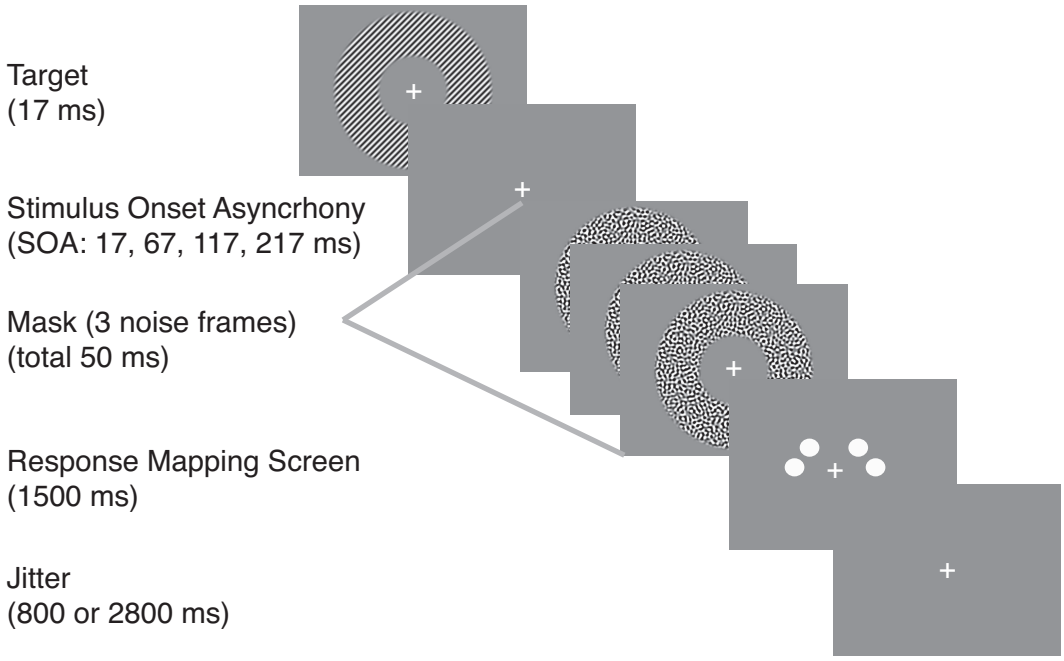


Figure 5.1.2: Experimental design.

in a pseudo-randomized sequence.

Prior to the ten experimental runs in the scanner, subjects participated in a practice run outside the scanner. During this practice run, subjects received acoustic feedback to train the masking procedure.

The stimuli were presented via a projector (resolution 1024×768 pixel, 60 Hz) that projected from the head-end of the scanner onto a screen placed within the scanner bore. Subjects viewed the projection via a mirror fixed onto the head coil.

5.1.4 FMRI data acquisition and preprocessing

A Siemens Trio 3-Tesla scanner equipped with a 12-channel head coil was used to acquire functional MRI volumes. T2*-weighted gradient-echo echo-planar images (EPI) containing 33 axial slices ($2\,mm$ thick, $1\,mm$ gap, ascending) resulting in a voxel size of $3 \times 3 \times 3\,mm^3$ were acquired with the following imaging parameters: $TR = 2000\,ms$, $TE = 30\,ms$, flip angle = 90°, matrix size= 64×64, field of view (FOV) = $192\,mm$. A high resolution T1-weighted structural data set was collected for anatomical localization, with $TR = 1900\,ms$, $TE = 2.52\,ms$, matrix size= 256×256, FOV= $256\,mm$, 192 slices ($1\,mm$ thick), and flip angle = 9°.

Preprocessing of functional scans was performed using SPM8[1]. Functional images were corrected for motion and slice-acquisition time, and were normalized using the unified segmentation

[1] Wellcome Department of Imaging Neuroscience, Institute of Neurology, London

method implemented in SPM8 (see also Section 2.1). Voxels were spatially smoothed with a $6\,mm$ FWHM Gaussian kernel.

5.1.5 FMRI data analysis

The goal of the first analysis was to identify the brain regions that were activated during different perceptual visibility levels. We first searched for regions that were active in any of the four visibility conditions (four SOAs). This resulted in a stimulation mask that we used in further analyses. In a second analysis, we searched for visibility dependent brain activations using a constant parametric contrast. At the single subject level, we created a general linear model (GLM) with four visibility levels (four SOAs) and additional six movement parameters as nuisance regressors. All experimental conditions were convolved with a canonical hemodynamic response function. After estimating each subject's GLM we conducted a group level (2nd level) fMRI analysis.

Unless I state otherwise, the coordinates of the voxels I report in this chapter are in the MNI standard brain space.

Stimulation mask

We first identified regions that were responsive to visual stimulus regardless of the different SOAs. We applied a between subject ANOVA group statistics and searched for brain regions that were activated in any of the four visibility levels (low to high visibility) using an F-contrast. We identified a network of brain regions that were significant ($p < 0.0001\,(FWE)$ with a cluster extent threshold $k = 50$ voxels) and used this network (stimulation mask) to further analyze task related modulations within this network (Figure 5.1.3 on page 72).

Visibility specific activity

We applied a one-way within subject ANOVA with visibility levels as conditions (SOA1 = 17 ms, SOA2 = 67 ms, SOA3 = 117 ms, and SOA4 = 217 ms) to identify the brain regions that were visbility-related. We defined a linear contrast that is ascending with increased visibility (contrast values $[-1.5, -0.5, 0.5, 1.5]$ for [SOA1, SOA2, SOA3, SOA4]). We further used an inclusive masking procedure (Henson and Friston, 2007; Li et al., 2007) to select those significant voxels that were also within the stimulation mask (Figure 5.1.3). Hence, we selected those voxels that were activated during at least one visibility condition (stimulation mask) and showed a parametric modulation to changes in subjects' perceived visibility.

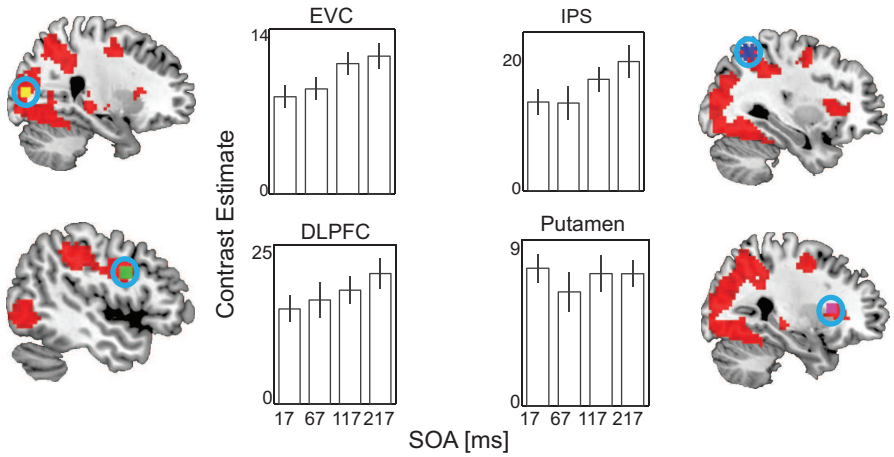


Figure 5.1.3: Stimulation mask. In red are regions that showed significant activation for at least one of the SOAs ($p < 0.0001$, FWE corrected, cluster extent threshold $k = 50$ voxels.). Colored regions within blue circles indicate the center voxel of sample regions within the stimulation mask for which the mean activation is plotted as a function of SOA (middle): early visual cortex (EVC, Left, [-6, -82, -8], $T_{40} = 19.34$; Right, [24, -85, 10], $T_{40} = 20.01$), bilateral dorsolateral prefrontal cortex (DLPFC, Left, [-48, 5, 28], $T_{40} = 20.27$; Right, [51, 8, 31], $T_{40} = 16.59$), bilateral intraparietal sulcus (IPS, Left, [-36, -52, 49], $T_{40} = 17.30$; Right, [33, -55, 55], $T_{40} = 13.57$), and the right putamen ([27, 17, 10], $T_{40} = 19.13$). Contrast estimates as a function of SOA is plotted for each region next to the brain images.

5.1.6 Correlation analysis between voxel activity and psychometric function

In this analysis, we were interested to identify significant activations that reflected individual subject's visibility profiles rather than a constant parametric contrast. We computed the similarity between an individual subject's psychometric visibility function and brain responses to the different visibility conditions. The behavioral and the mean brain responses at each voxel were treated as a four dimensional vector, where each entry corresponded to the measured SOA. We used the Pearson's correlation coefficient to measure the similarity between these two four-dimensional vectors. This resulted in a correlation coefficient $r(s, x)$ per voxel x that indicated how much this voxel was correlated with subject's s visibility. The resulting correlation coefficients were then Fisher Z-transformed, mapped onto a standard MNI brain template, and tested against the null hypothesis of zero correlation ($p < 0.05\,FWE$ corrected for multiple comparison). We used a t-test with eleven subjects and ten degrees of freedom. We further applied the inclusive masking procedure to select only those correlated voxels that were within the stimulation mask.

5.1.7 Functional connectivity analysis

In this analysis, we wanted to identify the functional coupling between brain regions that reflected subjects' perceived visibility. We computed functional connectivity between the visual cortex and the rest of the brain during changes in perceptual visibility. For this analysis, we first selected a seed region (physiological factor) from the nonspecific task-related brain activity (stimulation mask). We selected the peak voxel located within the right Middle Occipital Gyrus (MOG, $[39, -82, 10], F_{[3,30]} = 26.76$, $p < 0.05\, FWE$ corrected for multiple comparison) (Figure 5.2.6 on page 78 middle panel, 'PHYSIO'). This voxel coded overall visual activation and did not have any bias towards one or another visibility level (SOA1-4). We then draw a $6\,mm$ spherical cluster around this peak voxel. The mean signal of this seed region was used to compute the functional connectivity changes between this brain region and the rest of the brain using the psychophysiological interaction (PPI) analysis (Friston et al., 1997).

We used a parametric PPI analysis, which modeled the psychophysical performance of each subject as a linear (parametric) vector with weights $[-1.5, -0.5, 0.5, 1.5]$ corresponding for [SOA1, SOA2, SOA3, SOA4] (Figure 5.2.6 on page 78 left panel, 'PSYCHO'). The interaction factor was created by multiplication of the psychological and physiological factors. We included the psychological, physiological, and interaction factor as linear regressors into a voxel-wise GLM. Individual parameter estimates were then used for statistical inference (Macaluso et al., 2000; Imamoglu et al., 2012b) Using these three factors, we examined whether visibility changes effect the covariation between the visual cortex and other voxels in the brain.

5.2 Results

5.2.1 Behavioral results

Subjects performed an orientation discrimination task. They were instructed to report the orientation of the grating they had seen and to guess if they could not identify it. Figure 5.2.1 shows the behavioral results obtained from eleven valid subjects in the scanner. At the lowest visibility level (17 ms) subjects had about 50% correct responses (chance level for two orientations). In contrast, at the highest visibility level (217 ms) on average subjects responded correctly in almost all trials (100%).

In backward masking, the psychometric visibility function of subjects is in general an increasing function that increases with perceived visibility. Our results indicate that perceived visibility was successfully manipulated by backward masking in all subjects. However, in the two intermediate conditions (67 ms and 117 ms SOA) the variability observed across subjects was higher. In these cases, masking was more powerful for some subjects than others. In particular three out of the eleven subjects showed a lower performance (lower visibility) in the

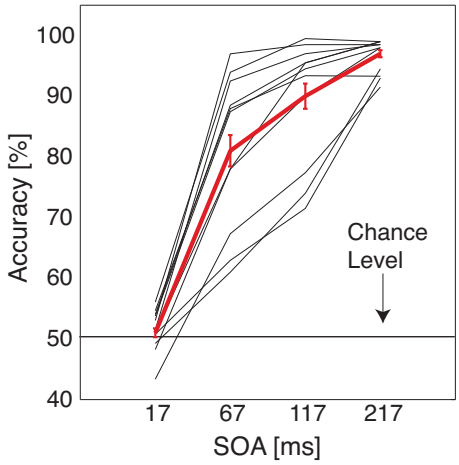


Figure 5.2.1: Behavioral results of the visual masking experiment. Black lines indicate the psychometric curve for eleven subjects. The thicker red line is the mean psychometric curve across subjects.

intermediate SOAs (Figure 5.2.1).

No systematic performance bias for left-tilted or right-tilted stimuli was observed. Over all 8800 trials, subjects failed to give a response in only 66 trials (0.8%). Of the remaining trials subjects responded in 79.2% correctly, and in 20% of the trials incorrectly.

5.2.2 FMRI results

Stimulation Mask

We identified brain regions that were responsive to any of the four visual stimulations (four SOAs). These regions consists visual, high-level cortical as well as subcortical activations. Bilateral early visual cortex (EVC, Left, $[-6, -82, -8]$, $T_{40} = 19.34$; Right, $[24, -85, 10]$, $T_{40} = 20.01$), bilateral dorsolateral prefrontal cortex (DLPFC, Left, $[-48, 5, 28]$, $T_{40} = 20.27$; Right, $[51, 8, 31]$, $T_{40} = 16.59$), bilateral intraparietal sulcus (IPS, Left, $[-36, -52, 49]$, $T_{40} = 17.30$; Right, $[33, -55, 55]$, $T_{40} = 13.57$), and the right putamen ($[27, 17, 10]$, $T_{40} = 19.13$) were significantly activate regions ($p < 0.0001\, FWE$ corrected with a cluster extent threshold $k = 50$ voxels). These regions formed the stimulation mask, which we further used as an inclusive mask in task related analyses (Figure 5.1.3 on page 72).

Visibility specific activity

We were further interested in brain responses that were related to subjects' perceived visibility. Hence, we computed a repeated measures ANOVA with the four visibility levels as factors and searched for brain regions that were significantly active when subjects' visibility was in-

creased. At the group level, we masked the results with the stimulation mask using the inclusive masking procedure. As a result, the bilateral fusiform gyrus (FUS, Left, $[-24, -73, -11]$, $T_{30} = 8.06$; Right, $[30, -70, -11]$, $T_{30} = 9.67$), and bilateral middle occipital gyrus (MOG, Left, $[-33, -88, 4]$, $T_{30} = 8.31$; Right, $[33, -76, 13]$, $T_{30} = 8.80$) were significantly active regions when perceived visibility was enhanced ($p < 0.05\, FWE$ corrected with a cluster extent threshold $k = 50$ voxels, Figure 5.2.2). The bar plots in Figure 5.2.2 represent across subject average of the contrast estimates computed separately for each SOA. Similar to our previous experiment the bilateral superior parietal cortex (SPC, Left, $[-27, -61, 55]$, $T_{30} = 5.81$; Right, $[24, -61, 52]$, $T_{30} = 5.89$) was activated when a lower threshold with cluster correction was used ($p < 10^{-5}\, (unc.)$ with a cluster extent threshold $k = 10$ voxels and only clusters surviving a cluster-correction $p < 0.05\, FWE$ are reported, please compare with Chapter 4). The dorsolateral prefrontal cortex was not significantly active using this procedure.

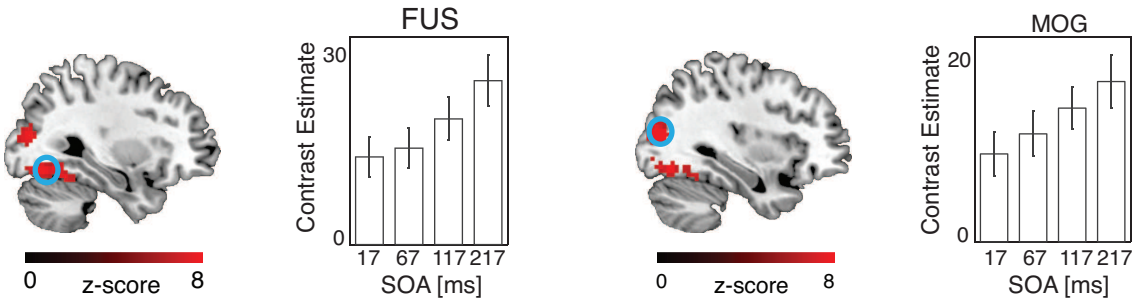


Figure 5.2.2: Visibility specific fMRI activity. Bilateral fusiform gyrus (FUS, Left, [-24, -73, -11], $T_{30} = 8.06$; Right, [30, -70, -11], $T_{30} = 9.67$) and bilateral middle occipital gyrus (MOG, Left, [-33, -88, 4], $T_{30} = 8.31$; Right, [33, -76, 13], $T_{30} = 8.80$) showed activity that was parametrically modulated with visibility, i.e. varying SOA ($p < 0.05$ FWE corrected with a cluster extent threshold $k = 50$ voxels).

When we further contrasted the highest visibility level (SOA4) with the lowest visibility level (SOA1) ($SOA\,4 > SOA\,1$) the same result was observed (Figure 5.2.3). This indicates that the differential activity between the highest and the lowest visibility is driving the activity in these regions.

5.2.3 Correlation results between voxel activity and psychometric function

In the previous section we computed the changes in visibility using a constant parametric linear contrast ($[-1.5, -0.5, 0.5, 1.5]$). This contrast does not take into account the variability in subjects' psychometric curves. However, as apparent in Figure 5.2.1, the temporal profile of subjects' visibility function differ across subjects. Threfore, to have a direct measurement between subjects' psychometric curves and their brain responses, we created the conditions according to subject's individual visibility function. Hence, we computed the similarity between

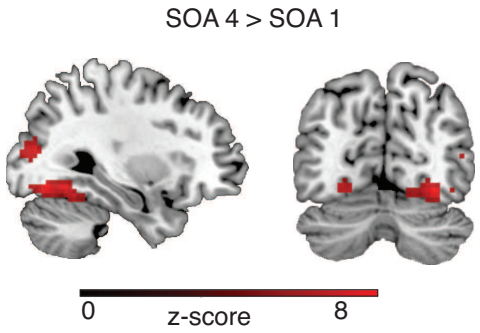


Figure 5.2.3: Repeated measures ANOVA result for the contrast SOA4 (highest visibility) > SOA1 (lowest visibility). As in the parametric contrast (see Methods and Figure 5.2.2) the bilateral fusiform gyrus (FUS, Right, [30, -70, 11], $T_{30} = 9.46$; Left, [-33, -52, -14], $T_{30} = 7.90$) and the bilateral middle occipital gyrus (MOG, Right, [33, -76, 19], $T_{30} = 8.77$; Left, [-39, -82, 4], $T_{30} = 8.25$) are significantly active ($p < 0.05$ FWE corrected with cluster extent threshold 50 voxels) when subjects' visibility is at highest (SOA4) in compared to when their visibility is at lowest (SOA1).

individual subject's psychometric visibility function and subject's voxel activity to the four visibility levels (SOAs) for each subject. Significant brain regions reflect the level of activity that was changing with each subject's individual visibility function.

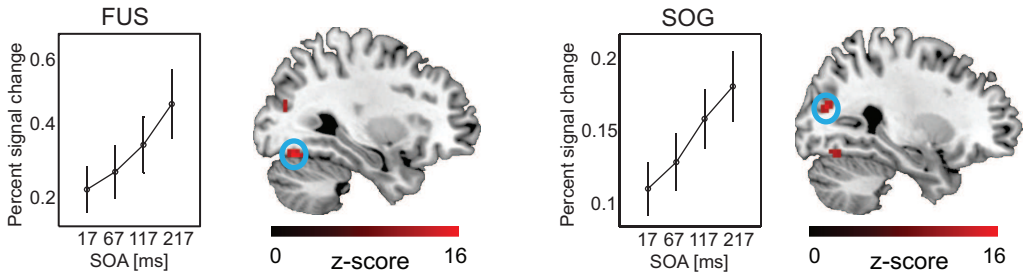


Figure 5.2.4: Results of the voxel-wise correlation analysis between voxel activity and subjects' visibility profiles. The right fusiform gyrus (FUS, $[30, -70, -17]$), and the right superior occipital gyrus (SOG, $[27, -76, 19]$) were significantly correlated ($p < 0.05\,FWE$, $k = 10$ voxels) with subject's individual visibility profiles.

The similarity was computed as described in Section 5.1.6. The computed correlation was significant ($p < 0.05\,FWE$, $k = 10$ voxels) in the right fusiform gyrus (FUS, $[30, -70, -17]$, $T_{10} = 16.19$), and the right superior occipital gyrus (SOG, $[27, -76, 19]$, $T_{10} = 14.42$) (Figure 5.2.4). The bar plots to the right of the brain images in Figure 5.2.4 represent the mean percent fMRI signal variation of these regions averaged across subjects (regions are depicted with blue circles).

When we furher applied a cluster corrected threshold, additional regions were significantly

correlated with visibility ($p < 0.05\,FWE$ corrected for multiple comparison). These are the bilateral superior parietal cortex (SPC, Left, $[-33, -58, 58]$, $T_{10} = 10.06$, Right, $[24, -58, 43]$, $T_{10} = 10.34$), and the bilateral superior frontal gyrus (SFG, Left $[-21, -10, 55]$, $T_{10} = 8.48$, Right $[27, -7, 58]$, $T_{10} = 10.18$). The percent fMRI signal changes from these regions, averaged across subjects and computed separately for each SOA, are plotted as a function of SOA and are shown as bar plots in Figure 5.2.5.

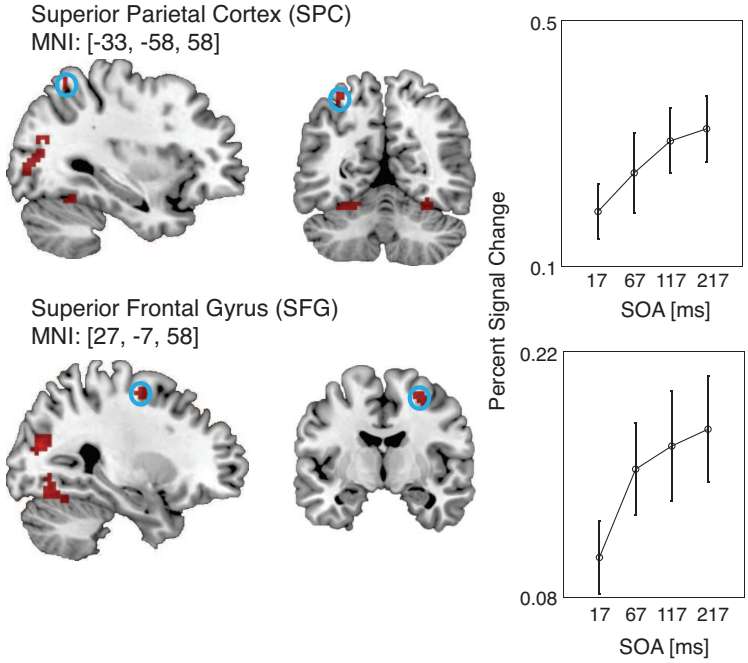


Figure 5.2.5: Results of the same voxel-wise correlation analysis between voxel activity and subjects' visibility profiles as in Figure 5.2.4. However, this time a cluster-level correction for significant testing was applied. The left superior parietal cortex (SPC, $[-33, -58, 58]$), and the right superior frontal gyrus (SFG, $[27, -7, 58]$) are further significantly correlated with visibility.

5.2.4 Functional connectivity results

We computed a functional connectivity analysis to investigate the dynamic interaction between the regions that were involved during changes in perceptual visibility. Therefore, we used a psychophysiological interaction (PPI) analysis to compute the functional connectivity between the visual cortex and the remaining voxels in the brain while taking into account changes in visibility. The psychological factor was a linear parametric function (Figure 5.2.6, left panel). The physiological factor was located within the middle occipital gyrus (MOG, $[39, -82, 10]$, $F_{[3,30]} = 26.76$) and was selected from the stimulation mask (Figure 5.2.6, middle panel). The interaction term was an element-by-element product of these two vectors (psychological factor

x physiological factor).

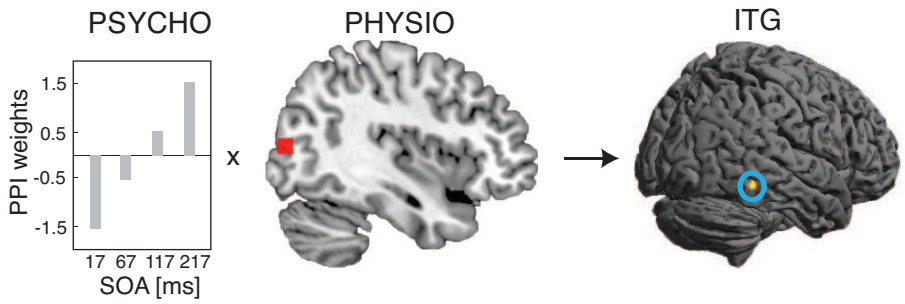


Figure 5.2.6: Psychophysiological interaction (PPI) result. A parametric PPI was computed between the right middle occipital gyrus (MOG, $[39, -82, 10]$) and the rest of the brain. The right inferior temporal gyrus (ITG, $[54, -52, -8]$) was significantly more connected with the right MOG when subjects' visibility was high than when it was low.

As a result, the right inferior temporal gyrus (ITG, $[54, -52, -8]$, $T_{10} = 7.89$) was significantly correlated with this interaction term ($p < 0.001$ (uncorrected) with a cluster extent threshold of $k = 20$ voxels. Only clusters surviving a cluster-level correction $p < 0.05\,(FWE)$ are reported). This suggests that there is a dynamical change in the functional coupling between the visual (MOG) and high-level visual areas (ITG) with increasing perceptual visibility.

We used additional seed regions located in the parietal cortex (IPS), fusiform gyrus (FUS), as well as dorsolateral prefrontal cortex (DLPFC) for the PPI analysis. However, these regions did not lead to any significant results. In Table 5.1 you can find the coordinates of these regions.

Region of interest	Right hemisphere			Left hemisphere		
	x	y	z	x	y	z
DLPFC	51	8	31	−48	5	28
IPS	33	−55	55	−36	−52	49
IFG	36	26	31	-	-	-
FUS	30	−70	−11	−24	−73	−11
EC	39	−82	10	−33	−88	4

Table 5.1: MNI coordinates of regions that were used in the PPI analysis.

5.3 Conclusion

The data I have presented in this experiment demonstrated that under backward masking conditions subjects' visibility profiles are significantly correlated with the right fusiform gyrus (FUS), the right superior occipital gyrus (SOG), the bilateral superior parietal cortex (SPC), and the bilateral superior frontal gyrus (SFG). Furthermore, enhanced functional coupling between

the right middle occipital gyrus and the right inferior temporal gyrus correlate with increase in perceptual visibility.

The involvement of high-level visual areas and the frontoparietal network in perceived visibility has been previously studied using backward masking (Grill-Spector et al., 2000; Green et al., 2005), sandwich masking of words (Dehaene et al., 2001), or metacontrast masking of complex stimuli (Haynes et al., 2005b; Lau and Passingham, 2006). In addition, studies using electrophysiological measures showed that neuronal responses to visual masking occur beyond early visual areas (Kovacs et al., 1995; von der Heydt et al., 1997; Kondo and Komatsu, 2000). The reentrant theory of visual masking similarly suggests that visibility changes during visual masking are represented in high-level visual brain regions (Bridgeman, 1980; Enns and Di Lollo, 1997; Lamme et al., 2002; Ro et al., 2003; Fahrenfort et al., 2007). In contrast, Tse et al. (2005) used a standing wave of invisibility experiment with low-level stimulus and demonstrated that the visual area V1/V2 correlated with subjects' perceived visibility. Hence, the role of high-level brain regions in perceived stimulus visibility (conscious perception) has been under debate. Therefore, our results fill an important gap in the visual masking literature by demonstrating the necessity of high-level visual areas and the frontoparietal network for increased perceptual visibility under backward masking of low-level grating stimuli. Furthermore, enhanced functional connectivity between visual and high-level visual areas during increased perceptual visibility has been previously shown using metacontrast masking (Haynes et al., 2005b) and sandwich masking (Dehaene et al., 2001). Our functional connectivity result supports these findings and further extends the findings reported by earlier studies that either have measured visibility changes only within localized brain regions (Macknik and Livingstone, 1998; Rolls et al., 1999; Grill-Spector et al., 2000; Beck et al., 2001; Ress and Heeger, 2003), or have examined brain responses to changes in visibility using other visibility measurements (Lumer and Rees, 1999; Imamoglu et al., 2012b). In addition, human EEG studies have demonstrated that early visual activation were preserved during visual masking (Dehaene et al., 2001; Del Cul et al., 2007; Fahrenfort et al., 2007; Melloni et al., 2007), which is in line with our negative results for detecting masking related responses in early visual areas.

To conclude, the frontoparietal network is involved in perceived visibility of low-level stimuli. Furthermore, functional coupling between the visual cortex and the temporal cortex is enhanced when perceptual visibility is increased. Thus, a distributed network of brain regions are important to give rise in conscious experience, and changes in functional connectivity between visual and high-level visual regions can be seen as an underlying brain mechanism that contributes to conscious perception.

Everything that has a beginning has an end.
The Oracle/The Matrix Revolutions

6

General Discussion and Future Directions

In this thesis, I have presented one behavioral and two fMRI experiments that I conducted in the course of my PhD studies to study the neural correlates of visual consciousness. Throughout this thesis, I investigated the underlying neural mechanisms that mediate conscious perception by asking two main questions:

- Does a distributed network of brain regions contribute to conscious vision? In particular, is frontoparietal network involved in conscious perception?
- Does functional connectivity between low-level and high-level brain regions support conscious visual perception?

I addressed these questions with two fMRI experiments. In the first fMRI experiment (Experiment 2, described in Chapter 4), I acquired whole brain images from healthy human subjects to investigate the neural correlates of conscious object recognition using two-tone, Mooney images. The stimuli were selected from the MoonBase image database that I have created (Experiment 1, described in Chapter 3). I demonstrated that a distributed network of brain regions, particularly the frontoparietal network, was involved in conscious object recognition. In addition, long-distance functional connectivity between the extrastriate visual cortex and the prefrontal cortex was enhanced during conscious recognition of objects. This functional connectivity was stronger in the feedforward direction than in the feedback direction as shown by Granger causality.

In the second fMRI experiment (Experiment 3, described in Chapter 5), I investigated for the neural correlates of perceptual visibility by using visual backward masking to manipulate perceived visibility. I demonstrated that activations in high-level visual regions and in the frontoparietal cortex reflected subjects' temporal visibility function. Furthermore, the functional connectivity between visual and high-level visual areas was enhanced when perceived visibility was increased.

In addition, prior to the fMRI experiments, I created an image database (MoonBase) that contained two-tone, Mooney images. These images are hard to recognize at first sight, but after some time, subjects abruptly recognize a *hidden* object in the image. One characteristic of these images is that they do not physically change throughout an image presentation but give rise to two different perceptions (object recognized vs. not recognized). Thus, these images are good candidates to be used as stimuli in visual consciousness experiments. A subset of these images was used in Experiment 2. I created a unique image database, i.e., to my knowledge, the largest Mooney image database available for the vision and cognitive science community. I provide this image database and a toolbox for research purposes[1].

It has been previously suggested that high-level visual and the prefrontal cortical areas are necessary to allow a person to report consciousness (Crick and Koch, 1995). Similarly, others argued that a global workspace that is mainly located in the frontoparietal cortex and the cingulate cortex, needs to be involved in order to be able to report conscious perception of a visual stimulus (Dehaene and Naccache, 2001). Following these hypotheses, other influential studies demonstrated extrastriate visual and frontoparietal cortex activations during conscious perception (Lumer et al., 1998; Dehaene et al., 2001; Haynes et al., 2005b; Tononi and Edelman, 1998; Green et al., 2005). The results of Experiment 2 and Experiment 3 are supporting these findings as well. For example, in Experiment 2, I have shown significant differential activation (conscious recognition vs. no conscious recognition) in bilateral object selective areas (e.g. LOC), bilateral intraparietal cortex (IPC) as well as left dorsolateral prefrontal cortex (DLPFC). Similarly, in Experiment 3, I have shown significant correlation between subjects' perceptual visibility and voxel activity in right fusiform gyrus, right superior occipital gyrus, bilateral superior parietal cortex (SPC), as well as bilateral superior frontal gyrus (SFG). Table 6.1 summarizes the brain regions that were significantly involved in conscious perception of objects (Experiment 2) and perceptual visibility (Experiment 3).

The activation patterns in Experiment 2 and Experiment 3 are overlapping to a large extent (Figure 6.0.1). Mainly the activations in right fusiform gyrus (located in the high-level visual pathway) and the activations in the bilateral parietal cortex are similar in both experiments (Figure 6.0.1, lower panel). In agreement with these findings, it has been previously shown that an interruption in the parietal cortex using TMS effects the number of perceptual transitions

[1] https://sites.google.com/site/hayneslab/links

Experiment 2 (Mooney and fMRI)				**Experiment 3 (Visual Masking)**			
Region	x	y	z	Region	x	y	z
L IPC	−42	−61	52	L SPC	−33	−58	58
R IPC	42	−58	55	R SPC	24	−58	43
L DLPFC	−45	26	37	L SFG	−21	−10	55
R Cerebellum	45	−73	−41	R SFG	27	−7	58
LOC	42	−68	−9	R FUS	30	−70	−17
FFA	37	−45	−17	R SOG	27	−76	19
OFA	39	−69	−18				

Table 6.1: MNI coordinates of brain regions that are involved in conscious perception. A summary of Experiment 2 (Chapter 4) and Experiment 3 (Chapter 5). Results of Experiment 2 demonstrate significant BOLD activations when objects were consciously recognized versus when they did not. The results of Experiment 3 demonstrate significant correlation between individual subject's voxel activity and their subjective visibility profiles. All results are FWE corrected for multiple comparison with $p < 0.05$.

during presentations of bi-stable stimuli (for TMS studies on the role of superior intraparietal cortex see Kanai et al., 2010; Carmel et al., 2010, and for a TMS study on the role of right intraparietal sulcus see Zaretskaya et al., 2010). This indicates the direct involvement of parietal cortex in conscious perception. However, a slight difference between Experiment 2 and Experiment 3 is evident when we consider the prefrontal cortical areas (Figure 6.0.1, upper panel). In Experiment 2, right DLPFC activation reflected the conscious recognition of objects, whereas bilateral SFG was significantly correlated with perceived visibility in Experiment 3. Others have shown right dorsolateral prefrontal cortex activity when subjects' detected a change in their visual field (Beck et al., 2001; Sahraie et al., 1997). Similarly, Lau and Passingham (2006) demonstrated using a metacontrast masking experiment that the activity in left mid-DLPFC (Broadmann area 46) was reflecting conscious perception. The DLPFC activation in Experiment 2 is directly supporting these results. However, it is unclear why there is a difference in lateralization of activations in these experiment. As apparent in Figure 6.0.1, the voxels that are active within the SFG are located on the border to premotor cortex, which might be related to premotor activation. Similar results have been presented previously under visual masking conditions (Dehaene et al., 2001; Lau and Passingham, 2006). Therefore, further experiments are needed to confirm the involvement of this region under backward masking while controlling for subjective behavioral reports. In addition, the high inter-subject variability in subjects' recognition can be seen as a limitation of Experiment 2. An ideal experiment would investigate trial-by-trial differences in brain activity prior and post time of recognition. However, the temporal resolution available in fMRI is relatively low making such an analysis futile. Such experiment might be better addressed using EEG, where data in milliseconds precision is available

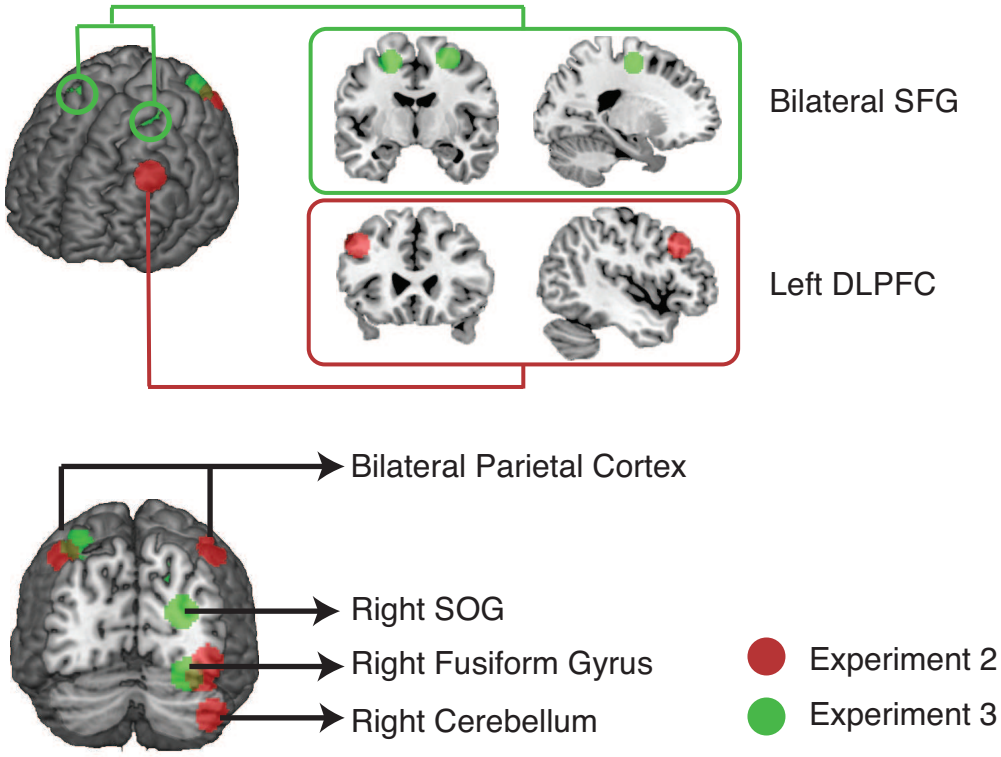


Figure 6.0.1: A summary of Experiment 2 (Chapter 4) and Experiment 3 (Chapter 5). The colored regions (red: Experiment 2, green: Experiment 3) represent brain regions that reflect conscious perception. SFG: Superior Frontal Gyrus, SOG: Superior Occipital Gyrus, DLPFC: Dorsolateral Prefrontal Cortex

for analysis.

Previous studies have suggested that interaction between high-level visual and non-visual areas is contributing to conscious perception (Lumer and Rees, 1999; Dehaene et al., 2001; Haynes et al., 2005b; Gaillard et al., 2009). Lumer and Rees (1999) reported that activity in the prestriate cortex (BA18/19) was significantly correlated with activity in bilateral fusiform gyrus, bilateral parietal cortex (inferior and superior), right frontal gyrus (superior, middle, and inferior frontal gyrus), and right insula, while controlling for subjective reports. Dehaene et al. (2001) demonstrated that left fusiform gyrus was differentially correlated with the posterior intraparietal sulcus, right dorsolateral prefrontal cortex, right anterior cingulate, and left inferior frontal regions (on the border to anterior temporal cortex). Furthermore, Haynes et al. (2005b) demonstrated strong voxel-wise functional coupling between the right early visual area V1 and the right fusiform gyrus during high perceptual visibility. Similar results were observed in sleep research (Massimini et al., 2005), vegetative state patients (Boly et al., 2011), as well as lesion studies (Del Cul et al., 2009). Our results from Experiment 2 and Experiment 3 are

in-line with these findings and support the notion that functional connectivity between distant brain regions is enhanced when subjects are conscious to a presented visual stimulus. In Experiment 2, functional connectivity between extrastriate visual cortex and the left dorsolateral prefrontal cortex was enhanced when *hidden* objects in the Mooney images were recognized by the subjects. Importantly, I presented connectivity results using physically constant stimuli with different response modalities ("button" and "memory"). Furthermore, when I computed directed functional connectivity using Granger causality, the feedforward connectivity (extrastriate to prefrontal cortex) was stronger than the feedback (prefrontal to extrastriate cortex) connectivity during conscious object recognition. This indicates that the sweep of information that travels from the back of the brain to high-level brain regions (feedforward sweep) is crucial for conscious recognition. Similar results have been previously shown by Gaillard et al. (2009). However, few limitations concerning the usage of Granger causality should be taken into account. That is, the causality measurement I used to asses the directionality, can not pick up reciprocal connectivities. Furthermore, downsampling of the BOLD signal has been recently shown to effect causality results (Seth et al., 2013). Therefore, I cannot exclude the possible role of top-down interactions between frontal regions and visual areas in conscious recognition of objects. Surprisingly, when I studied the functional connectivity changes within an attention related region (the parietal cortex), I did not find any significant differences in functional coupling during conscious perception. In Experiment 3, I observed stronger functional connectivity between a visual area, in the vicinity of area V2, and a high-level visual region (fusiform gyrus) during high visibility conditions. Due to the brief stimulation time of the visual stimulus (target) I could not assess any causality measurements to further discover the directionality of the functional connectivity.

When we study visual consciousness, it is important to use visual stimulus that remains constant over the course of the stimulus presentation, but gives rise to different perceptions (e.g. object recognized vs. not recognized, or stimulus perceptually visible vs. invisible). For example, Lumer et al. (1998) used binocular rivalry to study visual consciousness. However, to date, the underlying neural mechanisms and thus the use of binocular rivalry in consciousness studies is controversial (Andrews, 2001, but see Logothetis et al., 1996; Tong, 2001; Crick, 1996). The reason for such controversy emerges from inconsistent findings between human fMRI studies and animal electrophysiology studies. Based on these findings two views emerged concerning the mechanisms for binocular rivalry processing. One view, the interocular competition, suggests that competition in binocular rivalry occurs because of the competition of the signals in monocular neurons (Tong and Engel, 2001). In contrast, the pattern competition view suggests that binocular rivalry occurs because of the competition among incompatible representations at high-level visual areas (Logothetis et al., 1996). As the underlying mechanisms for binocular rivalry are in dispute, caution should be taken while interpreting results based on this approach.

In addition, Dehaene et al. (2001) used visual masking and compared unmasked word presentations to unmasked blank screen presentations (and masked words to masked blank). The interpretation of such analysis is rather difficult as it does not directly compare the stimulus in masked and unmasked conditions. Hence, the brain activation that results from such an analysis can be caused by changes in the stimulus. Therefore, in Experiment 2 and Experiment 3, I avoided such inconsistencies in the stimuli by using physically constant but perceptually changing stimuli. Hence, extended the results of previous studies.

In addition, when we study *content* consciousness and visual consciousness in particular, we rely on subjective reports. This means, when we associate brain signals with consciousness, we do not only exploit brain's response to a visual stimulus, but incorporate subjective criteria into the quantification (Rees, 2009). This can lead to possible confounds emerging from subject's behavioral responses that need to be controlled when visual consciousness is studied. However, studies (Dehaene et al., 2001; Lumer et al., 1998; Haynes et al., 2005b) that previously presented frontoparietal network activations during conscious perception rarely controlled for such motor confound (but see Lumer and Rees, 1999). In Experiment 2 (Imamoglu et al., 2012b), I avoided this confound by collecting subjective reports using two response modalities ("button" and "memory") and applying a conservative cognitive conjunction method. This ensured that significant brain activations were only driven by the cognitive task (conscious object recognition).

Until now, the results I presented in this thesis (Imamoglu et al., 2012a,b) and those described by others (Dehaene et al., 2001; Lumer and Rees, 1999; Haynes et al., 2005b) suggest functionally connected large-scale brain regions as one important underlying neuronal mechanism for conscious visual perception. These regions mainly incorporate high-level visual and the frontoparietal cortex. However, there have been views that challenge the necessity of high-level brain regions as an underlying mechanism for conscious perception (Pollen, 2007; Tong, 2003; Zeki and Bartels, 1999; Lamme et al., 1998). For example, neurological studies did not observe severe deficits in basic visual perception after bilateral prefrontal cortical lesions (Heath et al., 1949; Brickner, 1936). These have been taken as evidence for no or only an indirect involvement of prefrontal regions in conscious vision (Pollen, 2007; Lamme, 2006; Zeki and Bartels, 1999). In contrast, others recently demonstrated that patients with prefrontal cortex lesions required longer time to report a visual target as seen under backward masking (Del Cul et al., 2009), indicating an impairment in visual discrimination performance following prefrontal lesions. Furthermore, a TMS study by Ruff et al. (2006) showed that stimulating the frontal eye field (FEF) region in human subjects modulate perceived contrast and the occipital cortex activity, indicating a causal role from prefrontal regions to posterior regions of the brain. Similar resuts were observed by others (Libedinsky and Livingstone, 2011). Additionally, studies that demonstrated early visual cortex modulation during conscious perception have suggested that visual

consciousness is resolved in early stages of visual processing (Tong and Engel, 2001; Pins and Ffytche, 2003; Tse et al., 2005). In contrast, others have shown early visual activations in the absence of visual consciousness (He and MacLeod, 2001; Haynes and Rees, 2005). Similarly, the results I have presented in this thesis did not show any modulation in early visual areas that correlated directly with conscious perception. Furthermore, frontoparietal cortex activation has been related to attentional top-down modulation (Dehaene and Naccache, 2001; Macknik, 2006). Although in our experiments I cannot directly rule out attentional modulation, it has been previously shown that subjects can be conscious of an object in the absence of attention (Koch and Tsuchiya, 2007).

In regard to theoretical implications, the results of the experiments I presented in this thesis support the view that the involvement of high-level brain regions and the interaction between low-level and high-level brain regions to be important fo conscious perception (Crick and Koch, 1995; Dehaene and Changeux, 2003; Tononi, 2004). The results of Experiment 2 and Experiment 3 support the general idea of the distributed neural network that is involved in conscious perception in the global neural workspace theory (GNWT). A top-down (feedback) amplification from high-level brain regions, especially attention related regions, is one necessary element to access consciousness in GNWT. However, in my experiments, I did not observe any enhancement in functional connectivity between the visual cortex and the parietal cortex during conscious visual perception. Hence, the results of this thesis do not directly support the strict idea of a global workspace. The results of Experiment 3 (in particular the functional connectivity between visual and high-level visual areas) might be related to local reentrant processing of information that travels back to visual regions to prove stimulus recognition or visibility (Lamme and Roelfsema, 2000; Enns and Di Lollo, 2000). However, the functional connectivity results from Experiment 2 were demonstrating a long-distance connectivity (visual to frontal) that was directly involved in conscious perception. Thus, when the results of both experiments are considered together, the reentrant processing in localized brain regions cannot be the ultimate underlying neuronal mechanism that mediate visual consciousness. Information integration theory is another theoretical approach that lies at the heart of interacting global brain networks. Although, information integration does not specifically direct an importance to one or another brain region (such as the prefrontal cortex, parietal cortex, or others), our functional connectivity results support this theory. However, the applicability of this approach to neuroimaging data is computationally demanding and is under development, but merits further investigations.

In summary, I suggest that a distributed network of brain regions is part of the neural correlates of consciousness. These regions include high-level visual, parietal, and prefrontal cortical areas. This does not imply that the NCC is only located in these regions, but it is likely that they are the basis of a necessary building block for conscious perception. Furthermore, the

data presented in this thesis suggest that dynamic changes in functional connectivity between visual and high-level brain regions are contributing to conscious perception.

6.1 Future Directions

The results I have presented in this thesis demonstrated the importance of functional connectivity in conscious perception. An ultimate goal and important extension to these results would be to further investigate the causal relations between the different networks of brain regions that are functionally connected during conscious perception. I have applied Granger causality as a directed functional connectivity measurement to find out the necessary flow of information that mediated conscious object recognition (Experiment 2). However, it has been shown that Granger causality in combination with data of low temporal resolution (such as in fMRI) could confound causality results. Therefore, further applications are needed to understand the necessary and sufficient causal network structure of conscious systems. For example, EEG and two-tone images can be used to investigate directed functional connectivity before the time point of recognition versus after the time point of recognition. Although EEG has the disadvantage of low spatial resolution, it brings the advantage of high temporal resolution due to short sampling rates. Therefore, the model free Granger causality computation that is applied to such data set could provide reliable results. It would be certainly important to compare the results of this experiment with the results in Experiment 2.

Another potential experiment to study the causal role of the prefrontal cortex in consciousness can be performed using stimulation techniques such as TMS (as in Ruff et al., 2006). For example, a TMS pulse can be applied on the prefrontal cortex right before object recognition occurs using two-tone stimuli. A direct influence of prefrontal cortex in conscious object recognition would be observed when subjects could not report the recognition of the *hidden* object. However, the difficulty to successfully apply TMS stimulation on the prefrontal cortex makes such a study challenging.

A model based account (Dynamical Causal Modelling) with anatomically and functionally defined brain regions, can be further used to study effective connectivity that supports conscious perception.

Ambiguous images can be further used in an experiment to predict and/or reconstruct subjective experience from human brain activity using pattern classification (Haynes and Rees, 2006) or encoding models (Naselaris et al., 2011). This would mean to reconstruct not only what others see through their eyes but also what others subjectively perceive. Such reconstruction could help to understand the underlying mechanisms that mediate conscious perception. Furthermore, a machine that can reconstruct subjective content would be a helpful clinical application to facilitate the communication with patients with disorders of consciousness and improve the

quality of their life.

It is further important to merge neurobiological findings with theoretical investigations. It would be interesting to study the amount of *information integration (ø)* the subjects have when they are conscious of a particular content. For example, using a visual masking paradigm (Experiment 3) it would be crucial to see whether ϕ gradually changes between different visibility levels. Furthermore, ϕ might be able to predict the time point of an observer's conscious recognition of a *hidden* object in a two-tone experiment (Experiment 1 and 2). Although computing the amount of integrated information a natural systems has is very challenging, applying approximate computations of ϕ to neuroimaging data can shed an important light into our understanding of consciousness (Oizumi et al., 2012). Furthermore, the integrated information of a system can be considered as a useful measure in the context of artificial intelligence and machine consciousness (Koch and Tononi, 2008; Tononi and Koch, 2011).

As a final word, I would call the job of this thesis as finished, if the results of the experiments I introduced here, could shed some light on one of the most important experiences that a creature can have, namely consciousness.

A

Talairach Coordinates

In the following are the talairach coordinates of the regions that were illustrated in Figure 1.1.1 on page 9.

Study	**Cerebral region**	**Talairach Coordinate or MNI(*)**		
		x	y	z
Sahraie et al. 1997	DLPFC	40	31	20
Kleinschmidt et al. 1998	Parietal Cortex	21	-60	48
	Parietal Cortex	-18	-57	45
	Frontal Cortex	-57	12	9
Lumer et al. 1998	Inferior Parietal	67	-33	37
	Superior Parietal	36	-45	51
	Lateral extrastriate	42	-87	9
	Inferior Frontal	51	15	6
Tong et al. 1998	PPA	-6	18	-39
	FFA	40	-48	12
Tononi et al. 1998	Distributed	not reported		
Tong and Engel 2001	Blind spot	Calcarine Sulcus		

Kreiman et al. 2002	MTL		not reported	
Eriksson et al. 2004	MTL	-16	-13	-20
	MTL	20	-24	-19
	Parietal Cortex	-46	-41	44
	Parietal Cortex	51	-33	37
	VLPFC	55	32	9
Grill-Spector et al. 2000	Lateral occipital	42	-67	-4
	Lateral occipital	-44	-46	-29
Polonsky et al. 2000	V1		not reported	
	V2		not reported	
	V3		not reported	
	V4v		not reported	
	V3a		not reported	
Dehaene et al. 2001	Fusiform Gyrus	-40	-56	-24
	Inferior Frontal	-48	8	4
	Intraparietal Sulcus	32	-48	40
	Fusiform Gyrus	36	-60	-24
Haynes et al. 2005b (* MNI coordinates)	Temporal Parietal Junction	-54	-48	27
	Fusiform gyrus	48	-45	-21
	Calcarine Sulcus	15	-87	15
	V5/MT	-45	-66	15
	V5/MT	54	-63	15
	Middle frontal gyrus	-48	48	-6
	Middle frontal gyrus	-51	45	-3
	Middle frontal gyrus	30	15	45
Del Cul et al. 2007	Frontal Cortex		not reported	
	Parietal Cortex		not reported	

B

Mooney Images

B.1 Mooney Images

In the next six pages (the first three contains only animate images, the next three only inanimate images), I illustrate 120 Mooney images that I used in the first fMRI experiment (Chapter 4).

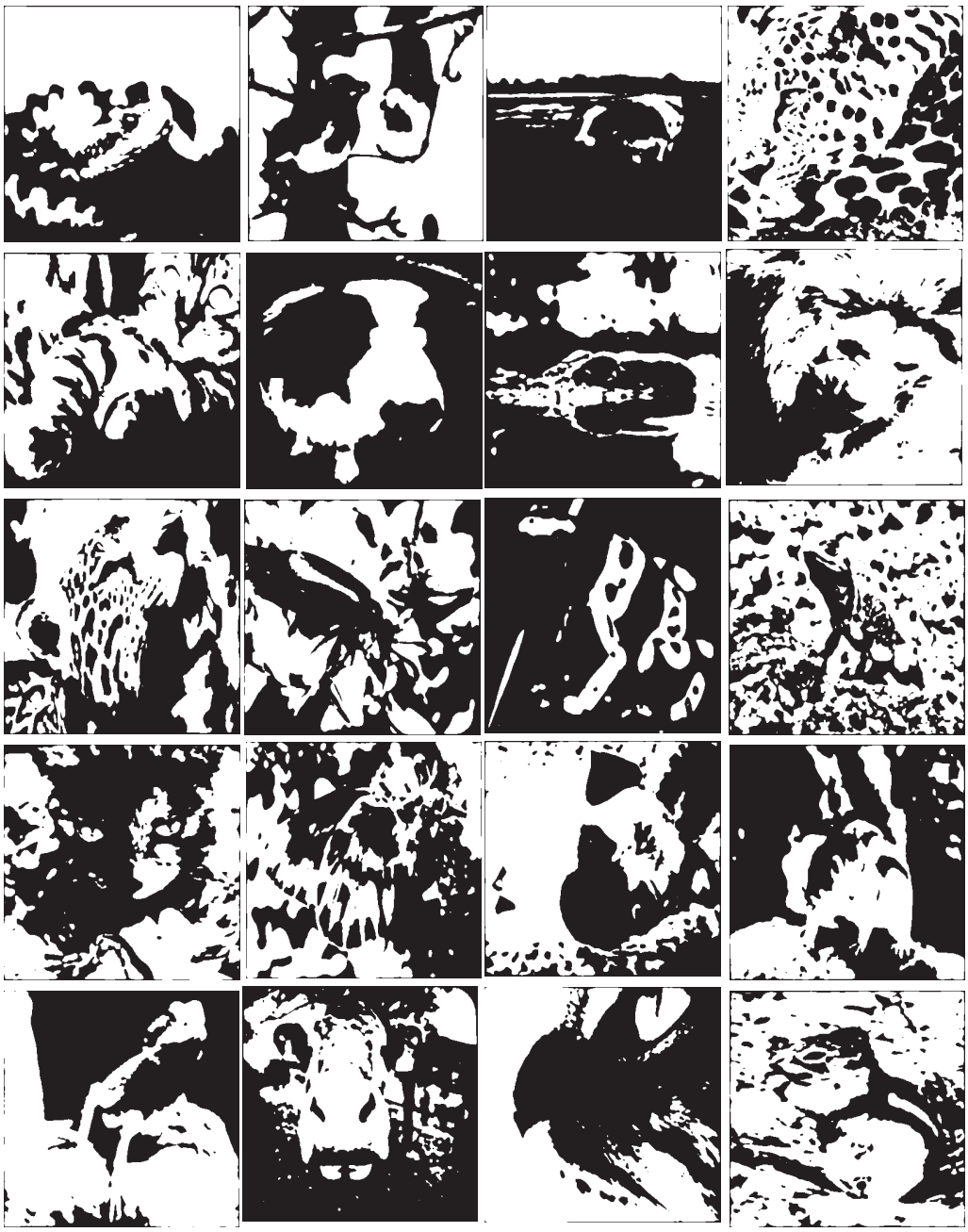


Figure B.1.1: Mooney images that were used in the fMRI experiment (imgs 1-20, animate).

Figure B.1.2: Mooney images that were used in the fMRI experiment (imgs 21-40, animate).

Figure B.1.3: Mooney images that were used in the fMRI experiment (imgs 41-60, animate).

Figure B.1.4: Mooney images that were used in the fMRI experiment (imgs 1-20, inanimate).

Figure B.1.5: Mooney images that were used in the fMRI experiment (imgs 21-40, inanimate).

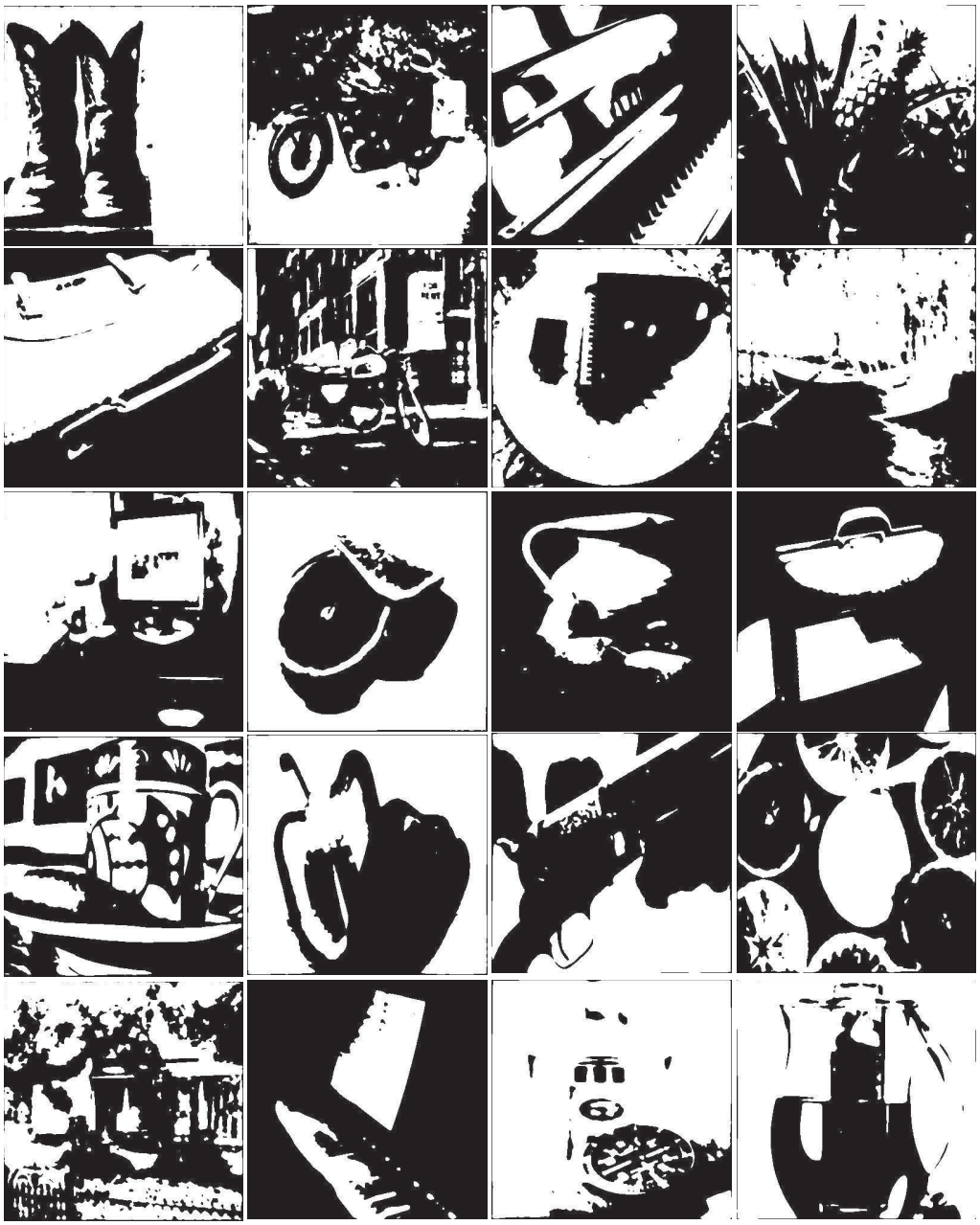


Figure B.1.6: Mooney images that were used in the fMRI experiment (imgs 41-60, inanimate).

B.2 Mooney Image Properties

Here, I present the data collected from six subjects in Experiment 1 (Chapter 3). Each row depicts one image out of the 120 Mooney images presented in Section B.1. These 120 images were used in the fMRI experiment described in Chapter 4. The six columns refer to the following: image identifier (Image Nr), mean time of recognition (T_{recog}), standard deviation of time of recognition ($std(T_{recog})$), mean confidence rating (Confidence), mean aha rating (Aha), mean difficulty rating (Difficulty). For each 120 Mooney image, I list the mean responses across six subjects.

Image Nr	Trecog	std(Trecog)	Confidence	Aha	Difficulty
3	5.21	3.51	3.50	3.00	2.00
12	5.52	2.66	3.83	3.33	1.83
13	4.53	3.02	3.00	2.83	1.83
21	5.08	3.13	3.50	1.83	2.50
38	2.77	0.72	4.17	1.83	2.17
39	6.00	3.78	2.17	3.67	2.00
45	3.31	0.91	2.67	3.17	2.17
58	4.41	3.12	2.67	2.50	2.50
62	3.79	3.10	2.67	2.83	2.67
64	2.88	1.55	3.83	2.83	1.50
87	3.59	1.92	2.33	2.67	2.33
123	5.96	2.17	3.00	3.33	3.00
127	4.05	2.61	3.33	3.17	1.33
128	4.75	3.05	4.00	2.50	1.83
129	6.12	3.01	3.17	3.00	2.83
137	7.40	4.65	3.33	2.50	2.83
148	6.13	3.41	3.50	3.33	3.17
155	8.63	3.60	2.33	3.17	3.00
184	3.26	1.25	3.17	2.00	1.50
191	5.90	4.05	2.17	3.50	2.67
192	5.29	3.24	3.50	3.17	1.83
209	9.57	6.01	2.00	4.00	4.17
243	8.54	6.48	2.17	3.17	2.33
260	3.74	2.27	3.50	2.33	2.00

267	3.29	1.89	4.17	2.17	2.17
273	6.58	5.85	4.00	2.83	2.33
292	3.94	2.79	4.33	2.50	2.17
298	7.91	4.46	3.00	2.83	2.50
350	10.09	4.29	1.33	3.33	2.67
359	7.40	2.22	3.17	3.17	2.83
365	5.63	3.53	3.00	2.67	2.50
367	7.54	4.24	3.17	3.00	2.67
378	8.43	6.19	2.67	3.50	3.17
394	3.62	2.18	3.00	1.83	1.83
399	4.66	2.32	4.17	2.83	1.83
400	6.53	5.58	3.17	2.50	2.50
404	6.24	2.69	2.33	2.50	1.83
417	3.87	1.58	3.50	2.33	2.17
422	6.88	4.46	2.17	3.83	3.67
429	5.52	1.08	2.67	2.67	2.67
442	3.71	1.79	4.33	2.50	1.50
449	8.10	4.14	2.00	3.67	2.33
450	3.84	2.91	3.33	1.67	1.83
451	5.15	4.75	3.67	2.67	2.17
469	6.93	2.34	2.50	2.33	2.33
474	3.94	3.39	4.00	1.83	2.00
475	4.61	3.07	3.17	2.67	2.17
478	8.22	4.79	2.33	2.83	2.33
487	7.43	5.40	2.83	3.17	2.83
506	4.36	3.13	4.33	2.33	2.17
522	4.48	3.63	4.00	2.00	1.83
525	7.77	4.94	3.67	2.50	2.00
526	2.77	2.08	4.50	2.50	1.33
533	9.15	6.77	1.83	3.33	2.50
549	7.11	3.15	1.83	3.17	2.50
555	4.00	2.44	4.50	2.83	1.83
556	6.29	7.01	2.83	2.17	2.33

558	3.68	0.78	3.50	3.50	2.00
559	10.25	6.52	2.83	3.17	2.67
562	5.34	5.20	4.00	2.33	2.17
573	4.55	3.42	3.50	2.33	1.83
576	3.79	1.78	2.17	3.17	2.33
582	9.19	2.65	2.83	2.83	2.00
585	9.22	7.15	2.00	3.33	2.83
586	4.34	2.26	4.17	3.17	2.33
588	9.77	3.00	2.00	2.33	3.17
590	8.13	7.63	2.50	2.33	2.50
593	6.89	3.80	2.17	2.33	3.00
595	5.29	3.04	3.83	3.17	2.50
606	4.86	2.36	3.17	2.00	2.33
609	4.93	3.27	2.67	2.67	2.67
623	4.05	3.42	4.00	2.50	1.83
626	6.14	4.07	2.50	2.83	2.50
627	2.13	1.46	4.17	3.00	1.67
629	7.50	6.01	2.83	2.50	2.83
635	3.39	2.36	4.17	2.00	2.50
639	3.97	3.72	3.83	2.33	2.67
640	3.48	2.80	4.67	2.83	1.67
641	2.81	1.41	4.33	2.33	1.83
643	6.35	3.90	3.67	3.00	2.17
650	5.38	3.29	3.67	3.00	2.50
652	2.91	1.35	4.33	1.83	1.67
655	6.66	3.66	3.67	3.50	2.17
661	4.50	3.12	4.33	3.17	2.33
676	7.37	4.55	3.50	2.67	2.33
683	3.50	1.26	3.33	2.50	1.83
686	2.94	2.20	4.33	1.83	1.50
691	6.56	2.38	3.00	4.00	2.83
694	6.13	3.46	3.33	2.67	3.00
698	3.99	2.95	4.00	2.17	1.83

713	2.36	1.64	4.83	1.17	1.33
716	4.45	2.81	4.50	2.17	1.33
721	7.67	2.09	2.17	3.50	2.50
723	7.05	4.65	3.17	2.17	2.00
728	4.33	2.10	2.17	3.00	2.67
729	4.98	2.41	3.17	2.67	1.50
738	3.80	1.86	4.17	2.17	2.00
742	2.65	1.45	4.33	2.50	2.00
745	7.55	6.24	4.00	2.67	1.83
747	4.69	4.74	4.33	1.67	1.67
751	6.90	3.86	1.83	3.00	3.33
752	4.35	1.94	3.83	2.67	1.83
760	1.97	1.03	4.83	2.00	1.17
773	3.95	1.92	3.33	3.17	2.50
777	3.18	1.17	4.00	2.33	2.50
779	8.95	4.80	2.50	3.83	3.33
781	7.65	5.58	2.67	3.33	2.50
787	3.08	2.59	4.33	2.33	2.67
792	3.32	1.50	4.17	1.83	1.67
797	4.13	2.96	4.50	2.50	2.17
801	6.74	3.35	2.83	3.17	2.50
802	3.95	2.29	4.67	2.17	1.50
805	7.30	5.76	3.33	2.17	2.17
807	3.42	2.26	4.50	1.33	1.33
809	10.06	3.90	2.33	3.67	3.33
812	3.73	1.63	3.00	2.83	2.33
817	6.32	3.92	1.50	3.83	2.83
823	3.62	1.94	3.67	3.83	2.33
824	4.49	2.72	4.00	2.33	1.50
826	5.66	4.13	3.67	2.67	2.67

Bibliography

Aguirre, G. K., Zarahn, E., and D'esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *Neuroimage*, 8(4):360–9.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Amassian, V. E. and Maccabee, P. J. (2006). Transcranial magnetic stimulation. *Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, 1:1620–1623.

Andrews, T. J. (2001). Binocular rivalry and visual awareness. *Trends in Cognitive Sciences*, 5(10):407–409.

Andrews, T. J., Halpern, S. D., and Purves, D. (1997). Correlated size variations in human visual cortex, lateral geniculate nucleus, and optic tract. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 17(8):2859–2868.

Andrews, T. J. and Schluppeck, D. (2004). Neural responses to mooney images reveal a modular representation of faces in human visual cortex. *Neuroimage*, 21(1):91–98.

Ardekani, B. A., Bachman, A. H., and Helpern, J. A. (2001). A quantitative comparison of motion detection algorithms in fMRI. *Magnetic resonance imaging*, 19(7):959–963.

Ashburner, J. and Friston, K. J. (1999). Nonlinear spatial normalization using basis functions. *Human brain mapping*, 7(4):254–266.

Ashburner, J. and Friston, K. J. (2005). Unified segmentation. *NeuroImage*, 26(3):839–851.

Ashburner, J. and Friston, K. J. (2007). Rigid body registration. In *Statistical parametric mapping: The analysis of functional brain images*, pages 49–62. Academic Press, London.

Ashby, F. G. (2011). *Statistical Analysis of fMRI Data*. MIT Press, Cambridge, Massachusetts, 1st edition.

Baars, B. (1988). *A Cognitive Theory of Consciousness*. Cambridge Univeristy Press, New York.

Barker, A., Jalinous, R., and Freeston, I. (1985). Non-invasive magnetic stimulation of human motor cortex. *The Lancet*, 325(8437):1106–1107.

Barnett, L. and Seth, A. K. (2011). Behaviour of granger causality under filtering: theoretical invariance and practical application. *Journal of neuroscience methods*, 201(2):404–19.

Bauby, J.-D. (1997). *The diving bell and the butterfly a memoir of life in death (Original Title: Le scaphandre et le papillon)*. Vintage, Princeton, N.J.

Beck, D. M., Rees, G., Frith, C. D., and Lavie, N. (2001). Neural correlates of change detection and change blindness. *Nature neuroscience*, 4(6):645–50.

Blake, R. and Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, 3:1–11.

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, 9(2):46–52.

Blumenfeld, H. (2009). The neurological examination of consciousness. In *The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology*, pages 15–30. Elsevier, Ltd, Amsterdam; Boston.

Boly, M., Garrido, M. I., Gosseries, O., Bruno, M.-A., Boveroux, P., Schnakers, C., Massimini, M., Litvak, V., Laureys, S., and Friston, K. (2011). Preserved feedforward but impaired top-down processes in the vegetative state. *Science*, 332(6031):858–862.

Boly, M., Massimini, M., Garrido, M. I., Gosseries, O., Noirhomme, Q., Laureys, S., and Soddu, A. (2012). Brain connectivity in disorders of consciousness. *Brain connectivity*, 2(1):1–10.

Bonneh, Y. S., Cooperman, A., and Sagi, D. (2001). Motion-induced blindness in normal observers. *Nature*, 411(6839):798–801.

Breitmeyer, B. and Ogmen, H. (2006). *Visual Masking: Time Slices Through Conscious and Unconscious Vision*. Oxford University Press, New York.

Bressler, S. L. and Seth, A. K. (2011). Wiener-granger causality: a well established methodology. *NeuroImage*, 58(2):323–329.

Bressler, S. L., Tang, W., Sylvester, C. M., Shulman, G. L., and Corbetta, M. (2008). Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. *The Journal of neuroscience*, 28(40):10056–61.

Brett, M., Anton, J. L., Valabregue, R., and Poline, J. B. (2002). Region of interest analysis using an SPM toolbox. *NeuroImage*, 16(2).

Brickner, R. M. (1936). *The intellectual functions of the frontal lobes.* Macmillian, Oxford, England.

Bridgeman, B. (1980). Temporal response characteristics of cells in monkey striate cortex measured with metacontrast masking and brightness discrimination. *Brain Research*, 196(2):347–364.

Bridgeman, B. (1988). Visual evoked potentials: Concomitants of metacontrast in late components. *Percept Psychophysics*, 43:401–403.

Carmel, D., Walsh, V., Lavie, N., and Rees, G. (2010). Right parietal TMS shortens dominance durations in binocular rivalry. *Current biology: CB*, 20(18):R799–800. PMID: 20869603.

Chalmers, D. J. (2010). *The Character of Consciousness.* Oxford University Press.

Clifford, C. W. (2010). Visual perception: Ambiguity involving parietal cortex. *Current Biology*, 20(18):R813–R815.

Colby, C. L. and Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, 22(1):319–49.

Cole, M. W., Bagic, A., Kass, R., and Schneider, W. (2010). Prefrontal dynamics underlying rapid instructed task learning reverse with practice. *Journal of Neuroscience*, 30(42):14245–14254.

Collins, D. L., Neelin, P., Peters, T. M., and Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized talairach space. *Journal of computer assisted tomography*, 18(2):192–205.

Crick, F. (1995). *The Astonishing Hypothesis.* Touchstone, New York, NY.

Crick, F. (1996). Visual perception: rivalry and consciousness. *Nature*, 379(6565):485–486.

Crick, F. and Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, 375(6527):121–123.

Crick, F. and Koch, C. (2003). A framework for consciousness. *Nature neuroscience*, 6(2):119–126.

Cumming, B. G. and Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, 389(6648):280–283.

David, O. (2011). fMRI connectivity, meaning and empiricism comments on: Roebroeck et al. the identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. *NeuroImage*, 58(2):306–309.

Dehaene, S. and Changeux, J.-P. (2003). Neural mechanisms for access to consciousness. pages 1145–1157.

Dehaene, S., Changeux, J.-P., and Naccache, L. (2011). The global neuronal workspace model of conscious access: From neuronal architectures to clinical applications. pages 55–84. Springer Berlin Heidelberg.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., and Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, 10(5):204–11.

Dehaene, S. and Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79(1-2):1–37.

Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J. F., Poline, J. B., and Riviere, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, 4(7):752–758.

Del Cul, A., Baillet, S., and Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS biology*, 5(10):2408–2423.

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., and Slachevsky, A. (2009). Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain*, 132(Pt 9):2531–40.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. pages 248–255.

Deshpande, G., Santhanam, P., and Hu, X. (2011). Instantaneous and causal connectivity in resting state brain networks derived from functional MRI data. *NeuroImage*, 54(2):1043–1052.

Deshpande, G., Sathian, K., and Hu, X. (2010). Effect of hemodynamic variability on granger causality analysis of fMRI. *Neuroimage*, 52(3):884–96.

Ding, M., Bressler, S. L., Yang, W., and Liang, H. (2000). Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Biological cybernetics*, 83(1):35–45.

Ding, M., Chen, Y., and Bressler, S. L. (2006). Granger causality: Basic theory and application to neuroscience. pages 451–474. Wiley-VCH.

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S., and Friston, K. J. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature*, 389(6651):596–599.

Durbin, J. and Watson, G. S. (1950). Testing for serial correlations in least squares regression. *Biometrica*, 37:409–428.

Eccles, J. C. (1994). *How the self controls its brain*. Springer-Verlag, Berlin; New York.

Edelman, D. B. and Seth, A. K. (2009). Animal consciousness: a synthetic approach. *Trends in neurosciences*, 32(9):476–84.

Edelman, G. M. (1992). *Bright air, brilliant fire : on the matter of the mind*. Basic Books, New York (N.Y.).

Edlund, J. A., Chaumont, N., Hintze, A., Koch, C., Tononi, G., and Adami, C. (2011). Integrated information increases with fitness in the evolution of animats. *PLoS Computational Biology*, 7(10):e1002236.

Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, New York.

Enns, J. T. and Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science*, 8(2):135–139.

Enns, J. T. and Di Lollo, V. (2000). What's new in visual masking? *Trends in cognitive sciences*, 4(9):345–352.

Eriksson, J., Larsson, A., Riklund Ahlstrom, K., and Nyberg, L. (2004). Visual consciousness: dissociating the neural correlates of perceptual transitions from sustained perception with fMRI. *Consciousness and cognition*, 13(1):61–72.

Fahrenfort, J. J., Scholte, H. S., and Lamme, V. A. F. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of cognitive neuroscience*, 19(9):1488–1497.

Fair, D. A., Schlaggar, B. L., Cohen, A. L., Miezin, F. M., Dosenbach, N. U. F., Wenger, K. K., Fox, M. D., Snyder, A. Z., Raichle, M. E., and Petersen, S. E. (2007). A method for using blocked and event-related fMRI data to study "resting state" functional connectivity. *NeuroImage*, 35(1):396–405.

Fang, F. and He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience*, 8(10):1380–1385.

Farzin, F., Rivera, S. M., and Whitney, D. (2009). Holistic crowding of mooney faces. *Journal of vision*, 9(6):18.1–15.

Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*. Cambridge, MA.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., and Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnetic Resonance in Medicine*, 33(5):636–647.

Fox, M. D. and Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Frontiers in Systems Neuroscience*.

Friston, K. J. (2011). Dynamic causal modeling and granger causality comments on: the identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. *NeuroImage*, 58(2):303–305.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., and Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*, 6(3):218–229.

Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, 19(4):1273–1302.

Friston, K. J., Holmes, A., Poline, J. B., Price, C. J., and Frith, C. D. (1996). Detecting activations in PET and fMRI: levels of inference and power. *NeuroImage*, 4(3):223–235.

Friston, K. J., Penny, W. D., and Glaser, D. E. (2005). Conjunction revisited. *Neuroimage*, 25(3):661–667.

Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., and Evans, A. C. (1993). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, 1(3):210–220.

Gaillard, R., Dehaene, S., Adam, C., Clemenceau, S., Hasboun, D., Baulac, M., Cohen, L., and Naccache, L. (2009). Converging intracranial markers of conscious access. *PLoS Biology*, 7(3):e61–e61.

Gegenfurtner, K. R., Kiper, D. C., and Levitt, J. B. (1997). Functional properties of neurons in macaque area v3. *Journal of neurophysiology*, 77(4):1906–1923.

Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77(378):304–313.

Gitelman, D. R., Penny, W. D., Ashburner, J., and Friston, K. J. (2003). Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. *Neuroimage*, 19(1):200–207.

Goebel, R., Roebroeck, A., Kim, D.-S., and Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and granger causality mapping. *Magnetic Resonance Imaging*, 21(10):1251–1261.

Goldfine, A. M. and Schiff, N. D. (2011). Consciousness: Its neurobiology and the major classes of impairment. *Neurologic Clinics*, 29(4):723–737.

Goodale, M. A. and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, 15(1):20–25.

Gosseries, O., Bruno, M., Vanhaudenhuyse, A., Schnakers, C., and Laureys, S. (2009). Consciousness in the locked-in syndrome. In *The Neurology of Consciousness*, pages 191–201. Elsevier Academic Press.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Green, M. F., Glahn, D., Engel, S. A., Nuechterlein, K. H., Sabb, F., Strojwas, M., and Cohen, M. S. (2005). Regional brain activity associated with visual backward masking. *Journal of cognitive neuroscience*, 17(1):13–23.

Greicius, M. D. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, 100(1):253–258.

Grill-Spector, K., Kushnir, T., Hendler, T., and Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature neuroscience*, 3(8):837–843.

Gur, M. and Snodderly, D. M. (1997). A dissociation between brain activity and perception: chromatically opponent cortical neurons signal chromatic flicker that is not perceived. *Vision research*, 37(4):377–382.

Guse, B., Falkai, P., and Wobrock, T. (2009). Cognitive effects of high-frequency repetitive transcranial magnetic stimulation: a systematic review. *Journal of Neural Transmission*, 117(1):105–122.

Gusnard, D. A., Raichle, M. E., and Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nature Reviews Neuroscience*, 2(10):685–694.

Hamilton, J. D. (1994). *Time Series Analyis*. Princeton University Press, Princeton, NJ.

Handwerker, D. A., Ollinger, J. M., and D'Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *NeuroImage*, 21(4):1639–1651.

Harris, J. A., Wu, C.-T., and Woldorff, M. G. (2011). Sandwich masking eliminates both visual awareness of faces and face-specific brain activity through a feedforward mechanism. *Journal of Vision*, 11(7):1–12.

Haynes, J.-D. (2009a). Decoding visual consciousness from human brain signals. *Trends in Cognitive Sciences*, 13(5):194–202.

Haynes, J.-D. (2009b). The neural basis of perceptual awareness. *Encyclopedia of Consciousness*, pages 75–86.

Haynes, J.-D., Deichmann, R., and Rees, G. (2005a). Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. *Nature*, 438(7067):496–499.

Haynes, J.-D., Driver, J., and Rees, G. (2005b). Visibility reflects dynamic changes of effective connectivity between v1 and fusiform cortex. *Neuron*, 46(5):811–821.

Haynes, J.-D. and Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, 8(5):686–91.

Haynes, J.-D. and Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7):523–534.

He, S. and MacLeod, D. I. (2001). Orientation-selective adaptation and tilt after-effect from invisible patterns. *Nature*, 411(6836):473–476.

Heath, R., Carpenter, M., Mettler, F., and Kline, N. (1949). Visual apparatus: visual fields and acuity, color vision, autokinesis. *Selective partial ablation of the frontal cortex, a correlative study of its effects on human psychotic subjects*, pages 489–491.

Hebart, M. N. and Hesselmann, G. (2012). What visual information is processed in the human dorsal stream? *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 32(24):8107–8109.

Henson, R. and Friston, K. (2007). Convolution models for fMRI. In Friston, K., Ashburner, J. T., Kiebel, S. J., Nichols, T. E., and Penny, W. D., editors, *Statistical parametric mapping*, pages 178–192. Academic Press.

Henson, R. N. A. and Penny, W. D. (2003). ANOVAs and SPM. Technical report, Wellcome Department of Imaging Neuroscience.

Hesselmann, G. and Malach, R. (2011). The link between fMRI-BOLD activation and perceptual awareness is "Stream-Invariant" in the human visual system. *Cerebral Cortex*, 21(12):2829–2837.

Holmes, A. P. and Friston, K. J. (1998). Generalizability, random effects, and population inference. *Neuroimage*, 7:754–754.

Honey, C. J., Kotter, R., Breakspear, M., and Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences*, 104(24):10240–10245.

Horwitz, B. (2003). The elusive concept of brain connectivity. *NeuroImage*, 19(2 Pt 1):466–470.

Hsieh, P.-J., Vul, E., and Kanwisher, N. (2010). Recognition alters the spatial pattern of FMRI activation in early retinotopic cortex. *Journal of neurophysiology*, 103(3):1501–1507.

Huettel, S. A., Song, A. W., and McCarthy, G. (2008). *Functional magnetic resonance imaging*. Sinauer Associates, Sunderland, Mass.

Hwang, K., Velanova, K., and Luna, B. (2010). Strengthening of top-down frontal cognitive control networks underlying the development of inhibitory control: A functional magnetic resonance imaging effective connectivity study. *Journal of Neuroscience*, 30(46):15535–15545.

Imamoglu, F., Heinzle, J., Imfeld, A., and Haynes, J.-D. (2012a). Changes in functional connectivity support changes in visibility. In *Society for Neuroscience*, New Orleans, USA. Poster Session 285.17.

Imamoglu, F., Kahnt, T., Koch, C., and Haynes, J.-D. (2012b). Changes in functional connectivity support conscious object recognition. *NeuroImage*, 63(4):1909–1917.

Imamoglu, F., Koch, C., and Haynes, J.-D. (2013). MoonBase: generating a database of two-tone "Mooney" images. In *Vision Science Society Annual Meeting*, Naples, Florida. Poster Session.

Ives, H. E. (1923). A chart of the flicker photometer. *Journal of the Optical Society of America*, 7(5):363–373.

Jezzard, P., Matthews, P. M., and Smith, M. S. (2002). *Functional MRI. An Introduction to Methods.* Oxford University Press, New York, NY, 2nd edition.

Jin, S.-H. and Chung, C. K. (2012). Messages from the brain connectivity regarding neural correlates of consciousness. *Experimental Neurobiology*, 21(3):113–122.

Kanai, R., Bahrami, B., and Rees, G. (2010). Human parietal cortex structure predicts individual differences in perceptual rivalry. *Current biology: CB*, 20(18):1626–1630.

Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J. Neurosci.*, 17(11):4302–4311.

Kleinschmidt, A., BÃŒchel, C., Zeki, S., and Frackowiak, R. S. (1998). Human brain activity during spontaneously reversing perception of ambiguous figures. *Proceedings of the Royal Society B: Biological Sciences*, 265(1413):2427–2433.

Koch, C. (2005). *Bewusstsein ein neurobiologisches Raetsel*. Elsevier - Spektrum, Muenchen.

Koch, C. and Tononi, G. (2008). Can machines be conscious. *IEEE Spectrum Inside Technology*.

Koch, C. and Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in cognitive sciences*, 11(1):16–22.

Kolers, P. A. (1962). Intensity and contour effects in visual masking. *Vision Research*, 2(9-10):277–294.

Kondo, H. and Komatsu, H. (2000). Suppression on neuronal responses by a metacontrast masking stimulus in monkey v4. *Neuroscience research*, 36(1):27–33.

Konen, C. S. and Kastner, S. (2008). Two hierarchically organized neural systems for object information in human visual cortex. *Nature Neuroscience*, 11(2):224–231.

Kouider, S. (2009). Neurobiological theories of consciousness. *Encyclopedia of Consciousness*, pages 87–100.

Kouider, S., Dehaene, S., Jobert, A., and Le Bihan, D. (2007). Cerebral bases of subliminal and supraliminal priming during reading. *Cerebral cortex*, 17(9):2019–2029.

Kovacs, G., Vogels, R., and Orban, G. A. (1995). Cortical correlate of pattern backward masking. *Proceedings of the National Academy of Sciences*, 92(12):5587–5591.

Kreiman, G., Fried, I., and Koch, C. (2002). Single-neuron correlates of subjective vision in the human medial temporal lobe. *Proceedings of the National Academy of Sciences*, 99(12):8378–8383.

Kriegeskorte, N. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences*, 103(10):3863–3868.

Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 15(2):1605–1615.

Lamme, V. A., Rodriguez-Rodriguez, V., and Spekreijse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Cortex*, 9(4):406–413.

Lamme, V. A. and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in neurosciences*, 23(11):571–579.

Lamme, V. A., Super, H., and Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current opinion in neurobiology*, 8(4):529–535.

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10(11):494–501.

Lamme, V. A. F., Zipser, K., and Spekreijse, H. (2002). Masking interrupts figure-ground signals in v1. *Journal of cognitive neuroscience*, 14(7):1044–1053.

Lau, H. C. and Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences*, 103(49):18763–18768.

Laureys, S. (2005). The neural correlate of (un)awareness: lessons from the vegetative state. *Trends in cognitive sciences*, 9(12):556–559.

Laureys, S., Pellas, F., Van Eeckhout, P., Ghorbel, S., Schnakers, C., Perrin, F., Berre, J., Faymonville, M.-E., Pantke, K.-H., Damas, F., Lamy, M., Moonen, G., and Goldman, S. (2005). The locked-in syndrome : what is it like to be conscious but paralyzed and voiceless? *Progress in brain research*, 150:495–511.

Laureys, S. and Tononi, G. (2009). *The neurology of consciousness : cognitive neuroscience and neuropathology*. Elsevier Academic Press, Amsterdam; Boston.

Leopold, D. A. and Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, 379(6565):549–553.

Li, X., Branch, C. A., Bertisch, H. C., Brown, K., Szulc, K. U., Ardekani, B. A., and DeLisi, L. E. (2007). An fMRI study of language processing in people at high genetic risk for schizophrenia. *Schizophrenia research*, 91(1-3):62–72.

Liao, W., Mantini, D., Zhang, Z., Pan, Z., Ding, J., Gong, Q., Yang, Y., and Chen, H. (2010). Evaluating the effective connectivity of resting state networks using conditional granger causality. *Biological cybernetics*, 102(1):57–69.

Libedinsky, C. and Livingstone, M. (2011). Role of prefrontal cortex in conscious visual perception. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 31(1):64–69.

Liu, Y. (2009). *A new method of threshold and gradient optimization using class uncertainty theory and its quantitative analysis*. PhD thesis, University of Iowa.

Livingstone, M. S. and Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 7(11):3416–3468.

Logothetis, N. and Schall, J. (1989). Neuronal correlates of subjective visual perception. *Science*, 245(4919):761–763.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, 453(7197):869–878.

Logothetis, N. K., Leopold, D. A., and Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? *Nature*, 380(6575):621–624.

Ludmer, R., Dudai, Y., and Rubin, N. (2011). Uncovering camouflage: amygdala activation predicts long-term memory of induced perceptual insight. *Neuron*, 69(5):1002–1014.

Lumer, E. D., Friston, K. J., and Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*, 280(5371):1930–1934.

Lumer, E. D. and Rees, G. (1999). Covariation of activity in visual and prefrontal cortex associated with subjective visual perception. *Proceedings of the National Academy of Sciences*, 96(4):1669–1673.

Macaluso, E., Frith, C. D., and Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, 289(5482):1206–1208.

Macknik, S. L. (2006). Visual masking approaches to visual awareness. *Progress in brain research*, 155:177–215.

Macknik, S. L. and Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nature neuroscience*, 1(2):144–149.

Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., and Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science*, 309(5744):2228–2232.

Matthews, P. M. (2009). An introduction to functional magnetic resonance imaging of the brain. In *Functional MRI. An Introduction to Methods.*, pages 3–34. Oxford University Press, New York, NY, 2 edition.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., and et al. (2001). A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (ICBM). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 356(1412):1293–1322.

McKeeff, T. J. and Tong, F. (2007). The timing of perceptual decisions for ambiguous face stimuli in the human ventral visual cortex. *Cerebral Cortex*, 17(3):669–678.

Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., and Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *The Journal of neuroscience*, 27(11):2858–2865.

Miller, G. A. (1995). WordNet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.

Mooney, C. M. (1957). Age in the development of closure ability in children. *Canadian Journal of Psychology*, 11(4):219–226.

Nakamura, R. and Mishkin, M. (1986). Chronic 'blindness' following lesions of nonvisual cortex in the monkey. *Experimental Brain Research*, 63(1):173–184.

Naselaris, T., Kay, K. N., Nishimoto, S., and Gallant, J. L. (2011). Encoding and decoding in fMRI. *NeuroImage*, 56(2):400–10.

Nichols, T., Brett, M., Andersson, J., Wager, T., and Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage*, 25(3):653–660.

Nichols, T. and Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical methods in medical research*, 12(5):419–446.

Nichols, T. E. and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human Brain Mapping*, 15(1):1–25.

Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in cognitive sciences*, 10(9):424–30.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., and Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, 89(13):5951–5955.

Oizumi, M., Yanagawa, T., Amari, S.-i., Tsuchiya, N., and Fujii, N. (2012). Measuring the level of consciousness based on the amount of integrated information. *Frontiers in Human Neuroscience*, (75).

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66.

Penny, W. D. and Holmes, A. P. (2007). Random effects analysis. In Friston, K. J., Ashburner, J., Penny, W. D., Kiebel, S. J., and Nichols, T. E., editors, *Statistical parametric mapping: The analysis of functional brain images*, pages 156–166. Academic Press.

Philiastides, M., Auksztulewicz, R., Heekeren, H., and Blankenburg, F. (2011). Causal role of dorsolateral prefrontal cortex in human perceptual decision making. *Current Biology*, 21(11):980–983.

Pins, D. and Ffytche, D. (2003). The neural correlates of conscious vision. *Cerebral cortex*, 13(5):461–474.

Pollen, D. A. (2007). Fundamental requirements for primary visual perception. *Cerebral Cortex*, 18(9):1991–1998.

Polonsky, A., Blake, R., Braun, J., and Heeger, D. J. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neuroscience*, 3(11):1153–1159.

Popper, K. R. and Eccles, J. C. (1996). *Das Ich und sein Gehirn*. Piper, Muenchen; Zuerich.

Portas, C. M., Strange, B. A., Friston, K. J., Dolan, R. J., and Frith, C. D. (2000). How does the brain sustain a visual percept? *Proceedings of the Royal Society B: Biological Sciences*, 267(1446):845–850.

Price, C. J. and Friston, K. J. (1997). Cognitive conjunction: a new approach to brain activation experiments. *Neuroimage*, 5(4):261–270.

Price, C. J., Wise, R. J., Warburton, E. A., Moore, C. J., Howard, D., Patterson, K., Frackowiak, R. S., and Friston, K. J. (1996). Hearing and saying. the functional neuro-anatomy of auditory word processing. *Brain : a journal of neurology*, 119(3):919–931.

Rees, G. (2009). Visual awareness. In Gazzaniga, M. S., editor, *The Cognitive Neurosciences*, pages 1151–1163. MIT Press, 4 edition.

Rees, G., Kreiman, G., and Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience*, 3(4):261–270.

Rees, G., Wojciulik, E., Clarke, K., Husain, M., Frith, C., and Driver, J. (2000). Unconscious activation of visual cortex in the damaged right hemisphere of a parietal patient with extinction. *Brain*, 123(8):1624–1633.

Ress, D. and Heeger, D. J. (2003). Neuronal correlates of perception in early visual cortex. *Nature Neuroscience*, 6(4):414–420.

Ricci, C. and Blundo, C. (1990). Perception of ambiguous figures after focal brain lesions. *Neuropsychologia*, 28(11):1163–1173.

Rissman, J., Gazzaley, A., and D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage*, 23(2):752–763.

Ro, T., Breitmeyer, B., Burton, P., Singhal, N. S., and Lane, D. (2003). Feedback contributions to visual awareness in human occipital cortex. *Current Biology*, 13(12):1038–1041.

Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using granger causality and fMRI. *Neuroimage*, 25(1):230–242.

Roebroeck, A., Formisano, E., and Goebel, R. (2011a). The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. *NeuroImage*, 58(2):296–302.

Roebroeck, A., Seth, A., and Valdes-Sosa, P. (2011b). Causal time series analysis of functional magnetic resonance imaging data. *Journal of Machine Learning Research*, 12:65–94.

Rolls, E. T. and Tovee, M. J. (1994). Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proceedings. Biological sciences / The Royal Society*, 257(1348):9–15.

Rolls, E. T., Tovee, M. J., and Panzeri, S. (1999). The neurophysiology of backward visual masking: information analysis. *Journal of cognitive neuroscience*, 11(3):300–11.

Rubin, E. (1921). *Visuell wahrgenommene Figuren*. Gyldendalske Boghandel, Berlin.

Ruff, C. C., Blankenburg, F., Bjoertomt, O., Bestmann, S., Freeman, E., Haynes, J.-D., Rees, G., Josephs, O., Deichmann, R., and Driver, J. (2006). Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. *Current biology: CB*, 16(15):1479–1488.

Sahraie, A., Weiskrantz, L., Barbur, J. L., Simmons, A., Williams, S. C. R., and Brammer, M. J. (1997). Pattern of neuronal activity associated with conscious and unconscious processing of visual signals. *Proceedings of the National Academy of Sciences*, 94(17):9406–9411.

Schroff, F., Criminisi, A., and Zisserman, A. (2007). Harvesting image databases from the web. pages 1–8. IEEE.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Searle, J. R. (1997). *The mystery of consciousness*. New York Review of Books, New York.

Seth, A. (2007). Granger causality. *Scholarpedia*, 2(7):1667.

Seth, A. (2010). A MATLAB toolbox for granger causal connectivity analysis. *Journal of Neuroscience Methods*, 186:262–273.

Seth, A. K., Chorley, P., and Barnett, L. C. (2013). Granger causality analysis of fMRI BOLD signals is invariant to hemodynamic convolution but not downsampling. *NeuroImage*, 65:540–555.

Sheinberg, D. L. and Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences*, 94(7):3408–3413.

Smith, S. (2004). Overview of fMRI analysis. *Br J Radiol.*, 77(2):S167–175.

Soon, C. S., Brass, M., Heinze, H.-J., and Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11(5):543–545.

Sorokin, A. and Forsyth, D. (2008). Utility data annotation with amazon mechanical turk. pages 1–8. IEEE.

Sporns, O. (2007). Brain connectivity. *Scholarpedia*, 2(10):4695.

Sridharan, D., Levitin, D. J., and Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences of the United States of America*, 105(34):12569–12574.

Stockman, A., MacLeod, D. I., and Lebrun, S. J. (1993). Faster than the eye can see: blue cones respond to rapid flicker. *Journal of the Optical Society of America. A, Optics and image science*, 10(6):1396–1402.

Stoerig, P. and Cowey, A. (1997). Blindsight in man and monkey. *Brain*, 120(3):535–559.

Super, H., Spekreijse, H., and Lamme, V. A. (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex (v1). *Nature neuroscience*, 4(3):304–310.

Tallon-Baudry, C. and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in cognitive sciences*, 3(4):151–162.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., and Permier, J. (1997). Oscillatory gamma-band (30-70 hz) activity induced by a visual search task in humans. *The Journal of Neuroscience*, 17(2):722–734.

Thompson, K. G. and Schall, J. D. (1999). The detection of visual signals by macaque frontal eye field during masking. *Nature neuroscience*, 2(3):283–288.

Tong, F. (2001). Competing theories of binocular rivalry: A possible resolution. *Brain and Mind*, 2(1):55–83.

Tong, F. (2003). Primary visual cortex and visual awareness. *Nature Reviews Neuroscience*, 4(3):219–229.

Tong, F. and Engel, S. A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature*, 411(6834):195–199.

Tong, F., Nakayama, K., Vaughan, J. T., and Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, 21(4):753–759.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5(1):42–42.

Tononi, G. and Balduzzi, D. (2009). Toward a theory of consciousness. In *The Cognitive Neurosciences*, pages 1201–1217. MIT Press, 4 edition.

Tononi, G. and Edelman, G. M. (1998). Consciousness and complexity. *Science*, 282(5395):1846–1851.

Tononi, G. and Koch, C. (2008). The neural correlates of consciousness: An update. *Annals of the New York Academy of Sciences*, 1124:239–261.

Tononi, G. and Koch, C. (2011). How simple photos could be used as a test for a conscious machine. *Scientific American*.

Tononi, G. and Sporns, O. (2003). Measuring information integration. *BMC Neuroscience*, 4(1):31–51.

Tononi, G., Srinivasan, R., Russell, P. D., and Edelman, G. M. (1998). Investigating neural correlates of conscious perception by frequency-tagged neuromagnetic responses. *Proceedings of the National Academy of Sciences*, 95(6):3198–3203.

Tse, P. U., Martinez-Conde, S., Schlegel, A. A., and Macknik, S. L. (2005). Visibility, visual awareness, and visual masking of simple unattended targets are confined to areas in the occipital cortex beyond human V1/V2. *Proceedings of the National Academy of Sciences of the United States of America*, 102(47):17178–17183.

Valdes-Sosa, P. A., Roebroeck, A., Daunizeau, J., and Friston, K. J. (2011). Effective connectivity: Influence, causality and biophysical modeling. *Neuroimage*, 58(2):339–361.

Vanhaudenhuyse, A., Noirhomme, Q., Tshibanda, L. J.-F., Bruno, M.-A., Boveroux, P., Schnakers, C., Soddu, A., Perlbarg, V., Ledoux, D., Brichant, J.-F., Moonen, G., Maquet, P., Greicius, M. D., Laureys, S., and Boly, M. (2010). Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients. *Brain*, 133(1):161–171.

von der Heydt, R., Friedman, H., Zhou, H., Komatsu, H., Hanazawa, A., and Murakami, I. (1997). Neuronal responses in monkey v1 and v2 unaffected by metacontrast. *Invest. Ophthalm. Vis. Sci.*, 38:2146–2146.

Weiskrantz, L. (1990). *Blindsight a case study and implications*. Clarendon, Oxford.

Wiggs, C. L., Weisberg, J., and Martin, A. (1999). Neural correlates of semantic and episodic memory retrieval. *Neuropsychologia*, 37(1):103–118.

Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, 20(1):6–10.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., and Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human brain mapping*, 4(1):58–73.

Zaretskaya, N., Thielscher, A., Logothetis, N. K., and Bartels, A. (2010). Disrupting parietal function prolongs dominance durations in binocular rivalry. *Current Biology*, 20(23):2106–2111.

Zeki, S. (2007). A theory of micro-consciousness. In Velmans, M. and Schneider, S., editors, *The Blackwell Companion to Consciousness*, pages 580–588. Blackwell Publishing, Malden, MA, USA.

Zeki, S. and Bartels, A. (1999). Toward a theory of visual consciousness. *Consciousness and cognition*, 8(2):225–259.

Zhang, D., Snyder, A. Z., Fox, M. D., Sansbury, M. W., Shimony, J. S., and Raichle, M. E. (2008). Intrinsic functional relations between human cerebral cortex and thalamus. *Journal of neurophysiology*, 100(4):1740–1748.

Zipser, K., Lamme, V. A., and Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 16(22):7376–7389.

Statement of Authorship

I hereby certify that this dissertation has been composed by me and is based on my own work, unless stated otherwise. Ideas and thoughts cited directly or indirectly from other work have been cited accordingly.

Berlin, 13.03.2013 _____

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle Ausführungen, die anderen veröffentlichten oder nicht veröffentlichten Schriften wörtlich oder sinngemäss entnommen wurden, habe ich kenntlich gemacht.

Die Arbeit hat in gleicher oder ähnlicher Fassung noch keiner anderen Prüfungsbehörde vorgelegen.

Berlin, 13.03.2013 _____

Curriculum Vitae

Fatma İmamoglu

Berlin Center for Advanced Neuroimaging
Chariteplatz 1, PF: 35
10117 Berlin, Germany
cell: +49 (0)163 436 36 59
skype: fatma.imamoglu
email: fatmaimamoglu@gmail.com

Education

Since 2008 Bernstein Center for Computational Neuroscience,
Charité –Universitätsmedizin Berlin, Germany
PhD on Visual Consciousness and Corticocortical Connectivity in the Human Brain
Co-advisors:

- *Prof. Dr. John-Dylan Haynes (Bernstein Center for Computational Neuroscience)*
- *Prof. Dr. Christof Koch (California Institute of Technology, Allen Institute for Brain Sciences)*

2008 Technical University Munich (TUM)
Graduation with summa cum laude in Computer Science (1.0, excellent)

2007–2008 California Institute Of Technology (Caltech)
Master Thesis on Visual Saliency and Biological Inspired Text Detection
Co-advisors:

- *Prof. Dr. Michael Beetz (TUM)*
- *Prof. Dr. Christof Koch (CALTECH)*

2001–2008 Technical University Munich (TUM)
Major: Computer Science, Minor: Computer Linguistic

1997–2001 High-School
Graduation from High School, Turkey (Ranked 1st) (5.00, excellent)

Work Experiences

2007 July-Sep Technical University Munich
Image Understanding & Knowledge-Based Systems Group
Student assistantship

2006 June-Sep Siemens Corporate Technologies
Internship
Development and Implementation of Dynamically Tunable Attractor Neural Networks

2004–2007 *Computer Science Centre Bavaria*
Planning Unix-Systems, Unix-Databases
Oracle Database Administrator

Publications

2013 İmamoglu F., Heinzle J., Imfeld A., and Haynes J.-D., High-level visual brain regions reflect visibility of low-level visual stimuli (in prep.)

2013 İmamoglu F., Koch C., and Haynes J.-D., Generating a two-tone image database from the web: “MoonBase” (in prep.)

2012 İmamoglu F., Kahnt T., Koch C. and Haynes J.-D., Changes in functional connectivity support conscious object recognition, NeuroImage, 63:1909-17

2008 İmamoglu, F., Visual Saliency and Biological Inspired Text Detection, Master Thesis

Selected Conference Talks and Posters

2013 İmamoglu F., Koch C., and Haynes J.-D., MoonBase: Generating a database of two-tone "Mooney" images, Vision Science Society, Annual Meeting, Poster Session

2012 İmamoglu F., Heinzle J., Imfeld A., and Haynes J.-D., Changes in functional connectivity support changes in visibility, Society for Neuroscience, Poster Session 285.17

2011 İmamoglu F., Kahnt T., Koch C., and Haynes J.-D., Changes in effective connectivity support conscious perception, Society for Neuroscience, Nanasymposium on Human Visual Perception, 125.2

2011 İmamoglu F., Kahnt T., Koch C., and Haynes J.-D., Changes in effective connectivity support conscious object recognition, Berlin Brain Days, Session 4

2010 İmamoglu F., Koch C., and Haynes J.-D., The moment of conscious perception, Perception 39 ECVP Abstract Supplement, page 123

Invited Talks

2012 Feb 17 Neural Information Processing Group, Technical University Berlin
Changes in effective connectivity support conscious object recognition

2010 Dec 16 Koch Lab, California Institute of Technology
Adventures with Mooney Images: Psychophysics & fMRI

2010 Nov 10 Modelling Cognitive Systems Lab, Technical University Berlin
From two-tone to many-tone images

Attended Summer Schools and Workshops

2011 Nov 10–11 Workshop - Single neuron studies of the human brain New York University, New York, USA

2011 May 25–28 Alfred Nobel Symposia - Machines, Molecules and Minds Stockholm, Sweden

2010 June 02–04 Brain Connectivity Workshop Berlin, Germany

2010 May 19–21 International Workshop - Neuro-cognitive mechanisms of conscious and unconscious visual perception Hanse-Wissenschaftskolleg, Delmenhorst, Germany

2009 June 15–26 Summer School - Consciousness and the Brain Cognitive Science Center of the University of Amsterdam

Computer Skills

Programming languages: Matlab, Python, Unix Scripts, Prolog

Databases: Oracle Databases, SQL

Markup languages: LaTeX, XML, HTML

Languages

Turkish	Native
German	Fluent
English	Advanced in reading, writing and speaking

Scholarships, Grants and Awards

2008–2013	Bernstein Center for Computational Neuroscience, Charité – Universitätsmedizin Berlin, Germany – PhD scholarship
2007–2008	California Institute of Technology, Pasadena, USA – GRA scholarship and grant
2008	Technical University Munich, Germany – Scholarship for international students
2001	Ranked 1st while graduating from High School
2000–2001	President of Student's Council in High School

Mentoring Program

Since 2008	Mentor in the mentoring program Club Lise for school children[1]

Association Memberships

2012-2013	Vision Science Society
2009-2012	Society for Neuroscience
2011-2012	Association for the Scientific Studies of Consciousness

Academic Interests

Computer Science (Machine Learning, Artificial Intelligence, Robotics, Graph Theory, Information Theory),
Neuroscience (Human Brain Imaging Techniques, Human Brain Connectivity)
Psychology (Conscious/Unconscious Visual Perception, Attention, Conscious/Unconscious Emotional Processing).

[1] `http://didaktik.physik.hu-berlin.de/club-lise/lise-mentoring.php`

Personal Interests

Running, Climbing, Writing, Baking, Music (classical guitar), Architecture

Berlin, 13.03.2013 _____