Aus dem
Max-Planck Institut für Hirnforschung
in Frankfurt am Main

Distilling the neural correlates of conscious perception

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt am Fachbereich Psychologie und
Sportwissenschaften der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Jaan Aru
aus Tartu, Estland

Frankfurt am Main (2013)
D (30)

vom Fachbereich Psychologie und Sportwissenschaften der Johann Wolfgang Goethe – Universität als Dissertation angenommen.

Dekan:
Prof. Dr. Rolf van Dick

Gutachter:
Prof. Dr. Christian Fiebach
Prof. Dr. Wolf Singer
Prof. Dr. Catherine Tallon-Baudry

Datum der Disputation: 24.06.2014

Table of contents

Zusammenfassung

In der vorliegenden Arbeit wird untersucht, wie das Gehirn Bewusstsein erzeugt. Diese Frage wird als eines der größten Rätsel der heutigen Wissenschaft angesehen: Wie kann es sein, dass aus der Aktivität der Nervenzellen unsere subjektive Welt entsteht?

Es ist offensichtlich nicht einfach, diese Frage wissenschaftlich zu untersuchen. Eine der vorgeschlagenen Strategien für die Untersuchung von Bewusstsein behauptet, dass man zunächst die neuronalen Korrelate des Bewusstseins finden sollte (Koch, 2004). Einer Definition zufolge sind die neuronalen Korrelate des Bewusstseins die kleinste Menge neuronaler Prozesse, die hinreichend für eine bestimmte bewusste Erfahrung sind (zum Beispiel für die bewusste Erfahrung des Blaubeergeschmacks). Manche behaupteten, die Entdeckung der neuronalen Korrelate des Bewusstseins würde es erlauben, dem Rätsel des Bewusstseins näher zu kommen (Crick & Koch, 1990). Nur wie soll man die neuronalen Korrelate des Bewusstseins finden?

Eine relativ einfache Strategie dafür wurde schon vor mehr als 20 Jahren beschrieben. Es sollten einfach experimentelle Bedingungen erschaffen werden, in welchen ein Reiz manchmal bewusst wahrgenommen wird und manchmal nicht (Baars, 1989). Solche Analysen, die Bedingungen mit und ohne bewusste Wahrnehmung vergleichen, werden als „Kontrastierungsanalyse“ bezeichnet (da zwei Bedingungen miteinander kontrastiert werden). Es existieren viele verschiedene experimentelle Paradigmen, bei welchen man den Reiz unter denselben Bedingungen präsentieren kann, so dass er bei manchen Versuchsdurchgängen bewusst wahrgenommen wird, bei anderen nicht (Kim & Blake, 2005). Mit solchen experimentellen Paradigmen kann man angeblich die

neuronalen Korrelate des Bewusstseins finden, wenn man a) bei jedem Durchgang die Versuchsperson fragt, ob oder was die Versuchsperson bei dem Durchgang wahrgenommen hat und b) gleichzeitig die neuronalen Prozesse misst (zum Beispiel mit EEG, MEG oder fMRT). Anschließend kann man die erhobenen neuronalen Daten unter den Bedingungen mit und ohne bewusste Wahrnehmung vergleichen.

Mittlerweile gibt es viele Studien, in denen solche experimentelle Paradigmen – und damit die Kontrastierungsanalyse – angewendet wurden. Insofern könnte man glauben, die neuronalen Korrelate des Bewusstseins seien schon gefunden worden. Allerdings ist dies nicht der Fall. Es existiert in der Literatur weiterhin Uneinigkeit darüber, ob die Korrelate des Bewusstseins früh oder spät in der Zeit liegen, und ob die Korrelate in sensorischen Arealen oder eher im hierarchisch höheren fronto-parietalen Kortex zu finden sind.

Nach unserer Meinung sind die experimentellen Paradigmen, die üblicherweise zum Auffinden der neuronalen Korrelate des Bewusstseins verwendet werden, nicht spezifisch genug, um diese eindeutig zu lokalisieren. Eher glauben wir, dass die klassische Kontrastierungsanalyse auch andere Prozesse als Ergebnisse hervorbringt und uns deshalb prinzipiell nicht zu den neuronalen Korrelaten des Bewusstseins führen kann.

Im Kapitel 2 wird erklärt, wieso die typischen experimentellen Paradigmen nicht die neuronalen Korrelate des Bewusstseins ausfindig machen können. Wir behaupten, dass der Vergleich neuronaler Daten aus experimentellen Bedingungen mit und ohne bewusste Wahrnehmung auch die neuronalen Prozesse widerspiegeln könnte, die bewussten Wahrnehmungen entweder vorausgehen oder folgen. Es ist beispielsweise bekannt, dass neuronale Prozesse vor Auftreten des Reizes darüber bestimmen können, ob der Reiz bewusst wahrgenommen wird oder nicht (Busch, Dubois, & VanRullen, 2009; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009).

Wenn man experimentelle Bedingungen mit und ohne bewusster Wahrnehmung miteinander vergleicht, werden auch solche Prozesse als Ergebnis auftauchen, obwohl diese zeitlich klar vor dem Reiz stattfinden und deshalb keine neuronalen Korrelate des Bewusstseins sein können. Es ist natürlich einfach zu entscheiden, dass diese Prozesse, die schon vor dem Reiz stattfinden, der bewussten Wahrnehmung vorausgehen müssen, aber es ist unmöglich zu sagen, ob ein neuronaler Prozess 100 oder 200 Millisekunden nach der Präsentation des Reizes immer noch ein Vorläuferprozess ist schon ein neuronales Korrelat des Bewusstseins darstellt. Deshalb ist die typische Kontrastierungsanalyse nicht spezifisch genug und wir wissen nicht, ob neuronale Prozesse, die durch die Kontrastierungsanalyse aufgedeckt werden, direkt die neuronalen Korrelate des Bewusstseins oder eher Prozesse vor der bewussten Wahrnehmung widerspiegeln.

Nicht nur die Vorläuferprozesse der bewussten Warnehmung stellen ein Problem dar. Auch Konsequenzen der bewussten Verarbeitung werden durch die Kontrastierungsanalyse gefunden. Beispielsweise wurden im medialen Temporallappen Neurone gefunden, die nur dann feuern, wenn ein Patient eine Person auf einem Bild bewusst erkennt, aber nicht feuern, wenn der Patient die Person auf dem Bild nicht bewusst wahrnimmt (Quiroga, Mukamel, Isham, Malach, & Fried, 2008). So könnte man vorerst meinen, dass das Feuern dieser Neurone das neuronale Korrelat des Bewusstseins sein könnte. Nach einer Läsion, sprich neuronalen Schädigung des medialen Temporallappens kann man die Welt jedoch weiterhin bewusst wahrnehmen (man hat jedoch Probleme mit dem Gedächtnis und Wiedererkennen). Insofern kann das Feuern dieser Neurone nicht das neuronale Korrelat des Bewusstseins sein und ist eher ein Beispiel für die Konsequenz der bewussten Verarbeitung. Wir behaupten, dass es noch viele andere solcher Vorläuferprozesse und Konsequenzen gibt, die notwendigerweise als Ergebnis bei der Kontrastierungsanalyse auftauchen, und also ist die typische

Kontrastierungsanalyse extrem unspezifisch bezüglich der neuronalen Korrelate des Bewusstseins. In anderen Worten: Die typische Kontrastierungsanalyse, bei welcher man experimentelle Bedingungen mit und ohne bewusste Wahrnehmung miteinander vergleicht, wird uns nicht helfen die neuronalen Korrelate des Bewusstseins zu finden.

Wir glauben, dass neue experimentelle Paradigmen entwickelt werden sollten, um die neuronalen Korrelate des Bewusstseins ausfindig zu machen. Wahrscheinlich gibt es kein einfaches Experiment, mit dem man die Vorläuferprozesse und Konsequenzen vollständig vermeiden kann, um damit direkt die neuronalen Korrelate des Bewusstseins zu bestimmen. Eher braucht man viele verschiedene Experimente, die Schritt für Schritt unser Wissen über die neuronalen Korrelate des Bewusstseins erweitern.

In der vorliegenden Arbeit (in Kapiteln 3, 4 und 5) wird ein neues experimentelles Paradigma angewandt. Dieses Paradigma wird nicht alle oben erwähnten Probleme lösen, wird aber hoffentlich erlauben, einige Vorläuferprozesse der bewussten Wahrnehmung von den neuronalen Korrelaten des Bewusstseins auseinanderzuhalten. Der Vorteil unseres experimentellen Paradigmas besteht darin, dass die bewusste Wahrnehmung durch zwei verschiedene Vorläuferprozesse beeinflusst wird. Die Versuchspersonen müssen auf schnell präsentierten und mittels Rauschens undeutlich gemachten Bildern eine Person detektieren. Die experimentellen Bedingungen sind derart gestaltet, dass die Versuchspersonen nicht bei jedem Durchgang die Person auf dem Bild wahrnehmen können. Damit können wir den Wahrnehmungsprozess manipulieren. Bei einer Manipulation variieren wir den Anteil des Rauschens auf dem Bild und damit die sensorische Evidenz. Je weniger Rauschen, desto besser können die Versuchspersonen die Bilder wahrnehmen und desto öfter sehen sie auch bewusst die Person auf dem Bild. Bei der anderen experimentellen

Manipulation der Wahrnehmung werden einige Bilder den Versuchspersonen vorher klar und ohne Rauschen gezeigt. Damit erschafft man Wissen über bestimmte Bilder, die später mit Rauschen präsentiert werden. Man kann zeigen, dass solch bestehendes Wissen tatsächlich die Wahrnehmung beeinflusst. Wenn die Versuchspersonen bestehendes Wissen über ein Bild haben, ist es wahrscheinlicher, dass sie die Person auf dem Bild bewusst wahrnehmen. Damit haben wir zwei verschiedene Vorläuferprozesse – sensorische Evidenz und bestehendes Wissen, die beide die bewusste Wahrnehmung beeinflussen. Beide Vorläuferprozesse erhöhen den Anteil der Durchgänge, in welchen die Versuchspersonen die Person auf dem Bild bewusst wahrnehmen.

Mit diesem experimentellen Paradigma möchten wir einige Aussagen über die neuronalen Korrelate des Bewusstseins testen. Wenn über einen neuronalen Prozess behauptet wird, dass er einem neuronalen Korrelat des Bewusstseins entspricht, müsste dieser Prozess von den beiden manipulierten Vorläuferprozessen in ähnlicher Weise beeinflusst werden, da bewusste Wahrnehmung durch beide manipulierten Vorläuferprozessen in ähnlicher Weise erleichtert wird. Wenn aber der Prozess, über den behauptet wird, er sei ein neuronales Korrelat des Bewusstseins, nicht durch beide Manipulationen geändert wird, kann dieser Prozess kein neuronales Korrelat des Bewusstseins sein, da er nicht beeinflusst wird, obwohl die bewusste Wahrnehmung geändert wurde.

Mit diesem experimentellen Paradigma und dieser Logik haben wir zwei unterschiedliche neuronale Prozesse getestet, von denen behauptet wird, dass sie den neuronalen Korrelaten des Bewusstseins entsprechen könnten. In Kapitel 3 wurde untersucht, ob lokale kategorienspezifische Gammabandaktivität die neuronalen Korrelate des Bewusstseins reflektieren könnte. In Kapitel 4 wurde mit diesem experimentellen Paradigma untersucht, ob die neuronale Synchronisierung dem neuronalen Korrelat des Bewusstseins entsprechen könnte.

Unsere Arbeit im Kapitel 3 baut auf der von Fisch und Kollegen (2009) auf. Fisch und Kollegen (2009) zogen aus ihrer experimentellen Arbeit den Schluss, dass lokale kategorienspezifische Gammabandaktivität die neuronalen Korrelate des Bewusstseins reflektieren könnte. Sie hatten Elektroden auf dem visuellen Kortex von Epilepsiepatienten implantiert und von diesen Elektroden die Gammabandaktivität abgeleitet. Im ersten Schritt suchten sie nach Elektroden, die kategorienspezifische Antworten zeigen. Bei den kategorienspezifischen Elektroden ist die Gammabandaktivität abhängig vom präsentierten Stimulusmaterial. Zum Beispiel kann man bei einer Elektrode auf dem Fusiform Face Area starke Gammabandaktivität nur dann messen, wenn ein Gesicht auf dem Bild zu sehen ist. Die Autoren benutzten solche kategorienspezifischen Elektroden, um nach den neuronalen Korrelaten des Bewusstseins zu suchen. Sie zeigten den Patienten Bilder von Gesichtern, Häusern und Objekten, die direkt nach der kurzen Präsentation maskiert wurden, so dass die Patienten nur bei manchen Durchgängen erkannten, was auf dem Bild war, bei anderen Durchgängen nicht. Dies entspricht der typischen Kontrastierungsanalyse. Die Ergebnisse haben klar gezeigt, dass bei diesen kategorienspezifischen Elektroden die Gammabandaktivität erhöht wurde, als die Patienten bewusst wahrnahmen, was auf dem Bild zu sehen war. Aus diesen Ergebnissen zogen die Autoren den Schluss, dass lokale kategorienspezifische Gammabandaktivität dem neuronalen Korrelat des Bewusstseins entspricht. Diese Aussage wollten wir mit unserem experimentellen Paradigma testen.

Um diese Behauptung zu untersuchen, erhoben wir sehr ähnliche Daten wie Fisch et al. (2009) und analysierten die Daten auf ähnliche Weise. Unsere experimentelle Frage war, ob die lokale kategorienspezifische Gammabandaktivität durch unsere beiden Manipulationen – sensorische Evidenz und bestehendes Wissen – in ähnlicher Weise erhöht wird. Dies sollte der Fall sein, wenn die lokale kategorienspezifische Gammabandaktivität dem neuronalen Korrelat des Bewusstseins entspricht, da sensorische Evidenz und bestehendes Wissen beide den Anteil der Durchgänge, in welchen die Versuchsperson die Person auf dem Bild bewusst wahrnimmt, erhöhen. Dieses Ergebnis wurde nicht gefunden. Stattdessen fanden wir, dass die lokale kategorien-spezifische Gammabandaktivität nur durch sensorische Evidenz erhöht wurde, bestehendes Wissen aber keinen Effekt auf diese Aktivierung hatte. Da bestehendes Wissen auch den Anteil der Durchgänge mit bewusster Wahrnehmung erhöht, die kategorienspezifische Gammabandaktivität aber nicht durch bestehendes Wissen erhöht wurde, kann man schlussfolgern, dass die kategorienspezifische Gammabandaktivität nicht die neuronalen Korrelate des Bewusstseins reflektieren kann.

Als nächstes (Kapitel 4) haben wir die Hypothese getestet, dass Synchronizität dem neuronalen Korrelat des Bewusstseins entspricht. Um diese Idee zu testen, maßen wir mittels Magnetoenzephalographie die magnetischen Felder des Gehirns, schätzten aus diesen Daten mittels Beamforming die neuronalen Aktivitätsquellen und quantifizierten die Synchronizität zwischen diesen Quellen. Wenn die interareale Synchronizität dem neuronalen Korrelat des Bewusstseins entspräche, sollte die Synchronizität für Bedingungen mit mehr sensorischer Evidenz und mit bestehendem Wissen erhöht sein. Dies wurde nicht beobachtet. Wir fanden, dass Synchronizität (gemittelt über die Quellen) nur bei den Bildern erhöht war, für die bestehendes Wissen vorlag. Ein ähnlicher Effekt für sensorische Evidenz wurde nicht gefunden. Insofern können wir sagen, dass unsere Befunde dagegen sprechen, dass neuronale Synchronizität den Mechanismus für Bewusstsein darstellt. Allerdings können wir das in diesem Fall auch nicht völlig ausschließen, denn Synchronizität könnte die Informationsverarbeitung auf einem kleineren Maßstab koordinieren als wir es mit dem MEG messen können (Singer, in press).

Im Kapitel 5 untersuchten wir, wie schnell bestehendes Wissen bewusste Verarbeitung beeinflussen kann. Um dies herauszufinden machten wir uns die intraindividuellen Unterschiede der perzeptuellen Leistung zu Nutze. Wir fanden, dass bestehendes Wissen bewusste Verarbeitung schon innerhalb der ersten 100 Millisekunden nach der Präsentation des Reizes beeinflusst. Wir beobachteten auch, dass ein größerer perzeptueller Effekt des bestehenden Wissens in geringerer neuronaler Aktivität in Durchgängen mit bestehendem Wissen hervorruft. Diese Ergebnisse sind im Einklang mit Theorien, die besagen, dass unsere Wahrnehmung bestehendes Wissen nutzt, um vorherzusagen, wie die visuelle Welt sich ändert und um die neuronalen Antworten zu verringern (Friston, 2010).

In der vorliegenden Arbeit wurde diskutiert, warum die typische Kontrastierungsanalyse uns nicht zu den neuronalen Korrelaten des Bewusstseins führen kann. Wir schlugen vor, dass neue experimentelle Paradigmen nötig sind, um näher an die neuronalen Korrelate des Bewusstseins heranzukommen. Es wurde ein neues Paradigma benutzt, um zwischen Vorläuferprozessen und neuronalen Korrelate des Bewusstseins zu unterscheiden. Mit diesem Paradigma wurden zwei sehr unterschiedliche Hypothesen getestet und gefunden, dass die kategorienspezifische Gammabandaktivität nicht die neuronalen Korrelate des Bewusstseins widerspiegeln kann. Wir hoffen, dass unsere Experimente eine Entwicklung von vielen weiteren und besseren experimentellen Paradigmen stimuliert, die zwischen den Vorläuferprozessen, den Konsequenzen und den eigentlichen Korrelaten des Bewusstseins unterscheiden können. Wenn man über die Kontrastierungsanalyse hinausgeht, kann man die gegenwärtigen Theorien des Bewusstseins testen und damit Schritt für Schritt näher an die neuronalen Grundlagen des Bewusstseins kommen.

1

General Introduction

The problem of consciousness

As a neuroscientist I have two very different perspectives on the brain. From the third person perspective, from the outside, I can see fluctuating membrane potentials, neurons firing, neurotransmitters being released to the synaptic cleft, postsynaptic receptors opening etc. The brain is like a machine. A very complicated machine, but nevertheless a machine. Yet from the first person perspective, from the inside of this brain, it does not feel machine-like – I see the lightning illuminating the dark sky, I hear the roaring thunder, I am a bit frightened. It feels like something to be me, I am conscious. How do these two perspectives fit together? How is the neuronal machinery related to being conscious of oneself and the surrounding world? Our current laws of nature give no explanation to the question how matter could become mind. Although consciousness is a central part of our everyday life, it is not known how it is produced by the neurobiological processes in the brain.

The problem of consciousness is not only a problem for a single graduating neuroscientist; it has been acknowledged to be one of the biggest challenges for modern science by many esteemed researchers. To express it in the words of Erwin Schrödinger, a Nobel Prize winner and one of the founders of quantum mechanics:

> "The world is a construct of our sensations, perceptions, memories. It is convenient to regard it as existing objectively on its own. But it certainly does not become manifest by its mere existence. Its becoming manifest is conditional on very special goings-on in very special parts of this very world, namely on certain events that happen in a brain. That is an inordinately peculiar kind of implication, which prompts the question: What particular properties distinguish these brain processes and enable them to produce the manifestation? Can we guess which material processes have this power, which not? Or simpler: What kind of material process is directly associated with consciousness?" (Schrödinger, 1958).

In his essay “Mind and matter” he even discusses a strategy how one could approach the problem of consciousness:

> “Not every nervous process, nay by no means every cerebral process, is accompanied by consciousness. Many of them are not, even though physiologically and biologically they are very much like the 'conscious' ones, both in frequency consisting of afferent impulses followed by efferent ones, and in their biological significance of regulating and timing reactions partly inside the system, partly towards a changing environment. /…/ By examining various representatives of physiologically very similar processes, all playing within our body, it ought not to be too difficult to find out by observation and reasoning the distinctive characteristics we are looking for.” (Schrödinger, 1958).

As will be evident in the next section, this strategy of contrasting processes with and without consciousness has been extensively used in consciousness research. In the current thesis I would like to propose that it is time to move beyond this simple contrast.

Scientific approach to study consciousness

One of the most important research goals since the beginning of the science of consciousness has been to unravel the neural correlates of consciousness (NCC) (Crick & Koch, 1990). The NCC are defined as the “minimal set of neuronal events jointly sufficient for a specific conscious experience” (Koch, 2004, p. 16). Reasons, why the NCC are important to pursue, are the following: 1) If one has experimentally isolated the NCC and posses the appropriate tools for manipulating these neural processes that correlate with consciousness, one can test how changing the NCC affects conscious sensation. This would allow one to take the step from mere correlates to factors that causally and lawfully change conscious experience;

2) after finding NCC instead of asking “which brain processes underlie consciousness”, it can be considered “why these particular neurobiological operations cause conscious experiences”. According to the NCC-strategy (Crick & Koch, 1990; Koch, 2004) one can explain and understand consciousness only if the NCC have been located and these two steps have been taken.

How can the NCC be revealed? As quoted in the introduction, Schrödinger hinted already 50 years ago at a strategy that Bernard Baars brought officially into consciousness research (Baars, 1989): the contrastive analysis. One simply has to compare pairs of events that are similar in all regards but differ in the conscious experience, i.e, events that are associated with conscious experience are contrasted with events that are not associated with conscious experience (see also Crick & Koch, 1990). Presumably, in the former case the NCC are elicited, in the latter not. Thus, comparing the neural activities of these two conditions can help to unravel the NCC. Many experimental paradigms have been developed that allow one to study subjective conscious perception independently of the objective stimulation in the vein of the contrastive analysis (Bachmann, Breitmeyer, & Öğmen, 2011; Kim & Blake, 2005). Although the method of contrastive analysis is a general strategy in brain research and can be applied to other research questions within consciousness research, the method has been most widely used to search for the NCC in studies of (visual) perception (for reviews Dehaene & Changeux, 2011; Rees, Kreiman, & Koch 2002; Tononi & Koch, 2008).

Maybe the most well known experimental paradigm illustrating this principle is the phenomenon of binocular rivalry (for a recent review see Blake & Logothetis, 2002). Despite the fact that different stimuli are presented to each eye, the subject has a conscious experience of only one of them at a given moment and the subjective experience changes spontaneously during the viewing of these two stimuli so that sometimes one image is perceived consciously and sometimes the other. Yet, the stimulation is actually the same during the whole viewing epoch – the eyes "receive" both images all the time (i.e. on the retina the stimulation pattern is the same). What changes is the conscious experience – sometimes the content presented to the left eye is consciously perceived, sometimes not (during the dominance of the stimulus presented to the right eye) (Blake & Logothetis, 2002). If the content presented to the left eye is consciously perceived, the corresponding NCC should be active. At the same time the NCC corresponding to the content presented to the right eye, which is currently suppressed from consciousness, should not be active. Thus, comparing the epochs where the particular content is consciously perceived with those where it is not can help us to unravel the NCC. In addition to binocular rivalry, many other experimental paradigms have been developed that allow one to study subjective conscious perception independently of the objective stimulation, for example visual masking, attentional blink, crowding, inattentional blindness and many others (for reviews see Bachmann, et al., 2011; Kim & Blake, 2005).

In this situation, where the subjective experience alternates under invariant stimulation conditions, another important step has to be made if one wants to unravel the neural processes underlying this subjective experience. Namely, one has to measure the conscious experience itself, one has to "take seriously the introspective phenomenological reports" as Dehaene & Naccache (2001) have put it (see also Varela, 1996). This is at odds with the behaviorist view that one should observe only behavior: in the case of consciousness, measures have to aim at experience. Therefore, in consciousness research the subject reports if he or she consciously perceived the target or not and/or how clear the conscious perception of the target was.

Finally, if one has a suitable experimental paradigm and has collected subjective reports, one ought to gather data about the neural activities to see, what is different in the brain, when the subject reports having a conscious experience of the target compared with the trials where the stimuli are not consciously perceived. Which neural processes correlate with the changes in subjective experience? By considering these neural differences between conscious and non-conscious stimulus events, the NCC have been presumably captured (Crick & Koch, 1990; Koch, 2004).

So, how far is the quest for NCC? Although the first crucial empirical works appeared more than 20 years ago (Logothetis & Schall, 1989), no NCC (in the sense of "minimal set of neuronal events jointly sufficient for a specific conscious experience" (Koch, 2004, p. 16)) has been unraveled until this day. Consider one of the best evidence so far. The work of Nikos Logothetis and colleagues (for review see Logothetis, 1998) has shown that during binocular rivalry the firing of the cells in primary visual cortex was not modulated by the conscious experience of the monkey (Leopold & Logothetis, 1996). In contrast, the firing of the neurons in inferior temporal cortex was stronger when the preferred stimulus was consciously perceived (Sheinberg & Logothetis, 1997). This result implies that the neural correlates of conscious vision probably involve the activity of neurons in the higher visual areas, in this case in inferior temporal cortex, and not in the primary visual cortex. However, this does not mean that the NCC have been captured as this work did not show which cell types are important, in which cortical layers they are located, if any specific type of firing is relevant, if recurrent processing between the different layers or the different brain areas is needed or not, if the neurons have to fire in synchrony or not, how long the activity of the specified neurons should last, where these neurons get their inputs from, where they project to, if and which other areas contribute to the NCC etc. Furthermore, other studies have cast doubt about one of the central claims of the work by Nikos Logothetis and colleagues: namely, several studies with single-cell recordings have observed correlates of conscious perception also in V1 (Fries, Roelfsema, Engel, Konig, & Singer, 1997; Super, Spekreijse, & Lamme, 2001).

Pascal Fries and colleagues recorded multiunit activity from the early visual cortex of strabismic cats under binocular rivalry (Fries, et al., 1997). These authors investigated the idea that perceptual dominance in early visual areas might be achieved by modulation of the synchrony rather than by the rate of neural firing. It was found that in the rivalry condition neurons representing the dominating stimulus increase their synchrony, while neurons processing the suppressed image became less synchronized. Importantly, no effects were found for neural firing rates. This study is one of the cornerstones for arguing that neural synchrony and not the amplitude of neural activity determines the contents of conscious experience (Engel, Fries, Konig, Brecht, & Singer, 1999; Engel & Singer, 2001).

In another revealing study, Super and colleagues (Super, et al., 2001) used a figure-ground segregation task to study the contextual modulation effects in primary visual cortex. The monkeys were trained so that the researchers could infer when the monkey did not perceive the figure. When comparing the trials where monkey did or did not perceive the figure, these authors found no effect at the early transient response but observed a later modulation of V1 firing rates, which has been interpreted as the effect of neural feedback (Lamme, Super, Landman, Roelfsema, & Spekreijse, 2000; Super, et al., 2001). Thus, this study suggests that not any kind of V1 activity is relevant for conscious perception, but it is especially the recurrent feedback to V1 that might be related to the NCC.

In summary, despite the fact that many great research groups have done single-cell and multi-unit studies about the NCC, these works have not unraveled undisputed NCCs – some claim that V1 is not related to conscious perception, others find that neural synchrony but not firing rates in V1 could be reflecting conscious perception and yet others observe that firing rate modulations indeed correlate with the perception of the animal. When it comes to the comprehensive body of human neuroimaging works related to the NCC the picture is even more confusing and controversial.

NCC in human neuroimaging studies

It is parsimonious to expect that for unraveling the NCC one needs to understand the structure and function of specific neural circuits (Crick & Koch, 1990, 2003). In human neuroimaging studies one cannot directly tap on such specific activity, but rather has to deal with the macro-scale signals such as the haemodynamic activity measured by functional magnetic resonance imaging (fMRI) or the electrical currents in the scalp measured by electroencephalography (EEG) or the magnetic fields as captured by magnetoencephalography (MEG). These measured signals either reflect the average activity of millions of neurons (EEG and MEG) or are too slow for capturing the quick information processing in the temporal domain (fMRI), thus these methods simply have no required specificity to directly address the NCC, the minimally sufficient neural mechanisms underlying a particular conscious experience. To understand why, consider watching a movie, where scenes change quickly and continuously – our conscious perception clearly tracks these changes, but this specific activity cannot be recovered from EEG or MEG. With fMRI one can indeed read out and reconstruct the contents of visual perception during movie scenes (Nishimoto, et al., 2011) and even visual dreams (Horikawa, Tamaki, Miyawaki, & Kamitani, 2013), but as information about timing, crucial for information processing in the brain (e.g. Singer, 1993 for an early review), is effectively lost in the BOLD signal, one cannot be looking at the minimally sufficient mechanisms of conscious perception with fMRI either. Thus, the NCC cannot be directly captured with human neuroimaging.

Therefore, the human neuroimaging studies first try to find the macro-scale markers of the NCC – processes that reflect the NCC on the level that can be measured with human neuroimaging. Although it might seem frustrating that another step has to be taken away from the NCC, this strategy is pragmatic – unraveling the macro-scale markers of NCC could help to constrain the possibilities of NCC. In particular, neuroimaging could tell the researcher coarsely where (fMRI) and when (EEG / MEG) one could search for NCC. After that one could "zoom in" (e.g. in the monkey brain) to understand even more precisely when and where NCC are located and the "magnification" process could be repeated all over again.

In 1998 as one of the first seminal studies about the markers of NCC with human neuroimaging, Eric Lumer and colleagues (Lumer, Friston, & Rees, 1998) used binocular rivalry and fMRI to demonstrate that the activity of the frontoparietal networks is involved in the changes of the content of consciousness (i.e. the dominating image in rivalry). In a later study using visual masking (Dehaene, et al., 2001) it was revealed with fMRI that visible words elicited a strong activation of the frontoparietal network that was not observed when the same words were made invisible through masking. These works and many later works (reviewed in Dehaene & Changeux, 2011) support the idea that the activity of the frontoparietal network is the marker of the access to conscious perception. The global neural workspace theory which is based on these findings claims that the activity of the sensory cortices is unconscious or preconscious (Dehaene & Changeux, 2011; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006).

However, there are also other findings in human neuroimaging which suggest that the neural processes associated with conscious perception take place in sensory cortices. For example, Grill-Spector and colleagues (Grill-Spector, Kushnir, Hendler, & Malach, 2000) showed that the fMRI activity in the lateral occipital cortex was directly related to the performance in a visual discrimination task. Later the same group demonstrated that subjective visibility of the visual stimulus is reliably represented in the higher order visual cortex (Hesselmann, Hebart, & Malach, 2011). Also, these researchers have demonstrated that if subjects are fully engaged in a demanding categorization task, the activity in prefrontal cortex is suppressed, although the subjects were presumably fully conscious of the visual stimuli (Goldberg, Harel, & Malach, 2006). Furthermore, it has been shown with intracranial recordings that conscious perception correlates with local gamma band responses in visual cortex (Fisch, et al., 2009). Therefore, these researchers conclude that it is the activity of the sensory cortices that corresponds to the neural correlates of conscious perception (Fisch, et al., 2009; Hesselmann, et al., 2011).

Although the previous two theoretical accounts try to argue that some specific brain areas are related to consciousness, human neuroimaging studies have also demonstrated that unconscious processing activates the same cerebral areas as conscious processing, including prefrontal cortex (Lau & Passingham, 2007; van Gaal, Ridderinkhof, Fahrenfort, Scholte, & Lamme, 2008). Thus, it has been proposed that instead of having special areas related to conscious processing, distributed processes could be dynamically bound together for conscious perception by changing their temporal relationships, i.e. synchrony in time (Singer & Gray, 1995). It has been argued that especially long-distance synchrony across brain areas could satisfy the requirements for being the mechanism of conscious perception (Melloni & Singer, 2010; Varela, Lachaux, Rodriguez, & Martinerie, 2001). Empirical support for this proposal comes from studies with visual masking (Gaillard, et al., 2009; Melloni, et al., 2007), attentional blink (Gross, et al., 2004) and binocular rivalry (Cosmelli, et al., 2004; Doesburg, Green, McDonald, & Ward, 2009; Srinivasan, Russell, Edelman, & Tononi, 1999).

There are also different results and interpretations when it comes to the temporal domain, i.e. to the question when in time after stimulus onset does the activity associated with conscious perception diverge from the activity associated with unconscious processing. Highly influential works from the group of Stanislas Dehaene showed both with attentional blink (Sergent, Baillet, & Dehaene, 2005) and visual masking (Del Cul, Baillet, & Dehaene, 2007) that the marker of consciousness is the late positive deflection of the ERP which only emerges late in time, after 270 ms after stimulus onset. However, the extensive work of Mika Koivisto and Anti Revonsuo (Koivisto, Kainulainen, & Revonsuo, 2009; Koivisto, et al., 2008; Koivisto & Revonsuo, 2003, 2007, 2008) has constantly shown that the earliest process correlating with conscious perception is a negative deflection around 150-250 ms after stimulus onset (reviewed in Koivisto & Revonsuo, 2010). Recently, Melloni and colleagues (Melloni, Schwiedrzik, Muller, Rodriguez, & Singer, 2011) provided an interesting insight to the controversy about the timing of conscious perception. These authors used the hysteresis paradigm, where a stimulus is first made gradually visible for the subject by increasing its contrast and is then gradually faded out of consciousness by decreasing its contrast. Using this paradigm it was shown that the late activity (around 300 ms) only correlates with consciousness in the phase where the stimulus becomes gradually visible but does not correlate anymore after the stimulus has been visible (i.e. the subject knows which stimulus to expect and has a working memory representation of it) and the contrast is decreased until the stimulus fades out of awareness. Enhanced negativity around 200 ms, however, correlated with conscious perception parametrically in both phases.

The above quoted studies show that although the task of locating the macro-scale markers of NCC does not seem difficult, it has not led to a consensus despite more than a decade of work and tens of published results. Even based on coarse fMRI and EEG results it should not be so complicated to decide whether conscious perception is produced in sensory cortices, in the frontoparietal network or whether it requires long-range integration across areas. However, as shown above there is still an ongoing controversy about whether conscious perception arises locally from the activity in the sensory cortices (Lamme, 2006; Malach, 2007; Zeki, 2001), from the global workspace areas in the prefrontal and parietal cortex (Dehaene & Naccache, 2001, Dehaene & Changeux, 2011) or whether experiencing something consciously requires integration across different brain areas (Bachmann, 2007; Melloni & Singer, 2010; Tononi & Koch, 2008). Also, it should be possible to infer from the EEG results when conscious experience arises but there has been an intense debate about the timing of conscious perception with some scholars arguing for rather early correlates around 150-200 ms (e.g. Bachmann, 2000; Koivisto & Revonsuo, 2010; Pins & Ffychte, 2003) and others fiercely claiming that the respective markers can be found not earlier than around 300 ms (Del Cul, et al., 2007; Gaillard, et al., 2009; Sergent, et al., 2005). This thesis will not solve these debates. The humble goal of this thesis is to describe a problem that explains why such controversies persist and to try to make a small step towards solving that problem.

Reason for the controversies: the methods are not specific enough

In this thesis it is argued that the main reason for the failure to find universally accepted markers of NCC is that the experimental methods typically used to study the NCC are actually not specific for NCC and also unravel unconscious processes that precede conscious experience as prerequisites or follow conscious experience as consequences. In other words, contrasting trials with and without conscious perception of a particular target reveals more processes than just the macro-scale markers of the NCC. This means that differences between the studies might stem from differences in these other processes. This problem with the contrastive method is described in detail with thorough examples in chapter 2.

It is believed that taking this problem into account can help to focus the research on finding the NCC. Based on this theoretical problem it is concluded that the experimental paradigms need to go beyond the simple contrast between trials with and without conscious perception, as that contrast will stay unspecific with regard to the NCC and its markers.

A new experimental paradigm: aims and scope of the thesis

For the empirical studies presented in this thesis a new experimental paradigm was designed which allows one to go beyond some aspects of the mentioned methodological problem. In that experimental paradigm conscious perception is modulated either by sensory evidence or by prior knowledge. The logic of this approach is the following: although conscious perception is manipulated through these two different prerequisites (sensory evidence and prior knowledge), conscious experience of the target changes similarly due to these manipulations and thus any proposed correlate of consciousness should change similarly due to these manipulations. In particular, we demonstrate with behavioral studies that the proportion of trials where the subject perceives the target and the quality of that perception increase both by providing more sensory evidence and when prior knowledge is available. Thus, if any proposed neural process should directly reflect conscious experience, it should be changed in a similar fashion both by sensory evidence and by prior knowledge.

This experimental paradigm is used throughout the empirical part of this thesis. As we manipulate the prerequisites of consciousness, it is obvious that this experimental setup is focused on and also limited to tackling the problem of processes that precede conscious perception. In particular, with this paradigm we seek to achieve two goals.

First, the main goal is to use this experimental paradigm to test two proposed correlates of conscious perception. If a proposed correlate indeed reflects the conscious perception, it should be modulated similarly by both sensory evidence and by prior knowledge, given that both factors modulate conscious perception similarly. If, however, the proposed correlate is modulated by only one of these factors, it would mean that the correlate reflects the respective prerequisite process preceding conscious perception and not conscious perception itself. In the current thesis, two such proposed correlates are tested. In chapter 3 we test the hypothesis that local category-specific gamma band responses in the human visual cortex reflect conscious perception (Fisch, et al., 2009). For that we have measured intracranial EEG from the surface of human higher-order visual cortex. In chapter 4 we use this experimental paradigm to test the hypothesis that long-range synchrony is the NCC (Melloni & Singer, 2010; Varela, et al., 2001). This hypothesis required capturing the relevant nodes constituting the synchrony network, which was not possible with the relatively sparse coverage of intracranial electrodes we had in our intracranial study (chapter 3). Therefore, this second hypothesis is tested with MEG.

The second goal of using this experimental paradigm is to gain a better understanding about how prior knowledge affects conscious perception. In our experiments prior knowledge is used as a prerequisite process for modifying access to conscious perception, but at present it is not known how this beneficial effect is implemented in the brain. Prior knowledge could affect neural processing at a fairly local scale or it could rely on feedback signals from higher brain areas (Bar, 2003; Grossberg, 1980; Ullman, 1995). It is of course out of the reach of this thesis to reveal the whole set of respective mechanisms, but in chapter 5 we use this experimental paradigm to ask how quickly prior knowledge affects perception.

Locating the NCC is a very intuitive and pragmatic strategy for approaching the problem of consciousness. Human neuroimaging can constrain the search for the NCC in space and time by providing adequate markers of the NCC. However, this quest has been largely unsuccessful as the experimental paradigms have been too unspecific for narrowing down on the NCC. The work presented in this thesis provides no silver bullet for solving all these problems, but it is hoped that the approach taken here can inspire other people to go beyond the experimental paradigms that are commonly used. Hopefully these new experimental paradigms allow the scientific community to distill the neurobiological machinery underlying conscious experience.

2

Theoretical problems with the contrastive method

As published in: Jaan Aru, Talis Bachmann, Wolf Singer and Lucia Melloni (2012). Distilling the neural correlates of consciousness. Neuroscience Biobehavioral Reviews, 36(2), 737-746.

Contrastive analysis as the key methodological strategy in consciousness research

The problem of consciousness could be approached scientifically only after concrete research strategies were suggested and research problems were defined (Baars, 1989; Crick & Koch, 1990). Since then, an important research goal has been to find the minimal set of neural processes that are together sufficient for the conscious experience of a particular *content* – the neural correlates of consciousness (NCC). This purpose is complementary to the research tradition about general *states* of awareness (e.g., awake, NREM sleep, vegetative state) where consciousness and lack of consciousness are contrasted (Koch, 2004; Laureys & Tononi, 2010). In this paper we deal with the former aspect where NCC are studied in relation to the content of perception as opposed to treating consciousness as a state variable.

When investigating the neural correlates that underlie the content of our consciousness a common and widely accepted strategy has been to hold the stimulus conditions similar while conscious perception varies as the dependent measure – on some trials the subject perceives the target consciously, on other trials not. This experimental approach, known as the contrastive analysis (Baars, 1989), is implemented in various experimental paradigms (Bachmann, et al., 2011; Kim & Blake, 2005) and combined with measures of neural activity has brought about important insights into the NCC (Dehaene & Changeux, 2011; Koivisto & Revonsuo, 2010; Rees, 2007; Tononi & Koch, 2008). The asserted advantage of the contrastive analysis is the following: by comparing conscious and non-conscious perception, the NCC can be unraveled without confounding it with unconscious processes involved in target perception that take place in both conditions. However, despite the apparently straightforward logic of this approach the results are inconclusive and contradictory. For instance, some studies report differences between conscious and non-conscious conditions starting already at stimulus onset or around 100 ms post-stimulus (Aru & Bachmann, 2009a, 2009b; Melloni, et al., 2007; Pins & Ffytche, 2003), others have only found differences in later phases (>300 ms) (Del Cul, et al., 2007; Gaillard, et al., 2009). These contradictory results can be understood when considering a severe methodological pitfall: the result of the contrast between trials with and without conscious perception of a target is not only the NCC proper but could also reflect processes that in a particular experiment paradigm, regularly and lawfully, precede and/or follow conscious perception without directly corresponding to the subjective experience (Bachmann, 2009; de Graaf, Hsieh, & Sack, 2012; Melloni & Singer, 2010). These processes need to be distinguished and disentangled from each other before we can understand the crucial neural mechanisms underlying conscious experience. Figure 2.1 illustrates the development of enthusiastic proposals as well as critical ideas in the science of consciousness. The recent criticism on the contrastive analysis is expected to pave the way for further positive developments and experimental paradigms that will eventually lead us closer to unraveling the neural processes directly associated with conscious experience.

The 3 NCCs

In what follows we recapitulate the two new types of NCCs proposed previously (Bachmann, 2009; de Graaf, et al., 2012; Melloni & Singer, 2010): the prerequisites for and the consequences of conscious perception, in short NCC-pr and NCC-co, respectively. We also highlight arguments that support the existence of these processes confounding the search for the NCC. For the sake of clarity we start with the existing definition of the NCC.

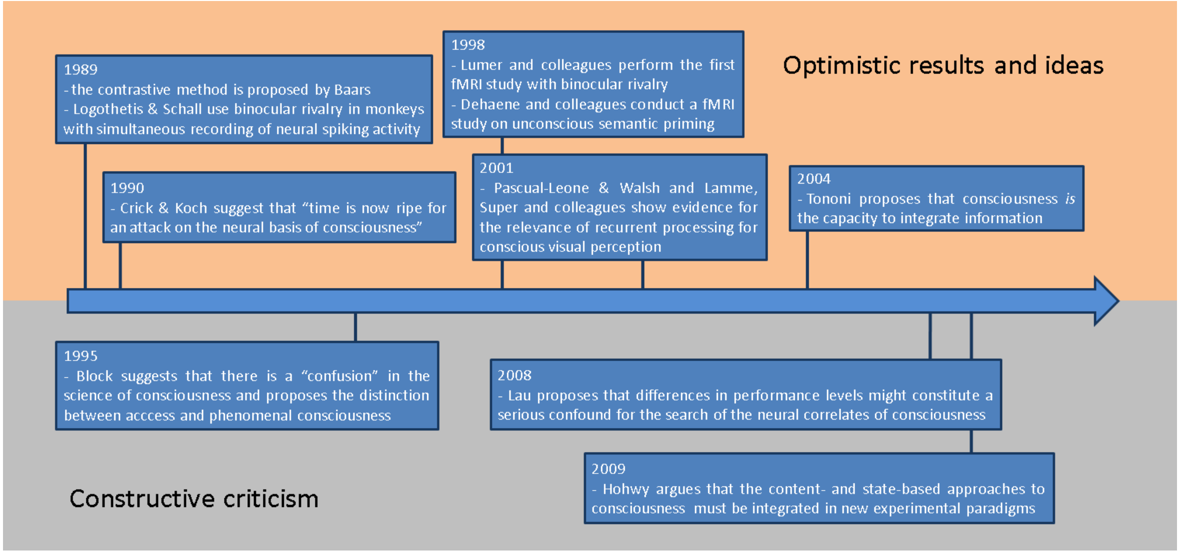


Figure 2.1: A time-line of the major events in the science of consciousness. The time-line illustrates that fresh optimistic theoretical and empirical insights are always followed by critical evaluation of these ideas and results. Therefore, the recognition of the problem regarding the prerequisites for and consequences of NCC is a logical corollary of the popularity of the contrastive method in the science of consciousness. The acknowledgement of the problem should be followed by clever experiments trying to separate the prerequisites for and consequences of conscious perception from the processes that reflect NCC directly. (The selection of the events for the time-line was necessarily subjective and we apologize to all the co-researchers who feel that their work has been neglected.)

NCC for processes directly corresponding to conscious experience

NCC is a neural process that directly corresponds to the phenomenal experience of the target. NCC is the "minimal set of neural events jointly sufficient for a specific conscious experience (given the appropriate enabling conditions)"(Koch, 2004). In other words, if we would stimulate or generate these neural events, a particular conscious experience would happen. Therefore, NCC is the process we need to study in order to understand how conscious experience of a particular content is related to the neuronal processes of the brain. Previously, it has been thought that the contrastive analysis, as commonly used, directly reveals the NCC (Figure 2.2a). Recently, however it has been argued that the result of the contrast between trials with and without conscious perception can also be processes that precede or follow conscious perception (Figure 2.2b).

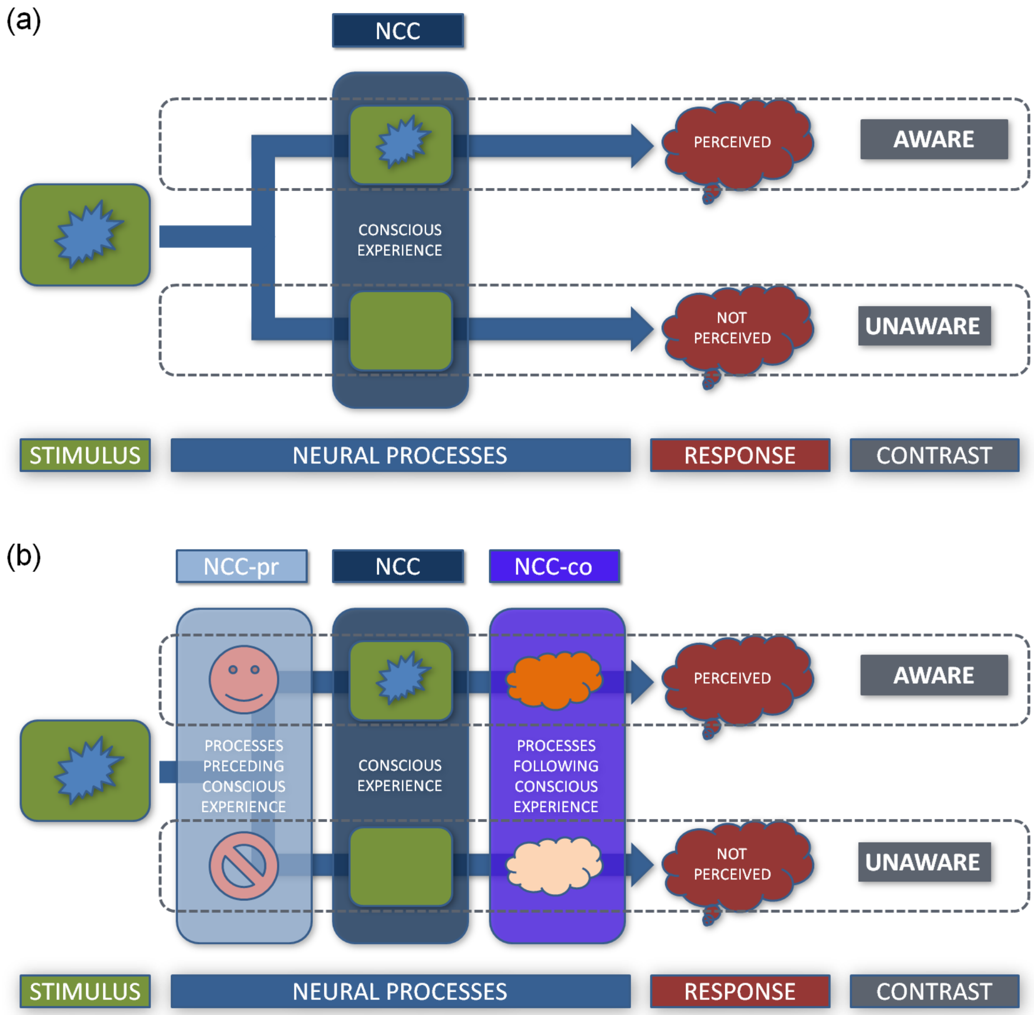


Figure 2.2: Neural processes revealed by the contrastive analysis. In the contrastive analysis, a stimulus (here a blue "splash") is presented so that it sometimes appears in conscious experience and sometimes not. On each trial the subject gives a response for indicating if the stimulus was consciously perceived or not. Neural processes are sorted according to the subjective responses to "aware" and "unaware" conditions, which are compared to each other. What neural processes can be the outcome of this contrast? A) The traditional view on contrastive analysis assumes that only difference between these conditions is the subjective experience of the stimulus. Thus, the contrastive analysis appears suitable for revealing the neural correlates of subjective experience (NCC). B) The proposed interpretation of contrastive analysis. According to this view the outcome

of contrastive analysis consists besides the NCC of two other processes – NCC-pr and NCC-co. NCC-pr corresponds to unconscious processes that though related to conscious perception of a stimulus (e.g., attention that enhances weak information to cross the threshold of consciousness) appear before any subjective experience emerges. NCC-co represents the consequences of consciously perceiving the stimulus. Consequently, if this view is correct, the traditional contrastive analysis by itself cannot exclusively reveal the NCC because it confounds NCC with these other processes that do not directly correspond to the conscious experience of the stimulus.

Prerequisite processes carrying no actual conscious contents (NCC-pr)

These processes are different in the conditions <target in consciousness> vs. <target not in consciousness> and participate in determining in the particular experimental setup if the target appears in consciousness, but are not part of the NCC (Bachmann, 2009; de Graaf, et al., 2012). There are many different processes that can act as NCC-prs. Hence, one particular NCC-pr need not be sufficient or even necessary for generating a conscious experience of the target, but some kind of NCC-pr might be necessary to bring about the very processes of NCC.

In a paradigm with transient stimulation an example of NCC-pr (Figure 2.3a and 2.3b) is the stochastic fluctuations in the excitability of neurons. For long, it has been recognized that the ability to perceive a weak signal fluctuates over time – an effect that has been exploited by the contrastive method to create conscious vs. non-conscious conditions. However, only recently it was shown that fluctuations of ongoing brain activity, as indexed by the phase of pre-stimulus oscillations, systematically determine these behavioral dynamics (Mathewson, et al., 2009; Busch, et al., 2009; Monto, Palva, Voipio, & Palva, 2008). The phase of ongoing alpha (10 Hz) oscillations also determines the effectiveness with which a single TMS pulse elicited a phosphene: with a 15% increase in the likelihood of a perceptual outcome (phosphene) between opposite phases (Dugue, Marque, & VanRullen, 2011). This effect was observed over an extended time period (~400 ms), showing that processes correlating with conscious perception of the target clearly precede the experience itself. As all this takes place in the pre-stimulus interval (being it either before stimulus presentation or TMS pulse application), where no conscious experience of the actual target could have emerged, it constitutes a NCC-pr. Other studies relating pre-stimulus activity to subjective reports show that NCC-pr can reflect spontaneous excitability as also indexed by oscillatory power (Ergenoglu, et al., 2004; Linkenkaer-Hansen, Nikulin, Palva, Ilmoniemi, & Palva, 2004; Romei, et al., 2008; van Dijk, Schoffelen, Oostenveld, & Jensen, 2008) and that the NCC-pr can be linked to attention (Thut, Nietzel, Brandt, & Pascual-Leone, 2006; Worden, Foxe, Wang, & Simpson, 2000; Wyart & Tallon-Baudry, 2009), decision bias (Wyart & Tallon-Baudry, 2009) and potentially other processes.

In the experimental paradigms with sustained epochs of perception (Figure 2.3c), NCC-prs can manifest themselves in various ways. For example, for binocular rivalry it was proposed more than a century ago and shown recently (Alais, Cass, O'Shea, & Blake, 2010) that adaptation leading to weakening of reciprocal inhibition determines the alternations between competing stimuli. More precisely, neurons coding for the dominant stimulus adapt over time, which in turn weakens the inhibition of the suppressed stimulus, increasing its neuronal responses and thus bringing that stimulus into consciousness. Importantly, (reciprocal) inhibition could be seen as NCC-pr, as it contributes to which target will be consciously perceived and is different between the two contrastive conditions (a particular target in consciousness vs. not in consciousness). However, it is not part of the neural processes sufficient for generating or maintaining conscious experience of that target.

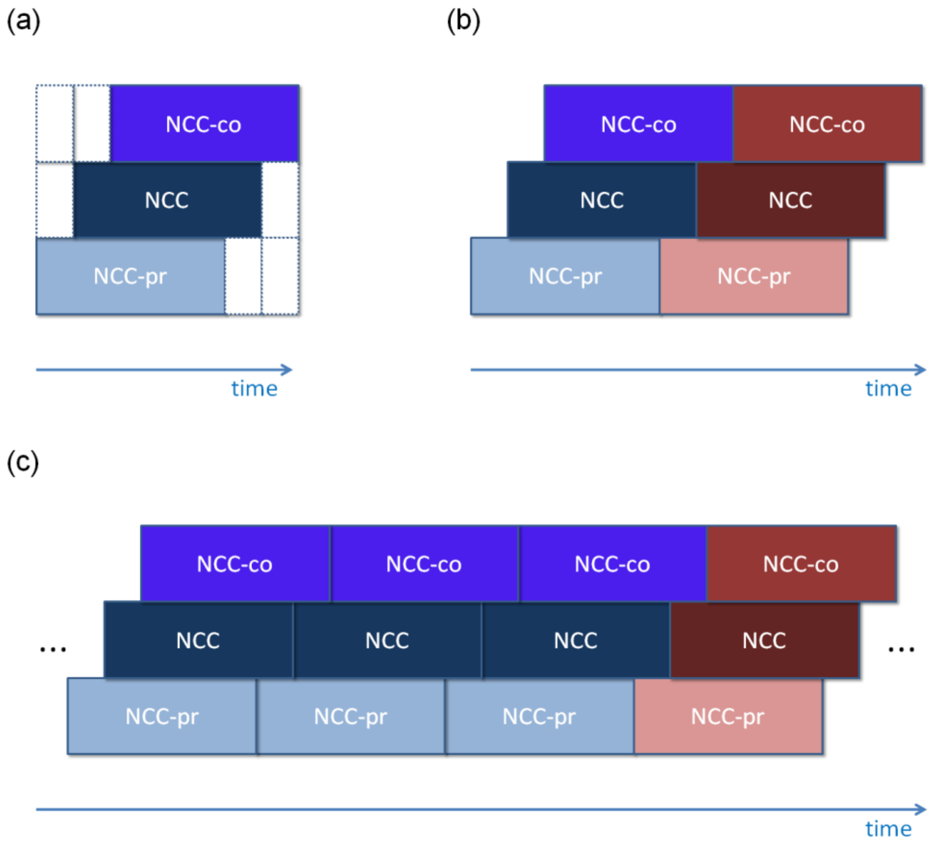


Figure 2.3: Possible temporal relationships between NCC, NCC-pr and NCC-co in various experimental situations. A) If a single transient stimulus is presented, NCC-pr, NCC and NCC-co are evoked. Areas with dotted outlines indicate that the temporal differences and overlap between the 3 NCCs is currently unknown. This scenario corresponds to experimental paradigms like threshold stimulation. B) If two (or more) transient stimuli are presented sequentially, the different NCCs of these different stimuli can be overlapping in time. This scenario corresponds to experimental paradigms like masking or attentional blink. C) In experimental situations with longer epochs of perception, each epoch can be understood as a succession of the transient events (as in A), where the NCC of time point t can be overlapping with NCC-co from time point t-1 and NCC-pr from time point t+1. This scenario corresponds to experimental paradigms like binocular rivalry or motion induced blindness. Similar ideas on the temporal overlap of the processes unfolding within pre-conscious and conscious stages of percept formation have been also proposed previously (Brown, 1988).

Consequences following actual conscious experience (NCC-co)

These processes appear as after-effects when NCC is or just has been present, but are not a part of the minimally sufficient mechanisms of the conscious experience of the target. These processes are the consequences of conscious perception (de Graaf, et al., 2012; Melloni & Singer, 2010). The existence of such processes is a logical consequence of assigning any function to conscious perception – if conscious perception enables certain processes that unconscious perception does not, these processes would inevitably appear in the contrast between trials with and without conscious perception, even if they are solely the consequences and not the direct correlates of consciousness. Importantly, most theories of consciousness do confer a function to consciousness in the chain of information processing (Seth, 2009), for instance sustained maintenance of information, access to long-term memory, novel combinations of operations, and intentional behavior.

For example, it is known that neurons in the medial temporal lobe (MTL) respond in all-or-none fashion, closely following the subjective report of the patient (Quiroga, et al., 2008). The fidelity with which MTL neurons follow the subjective visibility of the stimuli is so high that conscious versus non-conscious trials can be distinguished solely based on the neurons' firing rate: either the neuron responds when the subject reports to have consciously recognized the picture, or stay completely silent when the image did not reach consciousness. However, damage to the MTL-system (or even its complete resection) does not affect moment-to-moment conscious perception (Crick & Koch, 1990; Postle, 2009), but only the formation of a memory trace. That is, subjects will continue to have subjective experience but will have no memory of it. Therefore, such all-or-none responses in MTL, even though closely following the subject's report, do not correspond to the

NCC but are instead part of the NCC-co reflecting processes related to memory consolidation.

In addition, NCC-co processes can also reflect differences in performance. Usually, conditions in which subjects are conscious of the stimuli correlate with higher performance (for instance higher detectability or discriminability) than when subjects are not conscious of the stimuli. Thus, it has been claimed that a part of the brain activity observed using the contrast between conscious and non-conscious perception pertains to differences in performance instead of conscious experience (Lau, 2008). Direct evidence for this claim comes from a recent neuroimaging study (Lau & Passingham, 2006) reporting that when performance was equalized between conditions, so that only subjective experience varied, only dorsolateral prefrontal cortex distinguished conscious from unconscious conditions as opposed to the extensive fronto-parietal network typically reported in previous studies (Dehaene & Naccache, 2001). Recent studies in which performance was assessed alongside with subjective experience further confirm that these two processes can indeed dissociate in time, space and, importantly, in neural locus (Hesselmann, et al., 2011; Lamy, Salti, & Bar-Haim, 2009; Schwiedrzik, Singer, & Melloni, 2011). This highlights that performance and subjective experience are not interchangeable and that to investigate NCC proper, performance has to be controlled for.

The 3 NCCs: conclusions

Thus, despite the principal methodological importance of the contrastive analysis, the method, as commonly used, seems to lack the required specificity to unravel the neuronal processes exclusively related to the subjective experience. In fact, simply contrasting conscious to non-conscious conditions could also reflect processes preceding or following the NCC. The problem exists in all experimental paradigms currently used to investigate consciousness (Figure 2.3), regardless of whether differences in conscious perception result from internal switches in brain states, such as in binocular rivalry, or from external manipulations of visibility, as for example achieved by varying the SOA in masking experiments. It is conceivable that the distinction between the 3 NCCs also applies to studies in which different states of consciousness are contrasted (e.g. deep sleep vs. waking or vegetative state vs. minimally conscious states), as these conditions do not exclusively differ in the consciousness state (which is the intended manipulation), but could also do so in their prerequisite processes and especially in the consequences (in the conscious state people think, access long-term memory, have intentions etc.). It will be necessary to determine if and how much the 3 NCC problem applies to other areas of research where conscious experience is investigated (e.g. studies of conscious intentions, introspection, imagery, dreams, inner speech etc).

This conjecture implies that previous results based on the contrast between trials with and without conscious perception of a particular content cannot be unequivocally interpreted as reflecting the NCC, as they could also represent other processes. Given that the majority of studies investigating consciousness have used such contrasts (Dehaene & Changeux, 2011; Koivisto & Revonsuo, 2010; Rees, 2007; Rees, Kreiman, & Koch, 2002; Tononi & Koch, 2008) such reinterpretation might have a major impact on consciousness research and on current theories on the neural mechanisms of consciousness. As long as the NCC proper cannot be clearly dissociated from these other processes, we should be cautious when relating any experimental finding about neural processes correlating with conscious experience to the NCC. Therefore, if we want to proceed in understanding consciousness, the NCC has to be experimentally distilled from the prerequisites for and the consequences of conscious perception. As mentioned above, the problem of 3 NCCs may also apply to other areas in the study of conscious experience besides the investigation of conscious perception. Here we focus on the latter as it constitutes the best studied aspect of conscious experience. However, it might prove fruitful to apply the concept of 3 NCCs also to these other aspects of conscious experience as alternative solutions to disentangle the 3 NCC conundrum might emerge from that research.

Distilling consciousness: disentangling the NCCs

There is probably no single experiment with today's research techniques that would yield a clear separation between NCC and NCC-pr or NCC and NCC-co. As discussed above, the contrast between trials with and without conscious perception of a particular content does not dissociate these processes. In what follows, we propose research strategies to separate the NCC from the NCC-prs and NCC-cos (see also de Graaf, et al., 2012; Melloni & Singer, 2010). Success in understanding the neuronal processes directly underlying conscious experience rests upon distinguishing and disentangling its confounds.

Disentangling NCC-pr from NCC

There is no question that processes differentiating conscious from non-conscious stimuli already before stimuli onset correspond to NCC-pr. Beyond that, difficulties begin. As NCC-pr can potentially appear after stimulus onset but also concurrently with NCC, there is no simple time criterion that separates NCC-pr from NCC. Furthermore, an interesting theoretical but also empirically challenging problem arises when considering that the NCC-pr do not only have to appear before the NCC but can also stay active during the NCC (Figure 2.3). If NCC-pr is there only to "ignite" NCC by preceding it, the empirical study and measurement of NCC is tractable because later epochs of target experience in long-duration target stimuli are not confounded with NCC-pr. However, we currently lack the required knowledge about the NCC-prs to make any claims about their duration or temporal structure.

The most straightforward way to disentangle NCC-pr from NCC is to directly manipulate the NCC-pr processes, and compare the neural signatures that are common to all of them. We assume that different NCC-pr would elicit distinct neural activities, while neural process directly involved in consciousness would be invariantly present in all conditions. As an example, consider a "consciousness task" in which a stimulus under identical stimulation conditions (e.g., a masking experiment with invariant masks and SOAs) is sometimes consciously perceived and sometimes not. As discussed above, this approach would lead to the problem of 3 NCCs. However, now, in an additional step, in the same experimental setup with the same stimuli, we could vary potential NCC-prs such as for instance stimulus expectation, adaptation, working memory or allocation of attention independently. Specifically, in one condition stimuli could be brought to consciousness through expectation (Melloni, et al., 2011) and in another through attention (Wyart & Tallon-Baudry, 2008). Which neural processes resulting from the contrast <consciously perceived> vs. <not consciously perceived> are *similar* and which ones are *different* in these two conditions? Neural signatures that differ between conditions should belong to the NCC-pr; neural signatures common to both comparisons are likely to be related to the NCC. Using such an approach, it was recently shown that short-latency event-related potentials previously related to conscious perception (Pins & Ffytche, 2003) most likely represent signatures of the

NCC-pr (Melloni, et al., 2011) as opposed to NCC. Contrasting attention with expectations might be particularly revealing, as these processes are proposed to have opposite effects on neural activity – attention increases sensory responses whereas expectation decreases them (Summerfield & Egner, 2009). Thus, neural processes that increase with attention but decrease with expectation under similar subjective experience and objective performance are unlikely to reflect the NCC proper.

In addition to cognitive factors, one could affect NCC-pr in a similar vein with TMS. In a recent experiment it was observed that when the visual cortex was stimulated 100-120 ms after a near-threshold visual stimulus with TMS intensity below the phosphene threshold, the thresholds for explicit perception of the visual stimulus were decreased (Abrahamyan, Clifford, Arabzadeh, & Harris, 2011). In line with the experimental strategies proposed above, one could combine this TMS-related improvement of perception with an independent cognitive manipulation (e.g. of attention or expectation) that also has a beneficial effect on perception and measure the neural correlates of conscious perception with the same logic as presented previously: only neural signatures related to perceptual enhancement common to both manipulations are likely to be related to the NCC.

Disentangling NCC-co from NCC

As with NCC-pr, one cannot straightforwardly rely on a strict temporal distinction between NCC and NCC-co. Probably correlates that appear 2 seconds after the onset of a 30 millisecond stimulus indeed represent NCC-co, however, the critical question refers to neural correlates appearing earlier than 500 ms post-stimulus as the bulk of studies have related those to NCC (Del Cul, et al., 2007; Fisch, et al., 2009; Gaillard, et al., 2009; Koivisto, et al., 2008; Melloni, et al., 2011; Sergent, et al., 2005). Neurons in the MTL have a latency of about 300ms MTL (Quiroga, et al., 2008). If we are right in stating that the activity of MTL neurons, even when apparently closely following the subjective report, is not the NCC but constitutes the NCC-co, it becomes evident that the NCC-co can be present already in a time window that is often investigated for the NCC.

According to the definition, the NCC-co corresponds to the after-effects of conscious perception. Thus, it is logically possible to observe NCC even when NCC-co is not elicited. In the normal brain such a dissociation might not occur very often; however, lesion studies might be revealing: if an area can be lesioned or removed without any effect on conscious perception, then activation of this area can be regarded as NCC-co. This approach allowed, for instance, ruling out activity in MTL as a direct NCC (Crick & Koch, 1990; Postle, 2009). Importantly, the tests applied to identify NCC-cos have to be sensitive enough to capture subtle effects on conscious perception. For instance, although the view that prefrontal cortex (PFC) is causally relevant for conscious perception (Dehaene & Changeux, 2011) could at first have been rejected based on lesion studies (Pollen, 1999), it has gained support from carefully conducted studies exploiting the masking paradigm in patients with PFC lesions (Del Cul, Dehaene, Reyes, Bravo, & Slachevsky, 2009). Nonetheless, the fact that the psychophysical threshold for conscious access is elevated in patients with PFC lesions cannot be taken as direct evidence of PFC having a central role in conscious perception. Instead, PFC could provide top-down support for either strengthening or maintaining the (cognitive results of the) conscious percept (Gazzaley, et al., 2007; B. T. Miller, Vytlacil, Fegen, Pradhan, & D'Esposito, 2011). This is a likely possibility as patients could perform the difficult masking task relatively well and were reporting conscious perception despite their PFC lesion.

Lesion studies can be complemented with interventional techniques in healthy subjects, such as TMS (see also de Graaf, et al., 2012) and/or transcranial direct or alternating current stimulation (tDCS or tACS). This approach offers several advantages over lesion studies. First, it allows for comparisons within the same individual, as TMS can be applied in an on-off manner. Second, due to the transient nature of the intervention, profound reorganization of brain networks is not to be expected. This allows dissociating effects related to lesioning a particular area from alterations at the level of the network resulting from compensatory plasticity. Third, virtual lesions can be made with a precision not comparable even when performing probabilistic lesion mapping studies in patients. Furthermore, virtual lesions can be made not only of a single region, but also in several regions at the same or at varying times. Combined with neuroimaging, one could first find a neural signature X that correlates with conscious perception in the contrastive analysis (e.g. the activity in the frontal cortex) and then in a second step test whether perturbation of this signature X (e.g. by TMS) has an effect on conscious perception. If such perturbation produces no effects on conscious perception, the respective process or area is a prime candidate for NCC-co. However, to firmly conclude that, evidence has to be provided that such manipulation has a behavioral effect. Otherwise, it is unclear whether the effect was not observed because the perturbation is simply ineffective, or it specifically does not affect conscious perception. The feasibility of such selective manipulation has been recently shown: theta-burst TMS to bilateral dorsolateral PFC interfered with metacognitive aspects of visual awareness but not with discrimination performance (Rounis, Maniscalco, Rothwell, Passingham, & Lau, 2010).

Besides interventional approaches, much progress could be made by (i) obtaining comprehensive data on the sequence of cognitive events distinguishing conscious from unconscious processing, assuming that causes precede effects, and by (ii) considering theoretical approaches which clearly disclose some cognitive processes (e.g. working memory) as dependent on conscious perception (Baars, 1989). This is particularly relevant as it allows for contrasts in which not only consciousness is manipulated (e.g., aware vs. unaware trials) but also its consequences (e.g., conditions in which encoding in working memory is present or absent). Following such an approach, it was recently shown that late electrical signatures of consciousness, in particular the P3 event-related potential typically associated with conscious access (Dehaene & Changeux, 2011), does not follow conscious perception when subjects already have a conscious working memory representation of the target stimulus (Melloni, et al., 2011). This result points to the tantalizing possibility that late waves of EEG activity such as the P3 might reflect NCC-cos, and not the NCC itself.

The 3 NCCs and previous results

In light of the ideas proposed here, an important question is whether previous findings reflect NCC or rather NCC-pr, NCC-co or even a compound of these different processes. Since little is known about the NCC-pr and NCC-co and how they interact with the NCC proper, it is currently not straightforward to determine which results indeed relate to NCC and which ones do not. A first step is to identify and recognize the neuronal signatures of processes that might *not* reflect the NCC.

Importantly, many current theories of consciousness are founded on empirical findings, assuming that these genuinely reflect the NCC. As shown, this assumption might not hold true. Thus, it is possible that some theories about the NCC are actually based on NCC-prs or NCC-cos. Clarifying this issue will hopefully lead to better agreement among theories of consciousness. Note that we do not suggest that all previous findings are necessarily wrong or doubtable, nor do we claim that all theories are based on the wrong assumptions about the previous results; we simply point to the worrisome possibility that follows the theoretical distinction between the 3 NCCs. In the next paragraphs we link this theoretical problem to existing results from the contrastive analysis with the hope that the recognition of the 3 NCC problem will contribute constructively to the disputes about the timing and neuroanatomical locus of conscious perception.

The 3 NCCs and the temporal course of conscious perception

The distinction between the three NCCs could prove valuable to resolve some controversies regarding the question whether conscious perception happens early (Aru & Bachmann, 2009b; Melloni, et al., 2007; Pins & Ffytche, 2003) or late (Del Cul, et al., 2007; Gaillard, et al., 2009; Sergent, et al., 2005), as some of the studies observing early correlates might reflect NCC-pr while those reporting late correlates might reflect NCC-co. In a recent study, such an approach was purposely followed: the processes related to NCC-pr and NCC-co were manipulated in an attempt to more directly address neuronal processes related to the NCC itself (Melloni, et al., 2011). To investigate NCC-prs, Melloni and colleagues (2011) examined how sensory evidence and top-down expectations, respectively, influence the threshold of awareness, and whether the two factors modulate brain activity differently. This allowed contrasting brain states with and without expectations and with perceived and non-perceived stimuli for identical stimulation conditions. To investigate NCC-cos, conditions based on whether the target was or was not present in working memory were compared. As in previous studies (Mathewson, et al., 2009; Pins & Ffytche, 2003) it appeared that in the absence of expectations an early event-related potential (ERP), already present around 100 ms in occipital sensors, distinguished conscious from unconscious trials (Figure 2.4). Perception under those circumstances is known to depend on bottom-up stimulation (the amount of sensory evidence) and stochastic fluctuations in the prestimulus period which jointly determine whether a target will or will not be consciously perceived (Busch, et al., 2009; Mathewson, et al., 2009). However, in the presence of expectations, that is, when perception depends not only on bottom-up information but also on top-down expectations, this early component was not different between the consciously perceived and not perceived targets. As both conditions resulted in similar rates of visibility, this finding suggests that early processes differentiating seen from unseen stimuli - previously interpreted as a NCC - were actually not directly related to subjective experience, but reflect differences in NCC-pr (Figure 2.4). Along the same lines, a late ERP (P300) distinguishing consciously perceived from not consciously perceived trials was only observed in trials where the target had to be encoded in working memory but not when subjects already had a conscious working memory representation of the target. This suggests that late effects like the P300 can index processes related to the NCC-co instead of the NCC proper under certain conditions (Figure 2.4). At threshold, the component most consistently related to conscious perception was the P200, arising at about 200 ms over occipitoparietal sensors. This study also revealed that the electrophysiological signatures of conscious perception are not bound to processes with a strict latency, but depend on how consciousness comes about: earlier electrophysiological signatures were observed in the presence of top-down expectations (P200) than in their absence (P300). This further complicates the search of the NCC, as they might change in time depending on which NCC-pr determines perception.

Converging evidence for the proposed temporal relationships between NCC-pr, NCC and NCC-co comes from studies with intermittent binocular rivalry. In this paradigm, epochs of binocular rivalry stimulation alternate with blank intervals, which allows to study neural processes locked to the stimulus onset and to delineate the sequence of neural events related to bistable perception. Interestingly, Pitts & Britz (2011) when reviewing recent studies employing such methodology come to similar three time windows related to conscious perception as we did based on a different experimental paradigm described above (Melloni, et al., 2011): 1) early processes around 130-160 ms after stimulus onset that vary with conscious perception but probably constitute a preconscious state (in the current terms the NCC-pr), as although this component varies with the subjective percept (Pitts, Martinez, & Hillyard, 2010), the stable perceptual representation of the target has not emerged yet (Pitts & Britz, 2011), 2) the "reversal negativity" around 200-300 ms that is a "primary candidate for a neural correlate of awareness" (Pitts & Britz, 2011), as there the stable representation has been established and 3) the late positive complex around 400-600 ms that might reflect further processing of the perceptual information or processes related to working-memory maintenance (NCC-co in the current terms).

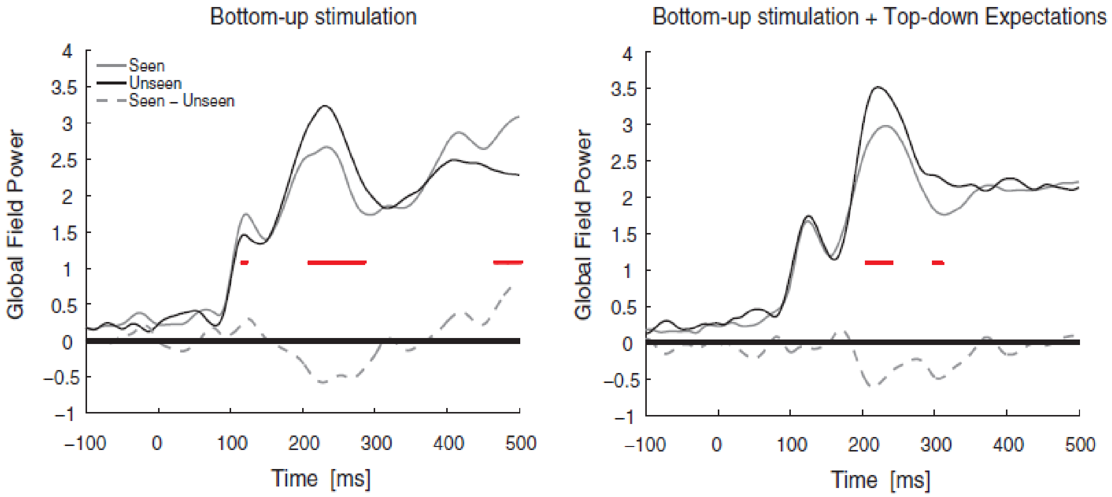


Figure 2.4: Neural signatures of conscious perception: NCC-pr, NCC and NCC-co. In the study of Melloni et al. (2011), both the prerequisites and the after-effects of conscious perception were manipulated. Subjects had to rate the visibility of a stimulus embedded in noise. Visibility could either arise because of the strength of sensory evidence

(left panel) or because of the presence of perceptual expectations (right panel). In these conditions, the prerequisites (NCC-pr) differed, but conscious perception was identical. Furthermore, a conscious working memory representation of the target could already be present (left panel) or absent (right panel). Here, conscious perception was again equalized, but the consequences (NCC-co) differed, since in the latter case, a new representation needed to be encoded in working memory. The red bars indicate the periods of significant difference between the aware and unaware trials. The only component differentiating consciously perceived from consciously not perceived trials independently of differences in NCC-pr or NCC-co was a mid latency component, likely reflecting neuronal processes directly involved in the NCC. In contrast, early processes differentiating seen from unseen stimuli were not present when consciousness resulted from an interaction between sensory evidence and top-down expectations. Also, the late ERP component was only present when a working memory representation had to be established. The fact that some neural signatures of consciousness disappeared when either the prerequisites or consequences of consciousness were manipulated questions the general involvement of these processes in conscious perception.

The 3 NCCs and the activity of different brain areas

Regarding the neural locus of consciousness it is also debated whether conscious perception is associated with activity in sensory areas (Hesselmann, et al., 2011; Lamme, 2006; Zeki, 2001) or whether higher non-sensory areas are directly involved in conscious perception (Dehaene, et al., 2006; Lau & Rosenthal, 2011). In this context, several previous studies are directly relevant for understanding whether particular brain areas such as the prefrontal cortex serve a crucial role in conscious perception, as proposed in some theories (Dehaene & Naccache, 2001; Lau, 2008), or whether they reflect other cognitive processes, such as top-down control, report, or performance on a task. The available evidence relating PFC activity with consciousness has been obtained by contrasting conscious vs. non-conscious trials (Dehaene & Changeux, 2011). The key message of our proposal is that by solely relying on the contrast between trials with and without conscious perception of a particular content it is undetermined whether such activation indeed constitutes an NCC or rather a consequence of conscious perception (NCC-co). In line with this conjecture, when subjects do not need to report the stimuli, differential activity between conscious and non-conscious stimuli is not observed in the prefrontal cortex but only in occipital visual areas downstream of V1/V2 (Tse, Martinez-Conde, Schlegel, & Macknik, 2005). In addition, activity in prefrontal cortex is suppressed when subjects are engaged in a demanding perceptual categorization task, although they are presumably fully conscious of their rapidly changing visual or auditory world during that time (Goldberg, et al., 2006). Finally, although TMS over frontal cortex affected voluntary control of bistable perception, TMS had no influence on passive bistable perception (de Graaf, de Jong, Goebel, van Ee, & Sack, 2011), which conforms nicely with the approach proposed above: TMS had a measurable effect on one process (voluntary control) but not on conscious perception. These results directly support the thesis that prefrontal activity might not be a part of the NCC but could rather represent processes that follow conscious perception as consequences.

In another relevant study, Eriksson and colleagues (Eriksson, Larsson, & Nyberg, 2008) found that prefrontal activity and coupling between prefrontal and occipital areas decreased after training in an identification task. This result might imply that at least part of the prefrontal activity observed in contrastive analysis could be attributable to the fact that in such paradigms, conscious recognition is often made difficult (e.g. by masking or by lowering the contrast of the stimuli) and PFC might be recruited to aid the recognition process under poor sensory evidence (Eriksson, et al., 2008). From this perspective PFC activity could constitute a NCC-pr, as prefrontal activity would determine whether the target reaches consciousness or not. This idea agrees with theories and empirical evidence suggesting that short-latency prefrontal activity facilitates perception (Bar, 2003; Bar, et al., 2006).

Alternatively, it could be argued that PFC activity represents NCC-co (instead or in addition to the NCC-pr) that change as a function of perceptual training. In particular, extensive experience with the stimulus could lead to established memory representations such that there is no need to create working memory representations on each trial, leading to decreased prefrontal activity with training.

All these results challenge the notion that prefrontal cortex is directly involved in visual awareness but suggest instead that it could reflect executive functions, top-down facilitation and/or top-down control under conditions of increased effort. As a number of theories of consciousness (Crick & Koch, 2003; Dehaene & Changeux, 2011; Dehaene & Naccache, 2001; Lau, 2008) assert that activation of prefrontal cortex is part of the NCC, it is clear that the question whether PFC activity is indeed a NCC or rather a NCC-pr or NCC-co is central for further research on the neural correlates of consciousness.

Similar caveats hold true for the involvement of parietal cortex in consciousness. In fact, when using the contrastive method in visual versus auditory tasks, it was found that activity in parietal cortex is correlated with visual but not with auditory conscious perception (Eriksson, Larsson, Ahlstrom, & Nyberg, 2007), questioning its general, modality-independent role in conscious perception. Even further, it is known from the neurological literature that bilateral lesions of the parietal cortex do not abolish conscious perception: although patients exhibiting such lesions are only able to perceive one object at a time (simultagnosia), this one object is nevertheless consciously experienced (Friedman-Hill, Robertson, & Treisman, 1995; Robertson, 2003). These results suggest that the activity of the parietal cortex might not be a part of the NCC. As substantial empirical evidence links the activity of the parietal cortex to perceptual alternations in paradigms with multistable perception (Kleinschmidt, Buchel, Zeki, & Frackowiak, 1998; Lumer, et al., 1998), it is likely that the processes in the parietal cortex constitute an essential prerequisite for NCC. This claim is substantiated by recent evidence that TMS to parietal cortex affects the dominance durations during binocular rivalry (Carmel, Walsh, Lavie, & Rees, 2010; Zaretskaya, Thielscher, Logothetis, & Bartels, 2010) and that the structure of the parietal cortex is correlated with the intraindividual differences in the perceptual alternation rate of a bistable stimulus (Kanai, Bahrami, & Rees, 2010).

Finally, there is a long-standing controversy about the role of primary visual cortex (V1) in conscious perception. Even if V1 lesions lead to loss of awareness in the corresponding parts of the visual field, it constitutes no strong argument for the NCC being in V1, as V1 lesions also disrupt information flow to higher order visual areas. Empirical and theoretical arguments in the mid-1990s suggested that V1 might not be necessary for visual conscious experience (Barbur, Watson, Frackowiak, & Zeki, 1993; Crick & Koch, 1995) but further research and theoretical positions generally favored the idea that cortical feedback to V1 is indeed necessary for visual conscious perception and therefore part of the NCC (Lamme, 2001; Pascual-Leone & Walsh, 2001; Tong, 2003). We believe that the division between the three types of NCC can be fruitful in solving such controversies, as it offers the possibility that although V1 activity correlates with conscious perception (Tong, 2003), it might constitute a NCC-pr rather than be a part of the NCC. Indeed, a recent experiment revisited the issue whether patients with V1 lesion lack visual consciousness (Ffytche & Zeki, 2011) and contrary to what is commonly believed the authors showed that these patients do have visual experiences. Furthermore, clinical studies have shown that conscious vision can recover after a V1 lesion (Silvanto & Rees, 2011). These results imply that neither activity in V1 nor the cortical feedback to V1 is necessary for visual conscious perception. In the current terminology, intact V1 would be an important prerequisite for visual conscious experience, but not part of the NCC. However, it is possible that V1 is not necessary for *all* visual experiences but those involving the *fine-tuned* spatial contrast distribution in the experienced image due to the functional capabilities of V1. Thus, whether some brain area is necessarily involved in NCC may depend not only on its neuroanatomical locus, but also on the specific details of the experience (Haynes, 2009).

Taken together, the tripartite distinction of NCCs calls for a reappraisal of previous results and theories, but also highlights the need of further studies that put emphasis on investigating and distinguishing the pre-conscious determinants of conscious perception (NCC-pr) and its consequences (NCC-co) from the NCC proper.

Conclusions

Despite its original simplicity and appeal, the traditional method, contrasting trials with and without conscious perception of a particular target, by itself does not appear to have the necessary specificity to reveal the NCC. Instead, processes that precede (NCC-pr) and follow (NCC-co) conscious perception are confounded with the NCC proper. Here we relied on the theoretical distinction between those processes (Bachmann, 2009; de Graaf, et al., 2012; Melloni & Singer, 2010) and suggested tentative experimental paradigms aimed at disentangling these processes from each other.

An important way to circumvent the specificity problem altogether is to move away from mere correlates and focus the search on mechanisms of conscious experience, as correlates which do not have explanatory power can be excluded a priori (Melloni & Singer, 2010). However, the key to success in unraveling the NCC will ultimately lie in combining such mechanistic models that generate testable predictions (Bachmann, 2007; Dehaene & Changeux, 2011; Lamme, 2006; Melloni & Singer, 2010) with suitable experimental paradigms. Only then we will be capable of putting theoretical proposals to the test. We thus hope that future work, adopting the distinction between NCC, NCC-pr and NCC-co and following the empirical strategies proposed here, will allow us to re-assess previous studies and theories, focus on the NCC, and distill the crucial neural processes that underlie our conscious experience.

3

Local category-specific gamma band responses in the visual cortex do not reflect conscious perception

As published in: Jaan Aru, Nikolai Axmacher, Anne T. Do Lam, Jürgen Fell, Christian E. Elger, Wolf Singer and Lucia Melloni (2012). Local category-specific gamma band responses in the visual cortex do not reflect conscious perception. Journal of Neuroscience, 32(43), 14909-14914.

Introduction

One central controversy in consciousness research is whether conscious perception arises locally from activity in sensory cortices (Malach, 2007; Zeki, 2001) or whether it requires integration across brain areas (Dehaene & Changeux, 2011; Melloni & Singer, 2010; Tononi & Koch, 2008). Empirical support for the idea that conscious perception is reflected in activity of sensory areas mostly stems from neuroimaging studies (Grill-Spector, et al., 2000; Hesselmann, et al., 2011), which are insensitive to the neural dynamics underlying the perceptual effects. Recently, however, pivotal support for a role of sensory cortices in conscious perception was gathered with intracranial EEG recordings that are precise in time and space. Fisch and colleagues (2009) showed that category-specific gamma band responses (GBR) in the visual cortex correlate with conscious perception. Conscious recognition led to an "ignition-like" enhancement of the GBR, supporting the idea that conscious perception arises locally within sensory cortices. However, other findings suggest that conscious perception is grounded on the coherent activation of widely distributed cortical networks, whereas local GBR reflects local processing that can happen unconsciously (Melloni & Singer, 2010). This idea is consistent with the classic findings showing strong local stimulus-specific gamma band activity even under anesthesia (Gray & Singer, 1989).

This obvious discrepancy might be resolved when considering a recently uncovered theoretical problem which states that the methods commonly used for studying the NCC are not specific for conscious experience but may reflect processes that precede or follow it (Aru, Bachmann, Singer, & Melloni, 2012; de Graaf, et al., 2012). Thus, an alternative explanation for the findings of Fisch et al. (2009) is that category-specific GBRs in sensory cortices reflect local processing which is a necessary prerequisite for but not the direct correlate of conscious perception. In other words, in a paradigm such as used by Fisch et al. (2009) it cannot be ruled out that the enhanced GBR just reflected more effective preprocessing of stimuli which enhanced the chances of these stimuli to get perceived consciously. It is possible that only for this reason local processing correlated with conscious perception.

To directly test this conjecture, we manipulated conscious perception in two different ways: either by changing the amount of sensory evidence, which directly modulates local processing, or by previously exposing pictures before the test phase (i.e., providing prior knowledge), which is a factor less dependent on local processing. As both factors enhance conscious perception to the same extent, a clear prediction is that if category-specific GBR in visual cortex reflect conscious perception, they should be modulated both by sensory evidence and prior knowledge. Instead, we found that GBR were only affected by sensory evidence, directly contradicting the proposal that local GBR in sensory cortices represent a direct neural correlate of consciousness.

Methods

Subjects

Six patients with pharmacoresistant epilepsy (three female, mean age 27, range 18-35, 5 right-handed) undergoing invasive monitoring for localization of epileptogenic foci participated in the study. For presurgical screening, all patients had lateral and basal strip electrodes (AD-Tech, 4-16 stainless steel contacts, 4 mm diameter, 10 mm inter-electrode spacing) covering the occipital and temporal visual areas. Recordings were performed at the Department of Epileptology, University of Bonn, Germany. The study was approved by the local ethics committee, and written informed consent was obtained from all patients. Electrode locations were determined based on individual MRIs (Table 3.1; see Figure 3.2 for an example). In total 62 electrodes free of epileptic activity were examined. In addition, twenty-four healthy human subjects (seventeen female, mean age 24, range 21-28, all right-handed) served as controls in a psychophysical experiment. All control subjects had normal or corrected-to-normal vision, reported no history of neurological or psychiatric disease, gave written informed consent and received monetary compensation for their participation.

Procedure and Stimuli

Stimuli were displayed to the patients on a CRT monitor and to the healthy controls on a translucent screen (both had a 60 Hz refresh rate) onto which the stimuli were projected from a liquid crystal display projector. Stimulus presentation and response collection were controlled by Presentation software (Neurobehavioral Systems, Inc.).

Stimuli consisted of 148 (for healthy controls: 276) gray-scale pictures containing a single person in the foreground with diverse backgrounds, and a set of 22 catch images (for healthy controls: 54) only containing background. To limit image visibility random noise was added parametrically (Figure 3.1a) while keeping contrast constant across degradation levels. Stimuli were edited in Matlab (R2008b, The MathWorks) using custom code. The noise levels yielding decreased visibility were determined in a pilot experiment and ranged from 60% to 90% in 5% steps. Stimuli were foveally presented at the center of the screen, spanned 4 x 3 degrees of visual angle in the horizontal and vertical plane (for healthy controls: 6 x

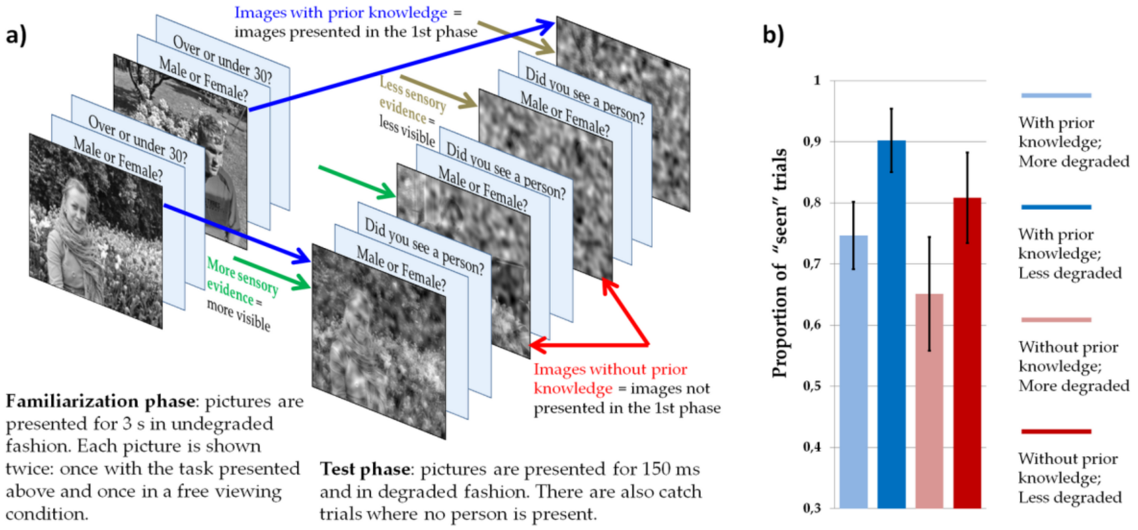
4.5 degrees of visual angle), and were surrounded by a gray background.

Prior to the experiment, each participant undertook a threshold experiment to determine two neighboring degradation levels that yielded a recognition performance around 70% in the male/female task. To that end, 60 degraded images were briefly presented (150 ms) at 2 different degradation levels. The individually determined degradation levels were subsequently used in the main experiment.

The main experiment consisted of 11 experimental blocks. Each block comprised two phases: a familiarization phase and a test phase (Figure 3.1a). Different set of images was used per block. The familiarization phase aimed at establishing prior knowledge by exposing half of the pictures subsequently shown in the test phase. Pictures without noise were presented twice for 3 seconds and subjects were asked to memorize them. To assure attention to and to facilitate encoding of the images, subjects indicated via button press first the gender of the person on the picture (male/female task) and then guessed their age (older or younger than 30 years). Subsequently, pictures were presented without an explicit task, and subjects had to freely explore and memorize them.

In the test phase, degraded images were presented briefly, thus limiting their visibility. Two experimental factors controlled the images' visibility: sensory evidence and prior knowledge (Figure 3.1a). Sensory evidence was varied by presenting images at two degradation levels (high and low noise). prior knowledge was varied by either showing previously exposed or new images, for which no pre-existing memory had been established. Additionally, to assess the reliability of the subjects' judgment, specifically for the subjective visibility task (see below), we included catch trials that only contained background on the higher degradation level. The low number of false alarms in the 44 catch trials (mean 1.5, SD 2.3) ensures that the subjects' visibility judgments were reliable.

Each trial consisted of a fixation cross on a gray background presented for a random duration (1.2-1.4 seconds) followed by the degraded image (150 ms). After 1 second, an objective and a subjective task were presented consecutively. In the objective task, subjects indicated the gender of the person in the picture (male/female judgment). In the subjective task, subjects reported whether they had seen a person in the picture. Occasionally, subjects were also asked to indicate whether the picture, now shown in the degraded fashion, had been presented in the familiarization phase. A block lasted 3-4 min intermixed with breaks between blocks. 44 trials per condition were acquired. In the control experiment, 108 trials per condition were acquired in 27 blocks.



Familiarization phase: pictures are presented for 3 s in undegraded fashion. Each picture is shown twice: once with the task presented above and once in a free viewing condition.

Test phase: pictures are presented for 150 ms and in degraded fashion. There are also catch trials where no person is present.

Figure 3.1: Experimental paradigm and behavioral results. A) Each block consisted of 2 phases: In the first phase half of the images are exposed. In the second phase pictures are degraded and shown briefly. Degraded pictures from phase 1 are presented together with new pictures (manipulation of prior knowledge). Pictures are also shown at two different degradation levels (manipulation of sensory evidence). **B**) Proportion of "seen" person trials in the four experimental conditions. Error bars indicate SEM.

To rule out any picture-specific effects in the neural measures, we pseudo-randomized the pictures containing a person such that across 4 patients each picture was assigned to every condition exactly 1 time. Although we report data from 6 patients, results are comparable when restricting the analysis to 4 subjects with a complete randomization.

Recording and analysis

Recordings were performed with Harmonie recording software (Stellate Systems, Inc), amplified (Schwarzer GmbH), sampled at 1 KHz with an analog filter (0.1-300 Hz), referenced to linked mastoids, and stored for off-line analysis. Data were analyzed using the open source Matlab toolbox Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). Preprocessing included line noise and harmonics (50, 100, 150, 200 and 250 Hz) removal by means of discrete Fourier transform, high-pass filtering (0.5 Hz [24 dB/octave]), and re-referencing to a bipolar montage. Continuous EEG was segmented into 2 sec long epochs starting 1 sec before the onset of the degraded stimuli. Segments were visually inspected for artifacts (e.g., epileptiform spikes), and channels near the epileptic focus, without any signal, or showing extensive artifacts were removed.

We analyzed gamma band response (GBR) in two frequency ranges: Low (30-70 Hz, lGBR) and high GBR (70-150 Hz, hGBR). Fisch et al (2009) had reported that low GBR (30-70 Hz) in visual cortex correlates with conscious recognition. We thus tested the same frequency range to allow direct comparisons between studies. We extended the analysis to high GBR (70-150 Hz) as recent work suggests that gamma band responses in ECoG data has a broad band profile beginning around 50 Hz with the strongest responses at frequencies higher than 70 Hz (e.g. Lachaux, et al., 2005). For each trial, bipolar derivations were analyzed in the time-frequency domain by convolution with complex Gaussian Morlet wavelets with a bandwidth parameter $f_0/\sigma_f=6$, i.e., wavelets of approximately six cycles length. Time-frequency power values were normalized per trial by dividing them by the average power across the baseline (-400 to -100 ms pre-stimulus). Subsequently, we separately averaged the relative power over the frequency ranges 30-70 Hz (lGBR) and 70-150 Hz (hGBR) and obtained two-dimensional signals of time and mean power.

Category-specific electrodes were determined following a procedure similar to Fisch et al. (2009). An electrode was considered person-selective if significantly higher GBR were observed for pictures containing a person in the foreground (from the condition low sensory evidence and without prior knowledge) than for those only containing the background (catch images), in the interval 100-500 ms post stimulus presentation. Importantly, this comparison was neither confounded by prior knowledge (as none of the picture groups had been seen before) nor by sensory evidence (as both picture groups shared the same level of degradation), enabling the extraction of person-specific signatures independently of the experimental factors. Figure 3.2 shows the person-specific power spectra obtained by averaging raw power values in the time-domain over the analyzed time window 100-500 ms post-stimulus.

Statistical analysis

Our analysis focuses on the subjective task as our goal was to address the neural profile that correlates with subjective reports about conscious perception. The percentage of “seen” responses in the subjective task was analyzed with repeated measures ANOVA with factors prior knowledge (previously exposed vs. novel) and sensory evidence (high vs. low degradation). For the electrophysiological data we restricted our analysis to the main effects (sensory evidence, prior knowledge) as the behavioral results consistently revealed no interaction (see results section), which enabled us to increase statistical power by pooling trials.

We also directly investigated power modulations in relation to subjective visibility by contrasting trials where subjects reported having perceived a person with those in which they did not. We restricted this analysis to the condition with lower sensory evidence, as it contained comparable and relatively high amounts of both “seen” and “unseen” person trials. We extracted the same amount of “seen” and “unseen” person trials from the conditions with and without prior knowledge, such that taken together the conscious recognition contrast was independent of both experimental factors.

Statistical analysis were performed at the single electrode level (Fisch, et al., 2009; Lachaux, et al., 2005) whereby the single-trial time courses of the lGBR and hGBR were subjected to a non-parametric Wilcoxon rank-sum test contrasting factors: sensory evidence, prior knowledge, and visibility. A population analysis was also performed whereby responses were averaged per condition, per electrode, and experimental factors were compared across the electrodes in Table 1. To limit the number of statistical tests, activity was averaged over 50 ms windows (no overlap) and compared over the period 100 - 500 ms after stimulus onset (8 points). Results were corrected for multiple comparisons using false discovery rate across all time windows and electrodes.

Post-hoc power analysis was done with Monte Carlo simulations. Trials were randomly assigned into two groups; 30% of the mean of the original distribution was added to one group to mimic the shift due to an experimental factor followed by a Wilcoxon test between the groups. This was repeated 10000 times. Here power corresponds to the proportion of comparisons the Wilcoxon test correctly identifies as "different".

Results

Behavioral results

In the group of 6 patients, sensory evidence and prior knowledge both modulated the visibility of the persons on the pictures, i.e., subjects reported to perceive the person on the degraded picture more often when pictures had either lower degradation (F(1,5) = 25.626, p = 0.004) or when subjects were previously exposed to the clear version of the picture (F(1,5) = 7.463, p = 0.041) (Figure 3.1b). To assess whether sensory evidence and prior knowledge enhance conscious perception of the person to a comparable degree, we contrasted the increase in “seen” responses resulting from changes in sensory evidence (% perceived low degradation - % perceived high degradation) with the corresponding increases resulting from prior knowledge (% perceived with prior knowledge - % perceived without prior knowledge). Both factors boosted subjective reports similarly (T(5) = 1.510, p = 0.191), which is also corroborated by the absence of an interaction of these factors (p > 0.9). We confirmed these results in a large sample of healthy subjects. Here, both prior knowledge (F(1,23) = 70.167, p = 1.93E-08) and sensory evidence (F(1,23) = 69.109, p = 2.21E-08) had a significant effect on subjective perception. Importantly, even with a larger sample we did not observe an interaction between both factors (p > 0.4).

We further confirmed that prior knowledge and sensory evidence similarly boost perception in an additional experiment (n=8) in which we used a fine grained 11-point perceptual rating scale (between “no perception of the person” and “clear perception of the person”). Even with this fine-scale perceptual task both prior knowledge (F(1,7) = 53.072, p = 1.65E-04) and sensory evidence (F(1,7) = 23.526, p = 0.002) modulated perception similarly (t test on the difference of the effects: T(7) = 0.237, p = 0.891).

The main question of the current work was whether the perceptual enhancement through sensory evidence and prior knowledge is reflected in corresponding increases of GBR. If category-specific GBR correlate with conscious experience, GBR should be affected both by sensory evidence and by prior knowledge as both modulate subjective reports about conscious perception.

Category-specific gamma band responses (GBR) in the visual cortex

We restricted our analysis to electrodes on the lateral (30) and ventral (32) surface of the occipital and temporal cortex for which extensive coverage existed. Importantly, GBR in those higher-order visual areas have been associated with conscious perception (Fisch, et al., 2009). First, we determined those electrodes showing selective GBR responses to the person on the picture (Fisch, et al., 2009). Person-specific enhancement of hGBR was observed on 7 electrodes belonging to 5 different patients, which all were localized on the surface of the lateral occipital cortex and fusiform gyrus, in good agreement with the localization of the body-selective areas of the human brain (Peelen & Downing, 2007) (Table 3.1). In the lGBR range, only one electrode showed person-specific effects that survived the correction for multiple comparisons. However, electrodes with (corrected) effects in hGBR mostly had congruent uncorrected effects in the lGBR (Table 3.1). Thus, hGBR and lGBR may reflect a similar process, which is stronger in higher frequencies. To confirm this possibility we analyzed the spectral changes over the whole gamma frequency band. Figure 3.2 shows that the differences between

		MNI coordinates			Category selectivity		Effects of conscious perception		Effects of sensory evidence		Effects of previous exposure	
Electrode	Patient	*x*	*y*	*z*	High GBR	Low GBR	High GBR	Low GBR	High GBR	Low GBR	High GBR	Low GBR
1	TM	51	−75	7	2.6E−8*	3.0E−3	3.4E−4*	1.1E−2	1.3E−6*	1.9E−3*	4.2E−2	1.5E−1
2	TM	57	−68	3	1.6E−8*	2.2E−2	3.7E−5*	4.0E−3	1.4E−6*	1.1E−4*	1.5E−1	9.2E−2
3	AM	54	−68	−7	1.2E−9*	8.6E−6*	2.9E−1	1.1E−1	3.4E−3*	1.6E−1	1.0E−1	5.5E−2
4	FC	52	−73	17	4.0E−4*	3.7E−1	6.5E−3*	4.4E−2	1.4E−1	1.4E−1	7.7E−2	2.1E−1
5	TS	44	−42	−24	4.1E−9*	9.0E−4	5.2E−4*	2.7E−1	7.2E−4*	3.8E−2	2.0E−1	5.1E−2
6	TS	−52	−75	−3	6.3E−5*	5.3E−4	1.4E−1	2.0E−1	1.3E−3*	8.5E−2	2.6E−1	1.8E−1
7	SH	−42	−36	−25	2.5E−5*	1.2E−1	2.8E−3*	7.8E−3	5.3E−3*	1.9E−2	1.6E−1	9.0E−2

Table 3.1: MNI coordinates and response properties of the category-specific electrodes. The response properties reflect the lowest p value of the 8 non-overlapping time windows between 100-500 ms post-stimulus. Asterisks indicate corrected effects across electrodes/times.

pictures containing a person and those with only background started at ~50 Hz and continued throughout higher frequencies, suggesting that lGBR and hGBR are part of the same broadband process. Interestingly, Fisch et al. (2009) also reported effects starting at 50 Hz that continued to the highest analyzed frequency (70 Hz). As there was no visible indication of attenuation in power in Fisch et al., it is likely that higher frequencies exhibited a similar response. We thus focus on the hGBR as the effects are stronger there. However, all effects were similar but smaller in the lGBR (see Table 3.1).

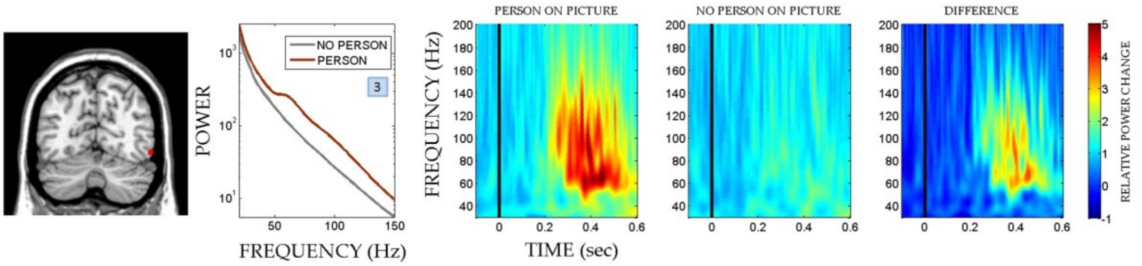


Figure 3.2: Location, spectral profile and time-frequency representation of an example electrode with person-specific responses in the human visual cortex. An example electrode (electrode 3 in Table 3.1 providing the MNI coordinates) exhibiting person-specific responses overlaid on a coronal view of patient's MRI. Person-specific electrodes were determined by contrasting conditions that included a person in the image with conditions that only included a natural background. Selective responses were observed around the body-selective areas of the human visual cortex. As evident from the power spectra and time-frequency plots, the effect is visible in a broad range of gamma frequencies starting around 50 Hz.

Sensory evidence but not prior knowledge modulates category-specific GBR

In line with Fisch et al. (2009) we observed a strong and sustained GBR that outlasted the stimulus presentation (150 ms) when subjects reported having perceived the person on the picture. This was observed in 70% (5/7) of the person-specific electrodes in 4 different patients. Figure 3.3a shows data from three exemplary electrodes. We then asked whether those responses are similarly modulated by sensory evidence and prior knowledge, as both factors enhance subjective perception of the persons.

We observed that hGBR was significantly higher when sensory evidence was stronger (Figure 3.3b). However, no difference in hGBR was observed between pictures with and without prior knowledge, despite the fact that prior knowledge enhanced conscious perception similarly to sensory evidence. hGBR was significantly modulated by sensory evidence in 85% (6/7) of the person-specific electrodes in 4 different patients, but none of them showed an effect of prior knowledge (Figure 3.3c). Table 1 shows that only one electrode (nr 1) barely approached the uncorrected level of significance for prior knowledge, while all other 5 electrodes with strong effects of sensory evidence exhibited uncorrected p-values > 0.1. These results were confirmed by the population analysis, where we only observed effects of sensory evidence ($p < 0.05$, 100 to 450 ms post-stimulus, except for 150-200 ms, where $p = 0.053$) but not of prior knowledge ($p > 0.2$ for all time windows). Power analysis revealed that the mean power of our test over the 7 electrodes and 8 time windows was 0.94 and 0.9 for a 30% change at the uncorrected and corrected level, respectively, confirming the robustness of our methods to detect a potential effect of prior knowledge.

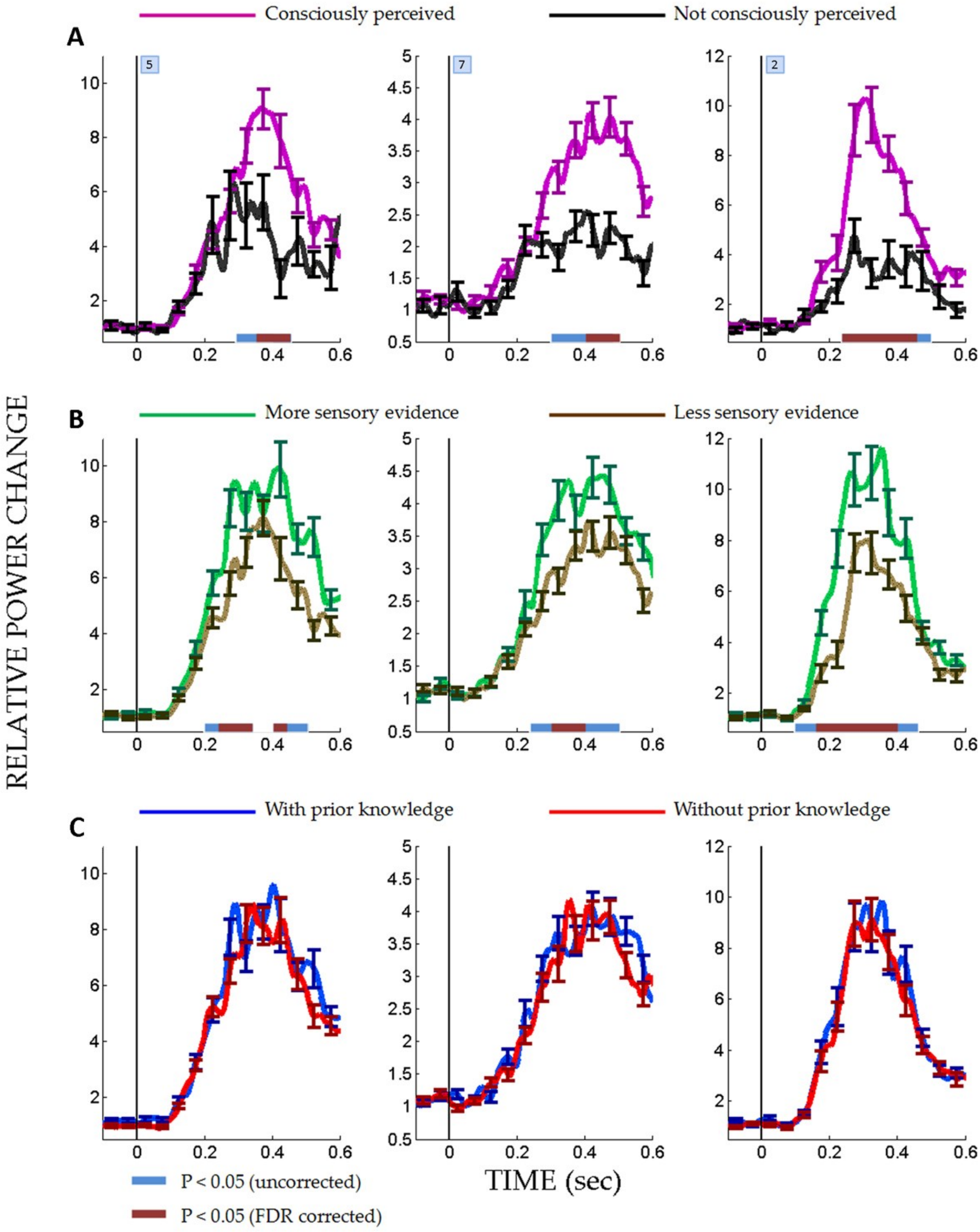


Figure 3.3: Time course of high-gamma band responses in human visual cortex. A) Trials where subjects reported to have consciously perceived the persons showed higher GBR. **B)** In the same electrodes a reliable modulation by sensory evidence is observed, i.e., trials where sensory evidence is stronger also show enhanced GBR. **C)** However, prior knowledge has no effect on GBR despite that subjective reports about conscious perception are increased to the same extent by sensory evidence and prior knowledge. Numbers refer to the electrode numbers in Table 3.1, which provides MNI coordinates and response properties. Error bars indicate SEM.

Discussion

We first demonstrated that conscious perception is similarly enhanced by increasing the amount of sensory evidence and by previously exposing the pictures. We argue that both factors indeed enhance perception rather than later memory processes, as our subjective task was about perception ("did you perceive the person on the picture") and we replicated our results when we used an 11-point perceptual rating scale. Establishing that both factors boost subjective reports about conscious perception to the same extent allowed us to directly examine whether category-specific GBR in visual cortex reflect conscious perception or unconscious local processing. We observed a clear effect of sensory evidence on GBR, with higher amplitudes when sensory evidence led to a higher proportion of seen trials. However, although prior knowledge also enhanced perception and the proportion of seen trials just like sensory evidence, it had no effect on the GBR (Figure 3.3c). The possibility that sensory evidence is simply more effective can be excluded as both factors had numerically similar effects with comparable gains in conscious perception. Moreover, even in those subjects where prior knowledge had a numerically stronger impact on perception than sensory evidence GBR was exclusively modulated by the latter.

Another possibility is that prior knowledge leads to reduced neural responses because of repetition suppression, masking a simultaneous enhancement of GBR. However, repetition suppression is typically observed only when the same, undegraded image is repeated, while repetition leads to enhancement when degraded images are repeated (Muller, Strumpf, Scholz, Baier, & Melloni, 2013). This is thought to be due to increased extraction of information. Thus, we should have observed enhanced responses as we employed degraded images, which was not the case. Taken together, perception appears to be modulated equally by prior knowledge and sensory evidence, but this effect manifests itself in the local GBR only in the case of sensory evidence. It is this lack of invariance of the local GBR which makes it unlikely that GBR reflect a mechanism directly responsible for conscious perception.

What could be the reasons for the discrepancy between our results and those obtained by Fisch and colleagues (2009) from which these authors concluded that local category-specific GBR are markers of conscious perception? It has been remarked previously that the contrastive method, i.e., the direct comparison between trials with and without conscious recognition does not exclusively reflect neural processes directly and uniquely associated with conscious experience of the target but also those that precede or follow conscious experience (Aru, Bachmann, et al., 2012; de Graaf, et al., 2012). In fact, early studies have shown that such local stimulus-specific increases in gamma band activity are observed even under anesthesia (Gray & Singer, 1989) and might thus reflect low-level unconscious processing that precedes conscious experience. In typical experimental paradigms differences in local unconscious processing might lead to different outcomes regarding conscious perception, and therefore, specific local responses can correlate with conscious perception without necessarily being direct correlates of conscious experience.

Numerous contemporary theories of consciousness propose that conscious experience is not related to the depth of local cortical processing but is grounded in long-range integration across brain areas (Dehaene & Changeux, 2011; Melloni & Singer, 2010; Tononi & Koch, 2008). It remains an open question whether sensory evidence and prior knowledge similarly enhance inter-areal integration (Gaillard, et al., 2009; Melloni, et al., 2007). Due to limitations in the electrode coverage we were unable to test this prediction in the current study. Until the relationship between long-range integration and conscious perception has not been disputed by experimental paradigms like ours, such inter-areal interactions remain a viable candidate for mediating conscious perception.

By going beyond the contrastive analysis we could directly test pivotal predictions arising from the experimental work of others (Fisch, et al., 2009) and dissect neural processes that precede consciousness from those that are unique for consciousness. We hope that this result encourages researchers in the field to develop novel experimental paradigms which will help to distill the neural events proper that underlie our conscious experience.

4

Testing the role of long-distance synchrony in conscious perception

Introduction

It has been proposed that long-range neural synchrony between the relevant neural populations is the key neural correlate of consciousness (Engel, et al., 1999; Llinas, Ribary, Contreras, & Pedroarena, 1998; Melloni & Singer, 2010; Singer, 1998). There is also evidence that long-range neural synchrony indeed correlates with conscious experience (Gaillard, et al., 2009; Melloni, et al., 2007).

However, recently it has been argued that the methods commonly used for investigating the NCC are not specific for conscious experience but may reflect unconscious processes that precede or follow it (Aru, Bachmann, et al., 2012; de Graaf, et al., 2012). Thus, previous results have to be re-evaluated in light of this methodological problem. In this thesis we developed a new experimental paradigm to move beyond the simple contrast between trials with and without conscious perception (Chapter 3: Aru, Axmacher, et al., 2012). In that experimental setup we manipulated conscious perception in two different ways: either by changing the amount of sensory evidence or by providing prior knowledge. As both factors increase conscious perception, any neural process that corresponds to conscious experience should change similarly with both manipulations. We used this experimental paradigm to test the hypothesis that local gamma band responses (GBR) are a direct correlate of conscious perception as had been claimed by previous works applying a contrast between trials with and without conscious perception (Fisch, et al., 2009). We observed that GBR only increased together with sensory evidence, although prior knowledge enhanced conscious perception in a similar fashion. Therefore, these results contradicted the idea that local GBR reflects conscious perception (Aru, Axmacher, et al., 2012).

Here we apply the same experimental paradigm and the same logic to test the theory that conscious perception is associated with long-range neural synchrony. Our experimental hypothesis is clear: if long-range synchrony is the mechanism responsible for conscious perception, it should increase under both manipulations, both with prior knowledge and with increased sensory evidence. To test this hypothesis, we estimated the activity of neural sources through beamforming, which allowed us to analyze synchrony in the source space (Hipp, Engel, & Siegel, 2011)

Our study had four specific goals. First, we wanted to assess whether we can replicate with MEG the local power effects obtained at the person-selective cortical areas with ECoG in chapter 3. Second, our aim was to analyze local power effects all over the brain. In our intracranial study (chapter 3) we had limited coverage of the visual cortex but the MEG beamforming approach provides in principle the possibility to investigate local power changes all over the brain. In our intracranial study we had observed effects of sensory evidence at the category-specific cortical regions. However, it could be that prior knowledge leads to power increases in some other areas that were not included in our previous study. Third, we wanted to assess with which areas the category-specific nodes synchronize. This is a goal which we hoped to achieve with our intracranial study (chapter 3) but could not due to the limited coverage of the electrodes. The key idea is that information about the persons on the picture computed locally should be propagated across the brain through long-distance synchrony. To investigate this, we computed synchrony between the category-specific node and each other source all over the brain. Finally, we wanted to estimate synchrony between all source pairs. Although category-specific nodes should be included in the synchrony networks many other sources and their dynamical interactions might be relevant for consciously perceiving the person on the picture.

Methods

Subjects

26 subjects (9 male, 17 female) participated in the experiment. The data of two male participants was excluded due to measurement problems and extensive blink artifacts. The age of the remaining 24 subjects ranged from 21 to 28 years (mean age 24.4 years, standard deviation 2.4 years). All subjects were right-handed and had normal or corrected to normal vision, and no history of neurological or psychiatric disorders. The study was conducted in accordance with the Declaration of Helsinki and all subjects gave written informed consent and received a payment of 15 Euros per hour for their participation.

Stimuli and procedure

The stimuli and the procedure are similar to the ones reported in chapter 3 (Aru, Axmacher, et al., 2012) but are here repeated for the reader's convenience. Stimuli were presented on a translucent screen (60 Hz refresh rate) onto which the stimuli were projected from a liquid crystal display projector (located outside the magnetically shielded MEG room) via two front-silvered mirrors inside the MEG. Stimulus presentation and response collection were controlled by Presentation software (Neurobehavioral systems, Inc).

Stimuli consisted of 276 gray-scale pictures containing a single person in the foreground with diverse backgrounds. In addition, a set of 54 catch images was also included, which shared a similar background but contained no person in the foreground. In order to decrease images' visibility random noise was added parametrically to all images (a reminder can be seen on Figure 4.1). Contrast was kept constant for all stimulus degradation levels. Stimuli were edited with custom code using Matlab (R2008b, The MathWorks). The noise level values that yielded decreased visibility were determined in a pilot experiment and ranged from 60% to 90% in 5% steps. Stimuli were displayed at the center of the screen, spanning 6 x 4.5 degrees of visual angle in the horizontal and vertical plane, surrounded by a gray background.

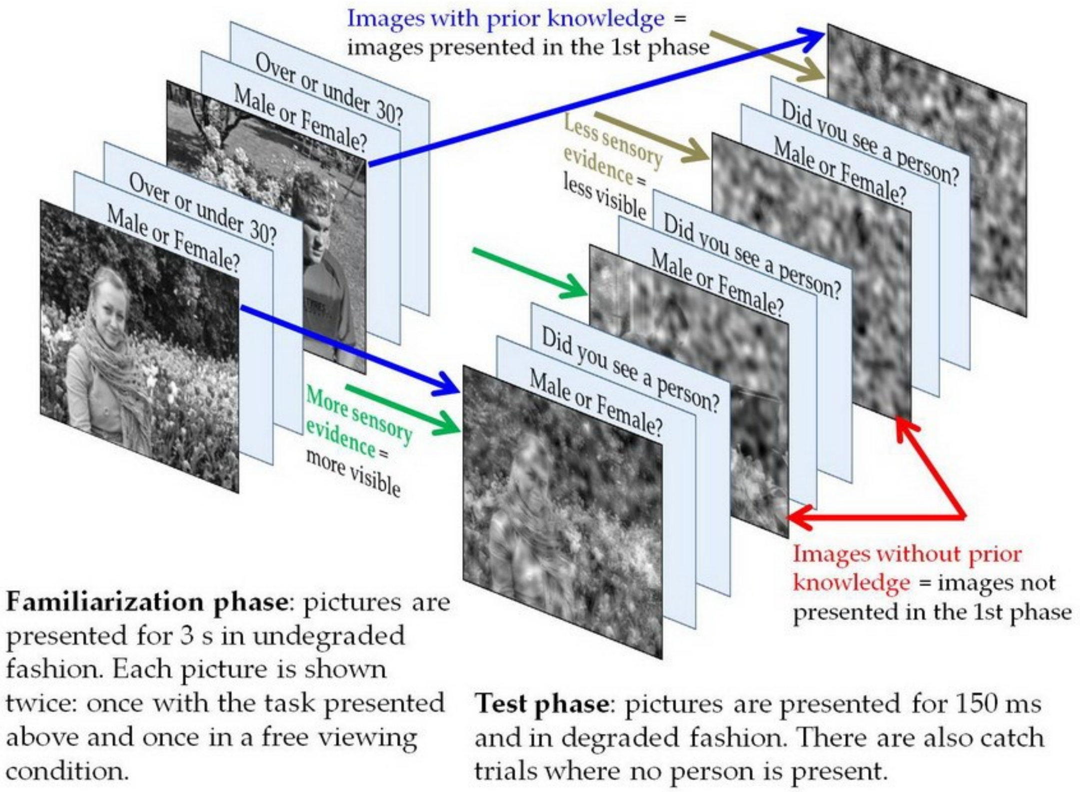


Figure 4.1: The experimental paradigm (the same as in chapter 3). Each block consisted of 2 phases: In the first phase half of the images are exposed. In the second phase pictures are degraded and shown briefly. Degraded pictures from phase 1 are presented together with new pictures (manipulation of prior knowledge). Pictures are also shown at two different degradation levels (manipulation of sensory evidence).

Each subject underwent two consecutive tests: a threshold experiment and a main experiment. The threshold experiment aimed at determining per subject two neighboring degradation levels (e.g., 80% noise and 75% noise) that yielded an optimal recognition performance around 70% in the male/female task. In the threshold experiment, 60 degraded images were briefly presented (150 ms) in randomized order at 2 different degradation levels. Those individually determined degradation levels were used for the main experiment. To avoid any effect of familiarization, new gray-scale pictures were used for the main experiment.

The main experiment consisted of 27 experimental blocks. Each block comprised two phases: a familiarization phase and a test phase (Figure 4.1). A different set of images was used per block. The familiarization phase aimed at establishing prior knowledge in half of the pictures subsequently shown in the test phase. To that end, four pictures without noise were presented twice for 3 seconds. Subjects were asked to commit those pictures to memory. To assure attention and encoding of the images, during the first presentation subjects were asked to indicate via button press first the gender of the person on the picture (male/female task) and then guess their age (older or younger than 30 years). Pictures were then presented for a second time without an explicit task, and subjects were asked to freely explore and memorize them.

In each test block, a total of 20 degraded images containing a person were briefly presented (150 ms) limiting their visibility. Two experimental factors controlled the perception of the persons on the picture: sensory evidence and prior knowledge. Sensory evidence was varied by presenting images in two degradation levels (high and low noise). Prior knowledge was varied by either showing familiarized (4) or new (4) images, for which no pre-existing memory was established. Each picture was presented twice. Additionally, to assess the reliability of the subjects' judgment, specifically for the subjective task (see below), we included 4 catch trials only containing background on the higher degradation level. The subjective performance was in fact reliable as the amount of false perceptions in the catch trials was very low for all subjects ranging from 0% to 38% (m = 8.5%, sd = 8.5%).

Each trial consisted of a fixation cross on a gray background presented for a random duration of 1.2-1.4 seconds and followed by the degraded image (150 ms). After 1 second, an objective and a subjective task were presented. In the objective task, a screen prompted the subjects to indicate via button press the gender of the person in the picture (male/female judgment). Subsequently, in the subjective task subjects reported whether they had indeed perceived a person on the picture. Occasionally, subjects were also asked to indicate whether the picture, now shown in the degraded fashion, had been presented in the familiarization phase. Each block lasted 3-4 minutes and subjects could take breaks between blocks. A total of 108 trials per condition were presented.

To rule out any picture-specific effects in the neural measures, pictures containing a person were randomized such that across 4 subjects each picture was assigned to every condition (2 sensory evidence X 2 prior knowledge) exactly 1 time.

Recordings and data analysis

Data acquisition

MEG data were acquired with a 275-channel whole-head system (Omega 2005, VSM MedTech Ltd., BC, Canada) at a rate of 1200 Hz with a hardware antialiasing filter at 300 Hz in a synthetic third order axial gradiometer configuration (Data

Acquisition Software Version 5.4.0, VSM MedTech Ltd., BC, Canada). The head of the subjects' was fixated with paper pads and head position was measured using three coils placed at the subject's nasion, and the preauricular points. The head position was checked before and after each run (i.e. 3 experimental blocks, approximately every 10 minutes) to make sure that the subjects' heads did not drift more than 5 mm form the original position at the beginning of the recording. All runs where the head movement exceeded 5 mm were excluded from further analysis (1 run from 2 subjects and 2 runs from 2 other subjects). We monitored eye movements and blinks during the recordings with two pairs of electrooculogram (EOG) electrodes, one pair placed vertically with one electrode above and the other below the left eye and the other pair placed horizontally, one electrode 1 cm lateral from the outer canthus of the left eye, the other 1 cm lateral from the outer canthus of the right eye. Behavioral responses were recorded using in-house fiber optic light barriers.

Individual high-resolution structural MRIs were acquired with a 3 Tesla Magnetom Allegra scanner (Siemens, Erlangen, Germany) using a T1-weigthed magnetization-prepared rapid acquisition gradient-echo (MPRAGE) sequence (160 slices; TR: 2300 ms; TE: 3.93 ms; FA: 12°; FOV: 256 mm; voxel size 1×1x1 mm). For the alignment of the MEG and MRI data we marked the position of the nasion and the preauricular points with Vitamin E capsules.

Data preprocessing

Data were analyzed using the open source Matlab toolbox Fieldtrip (Oostenveld, et al., 2011). As we focused here on frequencies above alpha frequency band, continuous data were filtered with high pass filter at 10 Hz and low pass filter at 200 Hz. Then, data epochs –500 to 700 ms with respect to the onset of the visual stimulus were defined from the continuously recorded MEG signals. Trials contaminated by muscle activity, or jump artifacts in the sensors were discarded using automatic artifact detection and rejection routines from the Fieldtrip toolbox. Additionally, data was visually inspected for remaining artefacts (e.g. eye blink and eye movement artifacts that were still visible despite the 10 Hz high pass filter).

Localizing the sources with beamforming

We used a linearly constrained minimum variance beamformer (Van Veen, van Drongelen, Yuchtman, & Suzuki, 1997) to estimate the power and the phase of the neural population activity at the source level. Beamformers are spatial filters that project sensor level activity to specified source locations in the brain using a linearly weighted sum of the sensor signals. For each source location power is reconstructed with unit gain while suppressing activity from all other source locations as much as possible. Spatial filters were calculated over a time-window stretching from 400 ms before the stimulus onset to 500 ms after stimulus onset. For each grid location three orthogonal filters were computed and the dominant dipole direction of the time courses was extracted using eigenvalue decomposition. In all experimental comparisons we calculated the common filters based on the similar amount of data from both conditions.

We constructed the set of potential source locations (i.e. the source grid) following the procedure described in Grutzner et al. (2010). First, we overlaid a regular three-dimensional dipole grid with an isotropic spacing of 10 mm on the T1 template of the SPM8 toolbox. Next, we transformed each individual subject's anatomical MRI onto this template using the linear transformation from SPM8 and recorded each subject's individual transformation matrix. Then we warped the regular dipole grid with the inverse of each individual transformation matrix. Thus, we obtained specific dipole grids for each subject, enabling us to perform beamformer source analysis with individual head models. The corresponding forward solution (lead field) for each subject was computed using its individual dipole grid and a realistic single-shell volume conductor model (Hamalainen & Sarvas, 1987, 1989) with an implementation described by Guido Nolte and colleagues (Nolte, 2003). Importantly, each location of the dipole grid in MNI space had a unique corresponding grid location in each subject.

After the pilot analysis were done with the full resolution of the 10 mm grid, we decided to downsample this source-space to a) increase the signal to noise ratio, b) be less endangered by interindividual variations of the exact location of sources, c) increase statistical power (less tests) and d) make the all-to-all synchrony analysis computationally feasible.

We first ran the frequency decomposition and the phase locking analysis on the raw time-courses of all grid-locations separately and only then averaged the power information or the phase locking values (see below). This averaging was done over the whole grid locally over 8 grid-points that constitute the vertices of a cube (with the edge length of 10 mm = grid spacing). Thus, the local power effects represent the power of these locally averaged sources and synchrony represents phase locking between these locally averaged source locations, 238 sources altogether. Throughout the manuscript we call these locally averaged sources simply sources.

Four goals of the experiment and the general statistical approach

As stated in the introduction this study had four specific goals: 1) try to replicate the local power effects reported in chapter 3 (Aru, Axmacher, et al., 2012), 2) examine the local power effects all over the brain, 3) compute synchrony between the category-specific node and the rest of the brain, 4) investigate synchrony between all source pairs.

To be maximally sensitive to the weak effects of local power and source synchrony we used two types of “localizers” (localizer contrasts) to find the interesting source-time-frequency combinations. These localizers were independent of the experimental contrasts and helped us to focus on certain points or clusters of points in the multidimensional space. Applying the experimental contrasts (trials with vs without prior knowledge / trials with more vs less sensory evidence) only at these specific points greatly reduces the amount of tests whose number has to be accounted for in the correction for multiple comparisons. We also used a general “source unspecific comparison” where we averaged over all the sources or source combinations and the experimental effects were tested over time and frequency on these source-averaged data.

Thus, we first computed the “category specific localizer” by contrasting trials with pictures containing a person in the foreground (from the condition less sensory evidence and without prior knowledge) versus those only containing the background (catch images). This localizer should provide source-time-frequency combinations that are sensitive to the processing of the target (the person on the picture). Although this is potentially a very revealing localizer, it must be kept in mind that for this comparison we have two times less trials than for the experimental contrasts and four times less trials than for the “task specific localizer” so not finding effects here could also be attributed to the low number of trials.

Second, we used the “task specific localizer” where we compared the baseline effect to the task-epoch effect, where all the 4 stimulus conditions were combined. Interesting source-time-frequency points were defined through these two types of localizers and the experimental effects were computed only at these points. This localizer is much less specific as all the areas related to simple visual processing will be activated too but it has been shown to be effective in previous studies (Hipp, et al., 2011).

Third, we computed the "source unspecific comparison" and tested the experimental effects only over time-frequency space. Any result from this comparison is unspecific with regard to the underlying sources or source combinations but this comparison is sensitive to big effects occurring over multiple sources or source combinations.

Finally, to be exhaustive, we ran the unconstrained comparisons where we directly contrasted the experimental conditions over all source-time-frequency points. This last comparison has the weakest statistical power but could potentially discover effects that are independent of the localizers but strong enough to survive correction for multiple comparisons across all data dimensions.

When comparing single source-time-frequency points we used the parametric t-test, which was validated with and gave very similar results to the monte-carlo permuted t-test. By the "category specific localizer" and direct comparisons of experimental conditions the single source-time-frequency points had too small effects to survive the false discovery rate (FDR) correction and thus we applied the cluster permutation method with 1000 randomizations to correct for multiple comparisons (Maris & Oostenveld, 2007; Hipp, et al., 2011, see specific sections for details). When the "category specific localizer" provided no significant clusters, we selected for interesting source-time-frequency points with uncorrected p-thresholds (see below). The "task-specific localizer" generally yields very strong effects, thus here we corrected for multiple comparisons with FDR. Also, as our baseline epoch was shorter than the task-epoch a cluster based comparison would not be mathematically correct in this case (Maris & Oostenveld, 2007). As the hypothesis tested in this study is specific about the *increase* of synchrony due to sensory evidence or prior knowledge all tests were one-tailed to increase statistical power.

Analyzing time-frequency data

Except for the pilot analysis on local population activity (see below) we analyzed frequencies from 15 to 105 Hz. Data were analyzed in the time-frequency domain by convolution with complex Gaussian Morlet wavelets with a bandwidth parameter $f_0/\sigma_f=4$ at 15 Hz linearly increasing in integers from $f_0/\sigma_f=4$ to $f_0/\sigma_f=8$ at 105 Hz in 20 Hz steps. To assess synchrony between the sources we computed the phase-locking value (PLV) (Lachaux, Rodriguez, Martinerie, & Varela, 1999). Except for the computation of the "task specific localizer", where the data are directly compared to the baseline data, post-stimulus data were baseline-normalized by dividing it by the mean activity of the baseline interval of -350 to -100 ms of a given frequency band. We analyzed the post-stimulus data in non-overlapping 50 ms steps from 50 ms to 500 ms and with a frequency step of 10 Hz from 15 to 105 Hz. Parts of the data were also analyzed with a frequency step of 4 Hz and very similar results were obtained.

Analyzing local population activity

We first sought to replicate the local power effects observed in chapter 3 (Aru, Axmacher, et al., 2012). For assessing whether similar results to our ECoG could be obtained with the present MEG setup, we ran the "category specific localizer" contrast in the source location that encompassed the extrastriate-body area (EBA). The right EBA is the most consistently activated node of the person-selective network (Downing, Jiang, Shuman, & Kanwisher, 2001; Peelen & Downing, 2005) and was also consistently person-selective in the ECoG study described in chapter

3. As noted above each source in the present study represents the average activity of 8 grid-locations forming a cube. For the middle point for the right EBA source the MNI coordinates were 55, -65 and 5. Although the left EBA has been shown to show weaker selectivity to human body parts, we also tested for category-specific effects there (MNI coordinates -55,-65 and 5).

In the pilot analysis of local population activity we assessed gamma band response (GBR) in two frequency ranges – low (30-70 Hz,) and high GBR (70-150 Hz) similarly to the analysis presented in chapter 3 – and over a time interval of 200-500 ms, where the key effects of person-selectivity emerged (Aru, Axmacher, et al., 2012). This pilot analysis showed that in MEG no person-selective responses could be measured at the high gamma band range whereas the lower gamma band was sensitive to the person-selective effects (see results).

Then, as in chapter 3, we assessed the effect of experimental factors on local EBA activity by contrasting the power of pictures with and without prior knowledge or pictures with higher or lower sensory evidence in the right and left EBA sources. This was done in the low gamma range where the person-selective responses were observed.

Next we searched for other person-selective sources over the whole time-frequency space. We used the cluster permutation method to find source-time-frequency points where the “category specific localizer” is significant. Here we used the cluster permutation method as implemented in Fieldtrip (Maris & Oostenveld, 2007). As we did not find any source-time-frequency points that would have survived the correction for multiple comparisons and as the “localizer” is orthogonal to the experimental contrast, we simply used a threshold $p < 0.001$ uncorrected to define the interesting source-time-frequency points and ran the experimental contrasts (with vs without prior knowledge; more vs less sensory evidence) on these source-time-frequency points.

Then we computed the "task specific localizer" by comparing the baseline power to the task-epoch power so that all 4 stimulus conditions were combined. In this comparison the power values were not baseline-normalized as the task-epoch power was directly contrasted with the baseline power. We used FDR to correct for multiple comparisons and to find the time epochs, frequencies and sources where task-epoch power is stronger than in the baseline. We cannot use the cluster permutation approach because the baseline epoch is shorter than the task-epoch (Maris & Oostenveld, 2007). Here after the FDR correction very many source-time-frequency points remained significant. As it would not be a statistically sensitive approach to test the experimental effects separately on all these points and as the source-time-frequency points showed a clear structure in source, time and frequency, we located the strongest cluster of these FDR corrected source-time-frequency points and tested the experimental effects on the averaged activity of this cluster. In fact, more than 4800 source-time-frequency points survived the initial FDR correction with $q = 0.05$. Therefore, to be more precise in relevant sources, times and frequencies, we ran the clustering algorithm on those source-time-frequency points that survived FDR correction with $q = 5.0$ e-6. Note that as this is an orthogonal contrast to the experimental comparisons we can select the threshold here to have a desired resolution in the source-time-frequency space. Clustering was done similarly to the approaches described in Maris & Oostenveld (2007) and Hipp et al. (2011) with the dimensions source position (3 dimensions), time and frequency. A cluster corresponds to a continuous component of this 5 dimensional space. To find continuous components we looked for the neighbors of each FDR corrected source-time-frequency point and evaluated whether it is also significant on the FDR level. A neighbor is a point that differs from the original point in exactly one step in one dimension (e.g. the same source and time but a neighboring frequency; or same time and frequency and same 2 source position dimensions, but a neighboring point in one source position dimension). As in Hipp et al. (2011) we defined the biggest cluster based on the sum of the t-scores across the source-time-frequency points of the cluster. Different from the approaches of Maris & Oostenveld (2007) and Hipp et al. (2011), where uncorrected p values are used and significance is tested on the cluster level, we defined the cluster from points that already were corrected for multiple comparisons with FDR. Therefore, we needed no extra significance testing of this cluster activity against the baseline. Then, we ran the experimental contrasts on the activity of these clusters.

Next we quantified the power changes with the "general source-unspecific" analysis where we averaged the power values over all sources and compared whether such average values are higher for pictures with prior knowledge than for pictures without prior knowledge or for pictures with more sensory evidence as compared to the pictures with less sensory evidence. As for this comparison synchrony is averaged over all sources, it is ran only over the time-frequency space and corrected for multiple comparisons over the time-frequency space with the cluster-permutation method.

Finally, we assessed the experimental effects over the whole brain without any constraints, i.e. simply by contrasting the experimental conditions at each source at each time step and each frequency. Here we corrected for multiple comparisons over the source-time-frequency space with the cluster permutation method as explained above.

Analysis of synchrony with the seed region in EBA

If long range synchrony is the key mechanism for conscious perception, having prior knowledge or more sensory evidence should both increase synchrony. In particular, the information about the target (the person on the picture) should be communicated to other brain areas more efficiently. Thus, we quantified the PLV of the time courses of the virtual electrodes of all sources with the seed region in the source corresponding to the right EBA, which was the only source in the visual cortex showing clear and robust person-selective responses (see results).

We first used the “category specific localizer” comparison between pictures where a person was on the foreground and catch pictures to find sources whose phase locking with EBA increases when a person is on the picture. We used the cluster permutation method as implemented in Fieldtrip (Maris & Oostenveld, 2007) to find source-time-frequency points where the “category specific localizer” is significant. As in this case we did not find any source-time-frequency points that would have been significant and as the “localizer” is orthogonal to the experimental contrast, we simply used a threshold $p < 0.001$ uncorrected to define the interesting source-time-frequency points. Then, synchrony to the EBA was tested for these localized source-time-frequency points between the experimental contrasts. Here we corrected for multiple comparisons over these source-time-frequency points with FDR.

Then we applied the “task-specific localizer”. In particular, we compared the baseline synchrony to the task-epoch synchrony, where all 4 stimulus conditions were combined. In this comparison the PLV values were not baseline-normalized, because the task-epoch synchrony was directly contrasted with the baseline synchrony. We used FDR to correct for multiple comparisons and find the time epochs, frequencies and sources where task-epoch synchrony is stronger than in the baseline. As in this case the FDR correction revealed a cluster of source-time-frequency points, we here analyzed the average synchrony of this cluster in the experimental conditions. The remaining source-time-frequency points that survived the FDR correction did not belong to any cluster, so we tested for experimental effects in them separately. We corrected for multiple comparisons over these source-time-frequency points with FDR.

Next we ran a general source-unspecific analysis where we averaged synchrony-to-EBA values over all sources and compared whether such overall synchrony values are higher for pictures with prior knowledge than for pictures without prior knowledge or for pictures with more sensory evidence as compared to the pictures with less sensory evidence. As for this comparison synchrony is averaged over all sources, it is ran only over the time-frequency space and corrected for multiple comparisons over the time-frequency space with the cluster-permutation method.

Finally we assessed the experimental effects of synchrony to EBA over the whole brain, i.e. without using the "localizer" contrast. Here we corrected for multiple comparisons with the cluster permutation method.

All-to-all synchronization

Finally we quantified changes in synchronization happening all over the brain (Hipp, et al., 2011). For this, PLV was computed between all pairs of the pairwise source space.

First we used the "category specific localizer" comparison between pictures where a person was on the foreground and catch pictures to find source pairs whose phase locking increases when a person is on the picture. We corrected for multiple comparison with the cluster permutation method. As we here have not simply sources but source-pairs, the cluster permutation is done in 8 dimensions similarly to Hipp et al. (2011): 3 dimensions for one source, 3 for the other source, time and frequency. A cluster corresponds to a continuous component of this 8 dimensional space. To find continuous components we thresholded the whole 8 dimensional space with the "category specific localizer" contrast at the threshold $p < 0.01$, looked for the neighbors of source-time-frequency point below this threshold and evaluated whether the neighbors are also below this threshold. A neighbor is a point that differs from the original point in exactly one step in one dimension (e.g. the same sources and times but a neighboring frequency; or same time and frequency and one source and a third source who is similar to the original source in two source position dimensions, but a neighboring point in one source position dimension). As in Hipp et al. (2011) we defined the biggest cluster based on the sum of the t-scores across the source-time-frequency points of the cluster. We then tested the significance of this cluster with the permutation method. Within the subjects we randomly exchanged the condition-labels and in this random permutation ran the same process of cluster identification as for the original cluster. We repeated this step 1000 times and by that we created a distribution of biggest clusters under the null-hypothesis that there is no difference between the conditions. As in this case we did not find any source-time-frequency points that would have survived the cluster permutation method and as the "localizer" is orthogonal to the experimental contrast, we simply used an uncorrected threshold to define the interesting source-combination-time-frequency points. As the number of source combinations is high, we used a more restrictive threshold ($p < 0.00001$) than in other comparisons (i.e. the source-combination-time-frequency space is roughly 100 times bigger than the source-time-frequency space used in other comparisons). Then, synchrony between these sources was tested across the experimental contrasts.

Then we ran the "task-specific localizer", where all 4 stimulus conditions were combined to find the sources, time epochs and frequencies where synchrony is significantly increased as compared to the baseline. We used FDR to correct for multiple comparisons and defined the biggest component of these FDR corrected ($q = 0.05$) source-combination-time-frequency points by the method described above. Then, we contrasted the experimental conditions on these clusters.

Next for the source-unspecific analysis we collapsed synchrony over all source pairs and thus analyzed whether pairwise synchrony increases in general when pictures with prior knowledge are compared to pictures without prior knowledge or when the two different levels of sensory evidence are contrasted. As for this comparison synchrony is averaged over all source combinations, it is ran only over the time-frequency space and corrected for multiple comparisons over the time-frequency space with the cluster-permutation method.

Finally we analyzed the experimental effects over the whole source-combination-time-frequency space without any restrictions. This approach corresponds to the 8-dimensional-clustering described by the "category specific localizer" of this section with the key difference that the compared conditions are the experimental contrasts (with vs without prior knowledge; more vs less sensory evidence).

Addressing the potential confound of SNR in the identified networks

It has been noted (Gross, et al., 2001; Hipp, et al., 2011) that the level of synchrony depends on the signal-to-noise ratio. Therefore, a significant effect in synchrony might be due to local activity changes. In order to control for such confounds we checked for effects of local power (see Hipp, et al., 2011). Hipp et al. (2011) only controlled for the power confounds in the overall level of activity over the whole network, but in reality this is insufficient: it could be that the power of one node is significantly stronger in condition 1 and that the power of another node is significantly stronger in condition 2 – averaged over the whole network there might be no changes in power, but these local changes would actually lead to confounds. Therefore, we controlled for the local power changes at each node separately by analyzing whether the power of the source-time-frequency points where the effects were found were also affected by the experimental conditions.

Further analysis of the identified networks

To find the networks that are more intimately related to conscious perception, we used Spearman correlation to quantify the relationship between the increase of synchrony with the perceptual effects across subjects. In particular, we correlated the effect of synchrony due to prior knowledge (<synchrony with prior knowledge> – <synchrony without prior knowledge>) with the perceptual effect (<percentage “person seen” responses with prior knowledge> - <percentage “person seen” responses without prior knowledge>). The same was done for sensory evidence, where we subtracted the synchrony and the perceptual effect of trials with lower sensory evidence from trials with higher sensory evidence.

Power and synchrony at the sensor level

We also analyzed both power and synchrony at the level of MEG sensors. Data were analyzed in non-overlapping 50 ms steps from 50 ms to 500 ms and with a frequency step of 10 Hz from 15 to 105 Hz similarly as was done for the source data. To assess synchrony between the sensors we computed the phase-locking value (PLV, Lachaux, et al., 1999). Here, at the sensor level, volume conduction could lead to spurious phase locking. Therefore we removed any synchrony at the angles 0 degrees and 180 degrees. Similarly to the source-level analysis post-stimulus data were baseline-normalized by dividing it by the mean activity of the baseline interval of -350 to -100 ms of a given frequency band. PLV was calculated between all pairwise sensors and then averaged over all these pairwise sensor combinations, yielding a time-frequency map of synchrony values. This map was corrected for multiple comparisons with the cluster permutation method. Power was computed separately for each sensor and analyzed statistically with the cluster permutation method over the source-time-frequency space. We also analyzed the sensor unspecific power by averaging the power values over all sensors. This yielded a time-frequency map of power values, which was corrected for multiple comparisons with the cluster permutation method.

Results

Behavioral results

Recognition was strongly modulated by the amount of sensory evidence as well as prior knowledge and the corresponding pattern of results was the same for the objective and the subjective measure of recognition (Figure 4.2).

Figure 4.2 (on the left) illustrates how the objective recognition performance was modulated by the amount of sensory evidence and prior knowledge, i.e. the subjects were more correct when the pictures had either lower degradation or when they had been familiarized with the clear version of the image. A two-way within subject ANOVA revealed a main effect of degradation ($F(1,23) = 116.71$, $P < 1.746E-10$) and a main effect of prior knowledge ($F(1,23) = 77.812$, $P < 7.71E-09$). The interaction between degradation and prior knowledge was not significant ($F(1,23) = 0.08$, $P = 0.7789$).

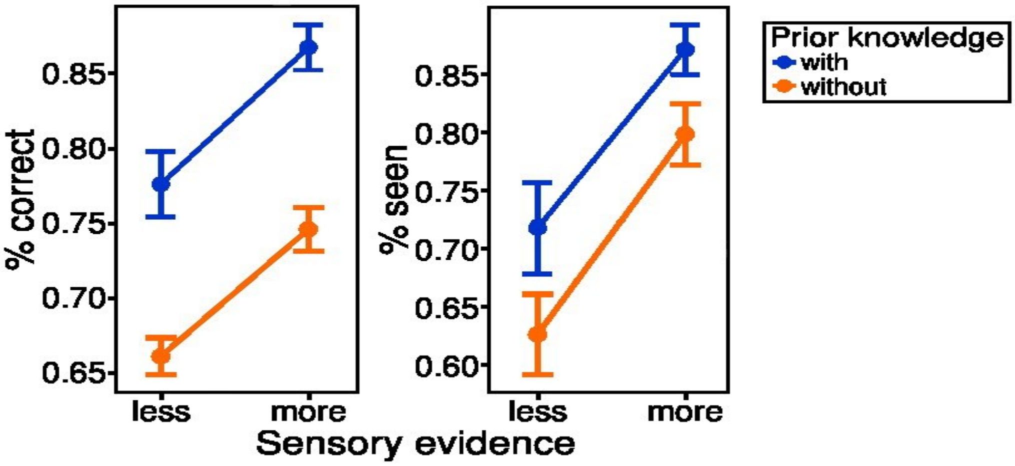


Figure 4.2: Behavioral results. Sensory evidence and prior knowledge both enhance accuracy (left) and conscious perception (right).

But not only the objective recognition performance improved with increasing sensory evidence or prior knowledge. An identical pattern of results was obtained for the subjective measure of recognition showing that our experimental factors also had an effect on the participants' subjective impression of visibility when the amount of degradation decreased or when a stimulus was familiar from before (Figure 4.2, right). As for the objective measure, a two-way within subject ANOVA revealed a highly significant main effect for degradation (F(1,23) = 69.109, P = 2.206E-08) and a highly significant main effect for prior knowledge (F(1,23) = 70.167, P = 1.931E-08). Again, the interaction was not significant (F(1,23) = 0.6376, P = 0.4328).

To assess whether sensory evidence and prior knowledge enhance conscious perception of the person to a comparable degree, we contrasted the increase in "seen" responses resulting from changes in sensory evidence (% perceived low deg - % perceived high deg) with the corresponding increases resulting from prior knowledge (% perceived with prior knowledge - % perceived without prior knowledge). In the current behavioral data sensory evidence had a stronger effect on perception than prior knowledge (T(23) = 3.8372, P = 0.0008). Therefore, any neural effect observed by sensory evidence but not by prior knowledge could simply reflect the fact that sensory evidence indeed modulated perception more than prior knowledge. To account for this confound, we equalized the perceptual effects of sensory evidence and prior knowledge by randomly leaving out "seen" and "unseen" trials within a subject until the perceptual effects of sensory evidence and prior knowledge were equal.

The main question of the current work was whether the perceptual enhancement through sensory evidence and prior knowledge is reflected in corresponding increases of synchrony. If long-range neural synchrony is a correlate of conscious experience, it should be affected both by sensory evidence and by prior knowledge as both modulate conscious perception.

Local population activity

As the goal of beamforming is to estimate the local population activity from the sensor-level MEG data, we first wanted to replicate our previous findings (Chapter 3, Aru, Axmacher, et al., 2012). In that study the same experimental paradigm was used but local cortical activity was measured directly from ECoG electrodes, which have a much better signal-to-noise ratio than the present MEG data. However, in the present experiment we had much more trials and subjects, which in principle could compensate the smaller signal-to-noise ratio.

For performing a similar analysis to the one conducted with ECoG we estimated the gamma band responses (GBR) for extrastriate body area (EBA) where we had previously observed our experimental effects (Aru, Axmacher, et al., 2012). In our intracranial study we had observed that local GBR only increased together with sensory evidence, although prior knowledge enhanced conscious perception in a similar fashion. We sought to replicate these findings with MEG beamforming.

A crucial first step in our logic was that EBA is the category-specific node that is involved in the task-relevant processing (Chapter 3, Aru, Axmacher, et al., 2012). In the previous chapter, we had delineated these nodes with the contrast “catch pictures” (i.e. pictures with no person on them) vs “pictures from the condition less sensory evidence and no prior knowledge” (i.e. pictures that have the same noise parameters as the catch pictures, but have a person on them), which we call the “category specific localizer” in the present work. In our ECoG data we had observed that these nodes are indeed category-specific, as the gamma band responses in them are significantly higher for the pictures with a person than for catch pictures with no person. In our ECoG data this effect was evident in the 30-70 Hz band but was much more pronounced in the range of high GBR, 70-150 Hz.

In the MEG beamformed data we have not been able to find the category-specific effect in the high gamma range: between 70-150 Hz the two conditions (with and without a person) are not different at the source corresponding to the EBA or at the sources around it or at any other source in the visual cortex. (As this is an important first step, we have played around with the parameters and confirmed the absence of this high gamma effect at the EBA very thoroughly: 1) We have used different time windows for the analysis and for computing the spatial filters, 2) We have used different regularization parameters (lambda = 1%, 5%, 10%), 3) We have used different beamformers – LCMV vs. DICS).

Of course MEG beamformed signals do not have to correspond to the intracranial potentials. It is possible that in MEG the high frequency part of the response is attenuated, so that the effects are more visible in the lower frequencies. This is indeed confirmed by the statistics: with almost every parameter combination the baseline-normalized low GBR (30-70 Hz) in the right EBA shows a person-specific effect (with the parameters described in the methods section p = 0.0047, Figure 4.3). Thus, it could be that the local response in MEG is mainly confined to the lower frequencies, which would also fit with the MEG literature on gamma band responses (see discussion). The category-selectivity of the left EBA was not so robust over the range of used analysis parameters (e.g. with the current parameters p = 0.045). In fact, the right EBA was the only source that discriminated between pictures with and without a person below p < 0.01. Thus, we used only the right EBA as the category-specific node for our synchrony analysis (see below).

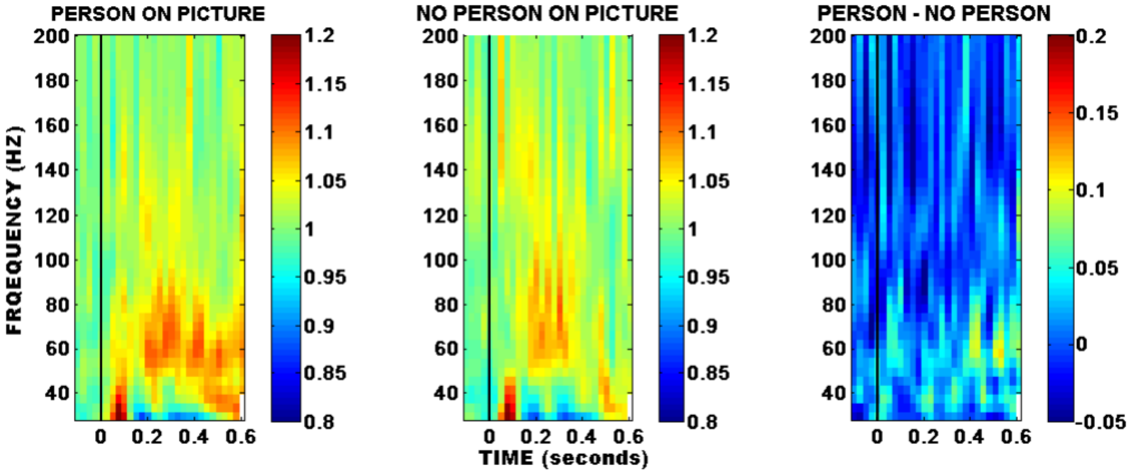


Figure 4.3: Category-selectivity at the person-specific source in the extrastriate body area (EBA). Time-frequency representations of images with persons are on them (left) are compared to the catch images without a person on the background (middle). Their activity difference can be seen in the right column. Category specific responses are confined to the lower gamma frequencies 30-70 Hz (compare to figure 3.2 where it is evident that the category specific responses in ECoG were mainly at higher frequencies above 70 Hz).

Of course the key question is, whether we can replicate our experimental effects from chapter 3. Based on the results from that study and the above results regarding the category-specific frequency response around 30-70 Hz, we expected that sensory evidence increases local population activity around the low GBR, while prior knowledge has no such effect. This result was indeed confirmed as increasing sensory evidence enhanced the low GBR (p = 0.022; top row on Figure 4.4) but prior knowledge had no effect on the low GBR (p > 0.2; bottom row on Figure 4.4).

In sum, we were able to replicate the findings from chapter 3, albeit only in the low gamma band: a) the right EBA showed person-selectivity in the power responses and b) this person-selective local activity was modulated by sensory evidence but not by prior knowledge.

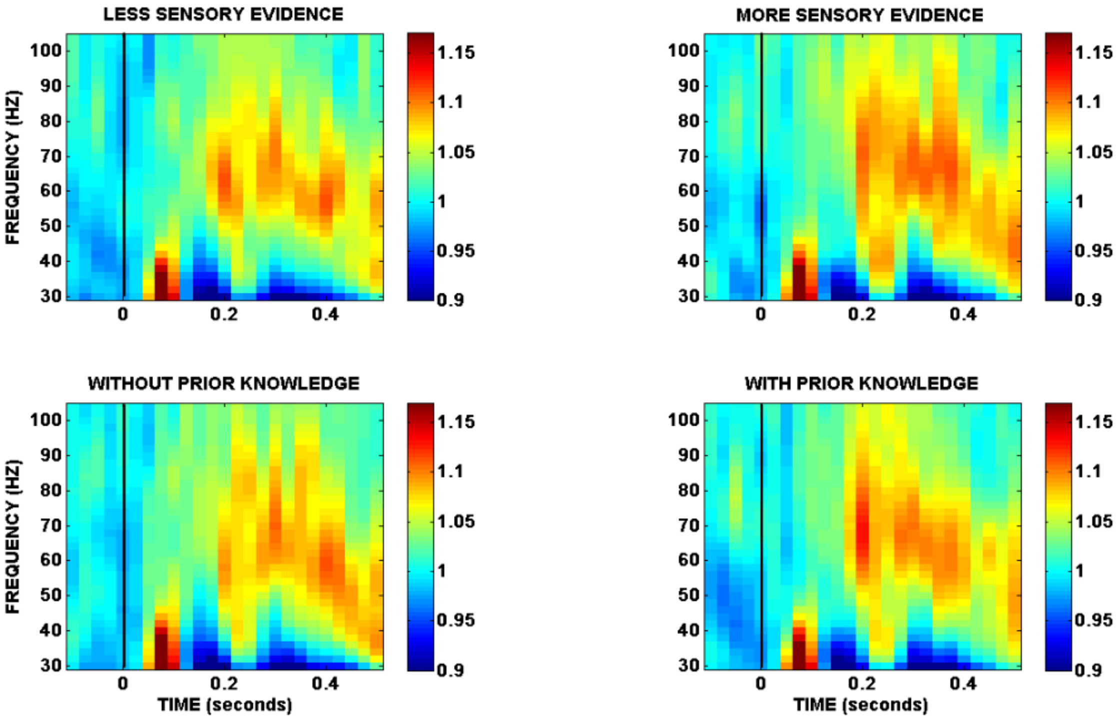


Figure 4.4: Effects of sensory evidence and prior knowledge on the category-selective responses at the person-specific source in the extrastriate body area (EBA). Top: sensory evidence has an effect on local gamma band responses so that more sensory evidence leads to stronger responses in the low gamma band (p = 0.022). Bottom: prior knowledge has no effect on local gamma band responses (p > 0.2). Note that for the statistical comparison only the low gamma band activity was analyzed, because the previous comparison (Figure 4.3) had revealed person-selective responses only in the low gamma band.

Local population activity: other effects

An advantage of the MEG-approach is of course that our analysis can cover the whole brain and not just the specific person-selective nodes. Hence, next we ran the “category specific localizer” contrast between catch pictures and pictures with persons on the foreground over the whole brain to search for other possible person-selective areas. The “localizer” comparison was conducted over the whole source-time-frequency space for selecting the interesting points in this multidimensional space. No source-time-frequency point survived the correction with the cluster permutation method in the “category specific localizer” contrast of local power. We thus used an uncorrected threshold of 0.001 for selecting the source-time-frequency points of interest for the experimental contrasts. This uncorrected threshold revealed 40 source-time-frequency points that were positioned in an unstructured way on the source-time-frequency space. We observed no experimental effects on these source-time-frequency combinations ($p > 0.1$ for both sensory evidence and prior knowledge).

Then we computed the “task-specific localizer” by comparing the task-epoch power values with the baseline power values. Here all 4 types of trials were combined for the task-epoch power, thus this comparison is again orthogonal with regard to our experimental contrasts. As expected, this “task-specific localizer” revealed a very strong network in the visual areas with p values reaching uncorrected $p < 1.0$ e-10 and around 4800 source-time-frequency points surviving the FDR correction for multiple comparisons. Computing the experimental effects separately on such an enormous number of source-time-frequency points would not be very fruitful. Therefore we formed a cluster in the source-time-frequency space where neighboring points were below FDR corrected value of $q < 5.0$ e-6 in the “task specific localizer”. This cluster was mainly confined to the early visual cortex but extended to the right EBA. As expected, this cluster was mainly active in the gamma range between 45 and 95 Hz and extended in time from 150 ms to 400 ms. The fact that the “task-specific localizer” revealed areas in the visual cortex active in the gamma range is expected based on the literature (for review Donner & Siegel, 2011) and therefore lends validity to our data analysis approach. We contrasted the average activity of this cluster in the experimental conditions but found effects of neither prior knowledge nor sensory evidence (both $p > 0.1$).

Next we assessed the “general source unspecific” effect by averaging over all sources and contrasting the experimental conditions on this averaged time-frequency data. We observed no experimental effects on these data ($p > 0.3$ for both prior knowledge and sensory evidence).

Finally we computed the experimental effects over the whole brain over all frequencies without any constraints. We found no further effects that survived correction for multiple comparisons ($p > 0.3$ for both prior knowledge and sensory evidence).

In conclusion, we observed no other effects of prior knowledge or sensory evidence on local power.

Synchrony to EBA

Our analysis of local population activity showed that the right EBA is the person-selective node. For processing the task, the information about the persons on the picture computed locally within the EBA should be propagated across the brain through long-distance synchrony. If long-range neural synchrony is a correlate of conscious experience, such synchrony between EBA and other brain areas should be increased both by sensory evidence and by prior knowledge as both modulate conscious perception. Thus, we fixed a seed region in the right EBA and searched for areas whose synchrony with the right EBA increases by increasing sensory evidence or by providing prior knowledge.

First we used the "category specific localizer" contrast to find sources whose phase locking with EBA increases when a person is on the picture and ran the experimental comparisons on those source-time-frequency points. No source-time-frequency point survived the cluster permutation correction for multiple comparisons. We thus used an uncorrected threshold of 0.001 for selecting the source-time-frequency points of interest for the experimental contrasts. This uncorrected threshold revealed 12 source-time-frequency points that were positioned in an unstructured way on the source-time-frequency space. We observed no experimental effects in these source-time-frequency points ($p > 0.3$ for both sensory evidence and prior knowledge).

Then we applied the "task-specific localizer" to find those source-time-frequency points where the synchrony to EBA increases as compared to the baseline period. We observed that compared to the baseline the stimulus led to synchronization between the right EBA and a cluster of neighboring sources in the left visual cortex (BA17 and BA18) at 15-25 Hz around 50-150 ms (FDR corrected). This result by itself is expected, as information should be propagated from extrastriate area to EBA around that time, but by showing up in our statistics this result again lends credibility to our beamforming synchrony approach. However, when we analyzed the experimental effects on this cluster, we did not observe any effects for prior knowledge or sensory evidence (both $p > 0.18$). We also observed no effects of experimental conditions on other source-time-frequency points that were significant after the task-specific localizer contrast but did not belong to this cluster.

Next we ran general source unspecific contrasts of experimental conditions by averaging the synchrony-to-EBA values over all sources to ask whether such general synchrony with the person-specific node increases due to prior knowledge and sensory evidence. We observed no experimental effects ($p > 0.4$ for both sensory evidence and prior knowledge).

Finally, we analyzed how prior knowledge and sensory evidence changed the synchrony-to-EBA pattern over the whole source-time-frequency space without any constraints. We observed no effects that survived the correction for multiple comparisons ($p > 0.4$ for both sensory evidence and prior knowledge).

All-to-all synchrony

Although EBA is the active node whose relevance for the present task has been verified in the present analysis and in our previous work (chapter 3; Aru, Axmacher, et al., 2012), many other sources and their dynamical interactions might be relevant for consciously perceiving the person on the picture. Therefore, in the final part of the analysis synchrony was computed between all pairs of the pairwise source space.

First we ran the “category specific localizer” between pictures where a person was on the foreground and catch pictures to find source pairs whose phase locking increases when a person is on the picture. This contrast yielded no results that would have survived the correction of multiple comparisons with the cluster permutation method. Therefore we used an uncorrected threshold of $p < 0.00001$ for selecting the source-time-frequency points of interest for the experimental contrasts. This yielded a set of 16 points that were unstructured over the source-combination-time-frequency space. When synchrony was tested on these source-combination-time-frequency-space across the experimental contrasts we observed no experimental results (all $p > 0.1$). We also built clusters of source-combination-time-frequency points that were below the uncorrected p threshold of 0.01 and tested only the mean activity of these clusters between the experimental conditions. Here we can report the incidental finding that in some visual cortex networks where "category-specific localizer" was accompanied by increased synchrony, experimental increase of conscious perception due to prior knowledge or sensory evidence led to decreased synchrony. However, no effects emerged in the expected direction of increased synchrony (all $p > 0.4$).

Next we computed the "task specific localizer" by contrasting source-synchrony between the baseline epoch and task-epoch. Almost 10 000 source-combination-time-frequency points survived the FDR correction for multiple comparisons. Computing the experimental effects separately on such enormous number of source-time-frequency points would not be very fruitful. Therefore we formed clusters from connected points in the source-combination-time-frequency space. The biggest and strongest cluster of these points was at the beta frequencies, mainly at 15 Hz between various visual cortical areas. No experimental effects were observed on this cluster ($p > 0.3$). We then looked for the next biggest cluster in the interaction space and found a network again mainly at the beta frequencies between medial temporal areas. However, in this network we also observed no experimental effects ($p > 0.4$).

We then collapsed synchrony over all source pairs and analyzed whether pairwise synchrony increases in general when pictures with prior knowledge are compared to pictures without prior knowledge or when the two different levels of sensory evidence are contrasted. We found that pictures with prior knowledge were associated with stronger synchrony than pictures without prior knowledge in the frequency range around 40-70 Hz and around 200-350 ms ($p < 0.0005$; corrected for multiple comparisons across time-frequency space with the cluster permutation method). These results were also confirmed with a larger number of 10 000 monte-carlo permutations. Figure 4.5 illustrates these results. We observed no comparable effect for sensory evidence (p > 0.13). These effects were not spuriously caused by changes in power as no effects for power were observed when the power was similarly averaged over all sources and the cluster of the respective time-frequency points was compared between the conditions with and without prior knowledge (p > 0.19). We also analyzed whether the activity of this time-frequency cluster averaged over all source combinations was correlated with the perceptual effect of prior knowledge but we did not observe a significant correlation (p > 0.2).

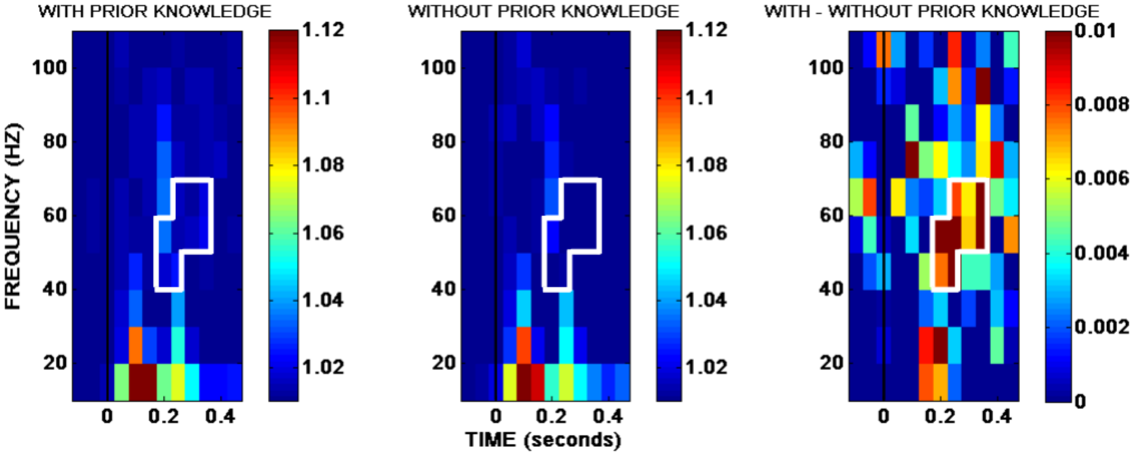


Figure 4.5: Source-unspecific effects of neural synchrony. Time-frequency representations of phase locking value (PLV) are averaged over all possible source combinations. Pictures with prior knowledge (on the left) are associated with higher synchrony than pictures without prior knowledge (in the middle). Their difference (on the right) is significant at 40-70 Hz between 200-350 ms (marked by white boxes). Note that the PLV values are relative to the baseline.

We cannot statistically analyze the synchrony underlying this source-unspecific synchrony effects as such analysis would be circular (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009). However, as a result of an exploratory analysis we can report that the strongest synchrony effects underlying the observed cluster were between the temporal lobes (medial and anterior parts) of both hemispheres. Further research focused on these networks needs to show whether these effects hold in proper statistical comparisons.

Finally, we conducted an analysis similar to Hipp et al. (2011) by running a cluster permutation test directly between the experimental conditions. We observed no experimental effects ($p > 0.2$).

Power and synchrony at the sensor level

We also computed the power and the synchrony effects of the experimental conditions on the sensor-level data. We observed effects neither for power nor for synchrony ($p > 0.3$ for both prior knowledge and sensory evidence).

Discussion

In this study we set out to investigate whether long-range synchrony is a viable candidate for being the mechanism of conscious perception. Several previous studies have found that conscious perception is indeed associated with stronger long-range synchrony (Gaillard, et al., 2009; Melloni, et al., 2007). However, recently it has been argued that the experimental paradigms commonly used might also reflect differences in early processing that determine access to consciousness or consequences of conscious processing (Aru, Bachmann, et al., 2012; de Graaf, et al., 2012). Therefore, previous findings and conclusions have to be re-evaluated. With a new experimental paradigm we have already demonstrated that local category specific gamma band responses thought to correlate with conscious perception (Fisch, et al., 2009) do not reflect conscious perception directly (Aru, Axmacher, et al., 2012). Here we used the same experimental setup (Aru, Axmacher, et al., 2012) with synchrony estimates computed between the source activities obtained from MEG beamforming to evaluate the role of long-range synchrony in conscious perception.

To investigate this question we studied phase locking synchrony between different sources. We investigated synchrony changes to the person-specific node EBA and also between all pairwise source combinations. For increasing statistical power we used two types of localizers, but in addition to that also analyzed the full range of source-combinations-time-frequency space without any constraints as done in Hipp et al. (2011).

We observed that averaged over all possible source combinations (i.e. in a source unspecific comparison) synchrony was stronger for pictures with prior knowledge as compared to the pictures without prior knowledge at the lower gamma frequency (40-70 Hz) and between 200-350 ms. This does not seem to be a general effect of increased perception as no such effects were observed for sensory evidence. Thus, this synchrony effect is specific for prior knowledge. Although far from conclusive, our correlational analysis showed that this general effect of synchrony is not directly related to perception, as no correlation was observed between the perceptual effect of prior knowledge and the synchrony effect. Although Gotts and colleagues (Gotts, Chow, & Martin, 2012) have claimed that in general stimulus repetition leads to improved processing due to synchrony of the respective neural populations, our results cannot fully confirm this hypothesis as we could not find a direct relationship between prior knowledge related increases in synchrony and enhanced perception. It is more likely that the present results that prior knowledge leads to increased gamma synchrony are explained by the relationship between gamma synchrony and memory processes (reviewed in Axmacher & Fell, 2011). In trials with prior knowledge memory traces exist about the pictures, thus these pictures are matched with the memory template and this recognition (“heureka!”) possibly leads to increased synchronization. This speculation is supported by the exploratory analysis showing that the source-combinations underlying this general synchrony enhancement were found to be in the temporal lobe, where memory processes and gamma synchrony relevant for memory processes has been observed (Axmacher & Fell, 2011). However, as the selection of these source-combinations was not independent (Kriegeskorte, et al., 2009) this result can only be taken as exploratory and has to be confirmed by further studies.

The logic of our experimental paradigm was that as conscious perception is facilitated by both sensory evidence and prior knowledge, a good candidate of NCC should change in a similar manner in a response to both of these manipulations. As we found these discussed source unspecific effects only for prior knowledge and observed no comparable effect of increasing general synchrony by sensory evidence, these results do not support the hypothesis that synchrony is the key mechanism underlying conscious perception.

These results reflect the synchrony effects averaged over all source combinations. We were not able to find synchrony between specific sources in the gamma frequency band. This negative effect might have various reasons. First, the algorithms of beamforming are designed so that they suppress the power of highly correlated sources (Van Veen, et al., 1997). It has been claimed that this suppression happens only in the range of source correlations that exceed any physiologically important synchrony (Hipp, et al., 2011). However, if in our study the two V1s or the two EBAs were both simultaneously active due to the input, the signal of both these crucial areas might have been reduced through beamforming and this might have affected the synchrony effects. A second reason for not observing strong synchrony effects is the fact that the physiological synchrony effects in the gamma band are numerically very small. For example in the work of Conrado Bosman and colleagues (Bosman, et al., 2012), the effect of attention on coherence is significant, but numerically only in the range of 0.04 coherence units (Figure 4 in Bosman et al. (2012)). Although we had 24 subjects, roughly 200 trials per condition per subject and the preliminary analysis of statistical power showed that we should be able to capture a small effect like that in our synchrony analysis, it is possible that the effects of gamma band synchrony are even much smaller than 0.04 coherence units in MEG because of two reasons: 1) the attentional effects in Bosman et al. (2012) were measured with special monkey ECoG, which has generally a much higher signal to noise ratio than MEG. This is for example evidenced when comparing Figure 3 from (Hoogenboom, Schoffelen, Oostenveld, Parkes, & Fries, 2006) to figure 1B from (Bosman, et al., 2012) – although Hoogenboom et al. (2006) used an optimal stimulus for generating gamma band responses, their gamma oscillations are roughly 10 times weaker than those measured in Bosman et al. (2012), 2) the coherence effects in Bosman et al. (2012) are found using optimal stimuli for eliciting gamma oscillations and gamma coherence, whereas our stimuli are pictures of natural scenes, for which there are yet no published studies convincingly showing gamma band synchrony. Thus, the effects of synchrony that are existent in the brain might simply have been too small to be measured in MEG.

However, both of these arguments are weakened by the fact that Hipp et al. (2011) have been able to recover gamma band synchrony successfully with EEG beamforming. The signal-to-noise ratio for the high gamma band is known to be weaker in EEG than in MEG because for EEG the skull acts as a low pass filter that diminishes the signal strength especially at the high frequencies (Pfurtscheller & Cooper, 1975) while magnetic fields pass through the skull unaffected. Despite that Hipp et al. (2011) were able to obtain a synchrony network of specific sources at the gamma band with a center frequency around 85 Hz where the experimental effect on coherence was also 0.04, just like in the study where recordings were performed directly from the cortex with the ECoG grid in monkeys (Bosman, et al., 2012). Although in the present study we were not able to find synchrony effects between specific source combinations at the gamma frequencies, the study of Hipp et al. (2011) suggests that at least under some circumstances this might be possible. Further research needs to show what these circumstances are and will hopefully help to better understand the results of the present work.

Besides the main goal of investigating long-distance synchrony, we also sought to replicate the local power effects that we had found in our previous ECoG study with the present MEG setup (Aru, Axmacher, et al., 2012). In our ECoG study we had observed that manipulating sensory evidence had a strong effect on gamma band responses, whereas prior knowledge had no such effect. In our ECoG study the effect of sensory evidence was pronounced at the high gamma range 70-150 Hz. Such broadband power change above 50 Hz is the typical gamma band response observed in intracranial recordings (Lachaux, Axmacher, Mormann, Halgren, & Crone, 2012). For example, faces, houses, words etc. generate strong category-specific gamma band responses in confined brain regions (Vidal, et al., 2010). In our ECoG study we found such high frequency category-specific effects of the target person on the picture around the person-specific area EBA (Aru, Axmacher, et al., 2012). We were not able to find such person-specific responses in the high gamma band in our present MEG recordings. We have tested two types of beamformers (DICS and LCMV), computed the filters on data with a different length, averaged over different amounts of data in space and time and varied the regularization parameters, but we have observed no category-specific responses to persons on picture in the high gamma band above 70 Hz. This negative finding applies to both EBA and the other visual cortical areas. However, in all these different parameter regimes we have robustly observed a person-specific effect in the low gamma band 30-70 Hz at the right EBA. The right EBA is the area where one would expect the person-specific effects based on previous fMRI studies (Downing, et al., 2001; Peelen & Downing, 2005, 2007) and our own intracranial results (Aru, Axmacher, et al., 2012). Moreover, we were able to replicate our findings from chapter 3 about the changes of local power due to the experimental effects in this lower frequency range: GBR in this frequency range was increased by sensory evidence but not by prior knowledge. Thus, both category selectivity and experimental effects that were in our ECoG study more pronounced at higher gamma frequencies between 70-150 Hz were observed between 30-70 Hz in the MEG recordings with no effects at higher frequencies. This result raises the questions whether the category-selective responses in MEG are at lower frequencies than those of the ECoG and what might be the mechanisms for this effect.

The best datasets for answering these questions come from studies that have used identical or very similar stimuli in both intracranial and MEG recordings. One such stimulus set are the black-and-white Mooney faces. In a classic intracranial study, Lachaux and colleagues (Lachaux, et al., 2005) investigated the neural responses to Mooney faces. They observed broadband gamma band responses between 50-150 Hz when the Mooney faces were presented. Sometimes these responses extend to as high as 200 Hz. As in our intracranial study, these broadband responses were reliably stronger when the subject reported perceiving the face (Lachaux, et al., 2005). In MEG sensor level the same stimuli lead to gamma band responses in lower frequencies, peaking around 80 Hz (Grutzner, et al., 2010; Grutzner, et al., 2013; L. Sun, et al., 2012), thus confirming that MEG responses to identical stimuli tend to be at lower frequencies than intracranial signals. In another study Dalal and colleagues (Dalal, et al., 2009) measured intracranial and MEG responses simultaneously and tried to reconstruct the intracranially measured sources with MEG beamforming. For high gamma band responses these authors concluded that “These results /…/ indicate that we were dealing with effects barely above noise level in two out of four patients.” (Dalal, et al., 2009) These authors proposed several reasons why the broadband gamma band responses that are so prominent in intracranial recordings might not be visible in the MEG. Most importantly, high frequency signals are synchronous only very locally and at the sensor level we might be measuring a mix of sources that are not synchronous and therefore tend to cancel each other out. In particular, in the high gamma band the period of the frequency gets shorter, for example being 10 ms at 100 Hz and 6.67 ms at 150 Hz. This can easily lead to signal cancellation if nearby sources have slight phase offsets, for example at 150 Hz a offset of 3.33 ms would lead to complete cancellation of the two signals when they arrive at the sensor level. Even worse, nearby sources, separated at the intracranial measurement level but measured together in MEG might have opposite experimental effects. For example in our intracranial measurements we have observed that signals from two neighboring electrodes exhibited opposite effects on the GBR, at one electrode GBR increased when a person was on the picture while at the other GBR decreased (unpublished results). Further, Dalal et al. (2009) discuss the role of the number of trials. Although in our MEG recordings we have roughly 5 times more trials per condition than in our intracranial study, this might still not be enough to assure high GBR reliably. In conclusion, it is possible that the real biophysical signals are exactly the same in intracranial recordings and MEG source reconstructed signals, but due to the propagation from the brain tissue to the MEG sensor the signal is mixed with other nearby signals and is therefore reduced especially at the high frequencies.

The main question of our study was whether long-distance synchrony is a good candidate for being the neural correlate of consciousness. It is possible that we did not obtain synchrony results because it is hard to find the weak synchrony patterns from the MEG signals. In any case the results obtained in this study do not support the view that synchrony is the key mechanism for conscious perception as we did not find that sensory evidence and prior knowledge would both have increased long-distance synchrony.

5

Early effects of prior knowledge on conscious perception

Introduction

It is well known that prior knowledge influences the way we perceive the world (Kersten, Mamassian, & Yuille, 2004; J. Sun & Perona, 1998; Whitaker & McGraw, 2000). It is however debated whether prior knowledge affects perception directly, i.e. facilitates the processes leading up to object-level perception, or only influences later post-perceptual processes (Henderson & Hollingworth, 1999; Hollingworth & Henderson, 1998; Pachella, 1975). The first group of models (e.g. Bar, 2003) proposes that prior knowledge facilitates object identification, whereas the other group of models (e.g. Hollingworth & Henderson, 1998) claims that object identification and prior knowledge are isolated and that the facilitatory effects of prior knowledge are due to response bias or educated guessing (Henderson & Hollingworth, 1999). It has been complicated to settle this issue with solely psychological experiments (Henderson & Hollingworth, 1999).

One possibility to answer whether prior knowledge affects perception directly or only the post-perceptual processes is to investigate with time-resolved neuroimaging methods when during the course of visual processing prior knowledge influences perception. This knowledge would also be important to constrain the computational theories of visual perception (DiCarlo, Zoccolan, & Rust, 2012). Recently, several studies have provided electrophysiological evidence that prior knowledge affects neural processes already before 100 ms (Chaumon, Drouet, & Tallon-Baudry, 2008; Chaumon, Hasboun, Baulac, Adam, & Tallon-Baudry, 2009; Gamond, et al., 2011), which indicates that prior knowledge can potentially impact perception rather early. However, these studies were interested in the effect of unconscious memory on perception and no evidence was provided that this early effect of prior knowledge directly affects the recognition process. Ghuman et al. (2008) went one step further by showing that early synchrony around 230 ms correlates with the reaction time benefit of repeated presentation of visual objects – those subjects who gained behaviorally more from the prior exposure of the objects had an earlier peak in beta range phase locking between the temporal and prefrontal areas (Ghuman, Bar, Dobbins, & Schnyer, 2008). However, the time frame around 190-270 ms is potentially already after the stage of visual object recognition happening around 200 ms post-stimulus (Bachmann, 2000). Also, reaction times (used in Ghuman et al. (2008)) are not necessarily good direct measures of perceptual effects, because they include variability of motor responses.

Here we investigated the early effects of prior knowledge on conscious perception by having targets close to the threshold of conscious perception and by having the subjects report whether they perceived the target or not. Thus, our measure was aimed directly at perceptual experience. Furthermore, to investigate the specificity of the effects of prior knowledge, we compared the beneficial effect of prior knowledge on conscious perception with that of sensory evidence by manipulating both factors independently in one experimental paradigm. To investigate the timing when prior knowledge influences perception we recorded MEG from 24 healthy subjects and correlated the perceptual benefits of prior knowledge with the MEG activity. With the help of MEG source reconstruction, we also located the sources of the neural effects of prior knowledge on conscious perception.

Methods

Subjects

For this chapter the data from the same subjects as in chapter 4 is analyzed. The final sample consisted of 24 subjects (mean age 24.4 years, standard deviation 2.4 years).

Stimuli and procedure

The stimuli and the procedure are similar to the ones used in chapter 3 and identical to the ones used in chapter 4. The basic experimental paradigm has been presented on figures 3.1 and 4.1

Recordings and data analysis

Data acquisition

The data acquisition procedures for MEG and MRI data correspond to those presented in the respective section of chapter 4.

ERF analysis

MEG data were analyzed with Fieldtrip (http://fieldtrip.fcdonders.nl; version 20-09-2010) an open-source MATLAB toolbox and custom made codes. The data were first high-pass and low-pass filtered between 0.1 and 30 Hz, respectively, and subsequently cut into trials from -300 ms to 700 ms relative to stimulus onset. The EOG recordings were manually checked for eye movements and blinks. All trials that were contaminated by artefacts were discarded from further analysis. Some additional trials were discarded due to squid jumps. The remaining trials were averaged according to the 4 experimental conditions and baseline corrected over a 100 ms window prior to the stimulus onset.

For quantifying the ERF effects we analyzed the global field power (GFP), which is a reference independent measure of response strength (Murray, Brunet, & Michel, 2008). GFP is equivalent to the spatial standard deviation of the magnetic field and is calculated as the square root of the mean of the squared value recorded at each sensor. GFP allows one to study the response strength differences between conditions without a prior selection of electrodes. Before the statistical analysis the GFPs were down-sampled to 200 Hz.

To study the timing of the effects of prior knowledge we used Spearman correlation to quantify the relationship between the perceptual effect in the behavioral measures and the difference of the GFP to pictures with and without prior knowledge across subjects over the time points 50 to 500 ms. The statistical effects were FDR corrected over the time-points.

Source reconstruction

Source activity was reconstructed with SPM8 (http://www.fil.ion.ucl.ac.uk/spm/; version 10-05-2010) an open-source MATLAB software package. First, individual forward models were created for each subject based on their MRIs. Second, the inverse reconstruction was performed for each experimental condition separately via a group inversion step where the condition specific ERFs of all subjects were inverted together to ensure consistency over the individual inverse models. We used the Multiple Sparse Priors (MSP) algorithm for the group inversion (Friston, et al., 2008), which has been shown to give the best results compared to other inversion algorithms (Henson, Mattout, Phillips, & Friston, 2009).

Results

Behavioral results

The behavioral results correspond to those from chapter 4, but as they are important for the main results of this chapter, they are illustrated again here. Recognition was strongly modulated by the amount of sensory evidence as well as prior knowledge and the corresponding pattern of results was the same for the objective and the subjective measure of recognition (Figure 5.1). The similarity between the two measures is also evident from a high positive correlation between them ($r = .82$).

Figure 5.1 (left) illustrates how the objective recognition performance was modulated by the amount of sensory evidence and prior knowledge, i.e. the subjects were more correct when the pictures had either lower degradation or when they had been familiarized with the clear version of the image. A two-way within subject ANOVA revealed a main effect of sensory evidence ($F(1,23) = 116.71$, $P < 1.746E\text{-}10$) and a main effect of prior knowledge ($F(1,23) = 77.812$, $P < 7.71E\text{-}09$). The interaction between sensory evidence and prior knowledge was not significant ($F(1,23) = 0.08$, $P = 0.7789$).

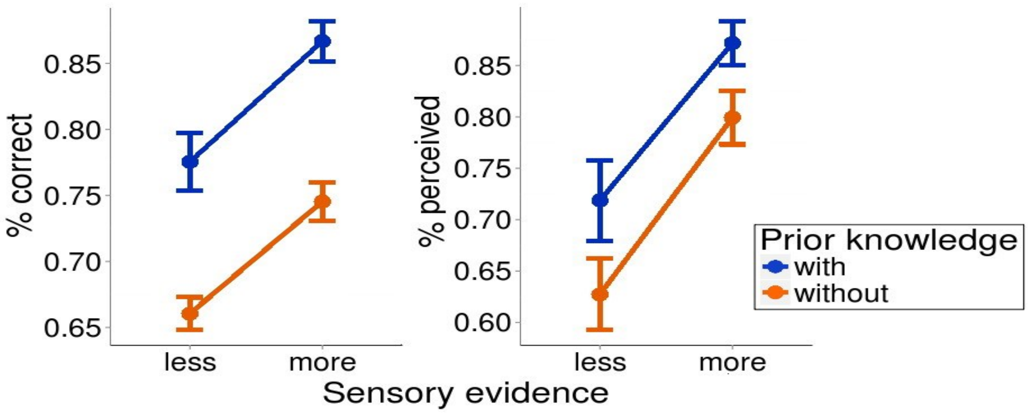


Figure 5.1: Behavioral results. Both sensory evidence and prior knowledge affect objective performance (left panel) and subjective perception (right panel).

But not only the objective recognition performance improved with increasing sensory evidence or prior knowledge. An identical pattern of results was obtained for the subjective measure of recognition showing that our experimental factors also had an effect on the participants' subjective impression of visibility when the amount of degradation decreased or when a stimulus was familiar from before (Figure 5.1 right). As for the objective measure, a two-way within subject ANOVA revealed a highly significant main effect for sensory evidence (F(1,23) = 69.109, P = 2.206E-08) and a highly significant main effect for prior knowledge (F(1,23) = 70.167, P = 1.931E-08). Again, the interaction was not significant (F(1,23) = 0.6376, P = 0.4328).

The results of the behavioral analysis allow for an insight into the mechanisms of recognition. Not surprisingly, recognition depends strongly on the quality of sensory input. The more impoverished sensory input is the less likely it becomes that the visual system can find enough information about the identity of the presented stimulus. In this sense a minimum amount of unambiguous sensory evidence is the basis for every successful recognition process. But our results line up with many previous works showing that this input-dependent threshold of recognition is not fixed (Kleinschmidt, Buchel, Hutton, Friston, & Frackowiak, 2002; Melloni, et al., 2011). Even if there is not enough evidence available in the sensory input itself, recognition can still be rendered possible if there exists reliable prior knowledge about the stimulus.

ERF results

The main aim of this study was to unravel at which time point during the recognition process perception is supported by prior knowledge. This time point cannot be identified with a direct comparison between trials with and without prior knowledge, as these conditions are associated with other differences that might not be related to the perceptual effects. Thus, to find the time when prior knowledge contributes to perception we relied on the fact that prior knowledge leads to improvements in perception that are different across subjects. This individual improvement can be quantified as the difference of the percentage of “person perceived” trials (subjective perception) or correct trials (accuracy) between the conditions with and without prior knowledge. Some subjects benefit more and some less from prior knowledge. This individual variation is then correlated with the neural measures. The idea is similar to the one applied in MRI research of interindividual differences (e.g. Genc, Bergmann, Singer, & Kohler, 2011; for review see Kanai & Rees, 2011), only we correlate the interindividual differences not over space (the characteristics of different brain areas) but rather over time. For the time variable we used the well defined measure of global field power (GFP, see methods), which captures the overall response strength (Murray, et al., 2008). We correlated the individual perceptual effects of accuracy and subjective perception due to prior knowledge with the difference of the GFP to pictures with and without prior knowledge across subjects. If prior knowledge affects perception, there should be a significant correlation and the time point of this significant correlation provides us with a measure of how early prior knowledge affects perception.

With that approach we observed a significant correlation between the difference of the GFP of pictures with and without prior knowledge and the perceptual effect of subjective perception of prior knowledge. This correlation was early in time, before 100 ms (uncorrected $p = 0.0007$, FDR corrected $p < 0.05$ over 80-95 ms), strongly suggesting that prior knowledge affects perception directly, i.e. that the effect is not post-perceptual as the subjective percept of a stimulus is thought to arise only around 200 ms after stimulus onset or even later (Bachmann, 2000; Dehaene & Changeux, 2011; Koivisto & Revonsuo, 2010).

We compared these effects with those of sensory evidence by also correlating the perceptual effect of sensory evidence with the difference of the GFPs to pictures with more and less sensory evidence. We found that the observed effects were specific for the enhancement of perception through prior knowledge as no comparable correlation was found for the enhancement of perception through providing more sensory evidence (over the time 50-500 ms the uncorrected minimum $p = 0.037$). Furthermore, our effect was specific for subjective perception, as we did not observe a correlation between the gain in accuracy and the GFP difference of pictures with and without prior knowledge (all $p > 0.1$).

Figure 5.2 (left) shows that the observed correlation between the perceptual effect and the difference of the GFP of pictures with and without prior knowledge was negative: the greater the difference between the subject's GFP response to the pictures without and with prior knowledge the bigger the perceptual effect of prior knowledge, i.e. pictures with prior knowledge lead to weaker GFP than pictures without prior knowledge and this difference correlated with the perceptual effect of providing prior knowledge. This result fits nicely with the theoretical framework that top-down information suppresses expected sensory input (Friston, 2005, 2010; Mumford, 1992; Rao & Ballard, 1999). Such top-down predictions should lead to weaker sensory responses to pictures with prior knowledge and at the same time be accompanied by more efficient perceptual processing of pictures with prior knowledge, exactly as observed in our experiment.

Our analysis first looked at the correlation between the perceptual effect and the difference of the GFP for pictures with and without prior knowledge. Thus, it could be that the correlation between the perceptual effect and the difference of the GFPs is not specific for the *difference* between the GFPs evoked by pictures with and without prior knowledge, but rather reflects differences between the GFP response to *either* the pictures with *or* without prior knowledge. However, when we only correlated the perceptual effect of the subjects with the GFPs to either the pictures with or without prior knowledge, we did not observe a significant correlation in the respective time window (Figure 5.2 right). Thus, the observed correlation between global brain responses and the effect of prior knowledge on perception is specific for the difference between responses to pictures with and without prior knowledge.

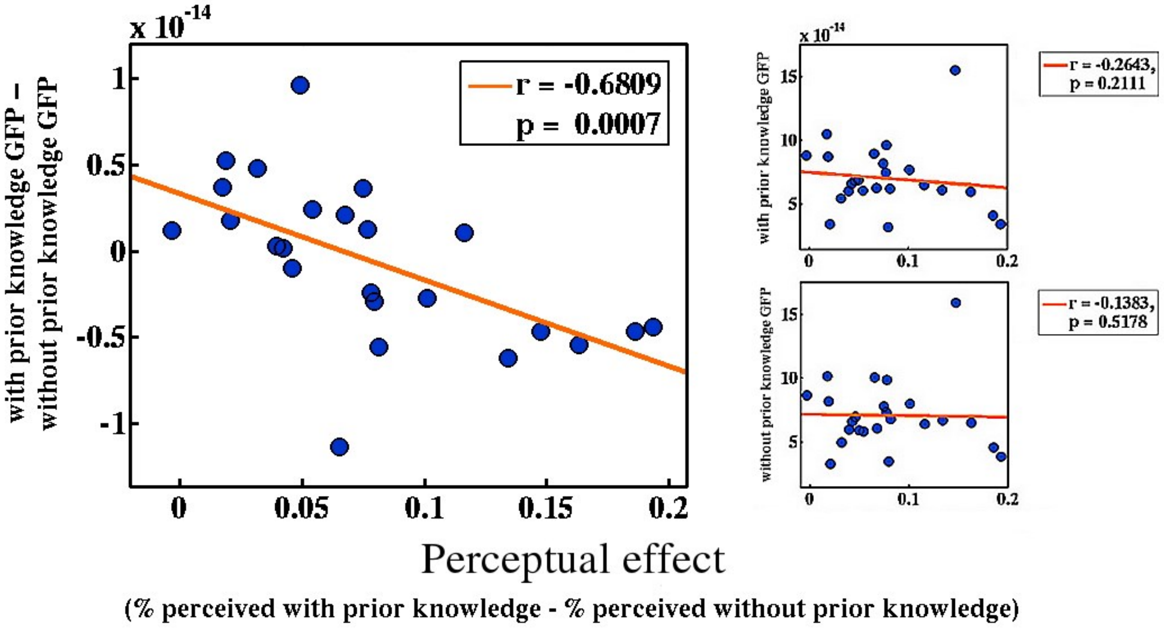


Figure 5.2: the correlation over the time frame 80-95 ms is specific for the difference of the GFPs between the conditions with and without prior knowledge. (Left) Scatter plot of the correlation between the GFP difference between trials with and without prior knowledge and the perceptual effect. (Top right) Scatter plot for the same time window only for the correlation between the perceptual effect and the GFP from the condition with prior knowledge. (Bottom right) Scatter plot for the same time window only for the correlation between the perceptual effect and the GFP for the condition without prior knowledge.

Source localization of the ERF effects

We next sought so search for the neural sources underlying the observed correlation between the perceptual effects of prior knowledge and the measured GFP responses. We localized the neural sources with the Multiple Sparse Priors (MSP) algorithm (Friston, et al., 2008) as implemented in SPM8. To locate the sources underlying the correlation described above, we source localized the MEG activity over the corresponding time window of 80-95 milliseconds and searched for neural sources where the activity difference between trials with and without prior knowledge would be negatively correlated with the perceptual effect of prior knowledge, just as we had observed over this time window in the GFP-analysis.

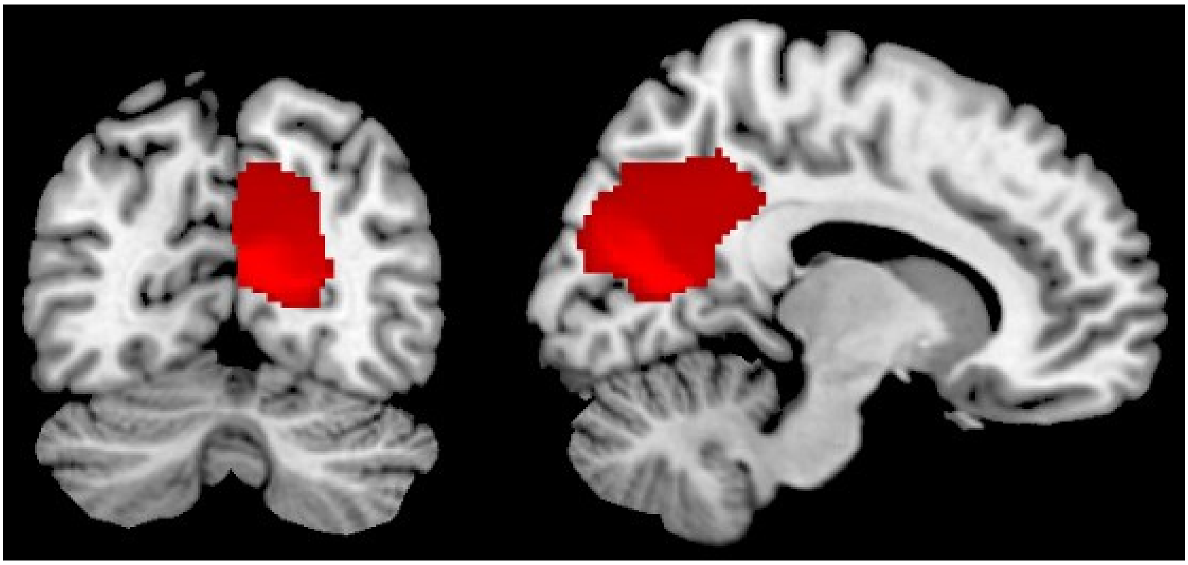


Figure 5.3: Neural sources underlying the temporal correlation observed in the GFP analysis. The plots (red) show the localization of source activity difference between trials with and without prior knowledge in the occipital and parietal lobe that are negatively correlated with the perceptual effect of prior knowledge in the time interval 80-95 ms after onset of the picture.

This analysis revealed a significant cluster (cluster level FDR corrected p value = 0.006) that included the early visual areas and the precuneus (peak in the calcarine sulcus at MNI 10; -76; 18) as the sources underlying the observed correlation (Figure 5.3).

Discussion

The aim of this study was to investigate when during the recognition process prior knowledge has an effect on the neural processes related to perception. To that end we made use of the interindividual variability of the perceptual effects to track the timing of the respective neural processes. We correlated the perceptual effects with the neural effects across subjects. We observed that prior knowledge affected subjective visibility of the targets early in time (80-95 ms post-stimulus) so that the greater the difference between the subject's brain response to the pictures without and with prior knowledge the bigger the perceptual effect of prior knowledge. This effect was specific for subjective visibility, i.e. there was no correlation with objective performance of detection. Also, no similar correlation was observed for perceptual enhancement (neither for objective performance nor for subjective perception) through sensory evidence. Hence, the early correlation between subjective perception and prior knowledge was not caused by a general perceptual enhancement, but rather provides evidence for an early specific effect of prior knowledge on conscious perception. In other words, although sensory evidence also led to perceptual enhancement, this effect was not manifested in the correlation with the GFP, which means that this early correlation between subjective perception and prior knowledge has to reflect specific effects of prior knowledge. We localized the source of this early effect to the early visual cortex and precuneus, areas which are important for processing visual information and memory content.

The central function of the brain might be prediction of the upcoming events in the environment (Friston, 2010; Hawkins 2004; Llinas, 2001). Therefore, it is not surprising that prior knowledge can support conscious perception early in the time course of object recognition. Maybe at the first sight it would seem more surprising that the relationship between the effect of perception and the neural activity was a negative correlation - the stronger the difference of neural activity to pictures without prior knowledge as compared to the pictures with prior knowledge the bigger the perceptual effect of prior knowledge. However, this result is less intriguing once one considers it within the framework of predictive coding (Friston, 2005, 2010; Mumford, 1992; Rao & Ballard, 1999).

Over the recent years it has been shown consistently that predicted stimuli elicit weaker neural responses than unpredicted stimuli (Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010; Kok, Jehee, & de Lange, 2012; Summerfield, Trittschuh, Monti, Mesulam, & Egner, 2008; Todorovic, van Ede, Maris, & de Lange, 2011). Neural modelling and theoretical work suggest that expected stimuli lead to quicker matching with the memory templates and therefore to smaller neural responses; on the other hand, unexpected stimuli lead to greater prediction errors in sensory cortices that cause stronger neural activity (Feldman & Friston, 2010; Friston, 2010; Wacongne, Changeux, & Dehaene, 2012). Our source localization results agree with these theoretical ideas and empirical results where prior knowledge leads to weaker neural responses in sensory cortices (Alink, et al., 2010; Kok, et al., 2012; Summerfield, et al., 2008; Todorovic, et al., 2011).

By familiarizing half of the pictures, prior knowledge about these pictures is created in the brain by forming some neural templates for these pictures and their contents. The familiarization phase is not very long, so there will be interindividual differences in the quality and stability of the memory templates that could be formed during that short time. Presumably, the better the memory template, the better the perception, as the memory template helps to perceive the contents of the otherwise degraded stimulus. Our central finding is that the size of the perceptual effect due to prior knowledge is correlated early in time with the difference of the responses in the early visual cortex to pictures with and without prior knowledge. Importantly, this correlation is negative, i.e. the weaker the neural response to pictures with prior knowledge as compared to the pictures without prior knowledge, the bigger the perceptual effect. As we know from the work of on the sensory systems, expected and predicted input leads to weaker neural responses already at early stages of visual processing (Alink, et al., 2010; Kok, et al., 2012; Summerfield, et al., 2008; Todorovic, et al., 2011). Thus, the weakening of the responses to pictures with prior knowledge could be taken as an index of the quality of the expectation. In turn, the quality of the expectation most likely reflects the quality and stability of the memory template that was created during familiarization. Thus, the weaker the response to pictures with prior knowledge as compared to the response to pictures without prior knowledge, the better the quality of the memory template that was formed during the familiarization phase. Considering this, it is not surprising that these subjects who have weaker response to pictures with prior knowledge as compared to the response to pictures without prior knowledge, benefit more perceptually – if the weakening of the responses to pictures with prior knowledge is an index of the quality of the memory template and the quality of the memory template determines the perceptual performance, it is only to be expected that there exists exactly the observed negative correlation. However, considering this explanation it *is* surprising that no such correlation was observed for the accuracy of the responses (the objective performance), as prior knowledge also increased them. Why did we find an early negative correlation of the prior knowledge effect only for subjective perception but not for accuracy?

Although it is tempting to think that subjective experience and objective performance should always go hand-in-hand, the research over the last decades has shown that they indeed can dissociate. It is well known that it is possible to achieve above chance performance in a task without consciously perceiving the stimuli (e.g. reviewed in Kouider & Dehaene, 2007). This shows that objective performance can change without the concomitant changes in subjective experience. In addition, several studies have demonstrated that even when objective performance is held constant, subjective experience can change (Lau & Passingham, 2006; Schwiedrzik, et al., 2011). Thus, objective performance and conscious experience can vary independently of each other and are most likely supported by different neural pathways (Lau & Passingham, 2006; Schwiedrzik, et al., 2011). Our result, that an early negative correlation with the neural effect of prior knowledge is found only for subjective experience but not for objective performance, adds to this list of growing evidence for the dissociation of subjective experience and objective performance. In particular, our results suggest that prior knowledge can facilitate conscious experience by affecting neural processes early in time in the early visual cortex and that this facilitation is specific for conscious experience as objective performance does not benefit from it.

To conclude, our behavioral results demonstrated that prior knowledge has a beneficial effect on conscious perception. The temporal analysis revealed that this effect of prior knowledge on conscious perception is early in time – between 80-95 milliseconds after stimulus onset. This result clearly indicates that prior knowledge does not only affect post-perceptual processes (Henderson & Hollingworth, 1999), but can have a direct swift impact on perception.

6

General discussion

This thesis had two main goals: 1) make it very explicit why the simple contrastive analysis will not lead to unraveling the neural correlates of consciousness (chapter 2) and with that, motivate the community to go beyond the simple contrastive analysis; 2) to derive clear predictions of two particular theories of consciousness and test them with a relatively straightforward approach (chapters 3 and 4).

The simple contrastive analysis has to be abandoned

Contrastive analysis, i.e. contrasting trials with and without conscious perception, has served consciousness research well (for reviews Dehaene & Changeux, 2011; Rees, et al., 2002). However, science advances by advancing its tools and scrutinizing its methodology. In this thesis it was argued that the main reason for why we have failed to find universally accepted markers of NCC is that the experimental methods typically used to study the NCC are not specific for NCC and also unravel unconscious processes that precede or follow conscious experience. In other words, contrasting trials with and without conscious perception of a particular target gives us more processes than just the NCC or the macro-scale markers of the NCC (Aru, Bachmann, et al., 2012; Bachmann, 2009; de Graaf, et al., 2012; Melloni & Singer, 2010).

It is legitimate to wonder whether it is really needed to distinguish between the different types of NCC or whether a generic concept of “processes correlating with conscious experience” is sufficient. For example, consequences of conscious perception are processes that are most likely directly related to the functions of consciousness (Seth, 2009). Therefore, studying NCC-co is an important endeavor for understanding the phenomenon of consciousness (and how and why it evolved).

However, if the goal is to uncover which neural mechanisms underlie conscious experience and are the NCC, "the minimal set" of neural processes underlying a particular conscious experience, it is clearly required that our experimental paradigms indeed aim at unraveling this "minimal set" of neural mechanisms. Therefore, if the present experimental paradigms are not specific for the NCC, the quest for the NCC is bound to fail.

If we take the possibility seriously that the methods we have used are not specific for revealing the NCC, we have to re-evaluate the whole literature where the contrastive analysis has been applied. It becomes obvious that we actually do not know *which* of the previous results reflect NCC and which correspond to these other confounding processes. Maybe the problem is not severe, but maybe most of the studies about the "neural correlates of consciousness" are reporting processes that in reality precede or follow conscious perception. In fact, we do not know if *any* of the reported results reflect NCC. Popular theories by prominent researchers are based on the results from the previous decades of work, but if we consider the problem with the contrastive analysis as commonly used, it becomes evident that some of these theories about conscious perception might be partly based on results that actually do not reflect correlates of conscious experience. And the worst issue is that we currently really do not know which results are indeed reflecting the NCC and therefore, which theories are more affected by this problem. It is crucial that researchers acknowledge that the simple contrasts between trials with and without conscious perception of a target are simply not informative for our purposes – such contrasts will stay unspecific and we will not know if the resultant neural activity is reflecting NCC or some other processes. The science of consciousness needs to move beyond this simple contrast and consider experimental paradigms that allow one to untangle the NCC from these other confounding processes.

Moving beyond the contrastive analysis: new paradigms

For the empirical studies presented in this thesis a new experimental paradigm was designed to resolve some of the methodological problems with the commonly used contrastive analysis. In that experimental paradigm conscious perception was modulated either by sensory evidence or by prior knowledge. The logic of our approach was the following: although conscious perception could be facilitated through these different prerequisites, conscious experience is enhanced similarly by both manipulations and thus the NCC should also change similarly by both sensory evidence and prior knowledge. As we manipulate the prerequisites for conscious perception, it is obvious that this experimental paradigm is focused on and also limited to the exclusion of confounding processes that precede conscious perception. Thus, such a paradigm is not a solution to the general problem but can only be a part of the “distillation process”. Furthermore, this paradigm is not exhaustive as there may be other prerequisites than sensory evidence and prior knowledge: it could be complemented by other manipulations of other other variables influencing conscious perception such as attention, working memory or TMS which are all known to modulate access to consciousness (Aru, Bachmann, et al., 2012). The processes that are equally enhanced by all these manipulations hand-in-hand with enhanced subjective perception are much more likely to be related to the real NCC. Conversely, processes that only show up by one manipulation (e.g. manipulation of sensory evidence as in chapter 3) are more likely to be prerequisites of conscious experience.

However, one must acknowledge that it is possible that all these influences converge on some similar processes that are still not the NCC but precede the NCC as prerequisites. In reality, it might be extremely complicated to dissociate the NCC from its prerequisites and consequences. Also, as only the prerequisites were manipulated in the present study, such paradigms would not solve the problem to distinguish the NCC proper from the consequences of conscious perception. Thus, it can be concluded that more stringent criteria for ruling out prerequisites and consequences are required.

A general approach for isolating the NCC could be the following: one should exploit the whole range of the phenomenal experience. In most experiments, conditions are often created by separating trials with and without conscious perception. Yet, conscious experience is much richer that this often simplified dichotomy, for instance: 1) across trials, the quality of perceiving a simple target stimulus varies gradually (Seth, Dienes, Cleeremans, Overgaard, & Pessoa, 2008), 2) within any given trial, the quality of subjective perception changes over time, a process termed microgenesis (Bachmann, 2000), 3) if more complex stimuli are used, the perceived contents and their attributes might vary over trials, 4) with learning, conscious perception of stimuli changes over time. Some of these points and the respective experimental approaches are illustrated and explained in the next paragraphs.

We feel that the community should aim at using such fine-grained information about subjective perception to identify neural processes whose changes go hand-in-hand with these fine-grained variations of conscious experience. There should be a tight mapping between nuanced changes of NCC and accordingly nuanced variations in conscious experience (Haynes, 2009). The key idea for distilling the NCC is that while these processes that correspond to NCC should vary according to the subjective perception, the confounding prerequisites for and the consequences of conscious experience do not necessarily change continuously. For example, while conscious experience is a bit clearer when target contrast is increased, the subsequent representation in working memory (NCC-co) might be similar in both cases (Melloni, et al., 2011). It is possible that some of the prerequisites and consequences indeed covary with conscious perception even in such experimental paradigms, but it is clear that neural signatures that vary hand-in-hand with nuances in conscious perception are more likely to be related to conscious experience than those that are revealed by comparing trials with and without conscious perception.

Exploiting the gradual variation of the quality of subjective experience

The gradual variation approach for studying the NCC has been implemented in paradigms with masking and hysteresis (Del Cul, et al., 2007; Melloni, et al., 2011) where visibility of the item was varied gradually, the subjects gave reports about their subjective perception and the researchers asked which components of the event-related potentials changed accordingly. In both cases visibility of the items was manipulated objectively either by contrast (Melloni, et al., 2011) or by changing the stimulus onset asynchrony (SOA; Del Cul, et al., 2007). Thus, changing neural responses could potentially also have been caused by changes in objective stimulation conditions. The trick used by these research groups was the following: whereas the objective change (contrast or SOA) was linear, the subjective change was sigmoidal (Del Cul, et al., 2007; Melloni, et al., 2011). This fact allowed the researchers to tease apart those neural responses that varied linearly together with objective stimulation from those that varied in a sigmoidal fashion hand-in-hand with subjective conscious experience. In addition to directly manipulating visibility, as done in the above mentioned studies, researchers can also take advantage of the known fact that subjective experience varies even under invariant stimulation conditions (Aru & Bachmann, 2009a, 2009b; Bar, et al., 2001). Furthermore, in the above cited studies the mean neural activity over trials was fitted to the mean visibility ratings over trials, but given that the visibility ratings vary over trials even under invariant stimulation conditions this variation could be used to track down the corresponding neural processes underlying this variability on the single trial basis (Rousselet & Pernet, 2011). Also, gradual increases of visibility are not a necessary condition because the the visibility ratings could also take the U-shape of typical metacontrast masking functions (Bachmann, 2000), which could be correlated with changes in neural dynamics (Haynes, Driver, & Rees, 2005). Finally, in the cited studies (Del Cul, et al., 2007; Melloni, et al., 2011) the ERPs were used as the neural measures, but this kind of analysis could well be expanded to analysis of spectral perturbations, phase synchrony, causality and other measures (Melloni, et al., 2007; Gaillard, et al., 2009) and performed with other measurement technologies such as fMRI (Christensen, Ramsoy, Lund, Madsen, & Rowe, 2006; Haynes, et al., 2005).

Let us have a look at some results of the above mentioned studies to illustrate some of the issues that arise when such experimental paradigms are used. Del Cul and colleagues (2007) manipulated the SOA between the target stimulus and the mask stimulus, obtained a sigmoidal function of visibility depending on the SOA and observed that the P300 component of the ERP mimicked the gradual changes of the visibility function. However, in an interesting twist to the story, Melloni and colleagues (2011) demonstrated that the P300 component could reflect a consequence of conscious perception – when the working memory representation of the target was already available, the P300 component did not follow the gradual changes of the visibility function. Del Cul and colleagues (2007) observed that P300 changes match the visibility function while the data of Melloni and colleagues

(2011) suggest that P300 reflects the transfer of perceived material into working memory. Does this imply that the consequences of consciousness are also gradual? This would mean that observing gradual changes of neural measures that follow the changes in subjective perception would not be helpful in disentangling the NCC from their consequences. A closer look explains why this is not necessarily so. Namely the visibility functions represent the average perceptual clarity for the particular SOA that is aggregated over many single trials. In a given single trial, however, the stimulus is either perceived consciously or not (in this experiment in a quite all-or-none manner; this might be different in other experiments). If the target is consciously perceived, the NCC-co are activated. Thus, in very simplified terms, if the stimulus is perceived consciously in 67% trials in a given SOA, the NCC-co are also activated in 67% of trials. Hence, in such experimental paradigms NCC-co processes can correspond to the visibility functions. This thought-exercise illustrates that the best conditions for distilling the NCC are when subjective perception varies gradually from trial-to-trial and these single-trial changes are correlated with the underlying neural activity.

Haynes and colleagues (2005) used a metacontrast masking paradigm, where the visibility depended on the SOA and the corresponding visibility function had a U-shape. This approach is also elegant because neuronal activities reflecting only the objective linear changes of SOA can be simply disregarded. The main take-home message from that study is, however, that not only simple activation patterns can be correlated with visibility but also the effective connectivity between areas (in their case between V1 and fusiform gyrus). Thus, we should not only look at ERP components and regional activation, but at more dynamic measures that could reflect the mechanisms of consciousness more directly. Another important aspect of this study is that the visibility function was based on the objective performance rather than subjective perception. Although we expect that the subjective perception depends on the SOA in a similar way as the objective performance (Bachmann, 2000), these two functions have been shown to differ at least slightly in metacontrast masking (Lau & Passingham, 2006).

Is there any evidence to suggest that the approach of gradual variations of conscious perception could help to distill the NCC from NCCpr and NCCco? For instance, in the above mentioned work of Quiroga and colleagues (Quiroga, et al., 2008) that we take as a prime example for consequences of conscious experience, it was observed that the responses were all-or-none and not modulated by stimulus duration when the stimuli were presented for either 33, 66 or 132 milliseconds, although conscious experience was most likely more vivid with longer stimulus durations. This latter claim is supported by a study by Christensen et al. (2006) where geometric figures were presented for either 33, 50, 83 or 100 ms and it was observed that the visibility ratings increased with the presentation time.

Exploiting the fine-grained variation of the contents of subjective experience

Haynes (2009) makes the clear point that our conscious experience varies along various dimensions and hierarchical levels (one can perceive objects invariant to low level features, but also the low level features themselves can be perceived). A NCC of a particular conscious experience of a complex scene should contain information about every aspect that is consciously perceived. Often experimental paradigms are used where simple objects (lines, line drawings, numbers) either are or are not consciously perceived, but as soon as more complex stimuli are used and more specific questions are asked about the perception of these stimuli, one can use these fine-grained variations of conscious experience to distill the NCC. As mentioned above, whereas slight changes in conscious perception go hand-in-hand with changes in the underlying NCC, it is unlikely that consequences of conscious perception track subtle changes in conscious perception. Various illusions and other experimental phenomena of consciousness offer a variety of possibilities to investigate, which neural processes correlate on the single trial level with the small changes in subjectively perceived brightness, color, size, shape etc (Bachmann, et al., 2011). For that, more complex stimuli have to be presented, the subjects have to be queried about the contents they perceived and machine learning algorithms need to be used to build the bridge between slight changes in neural activation patterns and corresponding variations of subjective experience (Haynes, 2009).

For example, consider the study of Haushofer and colleagues (Haushofer, Livingstone, & Kanwisher, 2008) where the authors studied whether distributed activation patterns in the human visual cortex reflect the perceived shape of the objects. An artificial shape space with known physical similarity was created and subjective reports were used to measure subjective similarity of these shapes. In addition, multivoxel activation patterns were measured for each stimulus, which allowed the authors to compute the neural similarity across the shapes. In the crucial step relevant for our discussion, the three types of similarities (physical, subjective and neural) were compared. Our view would predict that if the shapes are subjectively similar, their NCC should also be similar. And as the physical similarity and the subjective similarity were different, these neural underpinnings of subjective similarity were not confounded by objective physical similarity. The authors observed that subjective similarity correlated with the multivoxel activation patterns in the anterior lateral occipital cortex, whereas physical similarity was represented in the activity of the posterior part of the lateral occipital cortex. This study and similar studies (e.g. Mur, et al., 2013) demonstrate that nuances of subjective perception can be used to find the neural processes that explain these nuances. If neural activation patterns in some area or across many areas faithfully encode the subjective properties of complex stimuli, these patterns are more closely related to the NCC than the prerequisites or consequences of conscious perception.

Conscious perception and local cortical activity

In chapter 3 it was shown that local gamma band responses (GBR) are not the specific correlate of conscious perception. Our research was motivated by the work of Fisch and colleagues (2009) who elegantly demonstrated with intracranial recordings that conscious perception correlates with local GBR in visual cortex. However, in the light of the methodological problem described in this thesis, we conjectured that this local activity might not necessarily reflect conscious perception. Moreover, as early recordings had shown (e.g. Gray & Singer, 1989) local category-specific GBR can be obtained even in anesthetized animals, arguing that such processes are not specific for conscious perception. We thus evaluated the hypothesis that local GBR is the correlate of conscious perception with our experimental paradigm. As both sensory evidence and prior knowledge enhanced conscious perception to the same extent, a clear prediction was that if category-specific GBR in visual cortex indeed reflects conscious perception it should be modulated both by sensory evidence and prior knowledge. Instead, we found that GBR only correlated with sensory evidence, falsifying the hypothesis that local GBR in sensory cortices represents a neural marker of consciousness.

However, this result does not necessarily imply that local cortical activity has no role in conscious perception. Our result – that enhancement through prior knowledge is not related to enhancement of GBR – can be parsimoniously explained when taking into account that the broadband GBR reflects the average population firing rate and that neural activity patterns related to conscious perception do not necessarily have to be associated with stronger average population firing rates. In particular, it could be that increasing sensory evidence increases the average population firing, whereas prior knowledge leads to more specific firing patterns, which convey more information, but are not manifested as activity changes on the level of the average population firing rates (Figure 6.1).

This conjecture is directly supported by the findings from monkey visual cortex obtained by Rainer and colleagues (Rainer, Lee, & Logothetis, 2004). These authors also observed that prior knowledge enhances performance, but more importantly, they found that this enhancement was not evident in the average population firing rates (table 1 in Rainer, et al., 2004). If we agree that the intracranial GBR is a marker for the average population firing rate (Lachaux, et al., 2012; K. J. Miller, 2010; Ray & Maunsell, 2011; Vidal, et al., 2010), our results regarding the effects of prior knowledge in the human visual cortex directly match those of Rainer and colleagues from monkey V4 – increased performance is not evidenced in increased average firing rates. Furthermore, in their work these authors observed similarly to us that the average population firing rate increases monotonically with the increasing sensory evidence (Gregor Rainer, personal communication). Thus, these results nicely correspond to ours: the average population firing rate (and thus the GBR) is sensitive to changes in sensory evidence, but not so for prior knowledge, although both factors enhance performance behaviorally.

In the case of Rainer and colleagues (2004) the correlate of prior knowledge was revealed to be in the firing of a specific subset of neurons – these neurons indeed fired stronger when the picture was familiarized before and the fact that this was not reflected in the average population firing rate suggests that some other neurons were actually firing less as a response to the pictures with prior knowledge. Of course we have in the current experimental setup no possibility to test for such specific effects on the level of single neurons but it can well be that the same phenomenon is happening in our paradigm: due to prior knowledge the firing of some neurons is higher and this specific firing would be related to enhanced perception, but there is no effect on the level of the average population firing rate. This idea is illustrated in Figure 6.1.

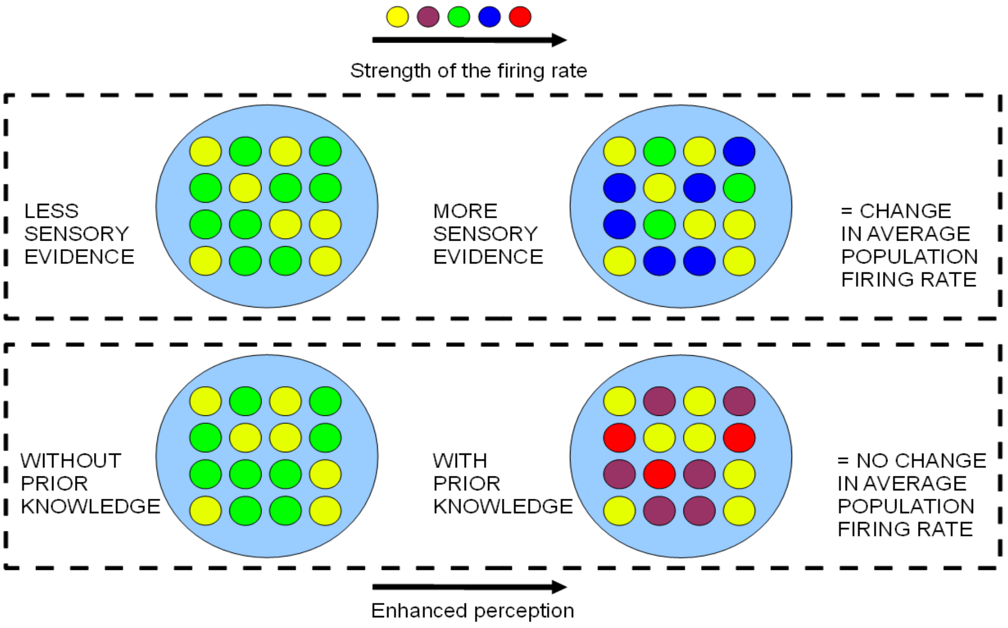


Figure 6.1: A toy model explaining the results reported in chapter 3. Each small circle represents a single neuron and the big circles illustrate our electrodes (each electrode measures the activity of approximately 5 x 10^5 neurons). Different colors indicate differences in the strength of the firing rate according to the scale presented on the top of the figure. As illustrated in the upper panel, increasing the amount of sensory evidence increases the activity of the whole population of neurons representing the corresponding stimulus. On the other hand, as illustrated in the bottom panel, providing prior knowledge about the stimulus will make the neural activation pattern more specific by increasing the activity of some neurons while reducing the activity of others, thus not changing the average population activity.

It has been demonstrated on the level of BOLD responses (Grill-Spector, et al., 2000; Bar, et al., 2001) and intracranial GBR (Fisch, et al., 2009) that the local activity of the visual cortex is related to conscious recognition. Our work agrees with these results by showing that GBR in these higher visual areas correlates with the subjective report about conscious recognition. However, these BOLD and GBRs reflect some kind of average activity of the underlying neural population. Therefore, believing that BOLD and GBRs are the sole correlate of conscious perception is equal to believing that it is the *average activity* of the neural populations that is relevant for conscious perception. We would disagree with that assumption and adhere to the fact that neural information is coded in the fine *temporo-spatial patterns of activity*, which may or may not be reflected in the average population activity. And if we believe that specific neural activity patterns are the key to understanding conscious perception, we need to conclude that the population activity, as captured by BOLD or GBR, does not always have to change hand in hand with conscious perception. There might be instances where the local population activity, and therefore BOLD or GBR, are dissociated from conscious experience. In the light of our results presented in chapter 4 our claim is that prior knowledge leads to specific changes in the patterns of neural activity, which is associated with increased informational content and that it is this increased informational content and not the average population activity that is relevant for enhanced perception (Figure 6.1). We next review recent literature on the neural effects of prior knowledge that provides experimental support for this idea.

From previous work we know that prior knowledge can indeed change the informational content without increasing the population firing rate. In the already discussed work of Rainer et al. (2004) the authors observed that despite the fact that prior knowledge is not associated with stronger population firing rates, the informational content is enhanced for the pictures with prior knowledge even on the population level. These authors used the approach of estimating mutual information, which is unfortunately impossible in the case of our data, because the estimation of mutual information requires much more trials than we had collected here. Nevertheless, in the work of Rainer and colleagues (2004) prior knowledge enhanced performance but the increase in performance was not accompanied by the increase of the average firing rate, but rather by the information that the firing rates conveyed. Furthermore, the work by Hsieh and colleagues (2010) supports the conjecture that prior knowledge changes the informational content rather than the average neural mass activity in human fMRI data: familiarization had an effect on the spatial patterns of fMRI activity (and made them more similar to the pictures from the familiarization phase) but importantly did not change the mean BOLD activity in the visual areas (Hsieh, Vul, & Kanwisher, 2010). Finally, Kok and colleagues (2012) recently showed that fMRI BOLD activity in V1 is reduced by prior knowledge (expectation) whereas the information content about the target object in V1 increases (Kok, et al., 2012). Taken together, these investigations (Hsieh, et al., 2010; Kok, et al., 2012; Rainer, et al., 2004) allow us to speculate that in our paradigm prior knowledge could be associated with enhanced perception because prior knowledge changes the neural firing patterns so that the information about the stimulus can be extracted more easily, whereas the average population firing rates remain unaffected. Given that the cost of every spike is high (Lennie, 2003), it is reasonable that conveying more information with the same average number of spikes in a given brain area is valuable for the brain.

This explanation resonates well with the ideas that our perception is predictive with the general goal to reduce the strength of the neural responses (Friston, 2005, 2010; Mumford, 1992; Rao & Ballard, 1999). Although in our experimental paradigm the subjects could predict neither the timing nor the identity of the upcoming stimuli, predictions also work within the hierarchical cortical processing of an unpredictable stimulus and can thus be facilitated through prior knowledge.

In summary, what do our present results (chapter 3) reveal about the relationship between the activity of visual cortex and conscious recognition? Although previous research has convincingly shown that stronger category specific activity of the visual cortex is associated with conscious perception of the respective content (Grill-Spector, et al., 2000; Bar, et al., 2001; Fisch, et al., 2009), our results suggest that average category specific population activity does not directly reflect conscious perception. This does not mean that previous results are not meaningful: our research agrees that providing more sensory evidence (Grill-Spector, et al., 2000) enhances neural firing and is thus related to increased signals on the population level (e.g., BOLD, GBR). Hence, when studying conscious recognition under invariant stimulation (Bar, et al., 2001; Fisch, et al., 2009), local population activity reflects perception, as in this case perception can be determined by the fluctuations of excitability in the earlier sensory areas (Mathewson, et al., 2009; Busch, et al., 2009), which modulate the amount of sensory evidence that is transmitted from these earlier sensory areas to the higher category-specific areas. Therefore, in both cases, enhanced perception will be related to enhanced average population activity, as also confirmed by our findings on the GBR (chapter 3). However, our results from chapter 3 show that if conscious perception is enhanced through prior knowledge, this perceptual enhancement is not reflected in changes of the average population signals (GBR). Based on previous research we suggest that the perceptual enhancement related to prior knowledge relies on reconfigurations in the fine spatio-temporal patterns of neural activity that increase the informational content conveyed by these activity patterns.

Conscious perception and long-distance communication

Many theories of consciousness hold a key role for long-distance neural communication in explaining consciousness (Singer, 1998; Singer & Melloni, 2010; Bachmann, 2007; Llinas, 1998). It is claimed that important characteristics of conscious experience are explained by synchronization among neural ensembles (Melloni & Singer, 2010). There is also converging evidence that long-range neural synchrony indeed correlates with conscious experience (Melloni, et al., 2007; Gaillard, et al., 2009; Hipp, et al., 2011). However, as was the case with the proposal that local category specific gamma band responses reflect conscious perception (Fisch, et al., 2009; chapter 3) these results have to be re-evaluated due to the methodological problem with the contrastive analysis that has been in the focus of this thesis. Therefore, one of the aims of the present work was to test the hypothesis that long-distance synchrony is the neural correlate of consciousness with our experimental paradigm.

In our study (chapter 4) we did not find support for the hypothesis that neural synchrony is the mechanism for consciousness. In particular we did not find that sensory evidence and prior knowledge would both increase long-distance synchrony between specific sources. We observed that prior knowledge increased source unspecific synchrony in the gamma range, but we found no similar effects for sensory evidence.

When we analyzed local category specific gamma band responses and found that they are increased neither by sensory evidence nor by prior knowledge we concluded that these activity patterns cannot reflect conscious perception. Now, in the case of long-range synchrony we are much softer, claiming that “our results do not support the hypothesis”. This difference in the strength of the conclusions stems from the fact that whereas the hypothesis that local category specific gamma band responses are the correlate of conscious perception is a very specific hypothesis – it defines the brain area and the type of activity (the frequency band), the hypothesis that long-distance synchrony is the correlate of conscious perception is much less concrete. In particular, this hypothesis defines neither the areas between which nor the frequencies where this synchrony effect should be found. Therefore, by exploring the full space of source-time-frequency combinations and correcting for multiple comparisons over that space we might have missed synchrony effects. However, this possibility is made less likely by the fact that even at the uncorrected level there were no clear synchrony effects. We have already discussed some possible reasons for why we did not observe synchrony patterns between specific sources in the gamma range (chapter 4). Namely, 1) beamforming could have suppressed some simultaneously active sources (e.g. early visual cortices or the body-specific areas of the both hemispheres) and 2) the synchrony effects in the brain are numerically too small to be captured in MEG recordings. Therefore, we cannot conclude from our results that synchrony is not the correlate of conscious perception. We can only conclude that our results do not support this hypothesis. Finally, it could be that the real synchrony effects take place at a finer temporal and spatial scale than the present methods for MEG analysis allow one to assess (Singer, in press).

In the previous section we have discussed that the information processing of the brain does not occur at the scale captured by the intracranial gamma band responses, but at a finer scale. This is even more so the case with MEG or EEG recordings, which record much more global signals than the intracranial recordings. For example if one compares the local gamma band responses from our intracranial study (Figure 3.2) and from the MEG source reconstruction (Figure

4.3), it is clear that the intracranial signals are an order of magnitude stronger than the source signals reconstructed from MEG recordings. This does not mean that nothing meaningful about neural synchrony can be unraveled from EEG and MEG as there are indeed many empirical studies finding effects of phase synchrony in EEG and MEG recordings (e.g. Cosmelli, et al., 2004; Ghuman, et al., 2008; Hipp, et al., 2011; Melloni, et al., 2007). However, it implies that the quest for understanding neural dynamics and the role of synchrony in brain functioning might be even more fruitful if one would search at a finer measurement scale (Singer, in press). This requires massive parallel recordings from single neurons and local populations to enhance the resolution spatially, but also novel data analysis tools for studying neural dynamics at a finer temporal scale. As Singer (in press) argues, the present analysis tools are only able to recover stationary and regular patterns of synchronization, but the interesting activation patterns reconfigure over a much quicker time scale and might change their frequencies dynamically. In the present analysis of MEG source synchrony these swift processes might have remained unseen, because if they are transient and change in frequency, they cannot be revealed by the traditional time-frequency decomposition approach. For example, the stimulus onset is associated with a non-stationarity of the measured signal and any non-stationarity is difficult to interpret in the frequency domain. Therefore, the frequency domain effects are often not analyzed for the first transients of the data even in good-quality intracranial or local field potential recordings (e.g. Bosman, et al., 2012). In fact, some of these early transient dynamics might be even better reflected in the classical event related potential (ERP) analysis than in the currently more popular analysis of MEG phase and power. This is because ERP has the full time resolution, but any frequency-domain analysis looses the temporal resolution due to the uncertainty principle of the harmonic analysis. And, indeed, in our own work (chapter 5) we have observed reliable effects of prior knowledge on conscious perception that were revealed with the analysis of global field power (GFP), a general measure of the response strength in the time domain. The fact that from the very same data we had difficulties observing synchrony effects while we obtained relatively strong effects from the GFP analysis support the view that at least under some circumstances sacrificing locality in time for the analysis of frequency-specific effects might not be advisable in MEG or EEG recordings. In general, new analysis tools are needed that go beyond the simple time-frequency decompositions and try to unravel the transient non-linear dynamics of the brain (Singer, in press). In other words, the present analysis methods severely underestimate the complexity of the brain. This complexity is daunting but if one wants to understand the brain, one has to fully embrace it!

Conclusions and outlook

In 1996 Francis Crick wrote: “Consciousness is now largely a scientific problem. It is not impossible, that with a little luck, we may glimpse the outline of the solution before the end of the century.” (Crick, 1996). Despite this early optimism, we are not close to solving the problem of consciousness. Since the 1990s researchers have been on the quest for the neural markers and correlates of conscious perception. However, the collective efforts have not even led to a consensus about the most reliable markers of consciousness, let alone the neurobiological mechanisms behind the phenomenon.

Here in this thesis it has been argued that one key reason for the ongoing debates is that the experimental paradigms are not specific enough for capturing the neural markers of conscious perception (Aru, Bachmann, et al., 2012; De Graaf, et al., 2012). Hence, what appears as a neuronal correlate of consciousness (NCC) in experiments comparing conscious and unconscious processing could also reflect differences in early processing that determine access to consciousness or else consequences of conscious processing such as transfer of contents to working memory. These various attributes of conscious processing are difficult to disentangle in experimental paradigms where trials with and without conscious perception are contrasted. It has been suggested, therefore, that the science of consciousness needs to move beyond such paradigms (Aru, Bachmann, et al., 2012).

Here in this thesis a relatively simple new experimental paradigm has been used, where conscious perception was manipulated either through sensory evidence or through prior knowledge. It is acknowledged that this experimental paradigm cannot solve the methodological problem of the prerequisites and consequences of conscious perception to the full extent, but it is hoped that it can at least contribute to the ongoing process of distilling the neural correlates of conscious perception. With our paradigm we have successfully demonstrated that local category-specific GBR do not reflect conscious perception directly. With that result we could falsify the idea that such local category-specific GBR is a faithful correlate of consciousness. Our research suggests that such local GBR rather reflect local sensory processing. With the same experimental paradigm we tried to test the idea that long-range synchrony is the correlate of conscious perception. Our results do not support the hypothesis that synchrony reflects conscious perception directly. Finally, with the help of this paradigm we were able to demonstrate that prior knowledge affects conscious perception early in time and has a specific effect on perception that is distinct from the effects of sensory evidence.

The problem of consciousness can obviously not be solved within one PhD thesis. It is currently even hard to estimate whether it will ever be solved or whether science will always have a limit when it comes to understanding conscious experience (McGinn, 1989). Are we just missing that “little luck” Francis Crick alluded to or something more fundamental?

References

Abrahamyan, A., Clifford, C. W., Arabzadeh, E., & Harris, J. A. (2011). Improving visual sensitivity with subthreshold transcranial magnetic stimulation. *J Neurosci, 31*(9), 3290-3294.

Alais, D., Cass, J., O'Shea, R. P., & Blake, R. (2010). Visual sensitivity underlying changes in visual consciousness. *Curr Biol, 20*(15), 1362-1367.

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *J Neurosci, 30*(8), 2960-2966.

Aru, J., Axmacher, N., Do Lam, A. T., Fell, J., Elger, C. E., Singer, W., et al. (2012). Local category-specific gamma band responses in the visual cortex do not reflect conscious perception. *J Neurosci, 32*(43), 14909-14914.

Aru, J., & Bachmann, T. (2009a). Boosting up gamma-band oscillations leaves target-stimulus in masking out of awareness: Explaining an apparent paradox. *Neurosci Lett, 450*(3), 351-355.

Aru, J., & Bachmann, T. (2009b). Occipital EEG correlates of conscious awareness when subjective target shine-through and effective visual masking are compared: bifocal early increase in gamma power and speed-up of P1. *Brain Res, 1271*, 60-73.

Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neurosci Biobehav Rev, 36*(2), 737-746.

Axmacher, N., & Fell, J. (2011). The role of phase synchronization in memory processes. *Nat Rev Neurosci,* 12(2), 105-118.

Baars, B. J. (1989). *A cognitive theory of consciousness*. New York: Cambridge University Press.

Bachmann, T. (2000). *Microgenetic Approach to the Conscious Mind*. Amsterdam / Philadelphia: John Benjamins.

Bachmann, T. (2007). Binding binding: Departure points for a different version of the perceptual retouch theory. *Adv Cogn Psychol, 3*(1-2), 41-55.

Bachmann, T. (2009). Finding ERP-signatures of target awareness: puzzle persists because of experimental co-variation of the objective and subjective variables. *Conscious Cogn, 18*(3), 804-808; discussion 809-810.

Bachmann, T., Breitmeyer, B. G., & Öğmen, H. (2011). *The Experimental Phenomena of Consciousnes: a brief dictionary revised edition*. New York: Oxford University Press.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cogn Neurosci, 15*(4), 600-609.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., et al. (2006). Top-down facilitation of visual recognition. *Proc Natl Acad Sci U S A, 103*(2), 449-454.

Bar, M., Tootell, R. B., Schacter, D. L., Greve, D. N., Fischl, B., Mendola, J. D., et al. (2001). Cortical mechanisms specific to explicit visual object recognition. *Neuron, 29*(2), 529-535.

Barbur, J. L., Watson, J. D., Frackowiak, R. S., & Zeki, S. (1993). Conscious visual perception without V1. *Brain, 116 (Pt 6)*, 1293-1302.

Blake, R., & Logothetis, N. (2002). Visual competition. *Nat Rev Neurosci, 3*(1), 13-21.

Bosman, C. A., Schoffelen, J. M., Brunet, N., Oostenveld, R., Bastos, A. M., Womelsdorf, T., et al. (2012). Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron, 75*(5), 875-888.

Brown, J.W. (1988). The Life of the Mind. Hillsdale: Erlbaum.

Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci, 29*(24), 7869-7876.

Carmel, D., Walsh, V., Lavie, N., & Rees, G. (2010). Right parietal TMS shortens dominance durations in binocular rivalry. *Curr Biol, 20*(18), R799-800.

Chaumon, M., Drouet, V., & Tallon-Baudry, C. (2008). Unconscious associative memory affects visual processing before 100 ms. *J Vis, 8*(3), 10 11-10.

Chaumon, M., Hasboun, D., Baulac, M., Adam, C., & Tallon-Baudry, C. (2009). Unconscious contextual memory affects early responses in the anterior temporal lobe. *Brain Res, 1285*, 77-87.

Christensen, M. S., Ramsoy, T. Z., Lund, T. E., Madsen, K. H., & Rowe, J. B. (2006). An fMRI study of the neural correlates of graded visual perception. *Neuroimage, 31*(4), 1711-1725.

Cosmelli, D., David, O., Lachaux, J. P., Martinerie, J., Garnero, L., Renault, B., et al. (2004). Waves of consciousness: ongoing cortical patterns during binocular rivalry. *Neuroimage, 23*(1), 128-140.

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

Crick, F., & Koch, C. (1990). Toward a neurobiological theory of consciousness. *Seminars in Neurosciences, 2*, 263-275.

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

Crick, F., & Koch, C. (2003). A framework for consciousness. *Nat Neurosci, 6*(2), 119-126.

Dalal, S. S., Baillet, S., Adam, C., Ducorps, A., Schwartz, D., Jerbi, K., et al. (2009). Simultaneous MEG and intracranial EEG recordings during attentive reading. *Neuroimage, 45*(4), 1289-1304.

de Graaf, T. A., de Jong, M. C., Goebel, R., van Ee, R., & Sack, A. T. (2011). On the functional relevance of frontal cortex for passive and voluntarily controlled bistable vision. *Cereb Cortex, 21*(10), 2322-2331.

de Graaf, T. A., Hsieh, P. J., & Sack, A. T. (2012). The 'correlates' in neural correlates of consciousness. *Neurosci Biobehav Rev, 36*(1), 191-197.

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70*(2), 200-227.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends Cogn Sci, 10*(5), 204-211.

Dehaene, S., & 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., et al. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nat Neurosci, 4*(7), 752-758.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biol, 5*(10), e260.

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., & Slachevsky, A. (2009). Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain, 132*(Pt 9), 2531-2540.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron, 73*(3), 415-434.

Doesburg, S. M., Green, J. J., McDonald, J. J., & Ward, L. M. (2009). Rhythms of consciousness: binocular rivalry reveals large-scale oscillatory network dynamics mediating visual perception. *PLoS One, 4*(7), e6142.

Donner, T. H., & Siegel, M. (2011). A framework for local cortical oscillation patterns. *Trends Cogn Sci,*15(5), 191-199.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science, 293*(5539), 2470-2473.

Dugue, L., Marque, P., & VanRullen, R. (2011). The phase of ongoing oscillations mediates the causal relation between brain excitation and visual perception. *J Neurosci, 31*(33), 11889-11893.

Engel, A. K., Fries, P., Konig, P., Brecht, M., & Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Conscious Cogn, 8*(2), 128-151.

Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn Sci, 5*(1), 16-25.

Ergenoglu, T., Demiralp, T., Bayraktaroglu, Z., Ergen, M., Beydagi, H., & Uresin, Y. (2004). Alpha rhythm of the EEG modulates visual detection performance in humans. *Brain Res Cogn Brain Res, 20*(3), 376-383.

Eriksson, J., Larsson, A., Ahlstrom, K. R., & Nyberg, L. (2007). Similar frontal and distinct posterior cortical regions mediate visual and auditory perceptual awareness. *Cereb Cortex, 17*(4), 760-765.

Eriksson, J., Larsson, A., & Nyberg, L. (2008). Item-specific training reduces prefrontal cortical involvement in perceptual awareness. *J Cogn Neurosci, 20*(10), 1777-1787.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Front Hum Neurosci, 4*, 215.

Ffytche, D. H., & Zeki, S. (2011). The primary visual cortex, and feedback to it, are not necessary for conscious vision. *Brain, 134*(Pt 1), 247-257.

Fisch, L., Privman, E., Ramot, M., Harel, M., Nir, Y., Kipervasser, S., et al. (2009). Neural "ignition": enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron, 64*(4), 562-574.

Friedman-Hill, S. R., Robertson, L. C., & Treisman, A. (1995). Parietal contributions to visual feature binding: evidence from a patient with bilateral lesions. *Science, 269*(5225), 853-855.

Fries, P., Roelfsema, P. R., Engel, A. K., Konig, P., & Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proc Natl Acad Sci U S A, 94*(23), 12699-12704.

Friston, K. (2005). A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci, 360*(1456), 815-836.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat Rev Neurosci, 11*(2), 127-138.

Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., et al. (2008). Multiple sparse priors for the M/EEG inverse problem. *Neuroimage, 39*(3), 1104-1120.

Gaillard, R., Dehaene, S., Adam, C., Clemenceau, S., Hasboun, D., Baulac, M., et al. (2009). Converging intracranial markers of conscious access. *PLoS Biol, 7*(3), e61.

Gamond, L., George, N., Lemarechal, J. D., Hugueville, L., Adam, C., & Tallon-Baudry, C. (2011). Early influence of prior experience on face perception. *Neuroimage, 54*(2), 1415-1426.

Gazzaley, A., Rissman, J., Cooney, J., Rutman, A., Seibert, T., Clapp, W., et al. (2007). Functional interactions between prefrontal and visual association cortex contribute to top-down modulation of visual processing. *Cereb Cortex, 17 Suppl 1*, i125-135.

Genc, E., Bergmann, J., Singer, W., & Kohler, A. (2011). Interhemispheric connections shape subjective experience of bistable motion. *Curr Biol, 21*(17), 1494-1499.

Ghuman, A. S., Bar, M., Dobbins, I. G., & Schnyer, D. M. (2008). The effects of priming on frontal-temporal communication. *Proc Natl Acad Sci U S A, 105*(24), 8405-8409.

Goldberg, II, Harel, M., & Malach, R. (2006). When the brain loses its self: prefrontal inactivation during sensorimotor processing. *Neuron, 50*(2), 329-339.

Gotts, S. J., Chow, C. C., & Martin, A. (2012). Repetition Priming and Repetition Suppression: A Case for Enhanced Efficiency Through Neural Synchronization. *Cogn Neurosci, 3*(3-4), 227-237.

Gray, C. M., & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc Natl Acad Sci U S A, 86*(5), 1698-1702.

Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nat Neurosci, 3*(8), 837-843.

Gross, J., Kujala, J., Hamalainen, M., Timmermann, L., Schnitzler, A., & Salmelin, R. (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proc Natl Acad Sci U S A, 98*(2), 694-699.

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., et al. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proc Natl Acad Sci U S A, 101*(35), 13050-13055.

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychol Rev, 87*(1), 1-51.

Grutzner, C., Uhlhaas, P. J., Genc, E., Kohler, A., Singer, W., & Wibral, M. (2010). Neuroelectromagnetic correlates of perceptual closure processes. *J Neurosci, 30*(24), 8342-8352.

Grutzner, C., Wibral, M., Sun, L., Rivolta, D., Singer, W., Maurer, K., et al. (2013). Deficits in high- (>60 Hz) gamma-band oscillations during visual processing in schizophrenia. *Front Hum Neurosci, 7*, 88.

Hamalainen, M. S., & Sarvas, J. (1987). Feasibility of the homogeneous head model in the interpretation of neuromagnetic fields. *Phys Med Biol, 32*(1), 91-97.

Hamalainen, M. S., & Sarvas, J. (1989). Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Trans Biomed Eng, 36*(2), 165-171.

Haushofer, J., Livingstone, M. S., & Kanwisher, N. (2008). Multivariate patterns in object-selective cortex dissociate perceptual and physical shape similarity. *PLoS Biol, 6*(7), e187.

Hawkins, J. (2004). *On Intelligence*. New York: Owl Books.

Haynes, J. D. (2009). Decoding visual consciousness from human brain signals. *Trends Cogn Sci, 13*(5), 194-202.

Haynes, J. D., Driver, J., & Rees, G. (2005). Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. *Neuron, 46*(5), 811-821.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annu Rev Psychol, 50*, 243-271.

Henson, R. N., Mattout, J., Phillips, C., & Friston, K. J. (2009). Selecting forward models for MEG source-reconstruction using model-evidence. *Neuroimage,* *46*(1), 168-176.

Hesselmann, G., Hebart, M., & Malach, R. (2011). Differential BOLD activity associated with subjective and objective reports during "blindsight" in normal observers. *J Neurosci, 31*(36), 12936-12944.

Hipp, J. F., Engel, A. K., & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron, 69*(2), 387-396.

Hoogenboom, N., Schoffelen, J. M., Oostenveld, R., Parkes, L. M., & Fries, P. (2006). Localizing human visual gamma-band activity in frequency, time and space. *Neuroimage, 29*(3), 764-773.

Horikawa, T., Tamaki, M., Miyawaki, Y., & Kamitani, Y. (2013). Neural decoding of visual imagery during sleep. *Science, 340*(6132), 639-642.

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

Kanai, R., Bahrami, B., & Rees, G. (2010). Human parietal cortex structure predicts individual differences in perceptual rivalry. *Curr Biol, 20*(18), 1626-1630.

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nat Rev Neurosci, 12*(4), 231-242.

Kayser, C., Salazar, R. F., & Konig, P. (2003). Responses to natural scenes in cat V1. *J Neurophysiol, 90*(3), 1910-1920.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annu Rev Psychol, 55*, 271-304.

Kim, C. Y., & Blake, R. (2005). Psychophysical magic: rendering the visible 'invisible'. *Trends Cogn Sci, 9*(8), 381-388.

Kleinschmidt, A., Buchel, C., Hutton, C., Friston, K. J., & Frackowiak, R. S. (2002). The neural structures expressing perceptual hysteresis in visual letter recognition. *Neuron, 34*(4), 659-666.

Kleinschmidt, A., Buchel, C., Zeki, S., & Frackowiak, R. S. (1998). Human brain activity during spontaneously reversing perception of ambiguous figures. *Proc Biol Sci, 265*(1413), 2427-2433.

Koch, C. (2004). *The quest for consciousness*. Englewood: Roberts & Co.

Koivisto, M., Kainulainen, P., & Revonsuo, A. (2009). The relationship between awareness and attention: evidence from ERP responses. *Neuropsychologia, 47*(13), 2891-2899.

Koivisto, M., Lahteenmaki, M., Sorensen, T. A., Vangkilde, S., Overgaard, M., & Revonsuo, A. (2008). The earliest electrophysiological correlate of visual awareness? *Brain Cogn, 66*(1), 91-103.

Koivisto, M., & Revonsuo, A. (2003). An ERP study of change detection, change blindness, and visual awareness. *Psychophysiology, 40*(3), 423-429.

Koivisto, M., & Revonsuo, A. (2007). Electrophysiological correlates of visual consciousness and selective attention. *Neuroreport, 18*(8), 753-756.

Koivisto, M., & Revonsuo, A. (2008). Comparison of event-related potentials in attentional blink and repetition blindness. *Brain Res, 1189*, 115-126.

Koivisto, M., & Revonsuo, A. (2010). Event-related brain potential correlates of visual awareness. *Neurosci Biobehav Rev, 34*(6), 922-934.

Kok, P., Jehee, J. F., & de Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron, 75*(2), 265-270.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philos Trans R Soc Lond B Biol Sci, 362*(1481), 857-875.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci, 12*(5), 535-540.

Lachaux, J. P., Axmacher, N., Mormann, F., Halgren, E., & Crone, N. E. (2012). High-frequency neural activity and human cognition: past, present and possible future of intracranial EEG research. *Prog Neurobiol, 98*(3), 279-301.

Lachaux, J. P., George, N., Tallon-Baudry, C., Martinerie, J., Hugueville, L., Minotti, L., et al. (2005). The many faces of the gamma band response to complex visual stimuli. *Neuroimage, 25*(2), 491-501.

Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Hum Brain Mapp, 8*(4), 194-208.

Lamme, V. A. (2001). Blindsight: the role of feedforward and feedback corticocortical connections. *Acta Psychol (Amst), 107*(1-3), 209-228.

Lamme, V. A. (2006). Towards a true neural stance on consciousness. *Trends Cogn Sci, 10*(11), 494-501.

Lamme, V. A., Super, H., Landman, R., Roelfsema, P. R., & Spekreijse, H. (2000). The role of primary visual cortex (V1) in visual awareness. *Vision Res, 40*(10-12), 1507-1521.

Lamy, D., Salti, M., & Bar-Haim, Y. (2009). Neural correlates of subjective awareness and unconscious processing: an ERP study. *J Cogn Neurosci, 21*(7), 1435-1446.

Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends Cogn Sci, 15*(8), 365-373.

Lau, H. C. (2008). A higher order Bayesian decision theory of consciousness. *Prog Brain Res, 168*, 35-48.

Lau, H. C., & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proc Natl Acad Sci U S A, 103*(49), 18763-18768.

Lau, H. C., & Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *J Neurosci, 27*(21), 5805-5811.

Laureys, S., & Tononi, G. (Eds.). (2010). *The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology*. Amsterdam: Elservier.

Lennie, P. (2003). The cost of cortical computation. *Curr Biol, 13*(6), 493-497.

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

Linkenkaer-Hansen, K., Nikulin, V. V., Palva, S., Ilmoniemi, R. J., & Palva, J. M. (2004). Prestimulus oscillations enhance psychophysical performance in humans. *J Neurosci, 24*(45), 10186-10190.

Llinás, R. (2001). *I of the Vortex: From Neurons to Self*. Cambridge, MA: MIT Press.

Llinas, R., Ribary, U., Contreras, D., & Pedroarena, C. (1998). The neuronal basis for consciousness. *Philos Trans R Soc Lond B Biol Sci, 353*(1377), 1841-1849.

Logothetis, N.K. (1998). Single units and conscious vision. *Philosophical Transactions of the Royal Society of London. B.,* 353, 1801–1818.

Logothetis, N. K., & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science, 245*(4919), 761-763.

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

Malach, R. (2007). The measurement problem in human consciousness research. *Behav Brain Sci, 30*, 516-517.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods, 164*(1), 177-190.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. *J Neurosci, 29*(9), 2725-2732.

McGinn, C. (1989). Can We Solve the Mind Body Problem. *Mind, 98*(391), 349-366.

Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., & Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *J Neurosci, 27*(11), 2858-2865.

Melloni, L., Schwiedrzik, C. M., Muller, N., Rodriguez, E., & Singer, W. (2011). Expectations change the signatures and timing of electrophysiological correlates of perceptual awareness. *J Neurosci, 31*(4), 1386-1396.

Melloni, L., & Singer, W. (2010). Distinct characteristics of conscious experience are met by large scale neuronal synchronization. In E. Perry, D. Collerton, F. E. N. LeBeau & H. Ashton (Eds.), *New horizons in the neuroscience of consciousness* (pp. 17-28). Amsterdam: John Benjamins.

Miller, B. T., Vytlacil, J., Fegen, D., Pradhan, S., & D'Esposito, M. (2011). The prefrontal cortex modulates category selectivity in human extrastriate cortex. *J Cogn Neurosci, 23*(1), 1-10.

Miller, K. J. (2010). Broadband spectral change: evidence for a macroscale correlate of population firing rate? *J Neurosci, 30*(19), 6477-6479.

Monto, S., Palva, S., Voipio, J., & Palva, J. M. (2008). Very slow EEG fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. *J Neurosci, 28*(33), 8268-8272.

Muller, N. G., Strumpf, H., Scholz, M., Baier, B., & Melloni, L. (2013). Repetition suppression versus enhancement--it's quantity that matters. *Cereb Cortex, 23*(2), 315-322.

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol Cybern, 66*(3), 241-251.

Mur, M., Meys, M., Bodurka, J., Goebel, R., Bandettini, P. A., & Kriegeskorte, N. (2013). Human Object-Similarity Judgments Reflect and Transcend the Primate-IT Object Representation. *Front Psychol, 4*, 128.

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr, 20*(4), 249-264.

Nishimoto, S., Vu, A. T., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. L. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. *Curr Biol, 21*(19), 1641-1646.

Nolte, G. (2003). The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys Med Biol, 48*(22), 3637-3652.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci, 2011*, 156869.

Pachella, R. G. (1975). The effect of set on the tachistoscopic recognition of pictures. In P.M.A. Rabbit & S. Domic (Eds.), *Attention & Performance V*. London: Academic Press.

Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science, 292*(5516), 510-512.

Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *J Neurophysiol, 93*(1), 603-608.

Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nat Rev Neurosci, 8*(8), 636-648.

Pfurtscheller, G., & Cooper, R. (1975). Frequency dependence of the transmission of the EEG from cortex to scalp. *Electroencephalogr Clin Neurophysiol, 38*(1), 93-96.

Pins, D., & Ffytche, D. (2003). The neural correlates of conscious vision. *Cereb Cortex, 13*(5), 461-474.

Pitts, M. A., & Britz, J. (2011). Insights from intermittent binocular rivalry and EEG. *Front Hum Neurosci, 5*, 107.

Pitts, M. A., Martinez, A., & Hillyard, S. A. (2010). When and where is binocular rivalry resolved in the visual cortex? *J Vis, 10*(14).

Pollen, D. A. (1999). On the neural correlates of visual perception. *Cereb Cortex, 9*(1), 4-19.

Postle, B. R. (2009). The hippocampus, memory, and consciousness. In S. Laureys & G. Tononi (Eds.), *The neurology of consciousness: Cognitive neuroscience and neuropathology* (pp. 326-338). London: Academic Press.

Quiroga, R. Q., Mukamel, R., Isham, E. A., Malach, R., & Fried, I. (2008). Human single-neuron responses at the threshold of conscious recognition. *Proc Natl Acad Sci U S A, 105*(9), 3599-3604.

Rainer, G., Lee, H., & Logothetis, N. K. (2004). The effect of learning on the function of monkey extrastriate visual cortex. *PLoS Biol, 2*(2), E44.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci, 2*(1), 79-87.

Ray, S., & Maunsell, J. H. (2011). Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol, 9*(4), e1000610.

Rees, G. (2007). Neural correlates of the contents of visual awareness in humans. *Philos Trans R Soc Lond B Biol Sci, 362*(1481), 877-886.

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nat Rev Neurosci, 3*(4), 261-270.

Robertson, L. C. (2003). Binding, spatial attention and perceptual awareness. *Nat Rev Neurosci, 4*(2), 93-102.

Romei, V., Brodbeck, V., Michel, C., Amedi, A., Pascual-Leone, A., & Thut, G. (2008). Spontaneous fluctuations in posterior alpha-band EEG activity reflect variability in excitability of human visual areas. *Cereb Cortex, 18*(9), 2010-2018.

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cogn Neurosci, 1*(3), 165-175.

Rousselet, G. A., & Pernet, C. R. (2011). Quantifying the Time Course of Visual Object Processing Using ERPs: It's Time to Up the Game. *Front Psychol, 2*, 107.

Schrödinger, E. (1958). *Mind and Matter*. Cambridge: Cambridge University press.

Schwiedrzik, C. M., Singer, W., & Melloni, L. (2011). Subjective and objective learning effects dissociate in space and in time. *Proc Natl Acad Sci U S A, 108*(11), 4506-4511.

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nat Neurosci, 8*(10), 1391-1400.

Seth, A. K. (2009). Functions of consciousness. In W. P. Banks (Ed.), *Encyclopedia of Consciousness* (Vol. 1, pp. 279-293). Amsterdam: Elservier.

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends Cogn Sci, 12*(8), 314-321.

Sheinberg, D. L., & Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proc Natl Acad Sci U S A, 94*(7), 3408-3413.

Silvanto, J., & Rees, G. (2011). What does Neural Plasticity Tell us about Role of

Primary Visual Cortex (V1) in Visual Awareness? *Front Psychol, 2*, 6.

Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annu Rev Physiol*, 55, 349-374.

Singer, W. (1998). Consciousness and the structure of neuronal representations. *Philos Trans R Soc Lond B Biol Sci, 353*(1377), 1829-1840.

Singer, W. (in press). The brain, a complex self-organizing system with non-linear dynamics? *Trends Cogn Sci.*

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annu Rev Neurosci, 18*, 555-586.

Srinivasan, R., Russell, D. P., Edelman, G. M., & Tononi, G. (1999). Increased synchronization of neuromagnetic responses during conscious perception. *J Neurosci, 19*(13), 5435-5448.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends Cogn Sci, 13*(9), 403-409.

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nat Neurosci, 11*(9), 1004-1006.

Sun, J., & Perona, P. (1998). Where is the sun? *Nat Neurosci, 1*(3), 183-184.

Sun, L., Castellanos, N., Grutzner, C., Koethe, D., Rivolta, D., Wibral, M., et al. (2012). Evidence for dysregulated high-frequency oscillations during sensory processing in medication-naive, first episode schizophrenia. *Schizophr Res*.

Super, H., Spekreijse, H., & Lamme, V. A. (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nat Neurosci, 4*(3), 304-310.

Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci, 26*(37), 9494-9502.

Todorovic, A., van Ede, F., Maris, E., & de Lange, F. P. (2011). Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. *J Neurosci, 31*(25), 9118-9123.

Tong, F. (2003). Primary visual cortex and visual awareness. *Nat Rev Neurosci, 4*(3), 219-229.

Tononi, G., & Koch, C. (2008). The neural correlates of consciousness: an update. *Ann N Y Acad Sci, 1124*, 239-261.

Tse, P. U., Martinez-Conde, S., Schlegel, A. A., & 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. *Proc Natl Acad Sci U S A, 102*(47), 17178-17183.

Ullman, S. (1995). Sequence seeking and counter streams: a computational model for bidirectional information flow in the visual cortex. *Cereb Cortex, 5*(1), 1-11.

van Dijk, H., Schoffelen, J. M., Oostenveld, R., & Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J Neurosci, 28*(8), 1816-1823.

van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *J Neurosci, 28*(32), 8053-8062.

Van Veen, B. D., van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng, 44*(9), 867-880.

Varela, F. (1996). Neurophenomenology: a methodological remedy for the hard problem. *J Consc Studies,* 3(4), 330-349.

Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nat Rev Neurosci, 2*(4), 229-239.

Vidal, J. R., Ossandon, T., Jerbi, K., Dalal, S. S., Minotti, L., Ryvlin, P., et al. (2010). Category-Specific Visual Responses: An Intracranial Study Comparing Gamma, Beta, Alpha, and ERP Response Selectivity. *Front Hum Neurosci, 4*, 195.

Wacongne, C., Changeux, J. P., & Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *J Neurosci, 32*(11), 3665-3678.

Whitaker, D., & McGraw, P. V. (2000). Long-term visual experience recalibrates human orientation perception. *Nat Neurosci, 3*(1), 13.

Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J Neurosci, 20*(6), RC63.

Wyart, V., & Tallon-Baudry, C. (2008). Neural dissociation between visual awareness and spatial attention. *J Neurosci, 28*(10), 2667-2679.

Wyart, V., & Tallon-Baudry, C. (2009). How ongoing fluctuations in human visual cortex predict perceptual awareness: baseline shift versus decision bias. *J Neurosci, 29*(27), 8715-8725.

Zaretskaya, N., Thielscher, A., Logothetis, N. K., & Bartels, A. (2010). Disrupting parietal function prolongs dominance durations in binocular rivalry. *Curr Biol, 20*(23), 2106-2111.

Zeki, S. (2001). Localization and globalization in conscious vision. *Annu Rev Neurosci, 24*, 57-86.

Erklärungen

ERKLÄRUNG

Ich erkläre hiermit, dass ich mich bisher keiner Doktorprüfung unterzogen habe.

Frankfurt am Main, den

EIDESSTATTLICHE VERSICHERUNG

Ich erkläre hiermit an Eides Statt, daß ich die vorgelegte Dissertation über *“Distilling the neural correlates of conscious perception”* selbständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, daß alle Entlehnungen aus anderen Schriften mit Angabe der betreffenden Schrift gekennzeichnet sind.
Ich versichere, nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben

Frankfurt am Main, den